Nonlinear dynamics & stochastic processes in cybersecurity applications

Pierce Ryan

University College Cork
Coláiste na hOllscoile Corcaigh

School of Mathematical Sciences
Discipline of Applied Mathematics

Thesis submitted for the degree of
Doctor of Philosophy

Awarded 19^{th} June 2023

Head of School:	Dr. Kevin Hayes
Head of Discipline:	Prof. Sebastian Wieczorek
Supervisors:	Dr. Andreas Amann
	Dr. Sorcha Healy

Contents

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism.

Pierce Ryan

For my grandfathers.

Acknowledgements

There are many people without whom I would not have finished this thesis, and many more without whom I would not have finished this thesis with my sanity largely still intact.

To start with, I thank my supervisors and mentors, Andreas, Sorcha and Niall, for their guidance. Five years ago, two old friends from the Tyndall Institute randomly ran into each other outside a sports centre. Sorcha had moved towards industry, becoming a data scientist at the cybersecurity provider McAfee. Andreas had moved towards academia, becoming a lecturer at University College Cork. Sorcha brought up a hybrid industry/academia funded postgraduate programme she wanted to pursue, and asked whether Andreas had any ideas for projects. A few months later, I was writing a proposal for the Irish Research Council Employment-Based Postgraduate Programme with Sorcha and Andreas. At McAfee, Sorcha and I worked side by side for over a year before Covid cleared us out of the office, while my weekly meetings with Andreas at UCC would range from an hour during the winter to 'until we run out of things to talk about' during the summer. When Sorcha's career took her from McAfee to Microsoft in the final year of the project, Niall stepped in to be my mentor at McAfee. Without their patience, guidance and humour, I wouldn't have gotten this far. I'd also like to thank my monitor, Kathleen, for her help and the use of her extensive library of statistics textbooks.

I thank my family for their support; my parents Eoin and Sheelagh, my sister Aisling and her husband Aidan. Throughout the COVID-19 pandemic, while we bubbled together in Baltimore, they ignored the sound of my head banging off my desk with tact and discretion. Sheelagh, as she coordinated local HSE efforts against the pandemic, inspired me with her unwavering patience. Eoin provided endless supplies of wisdom and outstanding food, while Aidan could always be relied on for a laugh. Aisling, who like me, followed Sheelagh in pursuing a PhD, even shared an office with me, which we christened the UCC Southwest-est Campus.

I thank my colleagues at McAfee for their acceptance. When I first began working at McAfee at the start of this project, there were plenty of people curious as to what my job actually was. I was one of them. Still, they welcomed me with open arms. Sorcha, Gerard, Niamh, Jill, Niall and I made up the Applied Data Science team. Together we weathered five changes in management, and countless sprint planning meetings. For a while, I practically had Gerard on speed dial to answer my 'data science for dummies' questions. Other standouts include Paul, Didier, Rachel, Jon, Crystal, and last but certainly not least, Brian, my final and longest serving/suffering manager.

Finally, I'd like to thank my friends. My fellow PhD students Andrew, Chris, Eoin, Rory and Conor, who brought that windowless office on the second floor to life, and my friends Ciara, Matt and Éadaoin, who made sure I went to the pub on a regular basis no matter how much or how little I protested.

This research was supported by the Irish Research Council and McAfee LLC through the Irish Research Council Employment-Based Postgraduate Programme (grant number EBPPG/2018/269).

Abstract

The Internet is an extremely complex system which has a significant impact on the world we live in. In this thesis, we formalise Internet-based problems as mathematical models to better understand their dynamics. Modelling these problems requires dynamical features such as time delay, periodic forcing, switching and stochasticity. We study several dynamical systems which employ a combination of these features from Internet applications, including targeted ransomware, data networks, and signal processing. We also study a climate science system which shares features with the signal processing system and exhibits similar dynamics. Stochasticity is found to be critical in the modelling of the negotiations involved in targeted ransomware, while time delay is a crucial feature in the modelling of data networks. The signal processing and climate science systems give rise to extremely rich dynamics, which we are able to study analytically due to the presence of switching. This yields further insights into related smooth systems.

Introduction

The Internet is the most complex man-made system in the world, with its dynamics driven by the activities of both man and machine. The basic infrastructure, which was established in the 1970s with ARPANET [1], has grown in capacity, complexity and utility through the iterations of Web 1.0, Web 2.0, and Web 3.0 [2, 3]. The Internet has expanded to support the broadest repository of information, the largest platform for international trade, and the greatest conduit of human interaction. But it is much more than just a system. The Internet is practically a world unto itself, a mirror of our own physical world which functions according to its own logic.

In the physical world, your social network is constrained by physical distance and language barriers. On the Internet, physical distance translates to a fraction of a second of latency and language barriers are broken down by instant translation, enabling your social circle to span the globe. At the same time, the Internet is subject to effects which do not exist in the real world. As you walk down the street, shops don't reshuffle to move those that paid higher advertising fees right in front of you. Your peripheral vision is not filled by advertisements for products you talked about yesterday. There are no 'bots' pasting fake reviews on the windows of businesses, and no algorithms subtly guiding you to places where you'll spend the most money.

The Internet is a world separate to our own, and yet they are irrevocably intertwined. The Internet is affected by events in the physical world. The basic infrastructure runs on physical hardware, chunks of doped silicon connected by wire and fibre optic, any piece of which may fail at any moment, taking fragments of the Internet with it. Many actions taken in the physical world, malicious or benign, start with, or are facilitated by, communication over the Internet. Online threats such as *ransomware* cause real-world harm through the leaking of private data, the extortion of ransoms, and the proliferation of criminal organisations financed by these crimes [4, 5, 6]. Even people who do not engage with the Internet are subject to its influence, and so understanding the dynamics of the digital world is crucial to understanding our own.

Due to the scale and complexity of the Internet, there is a truly vast amount to be understood. The Internet is a patchwork of systems and subsystems which fulfil a list of functions too long to fit in any thesis. As a result of this diverse array of functions, the Internet exhibits a broad range of dynamics across multiple scales.

At the large scale, we observe network phenomena and the collective dynamics of millions of devices connected by the Internet. When we focus on the connections between personal profiles on social media websites such as Facebook, we observe social networks. These networks have been found to exhibit small-world and scale-free properties [7, 8].

Less visible than social media websites are the background systems which silently propagate information between servers and personal computers. These *data networks* make use of various techniques to increase efficiency and minimise latency to the point that people forget these systems are even there. One of the most ubiquitous of these techniques is *caching*, where data sent to a computer is stored temporarily in case it is needed again in quick succession. Dynamically speaking, caching has two effects. The first is that it induces hysteresis, explicit dependence of a system on its history. The second is that the cost of using data from other computers changes discontinuously depending on whether or not the data is cached. Both hysteresis and discontinuous response give rise to rich dynamical behaviour, such as multistability [9, 10, 11, 12, 13] and border-collision bifurcations [14, 15, 16, 17, 18], respectively.

At the small scale, we can consider the activity of a single computer interacting with a data network. Here, the effects of the physical world become far more apparent. From an Internet perspective, the activity of humans can appear random; while some of their actions can be predicted based on their previous activity on the Internet, some actions are prompted by events in the real world. From the perspective of the Internet, this introduces stochasticity into their dynamical behaviour. In addition, while machines don't sleep, most people do, usually according to a regular sleep cycle. Therefore we can expect that any dynamical process which requires human interaction to feature both periodicity and stochasticity. Again, both periodicity [19, 20, 21, 22] and stochasticity [23, 24, 25] are known to induce dynamical behaviour, including synchronization and resonance.

The core principle of this thesis is to formalise small-scale problems on the Internet in mathematical terms to gain a deeper understanding of these problems and establish a foundation for further research at all scales. Specifically, we will study the developments in ransomware which have led to the current wave of targeted ransomware, and consider how dynamical features such as time delay, periodic forcing, stochasticity and switching can appear in data networks. We will also study dynamical systems from climate science and signal processing which share dynamical features with those seen in Internet systems.

Thesis structure

Chapter 1 is largely reproduced from P. Ryan, A. Keane and A. Amann, *Border-collision bifurcations in a driven time-delay system*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(2):023121, 2020. [26]. We study a piecewise-smooth time-delay system with periodic forcing derived from a phenomenological climate model [27] which demonstrates an extremely complex resonance structure. In particular, our system is piecewise-constant, as both the time-delayed feedback and the periodic forcing are switches, flipping back and forth between two values. This yields an extremely simple system which still demonstrates much of the same dynamics as the original model. We develop a symbolic representation for the non-smooth dynamics of the system, and explore the bifurcations of the system through bifurcations of Poincaré maps derived from the symbolic dynamics. These include both traditional bifurcations and border-collision bifurcations. This symbolic representation enables us to analytically calculate solutions and bifurcation curves while studying phenomena that were previously studied only numerically. This work has been found to be of use in more recent research as the study of non-smooth systems continues to develop [28, 29].

Chapter 2 forms the basis of a paper that I am co-authoring with Dr. Lucas Illing and Dr. Andreas Amann. We study a piecewise-linear second-order delay differential equation which is representative of feedback systems with relays that actuate after a fixed time delay. This system demonstrates a complex structure of classical and border-collision bifurcations. In contrast to the system considered in the previous chapter, this system is unforced and piecewise-linear between switching events. We extend the symbolic representation developed in the previous chapter to study this system through bifurcations of Poincaré maps, enabling us to analytically derive border-collision bifurcations, while classical bifurcations are characterised numerically.

Chapter 3 is largely reproduced from P. Ryan, J. Fokker, S. Healy and A. Amann, *Dynamics of targeted ransomware negotiation*, IEEE Access, 10:3283632844, 2022. [30]. Using game theory, we construct a stochastic model of negotiations between operators of targeted ransomware and their targets to better understand how to respond to ransomware attacks. In particular, our model considers the investments that a cybercriminal must make in order to conduct a successful targeted ransomware attack. We demonstrate how imperfect information is a crucial feature for replicating observed real-world behaviour. Furthermore, we present optimal strategies for both the perpetrator and the target, and demonstrate how imperfect information results in a non-trivial optimal strategy for the cybercriminal. This work has been found be of use in more recent research as the mathematical analysis of cybersecurity becomes more widely studied [31, 32, 33, 34, 35].

Chapter 4 presents unpublished material in which we consider the problem of estimating and improving the efficiency of the data networks commonly used by cybersecurity providers to deliver information to their customers. We develop a stochastic model of 'Time-To-Live' data caching systems commonly used in cybersecurity with reference to real-world data derived from a cybersecurity data network. The Time-To-Live mechanism induces a time-delay effect in the system. We use our model to estimate metrics for caching system activity based on observable data, and to estimate the optimal Time-To-Live. The deterministic limit of the stochastic model gives rise to a circle map with periodic solutions that are analysed in terms of a symbolic representation.

Chapter 1

Border-collision bifurcations in a driven time-delay system

The material in this chapter is largely reproduced from P. Ryan, A. Keane and A. Amann, *Border-collision bifurcations in a driven time-delay system*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(2):023121, 2020. [26]. My contribution to this paper consisted of exploring the parameter space and using the symbolic representation, which was proposed by Andreas, to calculate the existence and stability of solutions to study the bifurcations of the system.

1.1 Abstract

We show that a simple piecewise-linear system with time delay and periodic forcing gives rise to a rich bifurcation structure of torus bifurcations and Arnold tongues, as well as multistability across a significant portion of the parameter space. The simplicity of our model enables us to study the dynamical features analytically. Specifically, these features are explained in terms of border collision bifurcations of an associated Poincaré map. Given that time delay and periodic forcing are common ingredients in mathematical models, this analysis provides widely applicable insight.

1.2 Summary

Both time delay dynamical systems, and periodically driven dynamical systems have been thoroughly studied in the literature. This can be attributed to their great relevance to real-world problems. Time delay systems arise naturally in physical, biological or climate models due to finite propagation speed; periodic drive is ubiquitous in engineering applications and is known to generate complex resonance phenomena. However, systems that combine these two properties have received much less attention, despite being relevant in many real-world applications. In this paper, we study a simple piecewise-linear system with both time delay and periodic forcing, which exhibits interesting dynamical features as a nontrivial consequence of this combination. These features include multistabilities, Arnold tongues, and torus bifurcations. Since the system is piecewise linear and contains only two parameters, many phenomena can be interpreted through an analytically derived piecewise-smooth Poincaré map and an analysis of the associated border collision bifurcations. The analysis explains the origin of similar phenomena which has previously been observed numerically in more complicated related systems.

1.3 Introduction

Periodically driven systems appear in many real-world applications. Examples include optical injection in laser systems [21], vibration-driven energy harvesting devices [36], injection-locked frequency dividers in electronics [37], or seasonal forcing in climate systems [22]. They often give rise to interesting resonance behaviour in damped oscillators [38] and complex synchronization patterns in self-sustained oscillators [19, 20].

Similarly, time delay systems also arise in many experimental systems, for example in optics [39, 40], electronics [9], neuro-science [41], or climate systems [42], and also play an important role in chaos control, for example, through the use of time delayed feedback control [43]. From a mathematical point of view, time delay often leads to a formally infinite-dimensional phase space [44], which considerably complicates the analysis, but allows for a rich variety of phenomena.

The combination of external forcing and time delay has been studied, for example, in the context of the Duffing oscillator [45], the van der Pol oscillator [46] and more recently in the context of climate systems [27, 13]. However, a general understanding of this class of dynamical systems is not yet available. Here, our objective is to study the fundamental features of an elementary system with time delay and periodic forcing to obtain a broader insight into what phenomena are expected to arise as a consequence of this combination.

Let us consider a simple driven time-delay dynamical system introduced by Ghil et al. [27] as a model for a climate phenomenon known as the El Niño Southern Oscillation. The system of a real variable $x \in \mathbb{R}$ is defined by

$$\dot{x}(t) = -\tanh\left[\kappa x\left(t-\tau\right)\right] + b\sin\left(2\pi t\right) \tag{1.1}$$

$$x(t) \in C\left(\left[-\tau, 0\right]\right) \tag{1.2}$$

where $b \geq 0$ is the magnitude of the periodic forcing, $\tau > 0$ is the time delay of the delayed feedback, and $\kappa > 0$ is the linear slope of the delayed feedback at the origin. A solution of the system is a trajectory $x(t) \in \mathbb{R}$, $t \in \mathbb{R}$. A consequence of the reliance of the delayed feedback on a continuous function $x\left(t\right)$ over an interval $[-\tau, 0]$ is that the system has an infinite-dimensional phase-space. Another key feature of this system is that it has the symmetry $x\left(t\right) \rightarrow -x\left(t+\frac{1}{2}\right)$. This model has been studied extensively by Keane et al. [13, 47]. Numerical analysis of this system demonstrated an extremely complex resonance structure [13]. Furthermore, the autonomous system ($b = 0$) has been studied analytically [48, 49]. In this case the trivial solution $x \equiv 0$ is only stable for $\tau < \frac{\pi}{2\kappa}$. At $\tau = \frac{\pi}{2\kappa}$, it becomes unstable, and a family of stable 4τ-periodic solutions is born.

In order to analyse the phenomena seen in this model further, let us consider a further simplification of the system by taking $\kappa \rightarrow \infty$. This has the effect of changing the delayed feedback term from $-\tanh\left[\kappa x\left(t-\tau\right)\right]$ to $-\mathrm{sgn}\left[x\left(t-\tau\right)\right]$. We also apply the signum function to the periodic forcing to obtain the dynamical system

$$\dot{x}(t) = -\mathrm{sgn}\left[x\left(t-\tau\right)\right] + b\ \mathrm{sgn}\left(\sin\left(2\pi t\right)\right) \tag{1.3}$$

$$x(t) \in C\left(\left[-\tau, 0\right]\right) \tag{1.4}$$

where $b \geq 0$ and $\tau > 0$. Critically, this simplification of the system preserves the symmetry $x\left(t\right) \rightarrow -x\left(t+\frac{1}{2}\right)$. The feedback term $-\ \mathrm{sgn}\left[x\left(t-\tau\right)\right]$ takes values in $\{1, 0, -1\}$. However, while the feedback can in principle be 0, this occurs only under highly specific conditions which are not considered here. The forcing term $b\ \mathrm{sgn}\left[\sin\left(2\pi t\right)\right]$ takes values in $\{b, 0, -b\}$. As the forcing is 0 only at discrete times, we say that the forcing is positive for $t \bmod 1 \in [0, 0.5)$, and negative for $t \bmod 1 \in [0.5, 1)$. We now develop our method for solving the system.

1.4 Numerics

1.4.1 Iterative map

In order to solve Eqs. (1.3,1.4), we note that $\dot{x}(t)$ can only take discrete values in $\{1+b, -1+b, 1-b, -1-b\}$; therefore this continuous system can be modelled exactly as a discrete time iterative map, or Poincaré map. The state of the system at time t is a tuple of variable length

$$S(t) = (x; z_0, z_1, ..., z_{n-1}) \tag{1.5}$$

where $x \in \mathbb{R}$ is the position at time t and $t - \tau < z_0 < z_1 < ... < z_{n-1} \leq t$ are *zero elements*, which are the times at which the trajectory passed through $x = 0$ in $(t-\tau, t]$. Let n be the number of zero elements in $S(t)$. Thus, Eq. (1.3) may be rewritten as

$$\dot{x}(t) = -\mathrm{sgn}\left[x\left(t\right)\right]\left(-1\right)^n + b\mathrm{sgn}[\sin\left(2\pi t\right)] \tag{1.6}$$

Representing the state of the system in this manner shows that while the dimension of the system is infinite in principle, the dynamics of the system are restricted to a finite, but variable, set of dimensions. This will have a significant impact on our analysis.

There are three key events that occur in a trajectory that affect $S(t)$:

1. A zero element is added. When $x(t)$ passes through $x = 0$, a zero element equal to t is appended to S, increasing n by 1; this does not immediately cause the feedback to change as the sign of $x(t)$ also changes, but it will affect $\dot{x}$ at time $t + \tau$.

2. A zero element is removed. When $t = z_0 + \tau$, z_0 is deleted from S which decreases n by 1, resulting in a sign change of the feedback.

3. The forcing changes when $t \bmod 0.5 = 0$. While this does not affect the state directly, it changes $\dot{x}$, affecting the evolution of the state.

Between consecutive events, $\dot{x}$ is constant. We construct our iterative map to move the system forward in time to the next event. The step size Δt for the iterative map is variable. One iteration of the map consists of calculating $\dot{x}(t)$, calculating the step size Δt, then calculating $S(t + \Delta t)$. This process is used to simulate the trajectory of the system from any given initial state $S(0)$ up to a maximum time T.

1.4.2 Sample solutions

We now present sample solutions which demonstrate some of the characteristic features of this system. Fig. 1.1 shows eight stable solutions found by simulating the system from different initial conditions and parameters (b, τ).

There are periodic and aperiodic solutions present in the system. A periodic solution with period P is a trajectory that follows a cycle such that $S(t+P) = S(t)$. An aperiodic solution is considered to be a solution with infinite period. We label solutions by the *characteristic ratio* $P : R$, where R is the number of times the trajectory crosses from $x < 0$ to $x > 0$ in one cycle. Note that the characteristic ratio $P{:}R$ is not the same as the frequency ratio $p{:}q$ used in some previous literature [13, 47], where $\frac{p}{q}$ is the rotation number. The frequency ratio is less useful in this system because we do not observe any $p : q$ solutions where $p \neq 1$. The characteristic ratio $P : R$ is a more descriptive measure of solutions in this system, and we will make extensive use of it in our analysis.

Fig. 1.1(a) is an example of the stable solution to the unforced system ($b = 0$). The change in the feedback occurs at a time $t+\tau$ after the trajectory passes through $x = 0$ at time t, resulting in a 4τ-periodic solution that is stable for $\tau > 0$. This is consistent with the analytic results found for the unsimplified system. We consider 4τ to be the natural period of the feedback, as the solution to the unforced system is 4τ-periodic. The characteristic ratio of this solution is $4\tau : 1$.

Fig. 1.1(b) and Fig. 1.1(c) show a $4{:}2$ solution and a $3{:}1$ solution, respectively, that are stable for the same parameters. This is an example of bistability, where

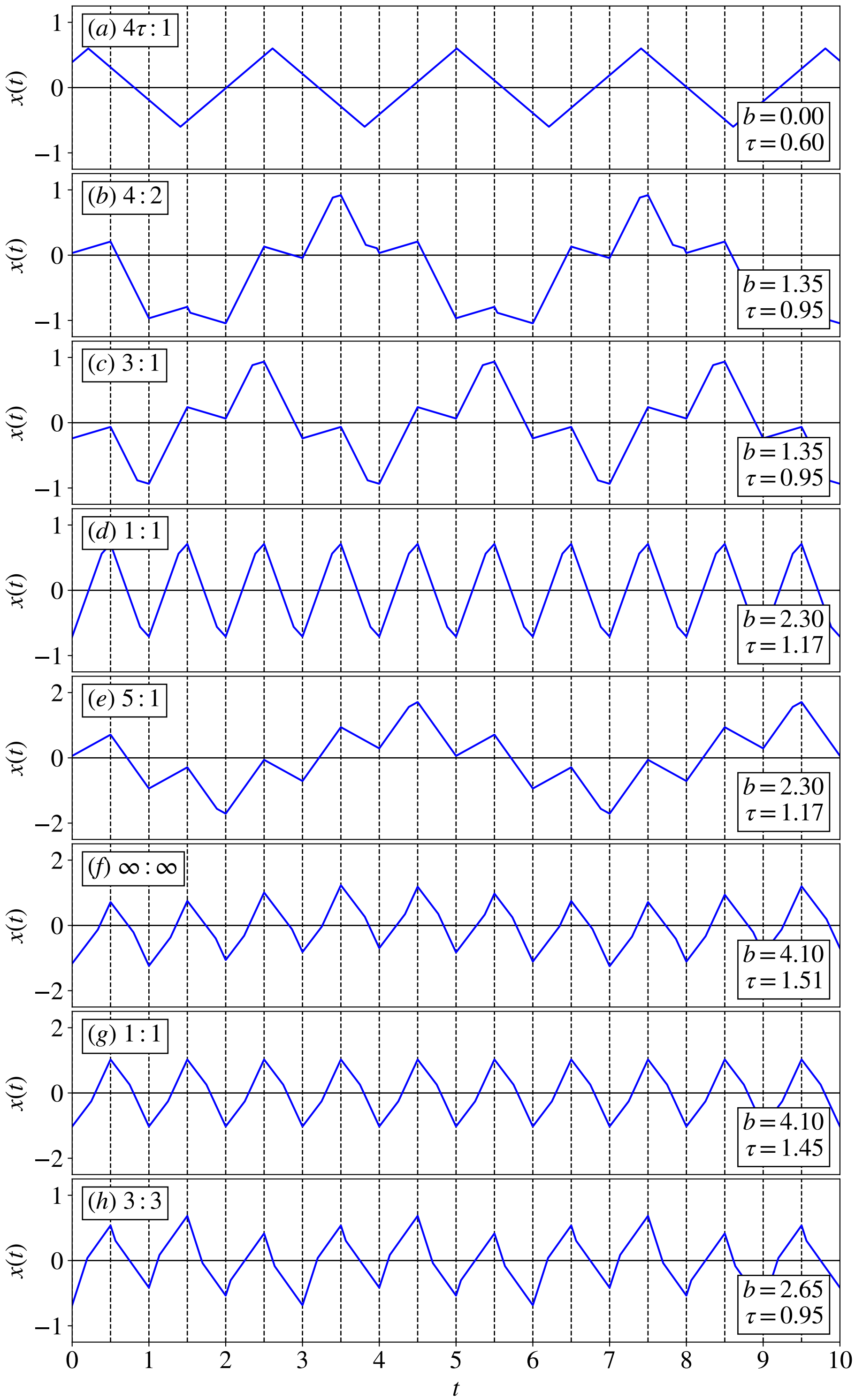

Figure 1.1: Sample solutions, excluding transients, obtained from simulating the system from different initial conditions and parameters (b, τ) shown in the bottom right corner of each plot. The ratio of the period of the solution to the number of crossings from $x < 0$ to $x > 0$ in one cycle is shown in the top left corner of each plot. The vertical dotted lines indicate times when the forcing changes.

there are two stable solutions for the same parameters; the solution to which the system converges depends on which solution's basin of attraction the initial conditions are in. A second example of bistability is seen in Fig. 1.1(d) and Fig. 1.1(e), which show a $1:1$ solution and a $5:1$ solution, respectively, that are stable for the same parameters.

The solution in Fig. 1.1(f) is assumed to be aperiodic, as the system does not converge to a periodic solution after running a simulation up to $T = 100000$. By comparing the aperiodic solution to the $1:1$ solution in Fig. 1.1(g), we may note that the aperiodic solution and the $1:1$ solution have the same number of $x = 0$ crossings per period of the forcing. However, the $x = 0$ crossings are evenly spaced in the $1:1$ solution, but are not in the aperiodic solution. This may be an indicator that the aperiodic solution is related to the $1:1$ solution in Fig. 1.1(g). A similar observation can be made for the $3:3$ solution in Fig. 1.1(h) and the $1:1$ solution in Fig. 1.1(d). In later sections of this chapter, we will investigate these relationships in detail.

1.4.3 Structure in the (b, τ) plane

Having seen some interesting features of the system, we move on to understanding the overall dynamics in the (b, τ) plane. Due to the bistability present in the system, simulating the system from arbitrary initial conditions across a (b, τ) mesh would produce an inconsistent picture, as we have no prior knowledge of the basins of attraction of bistable solutions. In order to circumvent this issue, for fixed τ, the solution is swept across a range of b by iteratively simulating the system, incrementing b slightly, then simulating again using the final state of the previous simulation as the initial state of the next one. This allows a stable solution to be followed until it loses stability or ceases to exist, at which point the system converges to a nearby stable solution. By taking multiple sweeps in b for a range of fixed τ values, we obtain the $P : R$ charts shown in Fig. 1.2. Fig. 1.2(a) shows the $P : R$ chart under an upward sweep in b, from left to right. Fig. 1.2(b) shows the $P : R$ chart under a downward sweep in b, from right to left. This figure demonstrates many striking features of the system which will be explored in more detail.

First, let us focus our attention on Fig. 1.2(a), where we sweep from $b = 0$ up to $b = 8$. We observe that there are regions in the (b, τ) plane in which particular $P : R$ solutions exist, such as the labelled $3:1$, $5:1$, $7:1$ and $9:1$ regions on the left side of the chart. These regions are sections of *Arnold tongues*. An Arnold tongue is a region of the (b, τ) plane, rooted on $b = 0$, within which the feedback and the forcing synchronise to produce a solution with period equal to a rational ratio of the forcing (and hence constant rotation number $\frac{p}{q}$). In piecewise-linear systems, an Arnold tongue can have shrinking points, at which the Arnold tongue has zero width. This phenomenon was first observed in the circle map [50], and analysed in detail in the

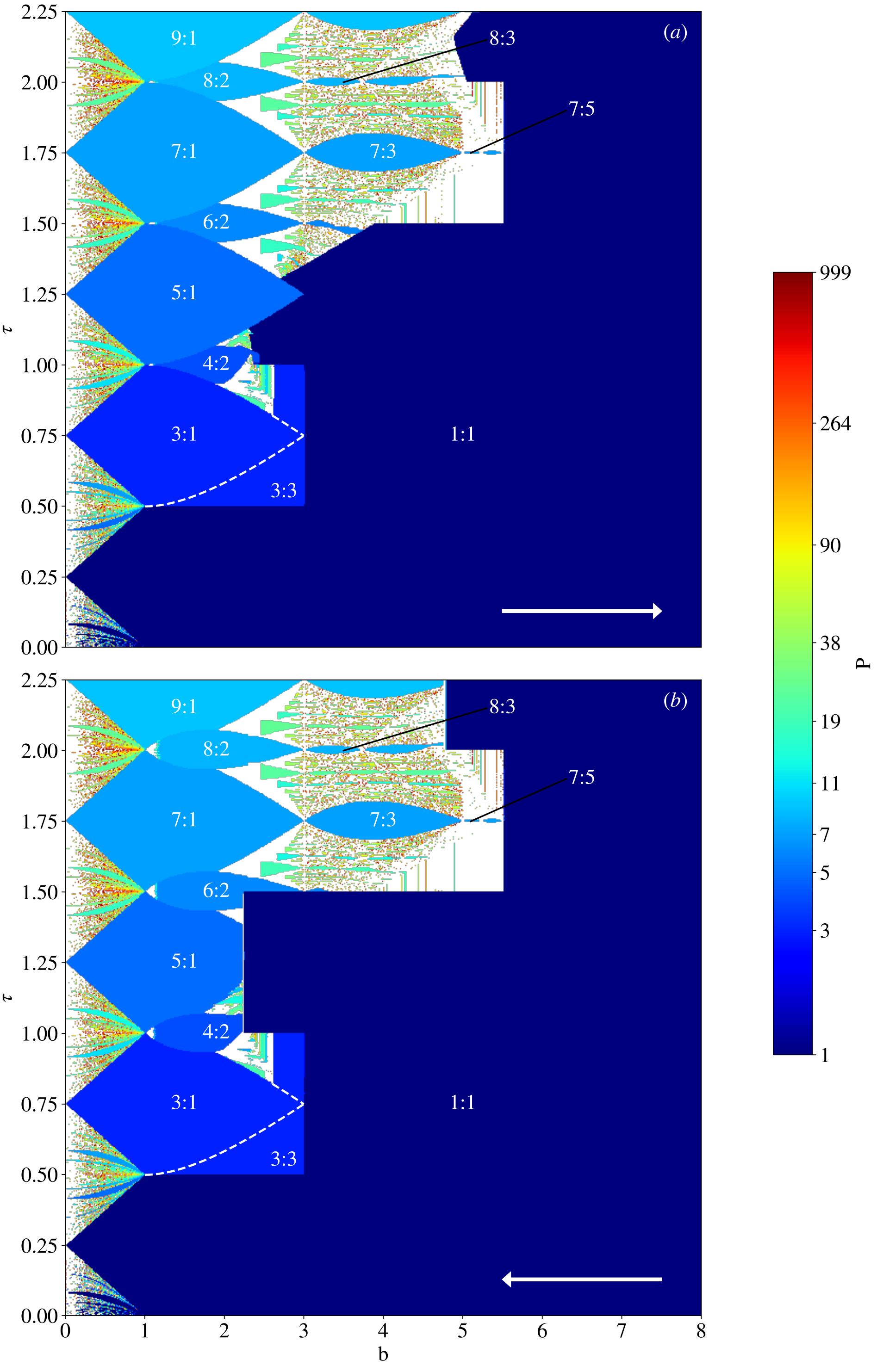

Figure 1.2: Colour maps showing the period of simulated solutions obtained by making 1124 sweeps of 1125 simulations of duration $T = 10000$ in $[0, 8] \times [0, 2.25]$. The white arrows indicate the direction of the sweep. The colour gradient shows the period P of the observed solution on a log scale; white space indicates where the solution was aperiodic, or with $P > 999$. The white text indicates the characteristic ratio $P : R$ of stable solutions found within the labelled tongues. The dashed white line shows the border between the regions where the $3:1$ and $3:3$ solutions occur.

context of piecewise-linear continuous maps with single switching mechanisms [51, 52, 53]. Such Arnold tongues have been compared to strings of sausages [50]. We will refer to an individual "sausage" as a *tongue*, and refer to a full "string" as an Arnold tongue. A notable feature of the Arnold tongues in this system is that the characteristic ratio is different in each tongue in the string. For example, the $P = 7$ Arnold tongue is rooted on $b = 0$ at $\tau = 1.75$, and consists of the $7:1$ tongue, the $7:3$ tongue, the $7:5$ tongue, and the unlabelled $7:7$ tongue. In each Arnold tongue, the leftmost tongue is a $P : R$ tongue that is rooted on $b = 0$ at $\tau = \frac{P}{4R}$. P is the same in every tongue in the chain, but R increases the further right the tongue lies in the string.

Tongues with similar characteristic ratios tend to have similar shape, with some variation. For example, the $6 : 2$ and $8 : 2$ tongues have identical shape; the $4 : 2$ tongue is different. The large $P : 1$ tongues noted earlier have identical shape for $P > 1$. The $1 : 1$ Arnold tongue is unlike any other Arnold tongue in shape. For large b, the forcing dominates the feedback, which results in the system converging to the $1 : 1$ solution exclusively. We refer to the region in which b is large enough as the locked region, where the system is locked to the period of the forcing. The boundary of the locked region is unusual, being made up of straight lines which meet at right angles. Branching off from the horizontal lines of the boundary, there are vertical stripes in which $P : P$ solutions exist. We will devote considerable attention to studying the $1 : 1$ solution later.

Now we compare the upward b sweep in Fig. 1.2(a) to the downward b sweep in Fig. 1.2(b). The most significant difference occurs around $(b, \tau) = (2.5, 1.25)$. In the upward sweep, the system follows the $5 : 1$ solution to the edge of the $5 : 1$ tongue, and there is a complicated region of smaller tongues and aperiodicity above the $5 : 1$ tongue. In the downward sweep, the system instead continues to follow the $1 : 1$ solution into the region where different features occurred in the upward sweep. This agrees with the bistability of the $1 : 1$ and $5 : 1$ solutions seen in Fig. 1.1(d,e). A less obvious difference is the bistability to the right of $b = 1$; note the apparent difference in shape of the $P : 2$ and $P : 1$ tongues near this line. This occurs because the $P : 1$ tongues overlap with the $P : 2$ tongues, in agreement with the bistability of the $4 : 2$ and $3 : 1$ solutions seen in Fig. 1.1(b,c). The absence of $P : 1$ tongues for even P is notable. We numerically observe such solutions in simulations, but only for small $b \lesssim 0.5$ and exactly $\tau = \frac{P}{4}$.

1.5 Dynamics

Fig. 1.3 shows maximum charts in the (b, τ) plane overlayed with bifurcation curves. The maximum is taken as the maximum value of a solution over an interval of length 1000 after a transient of length 9000 from the same simulations that were used to generate Fig. 1.2. Some of the larger tongues seen in Fig. 1.2 can be seen without difficulty in Fig. 1.3 due to a large jump in maxima at the boundaries of the tongues. We note that the even $P : R$ tongues are striped horizontally; this is most evident in the $4 : 2$ and $8 : 2$ tongues. It appears that solutions with even P or R are not invariant under the symmetry $x(t) \to -x\left(t + \frac{1}{2}\right)$; rather there exists a pair of symmetry-related counterpart solutions, each with a different maximum value. The system converges to one of these solutions depending on initial conditions, and remains at that solution until swept out of the tongue, resulting in stripes parallel to the direction of the sweep. This feature was also observed in the smooth system (1.1,1.2) by Keane, Krauskopf and Postlethwaite [13].

We will devote the rest of this section to deriving and characterising the bifurcation curves plotted in Fig. 1.3. In order to do so, we first need to establish a systematic method of analysing solutions. We apply this method to develop Poincaré maps and border collision maps through which we study the bifurcations present in this system.

1.5.1 Symbolic representation

We require a robust framework under which we can analyse the dynamics of this system. We note that the characteristic ratio does not distinguish between the two different forms of the $1 : 1$ solution shown in Fig. 1.1(d,g). Observing the order in which the feedback and forcing change after the trajectory passes through $x = 0$, we note that the feedback changes before the forcing in Fig. 1.1(d), and after the forcing in Fig. 1.1(g). In order to precisely capture these differences, a more explicit labelling system is required. We therefore adopt a symbolic representation of solutions. Symbolic representations have previously been used to great effect in the study of iterative maps [52, 54].

Let a solution be represented by a sequence of events $...X_1 X_2 ... X_n...$ where $X_i \in \left\{D, \bar{D}, Z, \bar{Z}, H, \bar{H}\right\}$. D denotes a transition of the forcing from $-b$ to b, Z denotes a transition from $x < 0$ to $x > 0$, and H denotes a transition of the feedback from -1 to 1. A bar over a symbol causes it to denote the opposite transition; a symbol with two bars over it is the same as the symbol unbarred. Fig. 1.4(a) shows the $5 : 1$ solution seen Fig. 1.1(e), labelled with the events that occur in the trajectory. As this solution is periodic, the sequence repeats, so we abbreviate the sequence of events representing the solution to a minimal repeating sequence $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}, \bar{Z}, D, \bar{D}, H, D, \bar{D}, D\right]$. In general, a $P{:}R$ solution is represented by a minimal repeating sequence of events $[X_1, X_2, ..., X_n]$ containing:

- P D events and P $\bar{D}$ events,
- R Z events and R $\bar{H}$ events,
- R $\bar{Z}$ events and R H events.

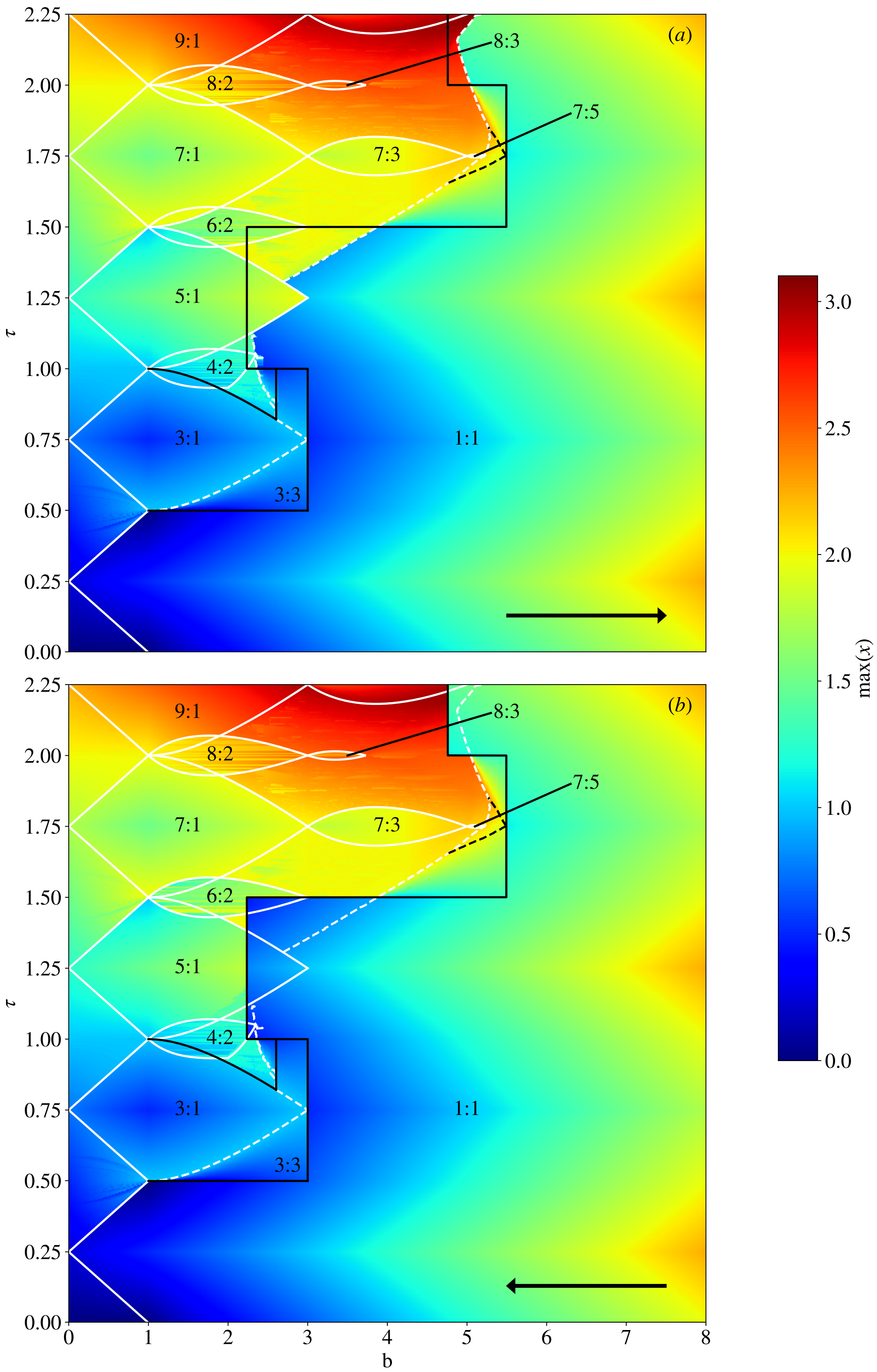

Figure 1.3: Maximum charts in the (b, τ) plane overlayed with bifurcation curves. The colour gradient indicates the maximum value of the observed solution. The black arrows indicate the direction in which solutions were swept. The black text indicates the ratio of the period of the solution to the number of Z symbols in the sequence for the stable solution within the associated tongue. BCSN bifurcations are shown in solid white, and T bifurcations are shown in solid black. The $D\bar{D}$ curve is shown in dashed white.

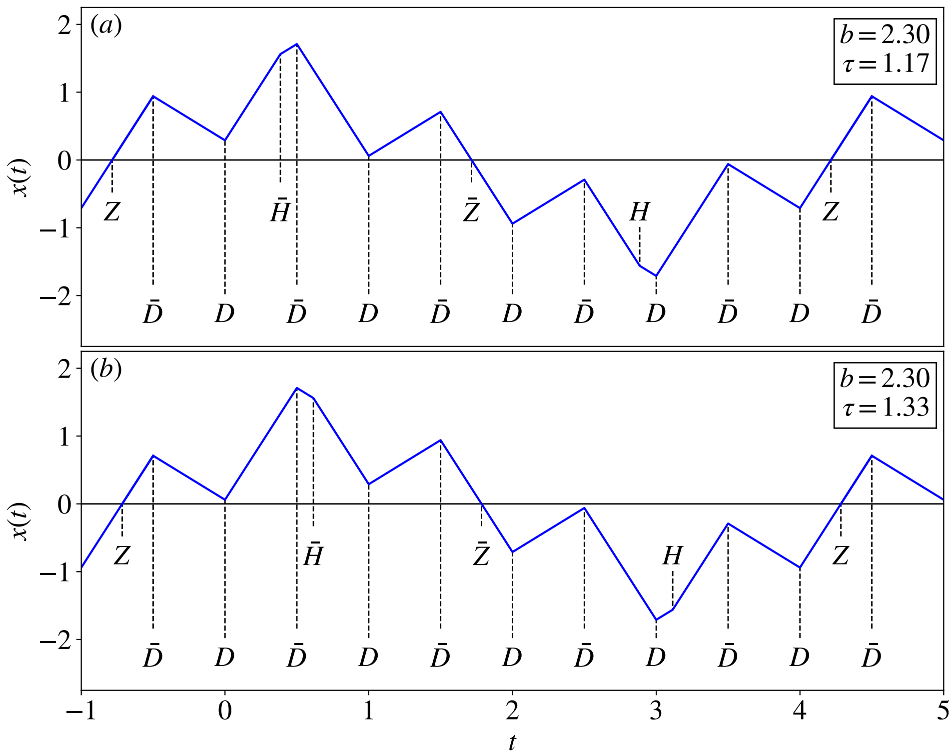

Figure 1.4: Derivation of the symbolic representations of the $5:1$ solution.

Every cyclic permutation of a sequence represents the same solution. For the sake of consistency, all sequences begins with a Z. We observe that the 5:1 solution shown in Fig. 1.4(a) is invariant under the symmetry $x(t) \to -x\left(t+\frac{1}{2}\right)$; this causes the second half of the sequence $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}, \bar{Z}, D, \bar{D}, H, D, \bar{D}, D\right]$ to be the same as the first half with all symbols barred. We abbreviate $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}, \bar{Z}, D, \bar{D}, H, D, \bar{D}, D\right]$ as $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}\right]^-$ for convenience. In general, a $P:R$ solution of the form $\left[X_1, X_2, ...X_n, \bar{X}_1, \bar{X}_2, ..., \bar{X}_n\right]$ can also be represented by a half-sequence $[X_1, X_2, ...X_n]^-$.

A sequence is considered *legal* within a subset of the (b, τ) plane if it represents a solution that exists within that subset. Each $P:R$ solution is represented by a set of legal sequences, each one existing in a unique subset of the (b, τ) plane. The union of these subsets is the region in which the solution exists. The $5:1$ solution exists within the $5:1$ tongue shown in Fig. 1.2(a). It is represented by the half-sequences $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}\right]^-$ for $\tau \leq 1.25$ and $\left[Z, \bar{D}, D, \bar{D}, \bar{H}, D, \bar{D}\right]^-$ for $\tau \geq 1.25$, as shown in Fig. 1.4. At $\tau = 1.25$, the increasing time delay between Z and $\bar{H}$ causes $\bar{H}$ to swap with $\bar{D}$, changing the sequence. Determining whether a given sequence is legal within a subset of the (b, τ) plane is a nontrivial problem. Appendix A.1.1 presents a general method to determine whether a sequence is legal for a given (b, τ), which can also be used to calculate a solution analytically from a sequence for a given (b, τ). This method was used to plot the solutions in Figs. 1.4,1.5 and the bifurcation curves in Fig. 1.3.

1.5.2 Dimension of a periodic solution

The rest of our analysis will primarily be concerned with the stability of periodic solutions. However, before we can analyse the dynamics of this system, we must understand the phase space in which these dynamics take place. Time-delay systems are generically of uncountably infinite dimension, with the phase space being a continuous function over a finite interval. Under certain circumstances, the effective dimension of smooth time-delay systems can be reduced through dimension reduction techniques [55, 56, 57, 58]. To avoid confusion, we note that such techniques are not being used here. Instead, we employ a somewhat novel approach inspired by our use of Poincaré maps to study this system.

As discussed previously, the state of the system $S(t)$ is a tuple whose length varies dynamically as the number of zero crossings in the last tau time varies. Consider the $5:1$ solution represented by $\left[Z, \bar{D}, D, \bar{D}, \bar{H}, D, \bar{D}\right]^{-}$ in Fig. 1.4(b). At any time after a $\bar{H}$ and before a $\bar{Z}$, or after a H and before a Z, there are no history elements in the state, and so $S(t) = (x; z_0)$. At any time after a Z and before the associated $\bar{H}$, or after a $\bar{Z}$ and before the associated H, there is one history element in the state, and so $S(t) = (x; z_0)$. However, by making use the sequence representation of the solution, if you know the displacement x, you can calculate the history element z_0, and so the state can be reduced to a single independent variable. To take a simple example, at time $t = 0$, a D event takes place at $x = x_D$, and the preceding Z event takes place at $x = 0$ at time $t = z_0$. Then

$$x_D = 0 + \left(-\frac{1}{2} - z_0\right)(1+b) + \frac{1}{2}(1-b) \tag{1.7}$$

which we can solve for z_0 to get

$$z_0 = -\frac{b + x_D}{b+1} \tag{1.8}$$

We say then that the $5:1$ solution is one-dimensional, as the state consists of either a single dynamic variable, or two linearly dependent dynamic variables.

Definition: The dimension of a periodic solution is the maximal number of independent dynamic variables in the state over a single cycle of the solution.

This is equivalent to saying that the dimension of a periodic solution is the minimal number of dynamic required to construct a Poincaré map which captures the dynamics of that solution. A one-dimensional Poincaré map is sufficient to capture the stability and existence of the $5:1$ solution, as we will demonstrate later. However, to fully consider the dynamics of the $5:1$ solution we will require additional concepts introduced later. Instead, we will begin with the $1:1$ solution which dominates so much of the parameter space.

1.5.3 Torus bifurcation of the $1:1$ solution

We now apply our sequence representation in analysing the $1:1$ solution; specifically, we determine what happens at the vertical black lines along the boundary of the locked region in Fig. 1.3. Consider the set of half-sequences that represent the $1:1$ solution. Each half-sequence contains one Z or $\bar{Z}$, one H or $\bar{H}$, and one D or $\bar{D}$. We need only consider half-sequences starting with Z, as $[X_1, X_2, X_3]^- = [X_3, X_1, X_2]^- = [X_2, X_3, X_1]^-$. There are only eight possibilities: $[Z, H, D]^-$, $[Z, D, H]^-$, $\left[Z, \bar{H}, D\right]^-$, $\left[Z, D, \bar{H}\right]^-$, $\left[Z, H, \bar{D}\right]^-$, $\left[Z, \bar{D}, H\right]^-$, $\left[Z, \bar{H}, \bar{D}\right]^-$, $\left[Z, \bar{D}, \bar{H}\right]^-$. $[Z, H, D]^-$and $[Z, D, H]^-$ are not legal, as they require Z to occur when $\dot{x} < 0$, which is impossible. Similarly, $\left[Z, \bar{H}, D\right]^-$and $\left[Z, D, \bar{H}\right]^-$are legal only for $b < 1$, and $\left[Z, \bar{D}, H\right]^-$and $\left[Z, H, \bar{D}\right]^-$are legal only for $b > 1$. The $1:1$ solutions shown in Fig. 1.1 are $\left[Z, \bar{H}, \bar{D}\right]^-$ in Fig. 1.1(d) and $\left[Z, \bar{D}, \bar{H}\right]^-$ in Fig. 1.1(g). In the case $b > 1$, the sequence representing the $1:1$ solution is restricted by τ in the following way:

$$\begin{aligned}
&\left[Z, \bar{H}, \bar{D}, \bar{Z}, H, D\right] \text{ for } \tau \bmod 1 \in [0, 0.25)\,, \\
&\left[Z, \bar{D}, \bar{H}, \bar{Z}, D, H\right] \text{ for } \tau \bmod 1 \in [0.25, 0.5)\,, \\
&\left[Z, H, \bar{D}, \bar{Z}, \bar{H}, D\right] \text{ for } \tau \bmod 1 \in [0.5, 0.75)\,, \\
&\left[Z, \bar{D}, H, \bar{Z}, D, \bar{H}\right] \text{ for } \tau \bmod 1 \in [0.75, 1)\,.
\end{aligned} \tag{1.9}$$

This can be explained by observing that for small τ, $\bar{H}$ must follow the Z that created it almost immediately. As τ increases, $\bar{H}$ drifts further away from Z in the sequence, drifting past $\bar{D}$ at $\tau = 0.25$. At $\tau = 0.5$, $\bar{H}$ drifts past the subsequent $\bar{Z}$. At $\tau = 0.75$, $\bar{H}$ drifts past the subsequent D. At $\tau = 1$, the $\bar{H}$ created at Z drifts past the subsequent Z, and the pattern repeats. At the same time, the same drift pattern occurs between $\bar{Z}$ and H. We now apply this information to analyse the $1:1$ solution.

We construct a Poincaré map on the $1:1$ solution. Let t_z be a time on the $1:1$ solution at which $x = 0$; then the state of the system at time t_z is

$$S(t_z) = (0; z_0, z_1, ..., z_{n-2}, t_z)^T \tag{1.10}$$

As $x = 0$ at $t = t_z$, we drop x and rewrite the state as

$$S_z = (z_0, z_1, ..., z_{n-2}, t_z)^T \tag{1.11}$$

Let t_z^* be the time on the $1:1$ solution at which $x = 0$ immediately after t_z. Then the state of the system at time t_z^* is

$$S_z^* = (z_1, z_2, ..., t_z, t_z^*)^T \tag{1.12}$$

As the $1:1$ solution is invariant under the symmetry $x(t) \to -x\left(t+\frac{1}{2}\right)$, it has the property

$$S_z = \begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_{n-2} \\ t_z \end{pmatrix} = \begin{pmatrix} z_1 - \frac{1}{2} \\ z_2 - \frac{1}{2} \\ \dots \\ t_z - \frac{1}{2} \\ t_z^* - \frac{1}{2} \end{pmatrix} = S_z^* - \frac{1}{2} \tag{1.13}$$

We define a Poincaré map $\mathbb{P} : S_z \to S_z^* - \frac{1}{2}$ so that the $1:1$ solution is a fixed point of $\mathbb{P}$. Therefore, $\mathbb{P}$ is defined by

$$\mathbb{P}\begin{pmatrix} z_0 \\ z_1 \\ \dots \\ z_{n-2} \\ t_z \end{pmatrix} = \begin{pmatrix} z_1 - \frac{1}{2} \\ z_2 - \frac{1}{2} \\ \dots \\ t_z - \frac{1}{2} \\ t_z^* - \frac{1}{2} \end{pmatrix} \tag{1.14}$$

For the purpose of deriving $\mathbb{P}$, we will assume w.l.o.g. that the trajectory is transitioning from $x < 0$ to $x > 0$ at $t_z \in [0, 0.5)$. Let $t_h \in [t_z, t_z^*)$ be the time at which the feedback changes. First, we consider the case where $\tau < 0.5$; then the zero element generated at t_z is consumed at $t_h = t_z + \tau$. Then $\mathbb{P}$ is one-dimensional as S_z has only one variable, t_z. We write the one-dimensional map as

$$\mathbb{P}(t_z) = t_z^* - \frac{1}{2} \tag{1.15}$$

If $\tau < 0.5$, then the solution is represented by the sequence $\left[Z, \bar{H}, \bar{D}\right]^-$; therefore, the feedback is positive for $t \in [t_z, t_h)$. We write an equation for the displacement of the trajectory between t_z and t_z^* as

$$b\left(\frac{1}{2} - t_z\right) - b\left(t_z^* - \frac{1}{2}\right) + (t_h - t_z) - (t_z^* - t_h) = 0 \tag{1.16}$$

We substitute $t_h = t_z + \tau$ and solve for $t_z^* - \frac{1}{2}$ to obtain

$$\mathbb{P}(t_z) = -\left(\frac{b-1}{b+1}\right) t_z + \frac{b+2\tau}{b+1} - \frac{1}{2} \tag{1.17}$$

As $\left|\frac{b-1}{b+1}\right| < 1$ for $b \geq 0$, $\mathbb{P}$ is stable for $\tau < \frac{1}{2}$.

If $\tau \in [0.5, 1)$, then t_h was generated, not at t_z, but at the previous $x = 0$ crossing z_0. Then $\mathbb{P}$ is two-dimensional as S_z has two variables, and the feedback is negative for $t \in [t_z, t_h)$, where $t_h = z_0 + \tau$. These conditions can be generalised for larger τ. The dimension of the system is $n = \lceil 2\tau \rceil$, where $\lceil 2\tau \rceil$ is the smallest integer greater than 2τ; then the feedback for $t \in [t_z, t_h)$ is $(-1)^{n-1}$. We can then generalise Eq.

((1.16)) for arbitrary n and solve for $t_z^* - \frac{1}{2}$ to obtain

$$t_z^* - \frac{1}{2} = -t_z + \frac{b + 2t_h(-1)^{n-1}}{b + (-1)^{n-1}} - \frac{1}{2} \tag{1.18}$$

For $\tau > \frac{1}{2}$, $\mathbb{P}$ can written as

$$\mathbb{P}\left(S_z\right) = AS_z + B \tag{1.19}$$

where A is an $n \times n$ matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ \frac{2(-1)^{n-1}}{b+(-1)^{n-1}} & 0 & \cdots & \cdots & -1 \end{pmatrix} \tag{1.20}$$

and

$$B = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \cdots \\ -\frac{1}{2} \\ \frac{b+2\tau(-1)^{n-1}}{b+(-1)^{n-1}} - \frac{1}{2} \end{pmatrix} \tag{1.21}$$

Note that A only depends on the parameter b and not on τ. This means that for fixed n, the value of of b at which the fixed point of $\mathbb{P}$ is bifurcating, b_{bif}, is constant. By solving the characteristic equation, we calculate

$$b_{\text{bif}}(n) = \frac{1}{\cos\left(\frac{\pi(n-1)}{2n-1}\right)} - (-1)^{n-1} \tag{1.22}$$

where $n = \lceil 2\tau \rceil$. Full calculations may be found in Appendix A.1.2. The curve $\left(b_{\text{bif}}\left(\lceil 2\tau \rceil\right), \tau\right)$ can be seen plotted as black vertical lines against maximum charts in Fig. 1.3. If we examine the maximum chart where b is swept from right to left, we see that the system ceases to converge to the $1:1$ solution along this curve. We now know that this is because the fixed point of $\mathbb{P}$, and hence the $1:1$ solution, loses stability at $b = b_{\text{bif}}$. By calculating the eigenvalues of $\mathbb{P}$ explicitly for $n = 2$ and $n = 3$, we find that the loss of stability occurs because a pair of complex conjugate eigenvalues $\lambda_{1,2}$ cross $|\lambda| = 1$. Therefore the fixed point of $\mathbb{P}$ loses stability due to a Neimark-Sacker (NS) bifurcation[59]. As $\mathbb{P}$ is a Poincaré map on a periodic orbit, a NS bifurcation of the fixed point of $\mathbb{P}$ corresponds to a torus (T) bifurcation of the $1:1$ solution. However, $\mathbb{P}$ only shows the existence of the T bifurcation along the vertical sections of the boundary of the locked region, where n is constant and $\mathbb{P}$ is smooth. To fully understand the horizontal sections of the boundary, we must look to non-smooth bifurcation theory.

1.6 Border-collision bifurcations

The Arnold tongues seen in Fig. 1.2 have sharply defined boundaries, made up of curves and straight lines. We seek to determine what happens to solutions at these boundaries and derive analytic expressions for the boundaries using Poincaré maps.

By simulating solutions near the boundaries of the Arnold tongues, we observe that moving closer to the boundaries causes a D or $\bar{D}$ to move closer to $x = 0$. An example of this is Fig. 1.4(a), where the D at $t = 0$ in the $5:1$ solution occurs for $x > 0$ near $x = 0$. If this D crossed $x = 0$ and occurred at $x < 0$, this would significantly impact the feedback. There would be two additional $x = 0$ crossings in the trajectory, changing the D to $\bar{Z}DZ$ and adding a H and a $\bar{H}$ elsewhere in the sequence. Therefore, when we construct a Poincaré map to describe the dynamics of the system close to the boundary of such a tongue, the map must have a border at $x = 0$, such that the map is continuous across the border but not differentiable at the border. Such maps, and the associated border collision bifurcations, have recently received systematic analysis in the literature [60, 15, 17, 16, 61, 62, 63, 64, 65].

1.6.1 Border-collision saddle-node bifurcation of the $5:1$ and $5:3$ solutions

To begin our analysis of the border-collision bifurcations in this system, we consider what happens to the observed $5:1$ solution at the boundary of the $5:1$ tongue. We begin with this solution for two reasons. Firstly, we observe that for odd $P \geq 5$, the $P:1$ tongues are all of the same shape, and the $P:1$ solution vanishes at the boundary of the tongue. This is in contrast to the $3:1$ solution, which may change smoothly to the $3:3$ solution at the boundary, and the $1:1$ solution, which behaves unlike any other periodic solution. Secondly, as an odd period solution, it has the symmetry $x(t) = -x(t + \frac{P}{2})$; this makes the construction of Poincaré maps easier than it is for the even period solutions.

We construct a Poincaré map $\mathbb{B}$ on the $5:1$ and $5:3$ solutions represented by $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}\right]^-$ and $\left[Z, \bar{D}, \bar{H}, H, D, \bar{H}, \bar{D}, \bar{Z}, D, Z, \bar{D}\right]^-$ respectively, which can be seen in Fig. 1.5(a), such that our fixed point is the state of the system at time $t = 0$, at the position x_D at which the D that will cross $x = 0$ occurs. As noted in Section 1.5.2, $\mathbb{B}$ is one-dimensional. We divide $\mathbb{B}$ into $\mathbb{B}^+$ and $\mathbb{B}^-$ for $x \geq 0$ and $x < 0$ respectively. The $5:1$ and $5:3$ solutions are invariant under the symmetry $x(t) = -x\left(t + \frac{1}{2}\right)$; therefore we construct our map $\mathbb{B} : x_D\left(0\right) \to -x_D\left(\frac{5}{2}\right)$. $\mathbb{B}^+$ is a map on the solution represented by $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}\right]^-$. Following the blue curve in Fig. 1.5(b), we can derive

$$\mathbb{B}^{+}(x_D) = -\left(\frac{b-1}{b+1}\right)x_D + \frac{2b}{b+1} - \frac{b+5}{2} + 2\tau \tag{1.23}$$

$\mathbb{B}^{-}$ is a map on the 5 : 3 solution represented by $\left[Z, \bar{D}, \bar{H}, H, D, \bar{H}, \bar{D}, \bar{Z}, D, Z, \bar{D}\right]^{-}$, which is shown in Fig. 1.5(b) in red. The feedback change due the trajectory dipping below $x = 0$ occurs at some $t \in (1, 1.5)$ has duration $\frac{-2bx}{b^2-1}$; $\mathbb{B}^{-}$ is otherwise identical to $\mathbb{B}^{+}$. Thus we derive $\mathbb{B}$ as

$$\mathbb{B}(x_D) = \begin{cases} -\left(\frac{b-1}{b+1}\right)x_D + \frac{2b}{b+1} - \frac{b+5}{2} + 2\tau & x_D \geq 0 \\ \left(\frac{4b}{b^2-1} - \frac{b-1}{b+1}\right)x_D + \frac{2b}{b+1} - \frac{b+5}{2} + 2\tau & x_D < 0 \end{cases} \tag{1.24}$$

By setting $x_D = 0$ and solving $\mathbb{B}$ for τ, we obtain the curve $\tau = \frac{b^2+2b+5}{4(b+1)}$. This matches the lower right boundary of the 5 : 1 tongue spanning from $(1, 1)$ to $(3, 1.25)$. For $b \in [1, 3]$ and $\tau \in \left[\frac{b^2+2b+5}{4(b+1)}, 1.25\right]$, $\mathbb{B}$ has two fixed points; a stable fixed point that exists for $x_D \geq 0$ and an unstable fixed point that exists for $x_D \leq 0$. These two fixed points collide and vanish at the border $x_D = 0$ at $\tau = \frac{b^2+2b+5}{4(b+1)}$ in a border collision saddle node (BCSN) bifurcation [15]. Hence the 5 : 1 tongue is bounded by a BCSN bifurcation. Fig. 1.5(b,e,f) shows the BCSN bifurcation of the 5 : 1 and 5 : 3 solutions.

We generalise $\mathbb{B}$ for every $P : 1$ tongue for odd $P \geq 5$ as

$$\mathbb{B}_P(x_D) = \begin{cases} -\left(\frac{b-1}{b+1}\right)x_D + \frac{2b}{b+1} - \frac{b+|P-4\tau|}{2} & x_D \geq 0 \\ \left(\frac{4b}{b^2-1} - \frac{b-1}{b+1}\right)x_D + \frac{2b}{b+1} - \frac{b+|P-4\tau|}{2} & x_D < 0 \end{cases} \tag{1.25}$$

where $|P - 4\tau|$ is the term that causes the $P : 1$ tongues to be symmetric across $\tau = \frac{P}{4}$. By setting $x_D = 0$, we calculate the boundaries of all $P : 1$ tongue for odd $P \geq 5$ and $b \in [1, 3]$ as

$$\tau_{\mathrm{P}}(b) = \frac{P}{4} \pm \frac{3b - b^2}{4\,(b+1)} \tag{1.26}$$

We find that this phenomenon persists throughout the system. The vast majority of tongues investigated are bounded by BCSN bifurcations occurring when a D crosses $x = 0$. For $b > 1$, a stable $P : R$ solution and an unstable $P : R + 2$ solution undergo a SN bifurcation when a D crosses $x = 0$ if the solution has the symmetry $x(t) = -x(t + \frac{P}{2})$. Otherwise, the stable $P : R$ solution undergoes a SN bifurcation with an unstable $P : R + 1$ solution when a D crosses $x = 0$.

For $b < 1$, the mechanism by which BCSN bifurcations occur is slightly different. Instead of a D crossing $x = 0$ and adding new symbols to the sequence, D instead crosses $x = 0$ by swapping order with an existing Z. The stable 5 : 1 solution $\left[Z, \bar{D}, D, \bar{H}, \bar{D}, D, \bar{D}\right]^{-}$ undergoes a BCSN bifurcation with the unstable 5 : 1 $\left[Z, D, \bar{D}, D, \bar{H}, \bar{D}, D\right]^{-}$ at $\left(b, \frac{5-b}{4}\right)$, $b < 1$, as shown in Fig. 1.5(a,c,d). While the

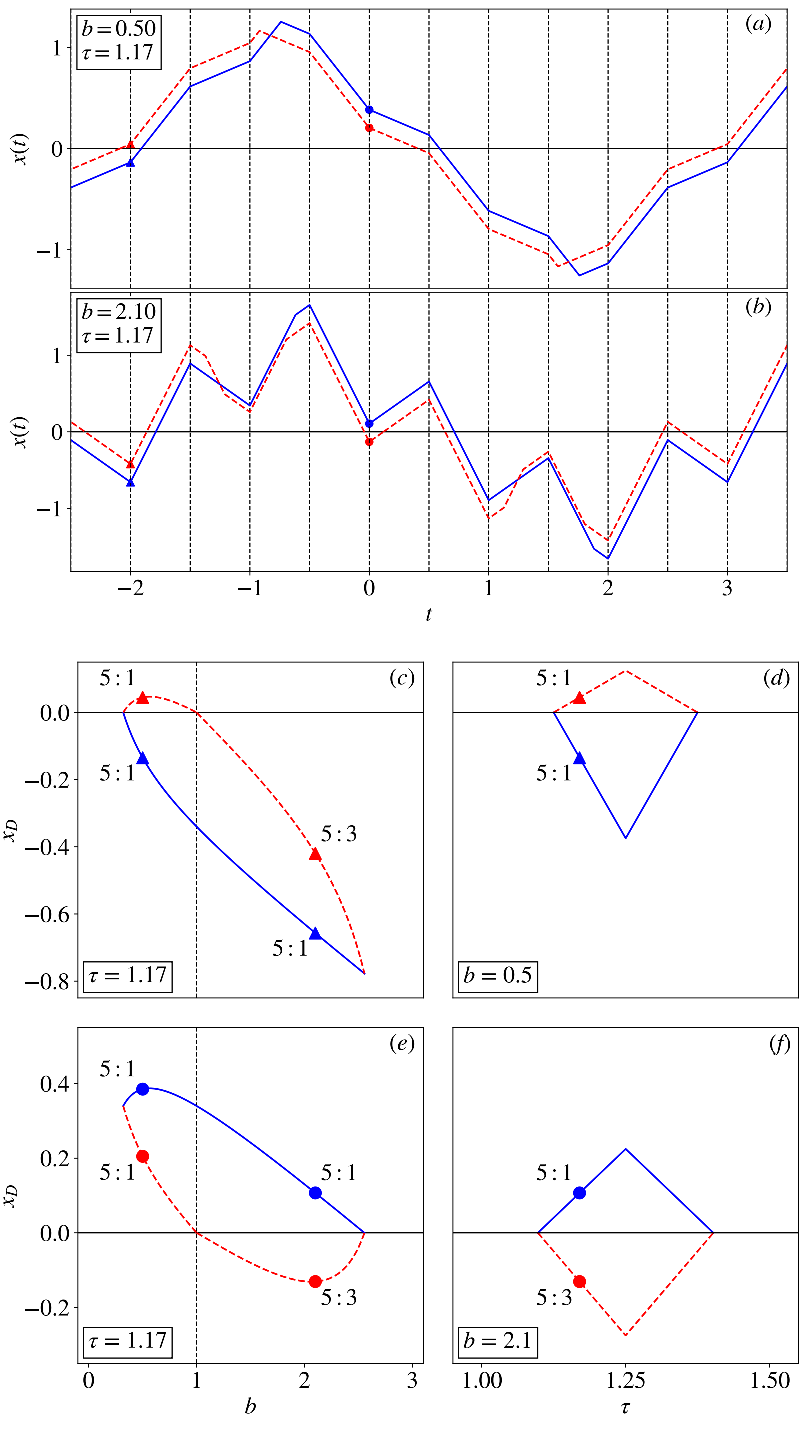

Figure 1.5: BCSN bifurcations at the boundary of the $5:1$ tongue. The stable solution is plotted in solid blue, while the unstable solution is plotted in dashed red. In (a), the stable $5:1$ solution $\left[Z,\bar{D},D,\bar{H},\bar{D},D,\bar{D}\right]^-$ and the unstable $5:1$ solution $\left[Z,D,\bar{D},D,\bar{H},\bar{D},D\right]^-$ are plotted close to $\left(b,\frac{5-b}{4}\right)$. In (b), the stable $5:1$ solution $\left[Z,\bar{D},D,\bar{H},\bar{D},D,\bar{D}\right]^-$ and the unstable $5:3$ solution $\left[Z,\bar{D},\bar{H},H,D,\bar{H},\bar{D},\bar{Z},D,Z,\bar{D}\right]^-$ are plotted close to $\left(b,\frac{b^2+2b+5}{4(b+1)}\right)$. The D events associated with the BCSN bifurcation for $b<1$ are plotted as triangles, while those associated with the BCSN bifurcation for $b>1$ are plotted as circles; (c), (d), (e) and (f) show their evolution as b and τ vary.

unstable $5:1$ and $5:3$ solutions shown in Fig. 1.5 are two different solutions under our sequence representation, they are identical at $b=1$, where a D passes through $x=0$ without a bifurcation. A noteworthy feature of the BCSN bifurcations at the boundary of the $5:1$ tongue is that the pair of D events which collide at $x=0$ is different for $b<1$ and $b>1$, as seen in Figure Fig. 1.5(a,b,c,e). Generally, for $b<1$, a stable $P:1$ solution undergoes a BCSN bifurcation with an unstable $P:1$ solution. Again, we produce a general form for the curves where a $P:1$ solution undergoes a BCSN bifurcation for odd $P \geq 1$ and and $b \in [0,1]$ as

$$\tau_{\mathrm{P}}(b) = \frac{P \pm b}{4} \tag{1.27}$$

The $1:1$ solution is a special case. The stable $\left[Z, \bar{H}, \bar{D}\right]^-$ solution and the unstable $\left[Z, D, \bar{H}\right]^-$ undergo a BCSN bifurcation along $\left(b, \frac{n}{2} + \frac{1-b}{4}\right)$, $b<1$, $n \in \mathbb{Z}^+$. The stable $\left[Z, \bar{D}, \bar{H}\right]^-$ solution and the unstable $\left[Z, \bar{H}, D\right]^-$ undergo a BCSN bifurcation along $\left(b, \frac{n}{2} + \frac{1+b}{4}\right)$, $b<1$, $n \in \mathbb{Z}^+$.

A selection of these curves can be seen plotted in solid white in Fig. 1.3. There is a conspicuous outlier to this pattern. The $3:1$ solution does not undergo a BCSN bifurcation at the right boundary of the 3:1 tongue for $b>1$. We will now investigate why this is the case.

1.6.2 Border-collision torus bifurcation of the $3:1$ solution

Fig. 1.2 shows that the $3:1$ tongue has the same shape as the other $P:1$ tongues, rooted at $\tau = 0.75$. Along some of the boundary, the $3:1$ solution changes to the $3:3$ solution without bifurcation. We test the observed sequences which represent the $3:3$ solution and find that the $3:3$ solution exists in the region $[1,3] \times [0.5,1]$ outside of the $3:1$ tongue. However, the $3:3$ solution is not always stable; this can be seen in the top half of the 3:3 region in Fig. 1.2. In order to investigate this phenomenon, we construct a Poincaré map $\mathbb{T}$ by a similar method as for the BCSN bifurcation. For $\tau > 0.75$, the $3:1$ solution is represented by $\left[Z, \bar{D}, D, \bar{H}, \bar{D}\right]^-$ from which we derive $\mathbb{T}^+$ for $x_D \geq 0$. For $x_D < 0$, the sequence changes to $\left[Z, \bar{H}, H, \bar{D}, \bar{Z}, D, Z, \bar{H}, \bar{D}\right]^-$, from which we derive $\mathbb{T}^-$. There is a significant difference however; the $3:1$ solution is one-dimensional, but for $\tau > 0.75$, the $3:3$ solution is two-dimensional. This occurs because the additional $\bar{H}$ and H events occur between the D that crosses $x=0$ and the Z present in both sequences. Therefore, we also need to know the time $t_{\bar{H}}$ at which the $\bar{H}$ present in both sequences occurs. We derive $\mathbb{T}:(x_D, t_{\bar{H}})^T \to$

$\left(x_D^*, t_{\bar{H}}^* - \frac{3}{2}\right)^T$ as

$$\mathbb{T}\begin{pmatrix} x_D \\ t_{\bar{H}} \end{pmatrix} = \begin{cases} \begin{pmatrix} -1 & -2 \\ \frac{1}{b+1} & \frac{2}{b+1} \end{pmatrix}\begin{pmatrix} x_D \\ t_{\bar{H}} \end{pmatrix} + C & x_D \geq 0 \\ \begin{pmatrix} \frac{4b}{b^2-1} - 1 & -2 \\ \frac{1}{b+1} & \frac{2}{b+1} \end{pmatrix}\begin{pmatrix} x_D \\ t_{\bar{H}} \end{pmatrix} + C & x_D < 0 \end{cases} \tag{1.28}$$

where $C = \begin{pmatrix} \frac{3-b}{2} \\ \frac{b}{b+1} + \tau - \frac{3}{2} \end{pmatrix}$. Note that $\mathbb{T}^+$ has two rows which are multiples of each other; this is due to the $3:1$ solution being one-dimensional, and results in a 0 eigenvalue. By setting $x_D = 0$ and solving $\mathbb{T}$ for τ, we obtain the curve $\tau = \frac{3}{4} + \frac{3b-b^2}{4(b+1)}$, which matches the upper right boundary of the $3:1$ tongue spanning from $(1,1)$ to $(3,0.75)$. $\mathbb{T}$ has a single fixed point which crosses $x_D = 0$ at the boundary of the $3:1$ tongue. For $x_D > 0$ the fixed point is always stable. For $x_D < 0$ the fixed point is stable for $b > 2.6038$, and loses stability when a pair of complex conjugate eigenvalues pass through $|\lambda| = 1$, resulting in a NS bifurcation at $b = 2.6038$. $|\lambda_\pm| < 1$ for $b > 2.6038$. Therefore, for $\tau > 0.75$ the $3:3$ solution is stable for $b > 2.6038$. This agrees with the shape of the region in which the $3:3$ solution is observed in Fig. 1.2.

For $b < 2.6038$, we observe an interesting interaction between $\mathbb{T}^+$ and $\mathbb{T}^-$. As $\mathbb{T}^+$ has a 0 eigenvalue, any point $(x_D, t_{\bar{H}})$ for $x_D > 0$ is mapped directly to a nullcline on which the fixed point sits for $x_D > 0$. The trajectory then undergoes decaying oscillations to the fixed point in the nullcline. However, if the fixed point is close to $x_D = 0$, $\mathbb{T}^+$ can map the point into $x_D < 0$, where the nullcline does not exist. In the absence of a stable fixed point at $x_D < 0$, the trajectory starts to spiral out to infinity. As this spiral must cross $x_D = 0$ eventually, the trajectory is caught by the nullcline again, causing an interesting half-spiral attraction. When the fixed point is at $x_D < 0$, the trajectory converges to a stable attractor that strikes the nullcline at multiple points. Fig. 1.6(a, b) show the interaction between $\mathbb{T}^+$ and $\mathbb{T}^-$. Fig. 1.6(c) shows the results of simulations in a (b, τ) sweep that crosses the boundary of the $3:1$ Arnold tongue. The simulated results agree with those derived from $\mathbb{T}$, and confirm that the closed invariant curve is born from the fixed point as it crosses $x_D = 0$. We consider this to be a border collision Neimark-Sacker (BCNS) bifurcation of $\mathbb{T}$, similar to that seen by Meiss [62]. It corresponds to a border collision torus (BCT) bifurcation of the $3:1$ solution in the continuous system. The BCT bifurcation is supercritical, as a stable closed invariant curve expands from the fixed point when it becomes unstable. The BCT bifurcation of the $3:1$ solution and the T bifurcation of the $3:3$ solution are plotted in black in Fig. 1.3.

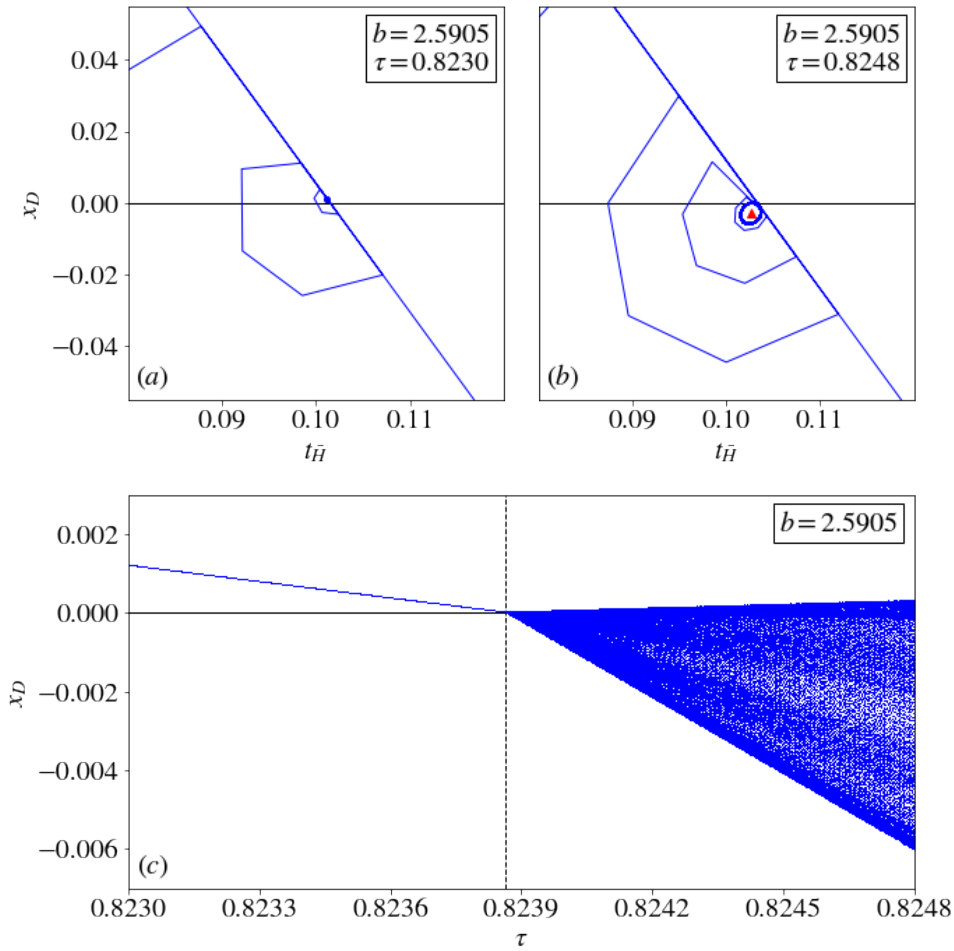

Figure 1.6: Border collision Neimark-Sacker (BCNS) bifurcation of the fixed point of the $3:1$ solution. The upper two plots show the BCNS bifurcation of the half-period map $\mathbb{T}$. The lower plot is based on simulation using our iterative map, and shows the stable closed invariant curve expanding from the fixed point from $x = 0$. The fixed point is plotted as a blue circle where it is stable, and as a red triangle where it is unstable.

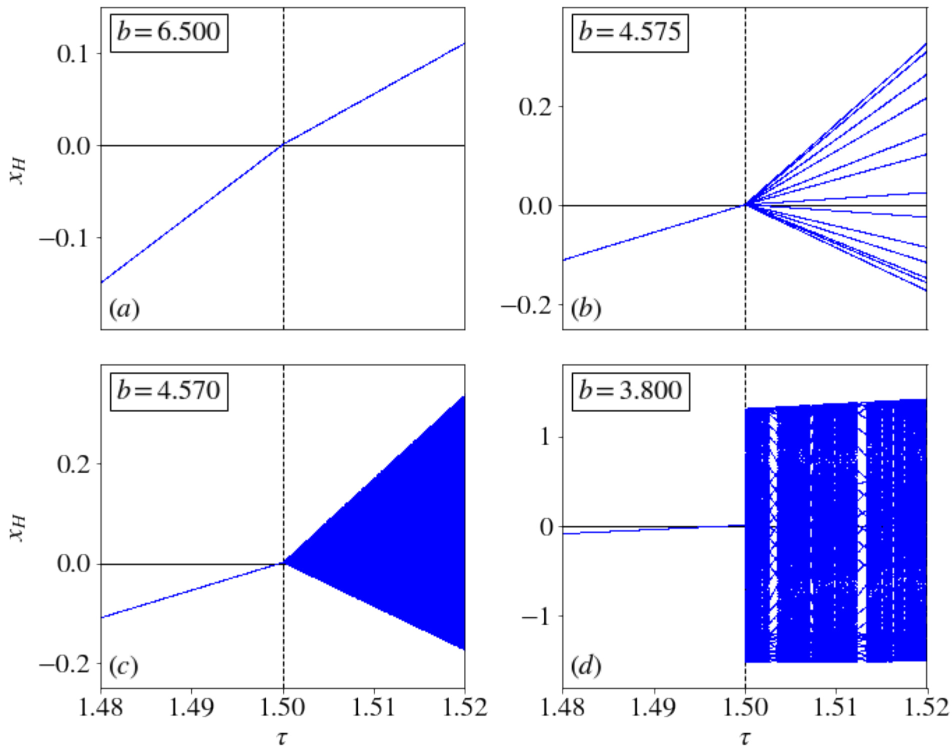

Figure 1.7: x-coordinates of H events from simulations. The four plots show the different cases of the $1:1$ solution at the border.

1.6.3 Border collision torus bifurcation of the $1:1$ solution

Now that we have studied the BCT bifurcation of the $3:1$ solution, we return to the $1:1$ solution. Previously we noted that the Poincaré map $\mathbb{P}$ only proved the existence of the torus bifurcation along the vertical boundaries of the locked region. We can now consider the horizontal boundaries in terms of border collisions. Recall that the $1:1$ solution is represented by four different sequences. At $\tau = \frac{n}{2}$, $n \in \mathbb{N}$, the sequence representing the $1:1$ solution changes when a H and a $\bar{H}$ cross $x = 0$. This results in a change in the dimension of $\mathbb{P}$. Let $\mathbb{P}_n$ denote $\mathbb{P}$ when $\mathbb{P}$ contains an $n \times n$ matrix. When τ increases past $\tau = \frac{n}{2}$, $\mathbb{P}$ changes from $\mathbb{P}_n$ to $\mathbb{P}_{\mathrm{n+1}}$. Let us consider x_H, the position of the H event, as the fixed point of the map. Then $x_H = 0$ when $\mathbb{P}$ changes from $\mathbb{P}_n$ to $\mathbb{P}_{\mathrm{n+1}}$, resulting in a border collision at $x_H = 0$. Fig. 1.7 shows four cases that occur in the border collision when we sweep across the border $\tau = 1.5$ for fixed b.

Fig. 1.7(a) shows the stable $1:1$ solution remaining stable, as both $\mathbb{P}_3$ and $\mathbb{P}_4$ are stable at $b = 6.5$. Fig. 1.7(b) shows the stable $1:1$ solution becoming unstable, generating a $13:13$ solution in a supercritical BCT bifurcation. This solution exists in one of the vertical stripes we noted in Fig. 1.2. Fig. 1.7(c) shows the stable $1:1$ solution becoming unstable, generating an aperiodic solution in a supercritical BCT bifurcation, showing that there are gaps between the vertical stripes. Fig. 1.7(d) shows the stable $1:1$ solution becoming unstable; however, the aperiodic solution that the system converges to afterwards is not generated at the border, indicating that the BCT bifurcation is subcritical at this point. Therefore the BCT bifurcation of the $1:1$ solution must change criticality at some point along $\tau = 1.5$. The BCT bifurcation of the $1:1$ solution produces the horizontal black lines in Fig. 1.3, completing the boundary of the locked region.

We examine the difference in maxima of nearby solutions on either side of $\tau = 1.5$ in Fig. 1.3(b). The point $(b, \tau) = (1.5, 3.86)$ stands out; the difference in maxima is sudden to the left of that point, and gradual to the right. The gradual change in maxima occurs where the BCT bifurcation is supercritical, and the abrupt change in maxima occurs where the BCT bifurcation is subcritical, as illustrated by Fig. 1.7(b-d). We note that $(b, \tau) = (1.5, 3.86)$ forms one corner of a roughly triangular region containing vertical $P:P$ stripes seen in Fig. 1.2; every $P:R$ solution in this region is $P:P$. This leads us to the dashed white curve in Fig. 1.3. We refer to this curve as the $D\bar{D}$ curve. To the right of the $D\bar{D}$ curve, all D events occur at $x < 0$ and all $\bar{D}$ events occur at $x > 0$, and so a legal sequence cannot contain a D adjacent to a $\bar{D}$; a $P:R$ solution in this region must, therefore, be $P:P$. To the left of the $D\bar{D}$ curve, a legal sequence representing a solution can contain a D adjacent to a $\bar{D}$; we refer to such a solution as a $D\bar{D}$ solution. The point at which the BCT bifurcation changes criticality occurs where the $D\bar{D}$ curve intersects the boundary of the locked region.

As noted previously, when sweeping from right to left, the system converges to the $1:1$ solution until it becomes unstable at the T bifurcation. Now, we see that when sweeping from left to right, the system converges to $D\bar{D}$ solutions until they vanish at the $D\bar{D}$ curve; we note that a $6:5$ and a $6:6$ solution undergo a BCSN bifurcation at the $D\bar{D}$ curve. If the T bifurcation lies to the left of the $D\bar{D}$ curve, then the $1:1$ solution remains stable while the $D\bar{D}$ solutions exist, resulting in regions of multistability. We observe that this behaviour, together with the change in criticality of the BCT bifurcation, is reminiscent of a Chenciner (generalized Neimark-Sacker) bifurcation, a co-dimension-2 bifurcation that was observed in the smooth system by Keane and Krauskopf [66] where it produced rich dynamics.

If the T bifurcation lies to the right of the $D\bar{D}$ curve, this produces regions where vertical stripes can occur between the BCT bifurcation of the $1:1$ solution and the $D\bar{D}$ curve. Note that the vertical stripes do not stretch all the way between the $D\bar{D}$ curve and the BCT bifurcation. There is a second condition that a region must satisfy for the existence of vertical stripes, which is shown by the dashed black curve in Fig. 1.3. Vertical stripes only occur where the order of D, $\bar{D}$, H, $\bar{H}$ symbols in the stable solution is consistent with the $1:1$ solution. The difference between stable $P:P$ solutions and the unstable $1:1$ solution they coexist with lies only in the order of Z H, $\bar{Z}$ and $\bar{H}$ symbols. This arises because these solutions are generated when Z, H, $\bar{Z}$ and $\bar{H}$ symbols swapped order in the sequence representation of the $1:1$ solution at $\tau = \frac{n}{2}$, $n \in \mathbb{N}$.

1.7 Conclusion

We thoroughly analysed an elementary two-parameter system which combines the effects of time delayed feedback and periodic forcing. In spite of its simplicity, it demonstrates a complex structure of Arnold tongues with zero-width shrinking points and a high degree of multistability. Due to the system being piecewise-linear, we are able to solve the system analytically using an iterative map. We investigate the existence and stability of solutions through the development of a symbolic representation of solutions and the analysis of the subsequently developed Poincaré and border collision maps. This analysis reveals that the Arnold tongues are bounded by curves of border collision saddle-node bifurcations of periodic orbits. Additionally, we find curves of torus bifurcations connected to curves of border collision torus bifurcations, and investigate changes in the criticality of these bifurcations.

Our analysis sheds new light onto previously obtained results in related smooth systems, particularly the El Niño Southern Oscillation climate model studied by Keane, Krauskopf and Postlethwaite [13]. Comparing the numerically calculated bifurcation structure found in that paper with the analytically calculated bifurcation structure found here reveals that the prominent features of the smooth system (1.1,1.2) are preserved in the non-smooth limit of $\kappa \to \infty$. Indeed, the solutions in the smooth system generally appear as "smoothed out" counterparts to the piecewise-linear solutions of the non-smooth system. There are some significant differences. The tongues in the smooth system do not feature shrinking points, which are expected only in non-smooth systems, according to Simpson and Meiss [67]. Additionally, the smooth system features period doubling bifurcations, which we do not observe in our system. This is likely due to the complete lack of smoothness in our system, which prevents small dynamical variations in feedback strength and forcing strength. However, the analysis presented here does provide new insights into dynamics previously observed numerically [27, 13].

Both delayed feedback and periodic forcing are very common mathematical model ingredients and can be found in a variety of models used to study, for example, laser dynamics [68] and chimera states [69]. The current work reveals the phenomena which are a genuine consequence of this combination. Therefore, we expect this work to be of interest to a wide readership.

Chapter 2

Dynamics of a band-pass filter system with switched time-delayed feedback

The material in this chapter is from a paper that I am co-authoring with Lucas Illing and Andreas Amann. My contribution to this paper consisted of deriving the map which analytically models the piecewise-linear system, exploring the parameter space and demonstrating the dynamics of the underdamped regime, proposing the mechanisms by which observed solutions cease to exist, and numerically determining the criticality of the subcritical torus bifurcations.

2.1 Abstract

We consider a piecewise-linear second-order delay differential equation which is representative of feedback systems with switches (relays) that actuate after a fixed delay. In particular, our model describes a single-input single-output system with delayed self-feedback in which the feedback is a band-pass filtered relay signal. We present a detailed study of periodic solutions and their bifurcations. Starting from an integro-differential model, we show how to reduce the system to a set of finite-dimensional maps. We demonstrate that the stability of solutions can be understood in terms of smooth bifurcations of maps, and the existence of solutions can be understood in terms of border-collision bifurcations corresponding to transitions between maps.

2.2 Introduction

Many naturally occurring and technological systems evolve based not only on their current state, but also on their state some time in the past, making them time-delay systems. The effect of time delay has been studied in a wide variety of systems, including photonic [39, 21], optoelectronic [70, 71, 72], electronic [73, 74], neuroscience

[75], and climate systems [26, 47, 13, 42]. Delay differential equations (DDEs) that arise from time-delay systems have, generically, an infinite dimensional state space [76], which makes analytic study of DDEs challenging, but introduces a wealth of complex dynamic behaviour on multiple time scales [77, 72, 70, 76, 78]. The ubiquity of time delay, and the complexity of the induced dynamics, have made it a phenomenon of great interest to researchers. Time delay plays an important role in applications such as chaos control [79, 80, 81, 82, 83, 84], communication [85] as well as computing [86].

The combination of time-delayed feedback and nonlinearity has been observed to give rise to a wealth of possible dynamics, such as multiple coexisting attractors and oscillatory behavior that ranges from periodic waveforms to high-dimensional chaos as well as hybrid states such as chaotic breathers [77, 72, 70, 78, 76]. In many applications, nonlinear terms containing time delays are well approximated by functions that take on discrete values. An important example are relay control systems [87, 88, 18]. In this approximation, DDEs become non-smooth. Often this leads to significant simplifications of the analytic treatment with the trade-off being that the interplay between the discontinuous nonlinearities (switching events) and delay in the switching functions leads to new types of bifurcation scenarios, referred to as discontinuity-induced bifurcations or border-collision bifurcations [14, 15, 16, 17, 18]. As the study of such bifurcation scenarios is still relatively recent, there is still much to be learned by studying the dynamics that can arise from simple non-smooth delay differential equations.

Let us consider a simple second-order delay differential equation derived from a widespread technology in communications and signal processing. A *band-pass filter* is a system or device that takes a input signal, passes frequencies within a bandwidth centred on a centre frequency, and attenuates (diminishes) frequencies outside that range. Such devices are used broadly in signal processing. The dynamics of a band-pass filter can be described by the system of equations [89, 90, 70]

$$\begin{aligned} Q\Omega^{-1}\dot{x} &= -x - y + f\left(t\right) \\ \dot{y} &= Q\Omega x \end{aligned} \tag{2.1}$$

where $f\left(t\right)$ is the input signal, Ω is the centre frequency, Q is the quality factor, defined as the ratio of the centre frequency divided by the bandwidth. A complete derivation of this system can be found in Appendix A.2.1. When we drive the band-pass filter Eq. (2.1) with input signal $f(t) = \cos\left(\omega t\right)$, the output $x\left(t\right)$ is a signal with frequency ω whose amplitude depends on Q, and ω relative to Ω, as shown in Fig. 2.1. Outside of a given bandwidth determined by Q, the amplitude of the output signal decreases according to a power law as ω moves away from Ω. To explore the effects of delayed relay feedback in the context of a band-pass filter, we

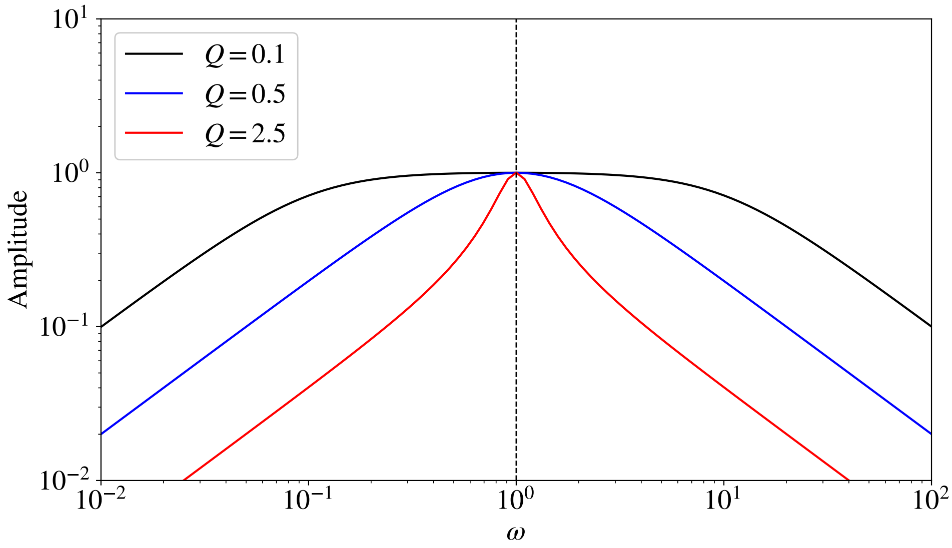

Figure 2.1: Band-pass filter output amplitude with input $f(t) = \cos(\omega t)$ for varying input frequency ω, where the centre frequency $\Omega = 1$.

drive Eq. (2.1) with time-delayed switched negative feedback

$$f(t) = -\mathrm{sgn}(x(t-\tau)) \tag{2.2}$$

where

$$\mathrm{sgn}(x) = \begin{cases} +1 & x > 0 \\ -1 & x < 0 \\ 0 & x = 0 \end{cases} \tag{2.3}$$

yielding the non-smooth delay-differential equation

$$\begin{aligned} Q\Omega^{-1}\dot{x} &= -x - y - \mathrm{sgn}(x(t-\tau)) \\ \dot{y} &= Q\Omega x \end{aligned} \tag{2.4}$$

As a consequence of the discontinuous input signal, Eq. (2.4) can be simplified by rescaling t w.r.t. τ, yielding the system that we study in this paper

$$\begin{aligned} Q\Omega^{-1}\dot{x} &= -x - y - \mathrm{sgn}(x(t-1)) \\ \dot{y} &= Q\Omega x \end{aligned} \tag{2.5}$$

The system is equivariant, possessing the symmetry

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \to \begin{pmatrix} -x(t) \\ -y(t) \end{pmatrix} \tag{2.6}$$

which is a prerequisite for some of the dynamical behaviour observed in this system. While $\mathrm{sgn}(x(t-1))$ is constant, the feedback is constant and the system is linear.

However, the system is discontinuous at changes in the feedback, and so the system is piecewise-linear. Therefore, while delay differential equations are generically infinite-dimensional, the state of this system is fully specified by knowing $x(t)$, $y(t)$, and the *zero elements*, which we define in the same manner as we did in Chapter 1 as the times at which the trajectory generated by Eq. (2.5) passed through $x = 0$ in the time interval $(t-1, t]$. Once again, we have no *a priori* assumptions of how many zero elements there can be at any given time t. Therefore, we let n be the number of zero elements, and define the state of the system $S(t)$ as

$$S(t) = (x, y; z_0, z_1, ..., z_{n-1}) \tag{2.7}$$

where $(x, y) \in \mathbb{R}^2$ is the position at time t and $t - \tau < z_0 < z_1 < ... < z_{n-1} \leq t$ are the zero elements. Again, while this system is infinite-dimensional, the dynamics of the system are restricted to a subspace of varying dimension $n + 2$ where $n \in \mathbb{Z}^+$.

2.3 Analysis

2.3.1 Dynamics of the linear subsystem

As the system is linear between changes in the time-delayed feedback, it is useful to study the dynamics of the linear subsystem where the feedback is constant. Eq. (2.5) can be written in matrix form as

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \Omega Q^{-1} \mathrm{sgn}\left(x\left(t-1\right)\right) \\ 0 \end{pmatrix} \tag{2.8}$$

where

$$A = \Omega \begin{pmatrix} -Q^{-1} & -Q^{-1} \\ Q & 0 \end{pmatrix} \tag{2.9}$$

Eq. (2.8) has a single fixed point

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ -\mathrm{sgn}\left(x\left(t-1\right)\right) \end{pmatrix} \tag{2.10}$$

which lies on the switching manifold $x = 0$. The fixed point flips between $(0, 1)^T$ and $(0, -1)^T$, depending on the state of the feedback. By taking $\det(A) = 0$, we obtain the eigenvalues of the fixed point

$$\lambda_{\pm} = \Omega \frac{-1 \pm \sqrt{1 - 4Q^2}}{2Q} \tag{2.11}$$

For $\Omega > 0$, $Re\left(\lambda_{\pm}\right) < 0$, so that the fixed point is always stable. In the over-damped regime $Q < \frac{1}{2}$, the eigenvalues are real and the fixed point is a sink. In the underdamped regime $Q > \frac{1}{2}$, the eigenvalues are complex and the fixed point is a stable spiral; this causes the trajectory to wind around the fixed point as it approaches. Thus, the fixed point is always attracting, and any more complicated dynamics must arise from a combination of switching and the transient behaviour of the linear system. In our analysis, we will use the fixed point of the linear subsystem as a point of reference; it is not a fixed point of the full system.

2.3.2 Reducing the system to a map

We can make use of the piecewise-linear nature of the system to simulate it analytically. A linear system

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} \tag{2.12}$$

with fixed point at the origin has solution

$$\boldsymbol{x}\left(t+T\right) = F_T\left(\boldsymbol{x}\left(t\right)\right) \tag{2.13}$$

for some $T > 0$ where F_τ is defined using the matrix exponential [91]

$$F_T\left(\boldsymbol{x}\right) = e^{AT}\boldsymbol{x} \tag{2.14}$$

which is calculated in Appendix A.2.3. We introduce a change of variables G that shifts the fixed point to the origin

$$G\left(\begin{array}{c} x \\ y \end{array}\right) = \left(\begin{array}{c} x \\ y + \operatorname{sgn}\left(x\left(t-1\right)\right) \end{array}\right) \tag{2.15}$$

G remains constant as long as the feedback remains constant. Thus, we can simulate the system analytically between changes in the feedback by

$$\left(\begin{array}{c} x\left(t+T\right) \\ y\left(t+T\right) \end{array}\right) = G^{-1}F_T G\left(\begin{array}{c} x\left(t\right) \\ y\left(t\right) \end{array}\right) \tag{2.16}$$

where $T = \tau + z_0$. This allows us simulate the system exactly as a map between changes in the feedback. However, to know when the changes in feedback occur, we also need to know when the trajectory generated by the system has passed through $x = 0$. This shows that the key events in this system are zero crossings and feedback changes [26]. Thus, we simulate the system as a map between Z events, or zero crossings, where the trajectory passes through $x - 0$ and a new history element is appended to S, and H events, where the oldest history element is removed from S and the feedback changes. The H events, in particular, are of great interest when we come to analyse the dynamics of the system.

2.3.3 Sample solutions

We now present sample solutions which demonstrate some of the characteristic features of this system. Fig. 2.2 shows eight stable solutions found by simulating the system from different initial conditions and parameters (Ω, Q). In each figure, t has been shifted so that a change in feedback occurs at $t = 0$ and thus, a zero crossing occurs at $t = -1$.

Solutions (a)-(g) are periodic solutions; a periodic solution with period P is a trajectory that follows a cycle such that $S(t+P) = S(t)$. Solution (a) demonstrates behaviour that is characteristic of the overdamped regime. The trajectory decays exponentially towards the fixed point until the feedback changes and the fixed point flips, then the trajectory decays towards the fixed point again, passing through $x = 0$ shortly after the feedback changes. This solution is multistable with higher-frequency solutions (b) and (c).

In order to better capture the difference between multistable solutions, we introduce the *mode* ν. ν can be defined as the number of times the solution passes through $x = 0$ in an open time interval of length 1 that begins with a zero crossing; in Fig. 2.2, the interval $(-1, 0)$ is convenient for this. Solutions (a)-(c) have mode $\nu = 0$, $\nu = 2$ and $\nu = 4$ respectively. This definition is equivalent to defining ν as the number of zero elements (n) in $S(t)$ at a time t when the feedback changes. Defining ν in this way allows us to compare these multistable solutions to the $1:1$ solution analysed in the previous chapter. Instead of having a solution whose appearance changed periodically with τ and whose stability changed according to a Poincaré map with $n = \lfloor 2\tau \rfloor$ variables, here we have multistable solutions of varying ν whose stability depends on a Poincaré map with $\nu + 2$ variables.

The underdamped regime features significant multistability, but also contains interesting dynamics not present in the overdamped regime. Solutions (d)-(f) show what happens to the $\nu = 0$ solution as Ω is varied with Q fixed. Solution (d) demonstrates behaviour that is characteristic of the underdamped regime; the trajectory decays towards the fixed point while winding around it until the feedback changes and the fixed point flips, and the trajectory winds around the fixed point again, passing through $x = 0$ along the way. As Ω increases, the frequency of the solution increases, and the trajectory winds further around the fixed point before the feedback changes; in (e), the feedback changes just before the trajectory passes through $x = 0$. In (f), Ω has increased further, and the trajectory passes through $x = 0$ *before* the feedback changes. This continous deformation of the solution as the parameters are varied causes ν to increase from 0 to 1. This pattern of behaviour is consistent for all solutions in the underdamped regime; for $m \in \mathbb{Z}^+$, a $\nu = 2m$ solution continuously changes to $\nu = 2m + 1$. This has a significant impact on the solution; when the change in feedback passes through $x = 0$, the dimension of the solution increases by 1. Due to the importance of feedback changes, we will make

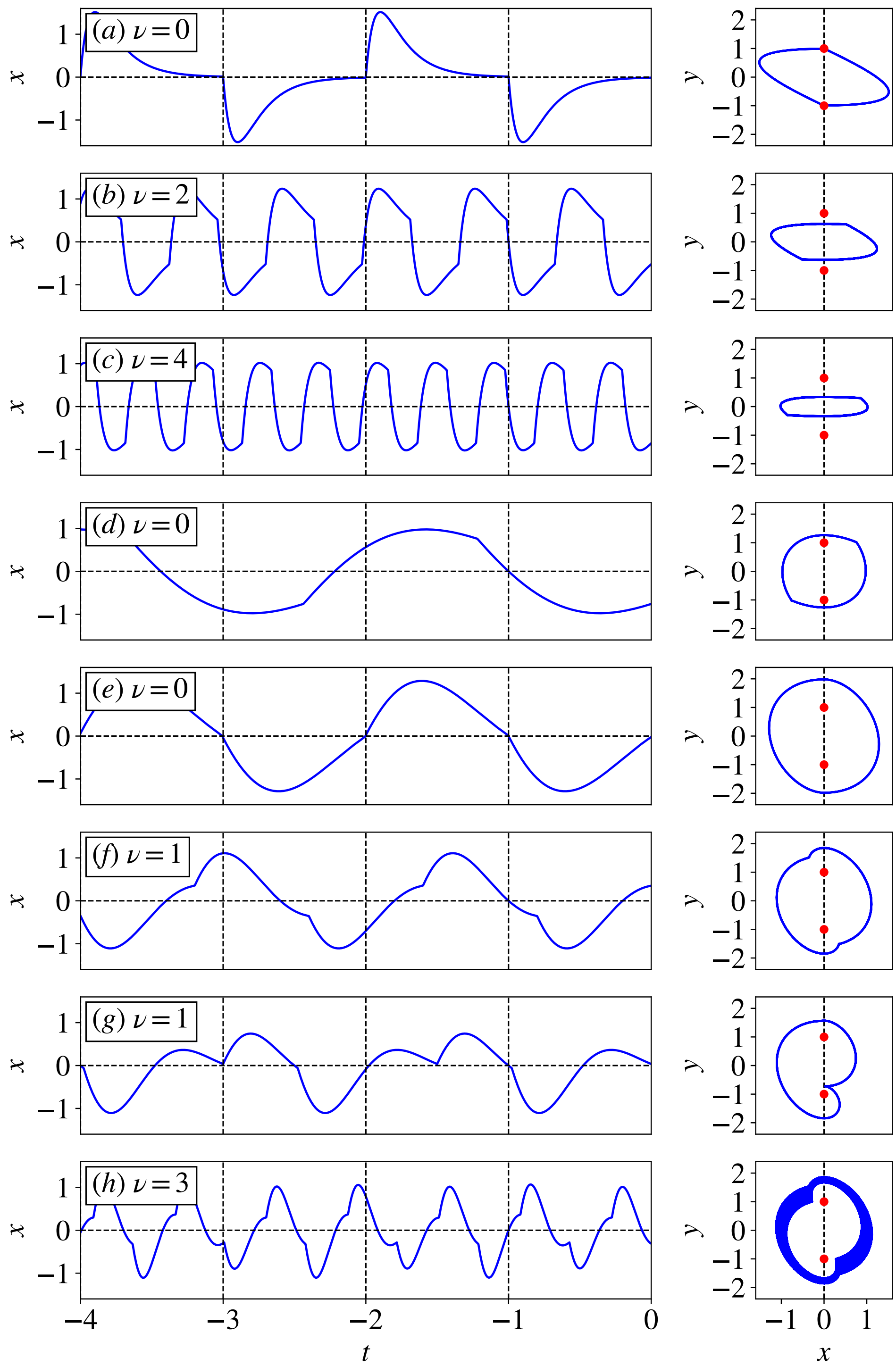

Figure 2.2: Sample solutions, excluding transients, obtained from simulating the system from different initial conditions. In (a)-(c), $(Q, \Omega) = (0.45, 3\pi)$. In (d)-(h), $Q = 1.5$, and Ω increases from $\Omega = 0.65\pi$ in (d) to $\Omega = 1.05\pi$ in (e), $\Omega = 1.65\pi$ in (f), $\Omega = 2.1\pi$ in (g) and $\Omega = 4.7138\pi$ in (h). In the (x, y) plots, the fixed point of the linear subsystem is plotted in red.

constant reference to the point at which it changes. We define $(x_H, y_H)^T$ as the position at which the feedback changes and a history element is deleted from the state S. We can calculate where in the (Ω, Q) plane $x_H = 0$ for a $\nu = 2m$ solution by noting that this occurs when the elapsed time between zero crossings is $\frac{1}{2m+1}$, as is visible in solution (e). Therefore, we can calculate curves in the (Ω, Q) plane along which solutions transition smoothly from $\nu = 2m$ to $\nu = 2m+1$ by equating

$$\frac{1}{2m+1} = \frac{\pi}{\frac{\Omega}{2Q}\sqrt{4Q^2-1}} \tag{2.17}$$

which yields

$$\Omega = \frac{2Q\pi\,(2m+1)}{\sqrt{4Q^2-1}} \tag{2.18}$$

Solutions (g) and (h) show more unusual types of stable solution. Solution (g) shows what happens as we continue to increase Ω from solution (f); the symmetry of the solution is broken, yielding an asymmetric solution. This asymmetric solution is observed for a small window in Ω before the trajectory falls off this solution and changes to a $\nu = 2$ solution. However, that same $\nu = 2$ solution, upon being swept up to larger Ω, does not break symmetry. Instead, it first transitions smoothly to a $\nu = 3$ solution as the frequency increases, before changing to an aperiodic solution (h), which is likewise observed briefly before the trajectory falls off this solution and changes to a $\nu = 4$ solution.

2.3.4 Multistability

We now examine the multistability observed in Fig. 2.2 in greater detail. Fig. 2.3 show the results of sweeping from $\Omega = 12\pi$ down to $\Omega = 0$ and back up for fixed Q, beginning at $\Omega = 12\pi$ with the stable $\nu = 12$ solution.

Consider Fig. 2.3 (a). As we sweep down in Ω, the frequency of the swept solution decreases in large jumps, between which the decrease in frequency is smooth and relatively small. These large jumps in frequency coincide with jumps in the observed value of ν. Note the inconsistent horizontal lines in the lower third of the plot; these artefacts occur where the swept solution drops from an unstable high ν solution to a stable low ν solution, skipping over an stable intermediate ν solution. For example, the largest value of Q at which this is observed is around $Q \simeq 0.6$,where the swept solution goes from $\nu = 12$ to $\nu = 8$, skipping $\nu = 10$ and leaving a yellow line surrounded by orange. This may occur because curves of bifurcations grow closer together for small Q, leaving the basin of attraction of the intermediate ν solution relatively small compared to that of the low ν solutions at the point where the high ν solution becomes unstable.

In Fig. 2.3 (b), we observe a significant difference between the underdamped

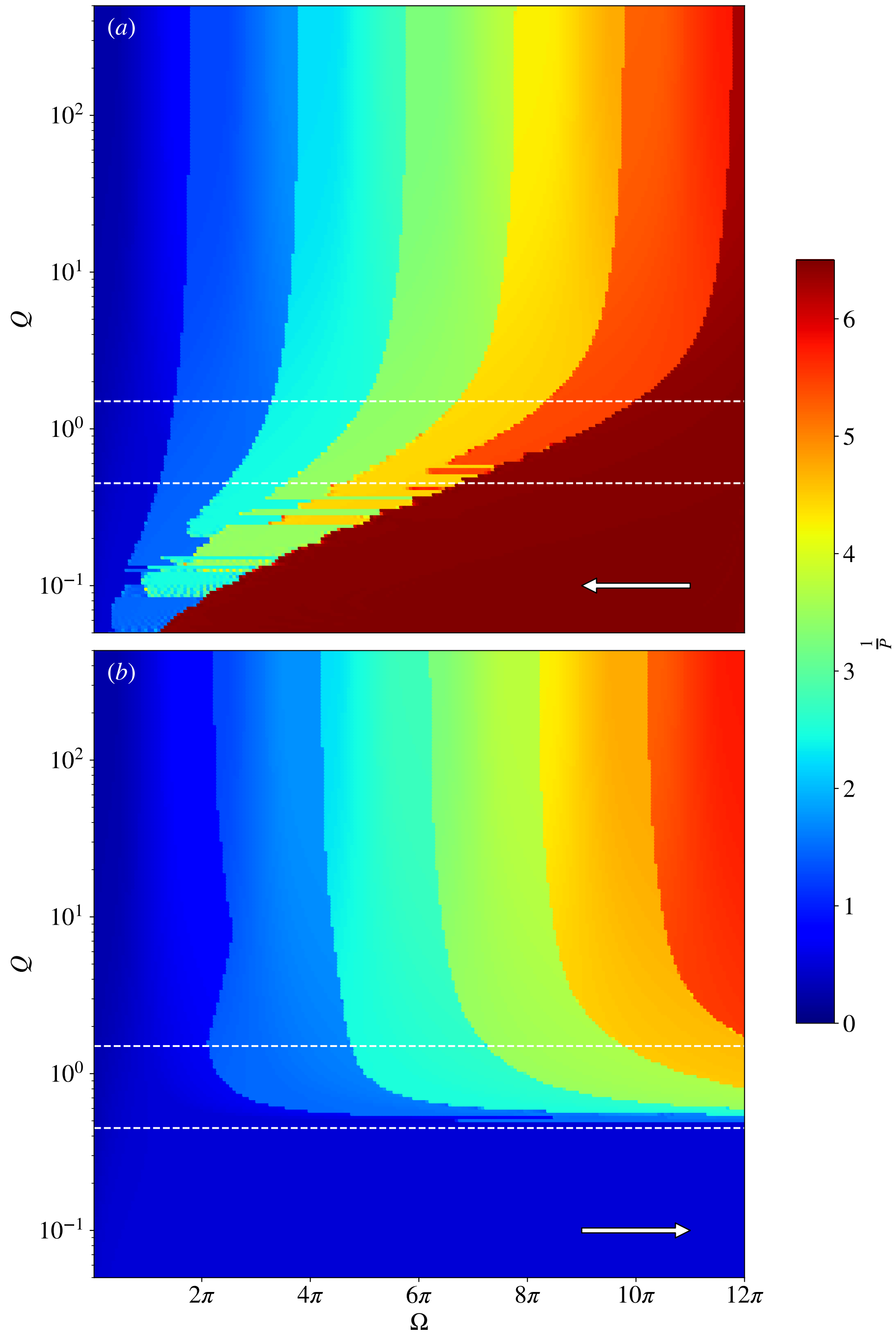

Figure 2.3: Numerical sweeps of the (Ω, Q) plane, where the colour indicates the frequency of the swept solution. The white arrow indicates the sweep direction. The dashed white lines show the Q values along which we conduct deeper analysis in the overdamped ($Q = 0.45$) and underdamped ($Q = 1.5$) regimes.

and overdamped regimes. In the underdamped regime $\left(Q > \frac{1}{2}\right)$, as we sweep up in Ω, the frequency of the swept solution increases in large jumps, between which the increase in frequency is smooth and relatively small. These large jumps in frequency coincide with jumps in the observed value of ν. If we combine this observation with our observations from the Fig. 2.3 (a), we can deduce that solutions of given ν (and hence are within a given frequency band) are stable within some band centred on some value of Ω, and that this band is wider for small Q. This makes some kind of intuitive sense; bandpass filters attenuate signals outside a given frequency band, so in our bandpass filter with self-feedback, we observe solutions outside a given frequency band becoming unstable. However, this does not happen in the overdamped regime! We observe that the $\nu = 0$ solution shown does not become unstable as Ω increases; as Ω increases in the overdamped regime, the multistability increases as more and more solutions become stable.

Fig. 2.3 demonstrates that the system exhibits extensive multistability and a variety of types of solution. However, we have not yet determined *why* the swept solution ceases to follow a particular solution. To find out whether this is due to a loss of stability or a loss of existence, we make use of our iterative map of the system to numerically continue a stable solution into regions where they are unstable. In Fig. 2.4, we plot the frequency of the periodic solutions observed through sweeping, where solid lines indicate stable solutions and dashed lines indicate unstable solutions. We observe that in the overdamped regime shown in Fig. 2.4(b), as we sweep down in Ω, the swept solution decreases in frequency when higher frequency solutions become unstable. The $\nu = 0$ solution is not unique in remaining stable as Ω increases, though it is unique in never becoming unstable as Ω decreases. We also note that these solutions continue to exist as $\Omega \to \infty$, and for a solution with $\nu = 2m$, the frequency tends to $\frac{2m+1}{2}$ as $\Omega \to \infty$.

Once again, the underdamped regime, shown in Fig. 2.4(a) is more complicated. Like the overdamped regime, as we sweep down in Ω, the swept solution jumps down in frequency when higher frequency solutions become unstable. However, as we sweep upwards, the swept solution jumps up in frequency as lower frequency solutions become unstable. Furthermore, as Ω increases further, the unstable solutions cease to exist.

We can show a more complete picture of the dynamics in the (Ω, Q) plane by focusing on solutions with specific values of ν. Fig. 2.5 shows the existence and stability for the $\nu = 0, 1$ solution and the $\nu = 2, 3$ solution. We find that there is a region in Fig. 2.5(a) where the symmetric $\nu = 1$ solution becomes asymmetric, as seen in Fig. 2.2(g). There should also be an extremely narrow band in Fig. 2.5(b) where the $\nu = 3$ solution becomes unstable and a stable closed invariant curve exists, as seen in Fig. 2.2(h); however, it is too narrow to plot clearly at this scale. To understand the cause of these changes in stability and existence, we consider the

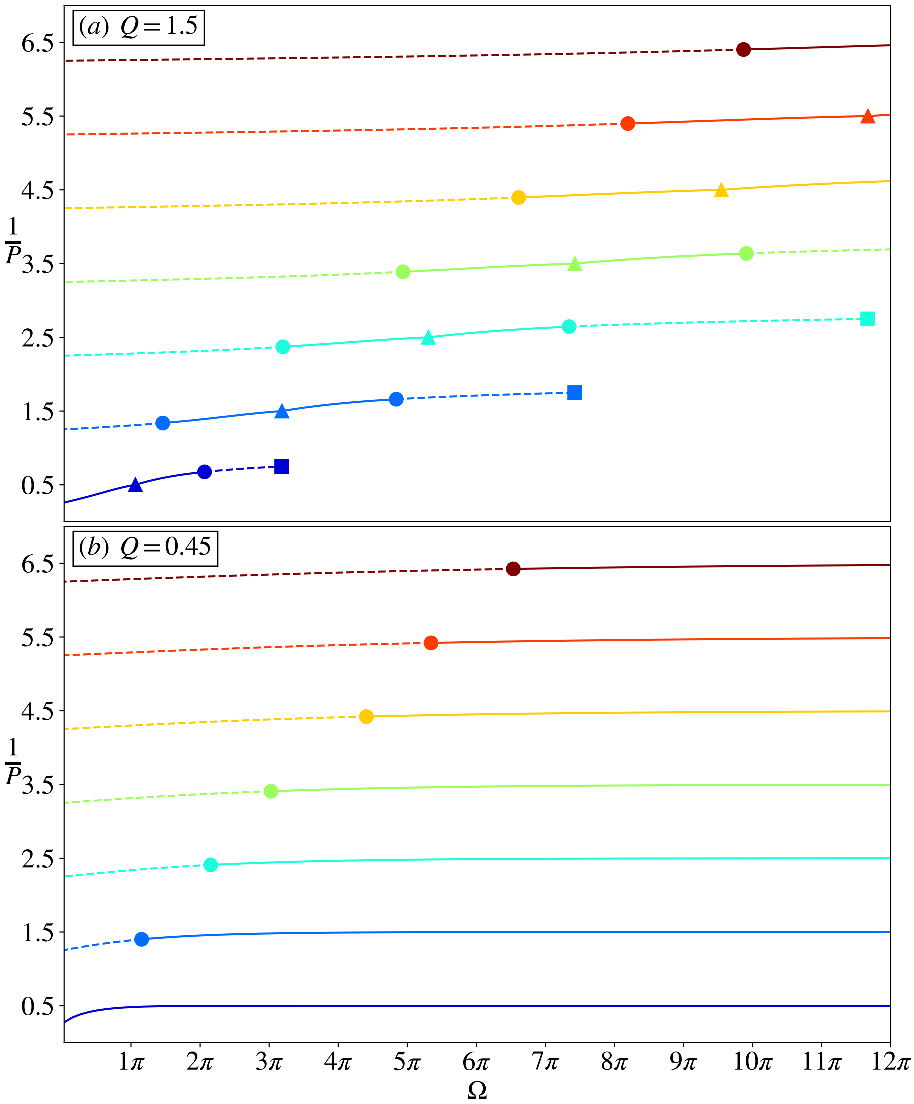

Figure 2.4: Frequency of solutions plotted against Ω for different ν solutions. Solid lines show where a solution is stable, dashed lines show where the solution is unstable. Triangles show where ν changes from even to odd, circles show where the solution undergoes a torus bifurcation, and squares show where a solution vanishes in a border-collision saddle-node bifurcation of periodic orbits.

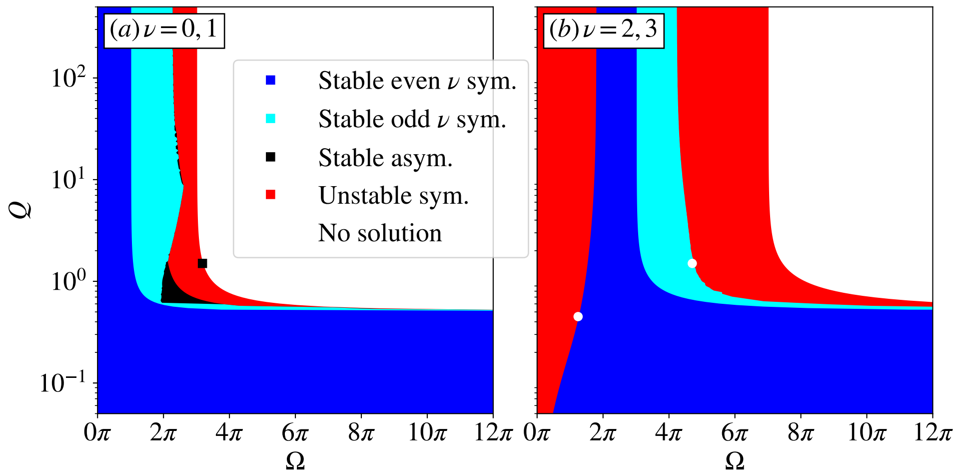

Figure 2.5: Existence and stability in the (Ω, Q) plane for (a) the $\nu = 0, 1$ solution and (b) the $\nu = 2, 3$ solution. The white circles indicate the location of the torus bifurcations shown in Fig. 2.6 and Fig. 2.7. The pitchfork bifurcation shown in Fig. 2.8 takes place near the top of the black triangular region. The black square indicates the location of the saddle-node bifurcation shown in Fig. 2.9.

dynamics of this system.

2.4 Smooth dynamics

2.4.1 Subcritical torus bifurcations of even $\nu \geq 2$ solutions

As Ω decreases, all even $\nu \geq 2$ solutions become unstable along curves in the (Ω, Q) plane that appear to branch out from the the origin, without generating any other observable phenomena. To understand why this occurs, we experimentally map out the basin of attraction of the stable $\nu = 2$ solution as we approach the curve along which it becomes unstable. We do this by simulating the $\nu = 2$ solution, perturbing the parameters just across the curve, allowing the system to evolve for a period of time, then resetting the parameters and observing whether the system restabilized to the $\nu = 2$ solution. Fig. 2.6 shows the results of this experiment as a Poincaré section through x_H [92]. The basin of attraction takes the characteristic shape of a subcritical Neimark-Sacker or secondary Hopf bifurcation; as we approach the bifurcation curve, an unstable closed invariant curve (inferred from our experiment and plotted in dashed red) collapses onto the stable solution, causing it to become unstable. Thus, as we sweep down in Ω, the periodic solutions become unstable through subcritical torus bifurcations.

2.4.2 Supercritical torus bifucations of odd $\nu \geq 3$ solutions

As Ω increases, all odd ν solutions eventually become unstable. For $\nu \geq 3$, when they become unstable, the Poincaré section through x_H generates a stable closed

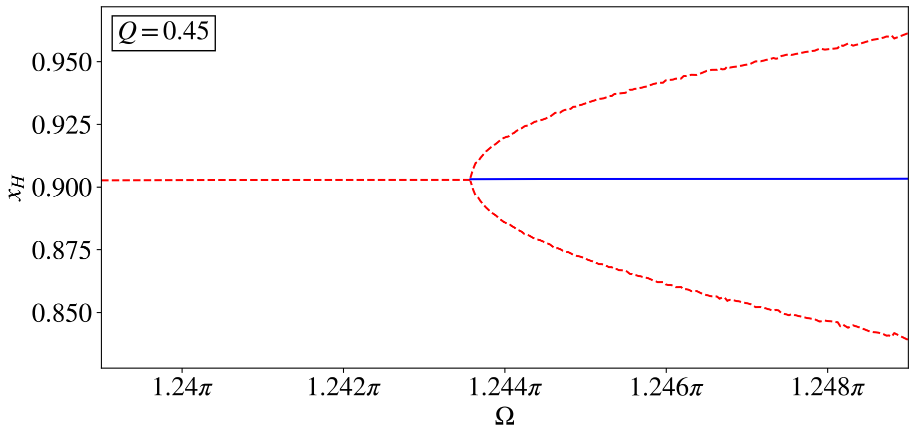

Figure 2.6: The subcritical torus bifurcation of the $\nu = 2$ solution in the overdamped regime indicated by a white circle in Fig. 2.5.

invariant curve, as shown in Fig. 2.7, where we follow the $\nu = 3$ solution which corresponds to the $\nu = 2$ solution shown in Fig. 2.6. This is consistent with a supercritical Neimark-Sacker bifurcation, and so the periodic solution undergoes a supercritical torus bifurcation, generating the aperiodic solution observed in Fig. 2.2(h).

2.4.3 Supercritical pitchfork bifurcation of the $\nu = 1$ solution

In the case of the $\nu = 1$ solution, we do not observe a torus bifurcation generating a closed invariant curve. Instead, a supercritical pitchfork bifurcation generates a pair of stable mirrored asymmetric solutions, one of which was shown in Fig. 2.2(g). The bifurcation diagram of the pitchfork bifurcation is shown in Fig. 2.8. The speed at which the asymmetric solutions move apart after the bifurcation makes it look superficially like a nonsmooth bifurcation; however, the inset plot in Fig. 2.8(a), which covers a much smaller range of $\Omega \in [6.5278, 6.528]$ shows that they do spread apart smoothly.

An interesting feature visible in Fig. 2.8(b) is that the asymmetric solutions cease to exist when $x_H = 0$. This is noteworthy because if x_H was to pass through $x = 0$, it would have significant consequences for the solution; there would be additional zero crossings, and hence, additional feedback changes, which would cause the solution to change in a manner that is continuous at $x_H = 0$, but not smooth across $x_H = 0$. In order to consider these effects, we must consider the non-smooth dynamics of the system.

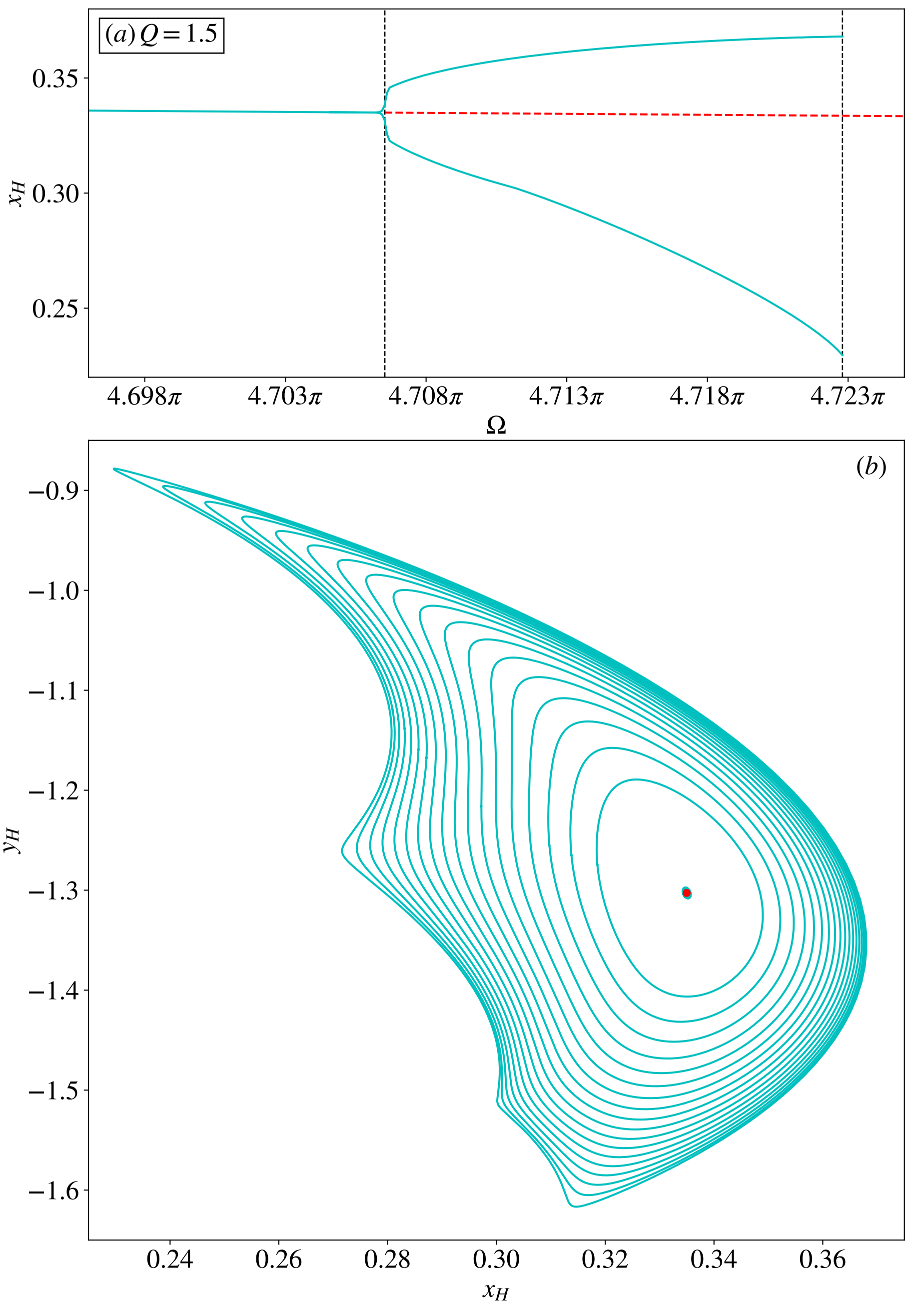

Figure 2.7: The supercritical torus bifurcation of the $\nu = 3$ solution indicated by a white circle in Fig. 2.5. In (b), the torus is plotted in the (x_H, y_H) plane for 20 evenly spaced values of Ω between the dotted black lines in (a).

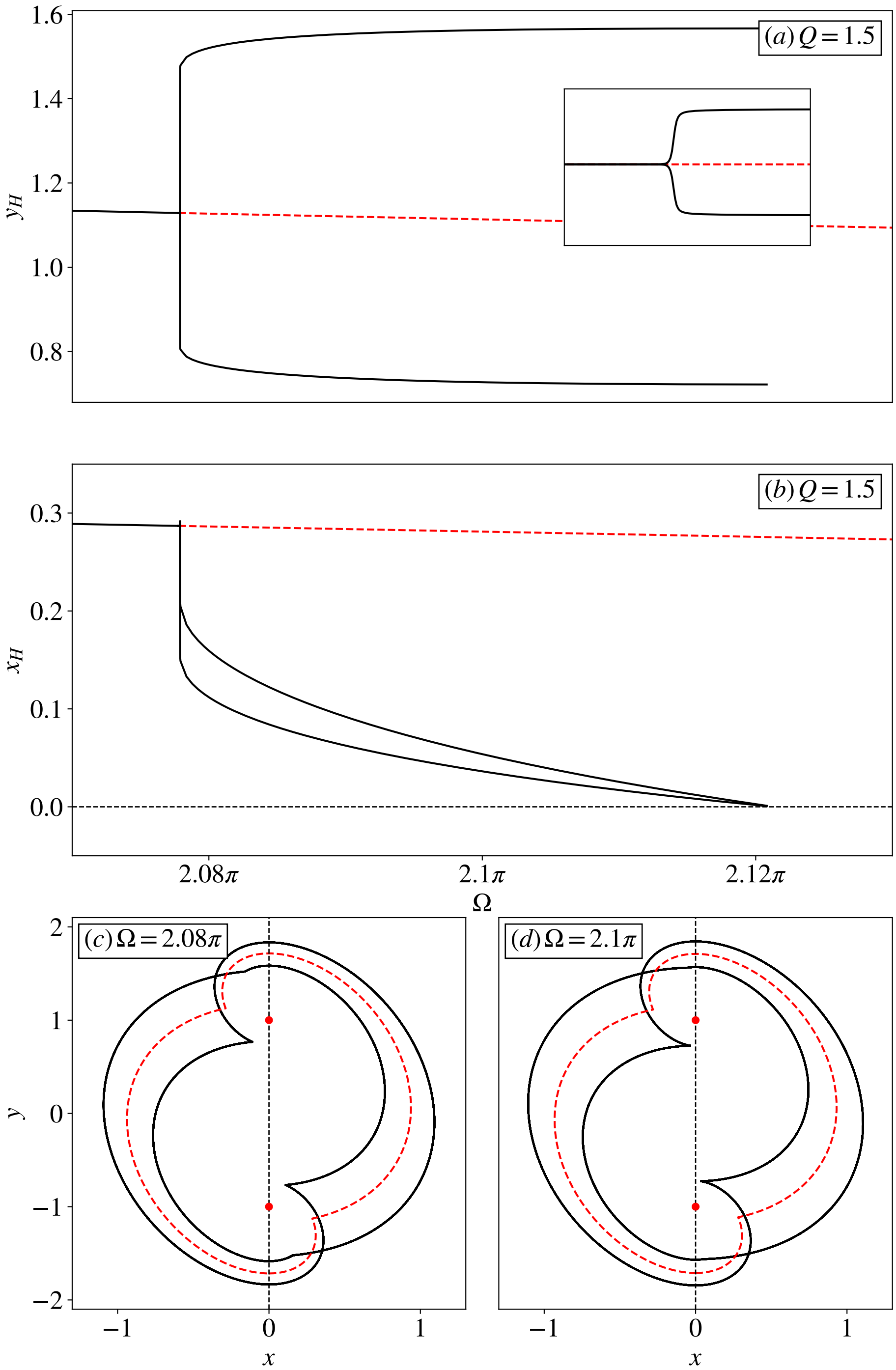

Figure 2.8: Pitchfork bifurcation of the $\nu = 1$ solution.

2.5 Non-smooth dynamics

2.5.1 Border-collision saddle-node bifurcations of symmetric solutions

In the underdamped regime, all odd ν solutions vanish when $x_H = 0$. Similar to how we calculated the smooth ν transitions in Eq. (2.18), we can calculate where this occurs for symmetric solutions by equating

$$\frac{1}{2(2m+1)+1} = \frac{\pi}{\frac{\Omega}{2Q}\sqrt{4Q^2-1}} \tag{2.19}$$

which yields

$$\Omega = \frac{2Q\pi(4m+3)}{\sqrt{4Q^2-1}} \tag{2.20}$$

However, in this case, the solution does not change smoothly when ν changes. Instead, the x_H that crosses over $x = 0$ forms the tip of a spike; when this spike reaches $x = 0$, it introduces two new zero crossings and hence, two new feedback changes. We use our iterative map to construct this 'crossed-over' $\nu = 5$ solution, and plot it in Fig. 2.9. The $\nu = 1$ solution shown here is the same solution that becomes unstable in a pitchfork bifurcation in Fig. 2.8. A close up of the spike crossing over is shown in Fig. 2.9(c) (recall that the solution winds anticlockwise in the (x, y) plane). Initially, the trajectory moves slowly as it decays close to the fixed point, before the feedback changes and the trajectory moves rapidly to decay towards the flipped fixed point. The corresponding change in feedback occurs close to $x = 0$ for $y < 0$. We find that as Ω increases, the crossed-over solution and the original solution collide and vanish in a degenerate border-collision saddle-node (BCSN) bifurcation of periodic orbits between two unstable solutions, as shown in Fig. 2.9(a).

As noted previously, Fig. 2.8 suggests that non-smooth dynamics are responsible for the vanishing $\nu = 1$ asymmetric solution. It would be reasonable to expect that a similar BCSN bifurcation of periodic orbits is the cause, and proving this would be of further interest.

2.5.2 Global bifurcation of the symmetric crossed-over solution

The discovery of the crossed-over $\nu = 5$ solution allows us to explain how the $\nu = 1$ solution vanishes in a BCSN bifurcation; we now consider how the crossed-over $\nu = 5$ solution might vanish. In Fig. 2.9, as Ω decreases, the x_H that reaches $x = 0$ at the BCSN bifurcation begins to curve back around towards $x = 0$, but the continuation

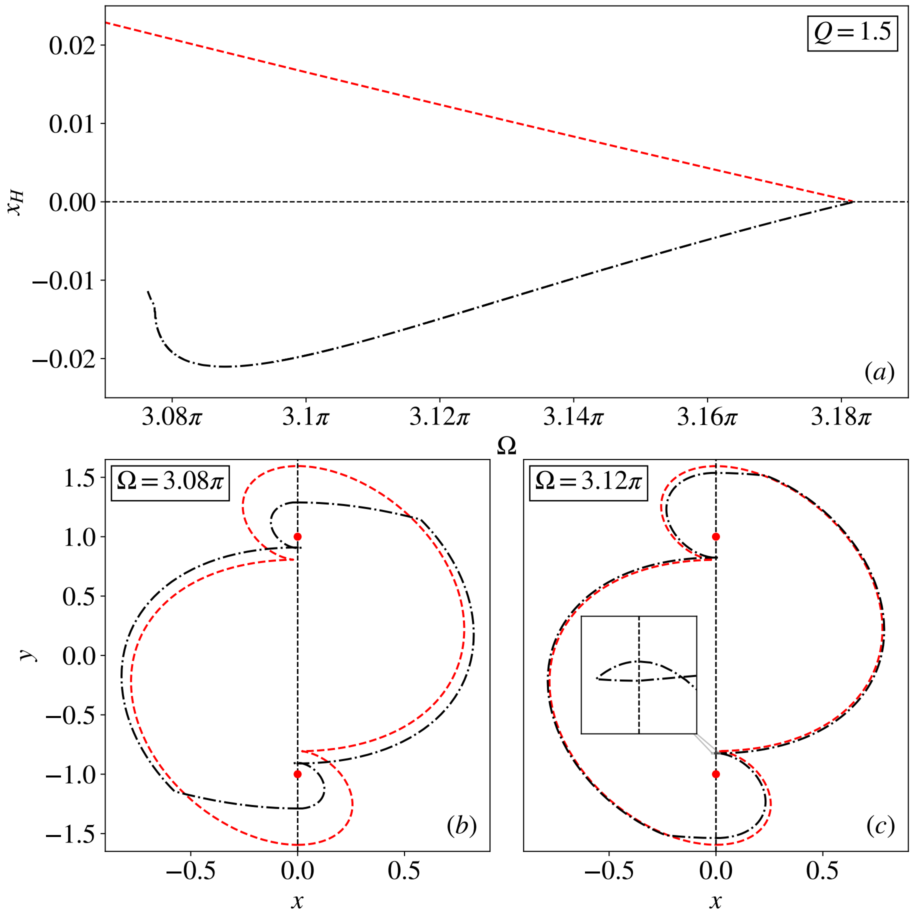

Figure 2.9: Degenerate border-collision saddle-node bifurcation of periodic orbits of the $\nu = 1$ periodic solution indicated by a black square in Fig. 2.5. The unstable solution observed while stable in Fig. 2.2$(d) - (f)$ is plotted in dashed red. The corresponding crossed-over $\nu = 5$ solution is plotted in dash-dot black.

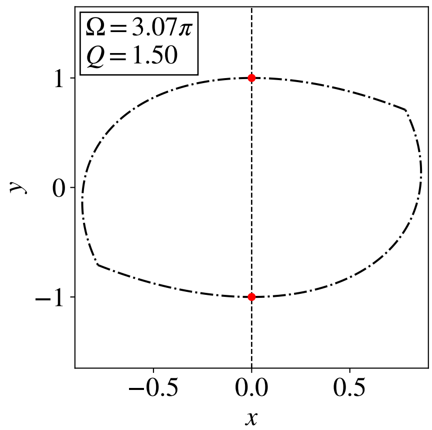

Figure 2.10: The limiting behaviour of the crossed-over $\nu = 5$ counterpart to the observed $\nu = 1$ symmetric solution shown in Fig. 2.9.

method employed here fails to complete the curve. However, we can note from Fig. 2.9 that as Ω decreases and we move away from the BCSN bifurcation in the parameter space, the x_H spike on the crossed-over solution moves towards the fixed point as $x_H \to 0$. While the spike is shorter, it occurs closer to the fixed point, and so the corresponding feedback change gets longer, as seen in Fig. 2.9(b). We can consider the limiting behaviour of this solution where the tip of the x_H spike reaches the fixed point, which we plot in Fig. 2.10.

Here, the interaction of the linear flow and the switched feedback cause the solution to intersect the points between which the fixed point switches. If Ω were to be decreased further, the solution would drop from $\nu = 5$ to $\nu = 1$ (and hence, the number of variables required to describe the state of the system drops from 7 to 3), but unlike the BCSN bifurcation above, this causes the feedback to change discontinuously, and so the solution ceases to exist.

It is difficult to attached a label to this bifurcation. The described behaviour is, of course, completely impossible in a smooth system. However, it may be a generic bifurcation in systems which feature smooth flow and switching fixed points (caused by switched feedback, or possibly, switched forcing), and so we should consider extending the nomenclature to piecewise-smooth (PWS) systems. The observed behaviour is comparable to both a homoclinic bifurcation, where a limit cycle vanishes after colliding with a saddle, and a heteroclinic bifurcation, where a limit cycle vanishes after colliding with a pair of saddles at two or more points on its orbit [93, 94]. However, in this system we do not have saddle points; instead, we having a stable spiral which switches dynamically between two points located on the switching manifold itself. Also, while the solution *intersects* the fixed point, it would not be accurate to say that it *collides* with it in any classical sense. The solution reaches the point $(0, 1)$ while spiralling into the point $(0, -1)$, but the feedback switches precisely when it reaches $(0, 1)$, just in time for it to be the fixed point, then switches back to allow the solution to continue spiralling into $(0, -1)$. This is a highly degenerate situation, but it is perhaps one in which a switching stable spiral might act like a pair of saddles.

It might be reasonable to refer to the observed behaviour as a *PWS heteroclinic bifurcation*. While we compared the behaviour to both a homoclinic and a heteroclinic bifurcation, heteroclinic bifurcations occur generically in systems with symmetry. This system has symmetry, and it is reasonable to expect that breaking that symmetry would cause the limit cycle to intersect the stable spiral at only one point, which might be better described as a *PWS homoclinic bifurcation*. However, a significant amount of further work will be required to determine the validity of these deductions.

2.6 Conclusion

In this chapter, we studied a piecewise-linear second-order delay-differential equation with delayed relay feedback. The system featured a complex bifurcation structure with extensive multistability. We simulated the system analytically using an iterative map with variable time-step, and analysed the bifurcations of the system through Poincaré sections. We found that the stability of solutions are governed by smooth bifurcations, and the existence of solutions are governed by non-smooth bifurcations.

There are two notably similar systems comparable to the one studied here. The first is an undamped harmonic oscillator with delayed relay feedback, derived from an optical model of pupil light reflex [95, 96, 18]. In this system, solutions consist of sections of circles, instead of the sections of spirals we observe. A key difference in this model is that while the feedback is relies on $x\,(t-\tau)$, it affects $\dot{y}$, rather than $\dot{x}$ as in our model. This causes the effect of the change in feedback to act parallel to the switching manifold, rather than perpendicular as in our model. Bayer and an der Heiden [96] observed saddle-node and pitchfork bifurcations in this system, in addition to extensive multistability. Barton, Krauskopf and Wilson [18] studied a slightly smoothed approximation of this system, and observed Neimark-Sacker bifurcations in addition to those previously found. They also broke the symmetry of the system to unfold the bifurcations.

Sieber [88] analysed an undamped oscillator with delayed relay feedback derived from an inverted pendulum. In this model, the feedback is comparable to our model, but with the addition of hysteresis. Hysteresis introduces a number of discontinuity-induced bifurcations not observed in our system, including corner-collision and grazing bifurcations. Sieber et al. [87] extended this analysis to a generic harmonic oscillator with an arbitrary switching manifold, rather than $x = 0$. The genericity of the system allowed it to yield a broader spectrum of smooth and non-smooth bifurcations.

There is signficant further work required to fully understand both this system, and whether the observed phenomena are generic to a wider class of PWS systems. Furthermore, the scenario in which the PWS heteroclinic bifurcation is described is extremely unusual, and perhaps of lesser impact because the limit cycle is unstable. It would be desirable to determine a simple normal form in which a stable limit cycle undergoes a PWS heteroclinic bifurcation. It would also be worthwhile to break the symmetry of this system to see what happens to the PWS heteroclinic bifurcation. We expect that further work in this area would be of great interest. Based on the observed dynamics, this system is worthy of further study, and may grant insight into a wider class of PWS system. As such, we hope that this work will be of interest to a wide audience.

Chapter 3

Dynamics of targeted ransomware negotiation

The material in this chapter is largely reproduced from P. Ryan, J. Fokker, S. Healy and A. Amann, *Dynamics of targeted ransomware negotiation*, IEEE Access, 10:3283632844, 2022. [30]. The Analysis section has been extended with additional material considering the economic conditions under which targeted ransomware attacks are viable. My contribution to this work consisted of developing, defining and analysing the model, while my co-authors authors provided guidance on which key features of ransomware negotiation should be replicated, and which dynamical features should be retained or removed to produce a model which is relatively realistic and simple.

3.1 Abstract

We consider how the development of targeted ransomware has affected the dynamics of ransomware negotiations to better understand how to respond to ransomware attacks. We construct a model of ransomware negotiations as an asymmetric non-cooperative two-player game. In particular, our model considers the investments that a cybercriminal must make in order to conduct a successful targeted ransomware attack. We demonstrate how imperfect information is a crucial feature for replicating observed real-world behaviour. Furthermore, we present optimal strategies for both the cybercriminal and the target, and demonstrate how imperfect information results in a non-trivial optimal strategy for the cybercriminal.

3.2 Introduction

Computer security is a rapidly developing field, with new threats emerging and evolving constantly. As computer security providers develop their methods for detecting malware (malicious software), the cybercriminals behind the various strains of malware are forced to refine their techniques for avoiding detection, prompting further development from the computer security industry. As a result of the interaction between these competing agendas, problems in computer security can give rise to rich dynamical behaviour. By analysing these dynamics, we can provide insights that assist in understanding phenomena observed in computer security. We seek to provide insight on recent developments in ransomware by using game theory to explore the dynamics they have introduced.

Ransomware is a type of malware designed to extort a ransom from the victim [6, 97], usually by denying the victim access to their computer or data until the ransom has been paid. In the past, ransomware relied on extorting a small amount of money from a large number of victims. The ransom itself would be fixed at a price low enough that nearly anyone could pay, but was typically non-negotiable, as negotiating a small ransom with many victims would not be worth the effort. In recent years, a new phenomenon known as *targeted ransomware* has emerged [4, 98]. Cybercriminals operating targeted ransomware specifically target large organisations, which can be extorted for significantly higher ransoms [5, 99]. However, they are also likely to have a higher level of computer security, and so the cybercriminals are forced to make significant investment into breaching their security. Given the effort involved in breaching security, the cybercriminals will invest further in calculating the highest ransom they believe their target will pay. As a result of the larger sums of money at stake, cybercriminals are willing to negotiate the ransom demand to facilitate payment. We consider these negotiations to be a crucial feature in targeted ransomware, as their outcome has an enormous impact on both the cybercriminals and the targeted organisation. In this chapter, we develop a model of targeted ransomware negotiations based on game theory. Game theory is a branch of mathematics which studies strategic interactions between rational decision-makers [100]. A game is a mathematical model with a clearly defined set of rules where two or more *players* make strategic decisions to influence the outcome of the game to their own personal benefit. By analysing the decisions available to players, optimal decision-making strategies can be determined which offer the best outcome for each player. Game theory can be applied to many real-life scenarios that involve competing interests by formulating a game as a mathematical abstraction of the given scenario. By analysing the decisions taken in the game logically, optimal strategies can be determined, granting insight into real-world behaviour. As such, game theory is highly applicable to fields such as ecology [101, 102, 103], economics [104, 105, 106, 107] and politics [108, 109]. Very recently, ideas from game theory have been found to be useful in the study of ransomware. While all ransomware follows the same fundamental principles, there is sufficient variety in observed phenomena to merit a variety of models, such as defence and deterrence [110, 111, 112, 113, 114], iterative negotiations [115], price discrimination [116], incentive to return encrypted data [117], and sale of stolen data [118]. Here, we examine how game theory can be applied to the growing threat of targeted ransomware. We construct a model of ransom negotiations as a game played between a cybercriminal and their target. The game focuses on the strategic behaviour and investments that the cybercriminal must commit to in order to implement a successful targeted ransomware attack. By analysing our model, we demonstrate how imperfect information is crucial for replicating observed real-world behaviour, and provide new insights into the real-world behaviour and strategy of cybercriminals operating strains of targeted ransomware.

3.3 Background

In order to construct a model of targeted ransomware negotiations, we first consider the key features of targeted ransomware. On a technical level, one of the main differences between untargeted and targeted ransomware is how it spreads [119]. Untargeted ransomware is distributed indiscriminately, relying on victims with weak security to make a mistake that allows the ransomware to infect their computer. Such a strategy may be appealing to a cybercriminal, as it requires a low investment of effort to infect victims. However, this indiscriminate strategy is relatively easy to defend against. A potential victim can reduce their chances of being infected through practices such as maintaining good security, careful internet usage, and maintaining up-to-date data backups. Large organisations with valuable data are likely to have such practices implemented across a complex network of computers. In order to target such organisations, a cybercriminal must invest significant effort in circumventing their security and spreading their ransomware across the computer network. This typically involves multiple attack vectors, such as targeted phishing, remote desktop protocol (RDP), and searching for weak passwords[120]. The cybercriminal may spend days or weeks escalating their level of access to the target's network; in 2019, the mean dwell time (the duration a threat is present in a system before it is detected) of ransomware was 43 days [121]. Only when the cybercriminal has a level of access high enough to compromise the target's backups will they start encrypting files, ensuring that the only way to restore the data is by paying the ransom. While this process of circumventing security requires a large investment of time and effort, it enables the cybercriminal to extort organisations for ransoms far larger than previously possible [122].

A crucial factor in any ransomware negotiation is the reliability of the cybercriminal. The victim's willingness to pay the ransom demand must be affected by the likelihood of getting their data back afterwards. There are two major reasons why a victim might get their data after paying. The first is that the cybercriminal chooses not to return the data, perhaps to avoid the cost incurred by doing so. This is not a sound business practice, as it results in a loss of perceived reliability that reduces a victim's willingness to pay, and hence, the cybercriminal's profits [117], which does not fit with the business-like stance that cybercriminals operating ransomware have adopted [123]. Therefore, we assume that our cybercriminal will always attempt to return the victim's data. But how does one unintentionally not return data? Modern ransomware operates by encrypting the victim's data using public-private key encryption, rendering the data unreadable until it has been decrypted [124]. In order to decrypt their data, the victim must pay the cybercriminal in exchange for the decryption key. As the encryption phase of the attack must proceed rapidly to avoid notice, flaws may arise during the process. If such flaws occur, then the decryption key will fail to decrypt the affected data. This results in some or all of the victim's files becoming permanently irretrievable, which is unlikely to be discovered until after the victim has paid the ransom. If a strain of ransomware is known to have a history of failure, the target's willingness to pay is reduced, and so the reliability of the cybercriminal's ransomware is an important factor in the negotiations. Ransomware strains can vary quite significantly in their reliability, depending on how heavily the cybercriminal invested in the development of the ransomware. They may even invest in providing "customer service" to their victims, walking them through the decryption process to further improve their image of reliability [125].

The potential for negotiation adds a significant feature from game theory to the dynamics of targeted ransomware; information asymmetry. The two parties to the negotiation have different degrees of information about each other, information that is crucial to determining the outcome of the negotiation. The cybercriminal does not know exactly how much the target's data is worth to them. This is a significant factor, as the cybercriminal seeks to set the ransom as high as possible in order to justify the significant investment they make in their attack. Therefore, estimating the value of the target's files accurately is of great importance. To aid them in this, the cybercriminal can make use of both publicly available data, including shareholder's reports and valuations, and private data found on the target's computer network, such as up-to-date finances and business plans. However, uncovering and interpreting an organisation's private documents is not effortless, and so producing an accurate estimate for the value of the target's data requires further investment from the cybercriminal. While investing greater effort can lead to a more accurate estimate, it is highly unlikely that the cybercriminal will ever achieve perfect accuracy. Even so, we might expect the target to be at a distinct disadvantage in terms of information asymmetry. However, as a result of developments in the computer security industry in response to targeted ransomware, this is not entirely true. The potential for negotiation of large ransoms has led to the emergence of organisations that offer professional ransomware negotiation services [126]. The negotiators have experience in dealing with cybercriminals behind the various strains of ransomware, allowing them to negotiate effectively on the behalf of targeted organisations. The negotiators also offer specialised knowledge. This can include statistics such as the reliability of a given strain of malware, but it can also include less quantifiable information, such as how aggressively the cybercriminal behind a particular strain will negotiate. Operators of targeted ransomware are typically willing to reduce their demand in the interest of getting paid, but not all are equally receptive to negotiations. Some cybercriminals are likely to perceive a low counteroffer to their demand as an insult, and may react aggressively to punish the target. The severity of the reaction depends on how aggressive the cybercriminal is, which varies depending on individual and cultural factors. In extreme cases, they may even react by abandoning the negotiations, and their ransom, so that future targets will be less likely to negotiate. This aggressive behaviour is a double-edged sword; if the cybercriminal is not aggressive enough, then their targets will not pay them a large ransom. If they are too aggressive, their inclination to punish their targets for perceived insults will cost them ransoms. In order to be successful, the cybercriminal must balance their aggression with their investments in their strain of targeted ransomware. In the next section, we construct a model of ransomware negotiation that is based on these key features.

3.4 Modelling

We propose to study the dynamics of targeted ransomware by modelling the negotiations as a two-player game. This approach was inspired by Selten's analysis of a two-player game modelling the interaction between a hostage taker and a hostage negotiator [23]. Our two players are the attacker A and the defender D. Player A is a cybercriminal (or group of cybercriminals) operating a strain of targeted ransomware. Player D is an organisation targeted by player A, assisted by a professional negotiator hired to negotiate the ransom.

3.4.1 Player A's investment

As noted in the previous section, there are three areas in which player A must invest in order to pull off a successful targeted attack:

- The circumvention of player D's security;
- the reliability of player A's ransomware; and
- the estimation of the value of player D's data.

We will not be considering player A's investment in circumventing player D's security here. While it is likely to be an important factor in the overall dynamics of the system, it has little bearing on the negotiations. Once player D's computer network has been infected, this investment only affects player A's net profit, and affects

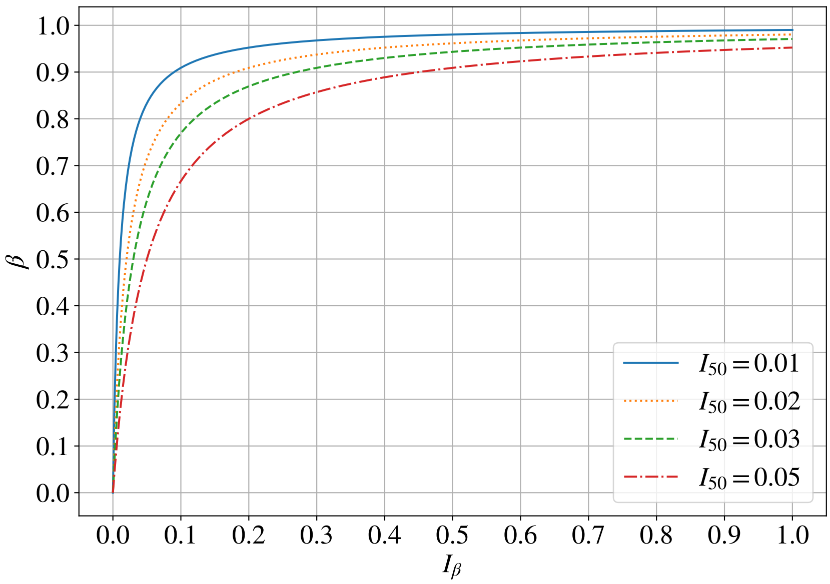

Figure 3.1: β, the probability of succesful decryption, depends on I_β and I_{50}.

player D not at all. However, the other two investments are very significant to the negotiations.

By the time player A launches their attack, they have already developed their strain of ransomware. In particular, they have invested in the reliability of their ransomware. Investing in reliability is important; if the decryption process is likely to fail, then player D will not be willing to pay very much for the decryption key. This is a significant up-front development cost; it does not scale with the number of targets player A attacks. For the sake of simplicity, we assume that player A can amortize this investment over the targets that they will infect with their strain of ransomware. We refer to this investment as I_β. As I_β increases, so does the reliability of player A's ransomware. Let β be the probability that the decryption key successfully decrypts data encrypted by the ransomware. We choose β such that

$$\beta = \frac{I_\beta}{I_\beta + I_{50}} \tag{3.1}$$

where I_{50} is the amount of investment required to achieve a reliability of 50%, or $\beta = 0.5$. I_{50} can be considered as an economic scaling factor, and determines the amount of development that player A can purchase for a given investment. This choice of β results in diminishing return for large I_β, so that $\beta \to 1$ as $I_\beta \to \infty$, as shown in Fig. 3.1. Note that I_{50} and I_β are dimensionless, so that our model remains generally applicable. $I_{50} = 0.02$ means that an investment of 2% of the value of the targeted files will yield a decryptor that is 50% reliable. While player D doesn't know the value of I_β, their negotiator can provide an estimate of β from their experience of individual ransomware strains.

The final area in which player A can invest is in their estimation of how much player D's data is worth. Without an accurate estimate, player A is unlikely to choose an optimal ransom demand, and so investing in producing an accurate estimate is an important aspect of their strategy. We let x be the value that player D attaches to their encrypted data, and let $\tilde{x}$ be player A's estimate of x. We refer to player A's investment in data value estimation as I_σ. As I_σ increases, so does the probability that $\tilde{x}$ will be close to x. In this analysis, we choose to model $\tilde{x}$ as random variable following a Lognormal(μ, σ^2) distribution [127] with probability density function

$$f(\tilde{x}, \mu, \sigma) = \frac{1}{\tilde{x}\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln \tilde{x} - \mu)^2}{2\sigma^2}\right) \tag{3.2}$$

for $\tilde{x} > 0$ and parameters

$$\begin{aligned} \mu &= \ln x \\ \sigma &= 1 - \frac{I_\sigma}{I_{50} + I_\sigma} \end{aligned} \tag{3.3}$$

With the chosen parameters, $\tilde{x}$ has median x. As $I_\sigma \to \infty$, $\sigma \to 0$, the variance $\left(e^{\sigma^2} - 1\right)\left(e^{2\mu+\sigma^2}\right) \to 0$ and mean $xe^{\frac{\sigma^2}{2}} \to x$. The Lognormal distribution has previously been used for modelling positive quantities that are determined by human behaviour [128]. As $\tilde{x}$ is player A's estimate of how much player D values their data, we believe that this is an appropriate choice. Fig. 3.2 shows the probability density function $f(\tilde{x}, \mu, \sigma)$ for $x = 1$ and varying levels of investment I_σ. I_σ is scaled similarly to I_β. As I_σ increases, the distribution narrows around x.

3.4.2 Negotiation

Once player A has infected player D's computer network with ransomware, the negotiation begins. Player A issues a ransom demand R, and player D must decide how to respond. We now examine the negotiation process between player A and player D that occurs if player D is willing to pay, but does not wish to pay the full demand. This is a time-sensitive issue, particularly for player D. Player D cannot conduct business while their data is encrypted, but they still incur costs, which can grow significant over a protracted negotiation. Player A does not suffer such ongoing costs, but they have invested significant resources in the attack, and the longer the negotiation continues, the more time player D has to consider their position. This makes a rapid negotiation process highly desirable, and so we make a simplification to the negotiation process and model it in the simplest way possible in the manner used by Selten [23]. Player A issues a ransom demand, player D responds with a counteroffer C, then player A decides whether to accept C and hand over the decryption key, or reject C and abandon the negotiations.

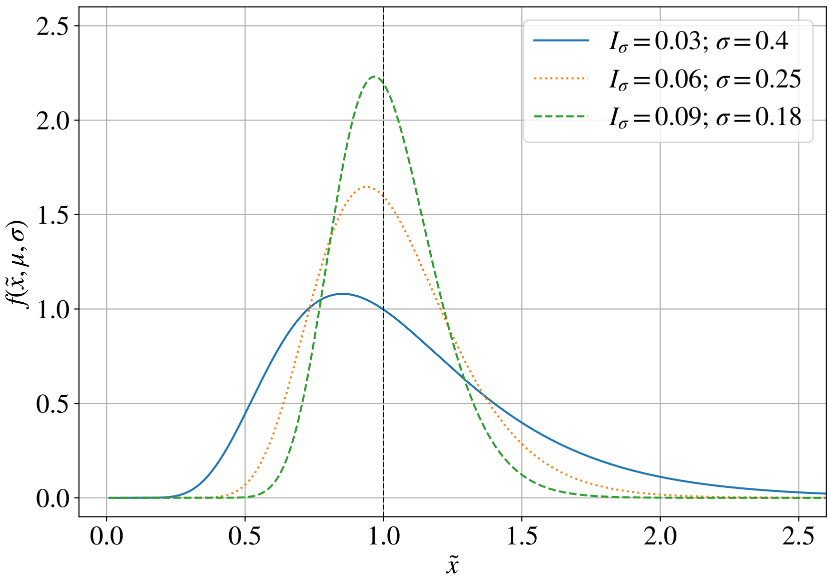

Figure 3.2: Probability density function $f(\tilde{x}, \mu, \sigma)$ for $x = 1$, $I_{50} = 0.02$ and varying levels of investment I_σ.

This is a substantially simplified description of the negotiation process and should not be taken literally. In reality, there may be a series of offers and counteroffers that take place over time. However, player D has finite capital with which they can absorb the costs incurred by not being able to conduct business. A drawn out negotiation for a lower ransom may be more costly than a prompt negotiation for a higher ransom.

Why would player A ever reject C? In doing so, they lose both their potential earnings and waste any investment they've made in the attack, which is clearly an undesirable outcome. However, player A must maintain their status as a threat; if they appear to be willing to accept low counteroffers, they will only receive low counteroffers. In order to maintain their credibility and their profits, player A may punish player D for making a low counteroffer. Therefore, we must expect that with a positive probability α, A will perceive a counteroffer $C < R$ as an affront, to which they react aggressively by abandoning the negotiations. It is reasonable to suppose that α will be greatest for $C = 0$ and lowest for $C = R$. We define α as

$$\alpha = 1 - \left(\frac{C}{R}\right)^a \tag{3.4}$$

where $a > 0$ is the aggression parameter of player A, quantifying their tendency to perceive a low counteroffer as an affront and react aggressively. This aggressive reaction is different to that implemented by Selten [23], where $\alpha = a\left(1 - \frac{C}{R}\right)$ and $a \in [0, 1]$ so that $\alpha \leq a \leq 1$. The reason for this is that in Selten's game, the aggressive reaction of the hostage taker is to kill the hostage, while in our game, the

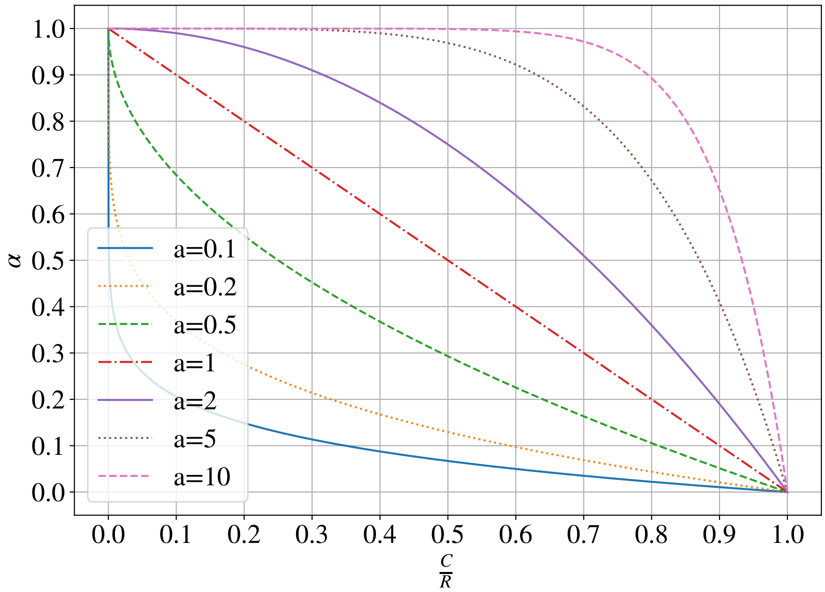

Figure 3.3: α, the probability of an aggressive reaction from player A, depends on a and the ratio of counteroffer to ransom demand $\frac{C}{R}$.

aggressive reaction is merely to not decrypt data. It is reasonable to assume that a hostage taker, faced with having their ransom demand being disregarded, might still refrain from killing their hostage. However, a cybercriminal, divorced from the consequences of their actions, would have no reason not to react aggressively and abandon the negotiation. This choice of α allows for a wide range of behaviour from player A, with very lenient negotiations for $a < 1$, scaling up to very aggressive negotiations as a increases. Larger a causes α to increase more rapidly as the difference between C and R increases, as shown in Fig. 3.3. As with β, player D can estimate a through the negotiator's experience of interacting with individual ransomware operators.

If player A does not react aggressively to a counteroffer $C < R$, player A receives payment C and player D receives the decryption key, which they then attempt to use to decrypt their data. The probability of successful decryption is β. If the decryption process is successful, player D retrieves their data. If the decryption is unsuccessful, their data is rendered permanently irretrievable. The payoff received by each player depending on the outcome of the game are detailed in Table 3.1. Player A loses their investment regardless of the outcome. If player A rejects player D's counteroffer, D loses the value of their data. If player A accepts the counteroffer, then they receive the value of the counteroffer to offset the loss of their investment. If player D's counteroffer is accepted and their data is decrypted successfully, they only lose the counteroffer. However, in the worst possible outcome for player D, their data is not decrypted successfully, and they lose both their data and their counteroffer! In the next section, we consider how these outcomes may arise through rational decisions

Table 3.1: Payoff table for the targeted ransomware negotiation game.

Outcome	Payoff	
	Player A	Player D
Counteroffer rejected	$-I_\beta - I_\sigma$	$-x$
Counteroffer accepted, data decrypted successfully	$C - I_\beta - I_\sigma$	$-C$
Counteroffer accepted, data not decrypted successfully	$C - I_\beta - I_\sigma$	$-x - C$

made by both players.

3.4.3 Summary of game rules

The model variables are summarised in Table 3.2. The rules of the game are summarised as follows:

1. Player A incurs cost $I_\beta + I_\sigma$ to infect player D's computer system and make a ransom demand R.
2. Player D makes a counteroffer C.
3. Player A rejects player D's counteroffer with probability α.
4. If player A does not reject the counteroffer, player A receives the counteroffer C and player D receives the decryption key.
5. Player D's data is successfully decrypted with probability β.

3.5 Analysis

3.5.1 Optimal choice of C

In the subgame beginning with player D's choice of C, player D knows that making a counteroffer $C < R$ will provoke an aggressive reaction from player A with probability α. Player D has no incentive to make a counteroffer $C > R$, so they rationally chooses C to maximize their expected utility U:

$$U = \begin{cases} -\left[R + (1-\beta)\,x\right] & C = R \\ -\left(1-\alpha\right)\left[C + (1-\beta)\,x\right] - \alpha x & C < R \end{cases} \tag{3.5}$$

The probability of successful decryption is β; therefore, the expected value of the encrypted data is βx, and so player D will not offer more than βx for the decryption key. We assume that there exists a critical value $C_{max} \in [0, \beta x]$ such that player D's expected utility is maximized by making counteroffer C_{max} if $R > C_{max}$. By substituting from Eq. (3.4), player D's expected utility for $C < R$ simplifies to

$$U = -\left(\frac{C}{R}\right)^a [C - \beta x] - x \tag{3.6}$$

Table 3.2: Table of game variables with detail on the information asymmetry in the targeted ransomware negotiation game.

Variable	Description	Known to player A?	Known to player D?
x	True value of D's files	No	Yes
$\tilde{x}$	A's estimate of x	Yes	No
R	A's ransom demand	Yes	Yes
C	D's counteroffer	Yes	Yes
a	A's aggression	Yes	Yes
I_β	A's investment in reliability	Yes	No
I_σ	A's investment in estimating x	Yes	No
α	Probability that A will react aggressively to $C < R$	Yes	Yes
β	Probability that the decryption key works	Yes	Yes
σ	Scale parameter of distribution of $\tilde{x}$	Yes	No

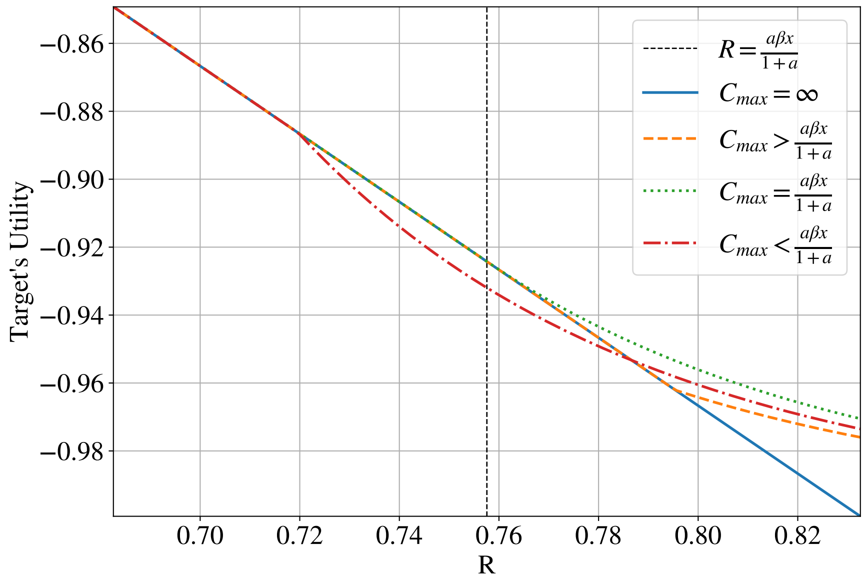

Figure 3.4: Player D's expected utility for varying R and different values of C_{max} where $a = 10$, $I_{50} = 0.02$ and $I_\beta = 0.1$.

To calculate the optimal value of C, we calculate

$$\frac{\partial U}{\partial C} = \frac{C^{a-1}}{R^a}\left(a\beta x - C\left(1+a\right)\right) \tag{3.7}$$

By setting $\frac{\partial U}{\partial C} = 0$ we find that for $0 < C < R$, U achieves its maximum at $C = \frac{a\beta x}{1+a}$. $\frac{\partial U}{\partial C} < 0$ for $C > \frac{a\beta x}{1+a}$; for any $R > \frac{a\beta x}{1+a}$, player D's optimal counteroffer is $C = \frac{a\beta x}{1+a}$. $\frac{\partial U}{\partial C} > 0$ for $C < \frac{a\beta x}{1+a}$; however, player D will never make a counteroffer $C > R$. If $R < \frac{a\beta x}{1+a}$, player D's optimal counteroffer is $C = R$. Thus, $C_{max} = \frac{a\beta x}{1+a}$, yielding player D's optimal counteroffer $\hat{C}$:

$$\hat{C} = \begin{cases} R & R \leq \frac{a\beta x}{1+a} \\ \frac{a\beta x}{1+a} & R > \frac{a\beta x}{1+a} \end{cases} \tag{3.8}$$

This scenario is demonstrated graphically in Fig. 3.4. Any choice of $C_{max} \neq \frac{a\beta x}{1+a}$ results in a drop in player D's expected utility.

3.5.2 Optimal choice of R

In the subgame beginning with player A's choice of R, player A knows that under rational decision-making, player D will optimally make counteroffer $\hat{C}$. Player A

rationally chooses R to maximize their expected profit P:

$$P = \begin{cases} R - I_\beta - I_\sigma & R \leq C \\ (1-\alpha)\, C - I_\beta - I_\sigma & R > C \end{cases} \tag{3.9}$$

Substituting from Eq. (3.8) yields

$$P = \begin{cases} R - I_\beta - I_\sigma & R \leq \frac{a\beta x}{1+a} \\ \left(\frac{\frac{a\beta x}{1+a}}{R}\right)^a \left(\frac{a\beta x}{1+a}\right) - I_\beta - I_\sigma & R > \frac{a\beta x}{1+a} \end{cases} \tag{3.10}$$

By differentiating with respect to R we find that $\frac{\partial P}{\partial R} > 0$ for $R < \frac{a\beta x}{1+a}$ and $\frac{\partial P}{\partial R} < 0$ for $R > \frac{a\beta x}{1+a}$. Therefore, the optimal ransom demand $\hat{R} = \frac{a\beta x}{1+a}$ is the highest ransom that player D is willing to pay. If player A can reliably make demand $\hat{R}$, player D will always pay. Player A will always make their maximum profit, and there is no risk of an aggressive reaction from player A, trivialising the negotiation. Under such conditions, player A's profit is $\frac{a\beta x}{1+a} - I_\beta - I_\sigma$. This would suggest that, in order to maximise their profit, player A should be infinitely aggressive, rejecting any counteroffer even slightly lower than their demand (i.e. $a \to \infty$).

Of course, this "ideal" scenario is unrealistic, as it ignores the often-significant effect of imperfect information [129, 130, 131]. In reality, negotiations are not trivial affairs, and the risk of an aggressive reaction is always present. Player A does not know x, only $\tilde{x}$, which, lacking any alternative, is what they use to calculate their ransom demand. Therefore, under optimal play while accounting for imperfect information, player A's ransom demand is $R = \frac{a\beta\tilde{x}}{1+a}$. The potential for error in player A's estimate $\tilde{x}$ gives rise to the necessity of negotiations that may result in an aggressive reaction. We can illustrate this by substituting $R = \frac{a\beta\tilde{x}}{1+a}$ into Eq. (3.10). Under optimal play,

$$P = \begin{cases} \frac{a\beta\tilde{x}}{1+a} - I_\beta - I_\sigma & \frac{a\beta\tilde{x}}{1+a} \leq \frac{a\beta x}{1+a} \\ \left(\frac{\frac{a\beta x}{1+a}}{\frac{a\beta\tilde{x}}{1+a}}\right)^a \left(\frac{a\beta x}{1+a}\right) - I_\beta - I_\sigma & \frac{a\beta\tilde{x}}{1+a} > \frac{a\beta x}{1+a} \end{cases} \tag{3.11}$$

which factors to

$$P = \frac{a\beta}{1+a} \left.\begin{cases} \tilde{x} & \tilde{x} \leq x \\ x\left(\frac{x}{\tilde{x}}\right)^a & \tilde{x} > x \end{cases}\right] - I_\beta - I_\sigma \tag{3.12}$$

Thus, it is the potential for error in the estimate $\tilde{x}$ which prevents player A from playing optimally. The effect of error in $\tilde{x}$ on player A's profit is shown in Fig. 3.5. Here, for investment levels are fixed at $I_\beta = I_\sigma = 0.1$, so that the maximum profit depends on a. The effect of error differs depending on whether player A underestimates or overestimates the value of the data. If they underestimate x, then their profit decreases linearly with $\tilde{x}$. If they overestimate x, then the possibility of an aggressive reaction emerges, which increases with both a, and the error in $\tilde{x}$. High aggression might increase player A's capacity for demanding large ransoms, but at an increased risk of an aggressive reaction.

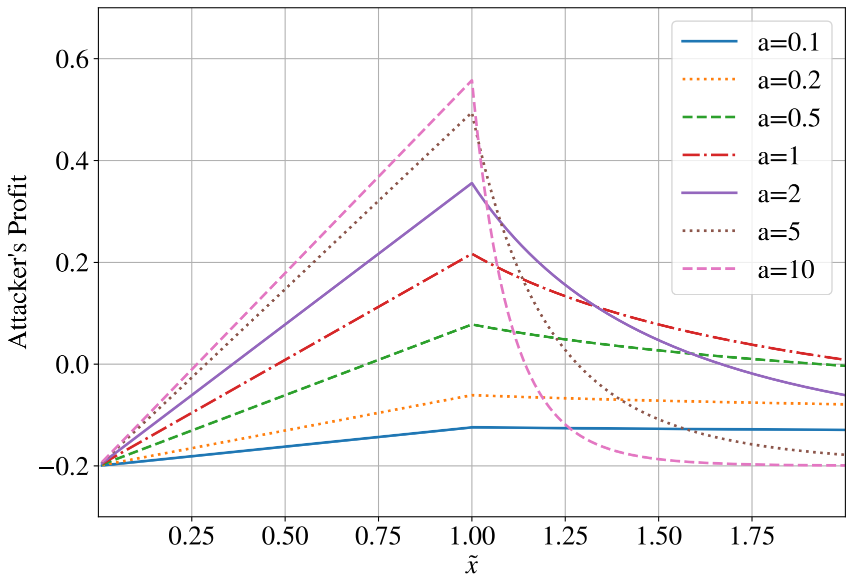

Figure 3.5: Player A's expected profit as a function of $\tilde{x}$ for varying a when $x = 1$, $I_{50} = 0.02$, and $I_\beta = I_\sigma = 0.1$.

3.5.3 Attacker's expected profit

In order to further understand how error and aggression interact, we consider player A's expected net profit P. First, let player A's gross profit be

$$GP\left(\tilde{x}, x; a, \beta\right) = \frac{a\beta}{a+1} \begin{cases} \tilde{x} & \tilde{x} \leq x \\ x\left(\frac{x}{\tilde{x}}\right)^a & \tilde{x} > x \end{cases} \tag{3.13}$$

We let x follow a distribution with probability density function $g\left(x, M\right)$ for $x > 0$ and mean M; the shape of the distribution is unimportant in this case. With probability density functions $g\left(x, M\right)$ for x and $f\left(\tilde{x}, \mu, \sigma\right)$ for $\tilde{x}$, Player A's net profit for a given combination of aggression and investments can be written as a convolution of x and $\tilde{x}$ in double integral form. However, due to our choice of α and Lognormal $\tilde{x}$, by making the substitution $y = \frac{\tilde{x}}{x}$ we can reduce the double integral to a single integral.

By making this substitution, we can see that what appears to be a convolution of x and $\tilde{x}$ depends merely on the ratio $\frac{\tilde{x}}{x}$. Player A's strategy is consistent across all values of x, so it is only the mean M of the distribution of x that remains in the final expression. In this form, we can clearly see where each element of player A's strategy (a, I_β, I_σ) comes into play. As aggression a increases, the multiplicative term $\frac{a}{a+1}$ increases, but the second integral in the sum, where player A has overestimated x,

$$
\begin{aligned}
P\left(a, I_\beta, I_\sigma\right) &= \int_{x=0}^{x=\infty} \int_{y=0}^{y=\infty} GP\left(xy, x; a, \beta\right) f\left(xy, \mu, \sigma^2\right) g\left(x, M\right) xdydx - I_\beta - I_\sigma \\
&= \int_{x=0}^{x=\infty} \int_{y=0}^{y=\infty} \left(\frac{a\beta}{a+1}\right) \left.\begin{cases} xy & xy \leq x \\ x\left(\frac{1}{y}\right)^a & xy > x \end{cases}\right] \frac{1}{xy\sigma\sqrt{2\pi}} e^{-\frac{\left[\ln(xy)-\ln(x)\right]^2}{2\sigma^2}} g\left(x, M\right) xdydx - I_\beta - I_\sigma \\
&= \left(\frac{a\beta}{a+1}\right) \int_{x=0}^{x=\infty} g\left(x, M\right) xdx \int_{y=0}^{y=\infty} \left.\begin{cases} y & y \leq 1 \\ y^{-a} & y > 1 \end{cases}\right] \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{\ln^2 y}{2\sigma^2}} dy - I_\beta - I_\sigma \\
&= \left(\frac{a\beta}{a+1}\right) M \left[\int_{y=0}^{y=1} y \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{\ln^2 y}{2\sigma^2}} dy + \int_{y=1}^{y=\infty} y^{-a} \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{\ln^2 y}{2\sigma^2}} dy\right] - I_\beta - I_\sigma
\end{aligned}
\tag{3.14}
$$

Table 3.3: Model variables for optimal and various sub-optimal strategies.

Strategy Type	(a, I_β, I_σ)	C	P
Optimal	$(4.68, 0.091, 0.104)$	0.675	0.304
Low Aggression	$(2.34, 0.091, 0.104)$	0.574	0.276
High Aggression	$(9.36, 0.091, 0.104)$	0.741	0.284
Low Reliability	$(4.68, 0.041, 0.104)$	0.554	0.265
High Reliability	$(4.68, 0.182, 0.104)$	0.742	0.264
Low Accuracy	$(4.68, 0.091, 0.052)$	0.675	0.283
High Accuracy	$(4.68, 0.091, 0.208)$	0.675	0.267

converges to 0. Increasing I_β increases costs, but leads to increased β which may increase profit. Increasing I_σ also increases costs, but narrows the distribution of $\tilde{x}$ around x, allowing for greater aggression at decreased risk of aggressive reaction. Thus, through our choice of α and $\tilde{x}$, we can more clearly demonstrate how player A's strategy depends on the interaction between the various elements of their strategy. The optimal counteroffer and expected net profit P for parameters corresponding to sample strategies are shown in Table 3.3. Fig. 3.6 shows P for varying parameters a, I_β and I_σ. The left column shows results calculated via numerical integration of Eq. (3.14) with $M = 1$, while the right column was calculated using an agent-based simulation, where results are averaged over $n = 10^4$ simulations of the game with constant target data value $x = 1$. These figures demonstrate that player A's optimal strategy is $(a, I_\beta, I_\sigma) = (4.68, 0.091, 0.104)$, rather than the naive maximal aggression $a \to \infty$ noted previously; in order to realise their potential profits, the attacker must be willing to negotiate.

We also observe that the results derived from agent-based simulation are quite noisy, despite being averaged over $n = 10^4$ simulations. While increasing n would reduce noise, this would come at the cost of increasing computation time. In Fig. 3.7, we consider how agent-based simulation compares to numerical integration as we vary n with parameters $(a, I_\beta, I_\sigma) = (4.68, 0.091, 0.104)$. To avoid confusion, let P_∞ be P calculated from numerical integration using the Python function `scipy.integrate.quad` [132, 133], and let P_n be P calculated from averaging over n simulations. The estimated absolute error of P_∞ is of order $O\left(10^{-11}\right)$. In Fig. 3.7(a) and Fig. 3.7(b), we observe that the distribution of P_n is well approximated by a Normal distribution, as predicted by the Central Limit Theorem (CLT) [134], and that the standard deviation decreases (and the distribution narrows) as n increases. In Fig. 3.7(c), we plot the standard deviation of P_n for varying n on a log-log scale, where the relationship between the two is linear with a slope of $-\frac{1}{2}$, so that the standard deviation is proportional to $n^{-\frac{1}{2}}$, which is again consistent with the CLT. In theory, we can simply increase n until the standard deviation of P_n is of lower order as the maximum error of P_∞. However in practice, reducing the standard deviation of P_n to order $O\left(10^{-3}\right)$ takes more computation time than numerically

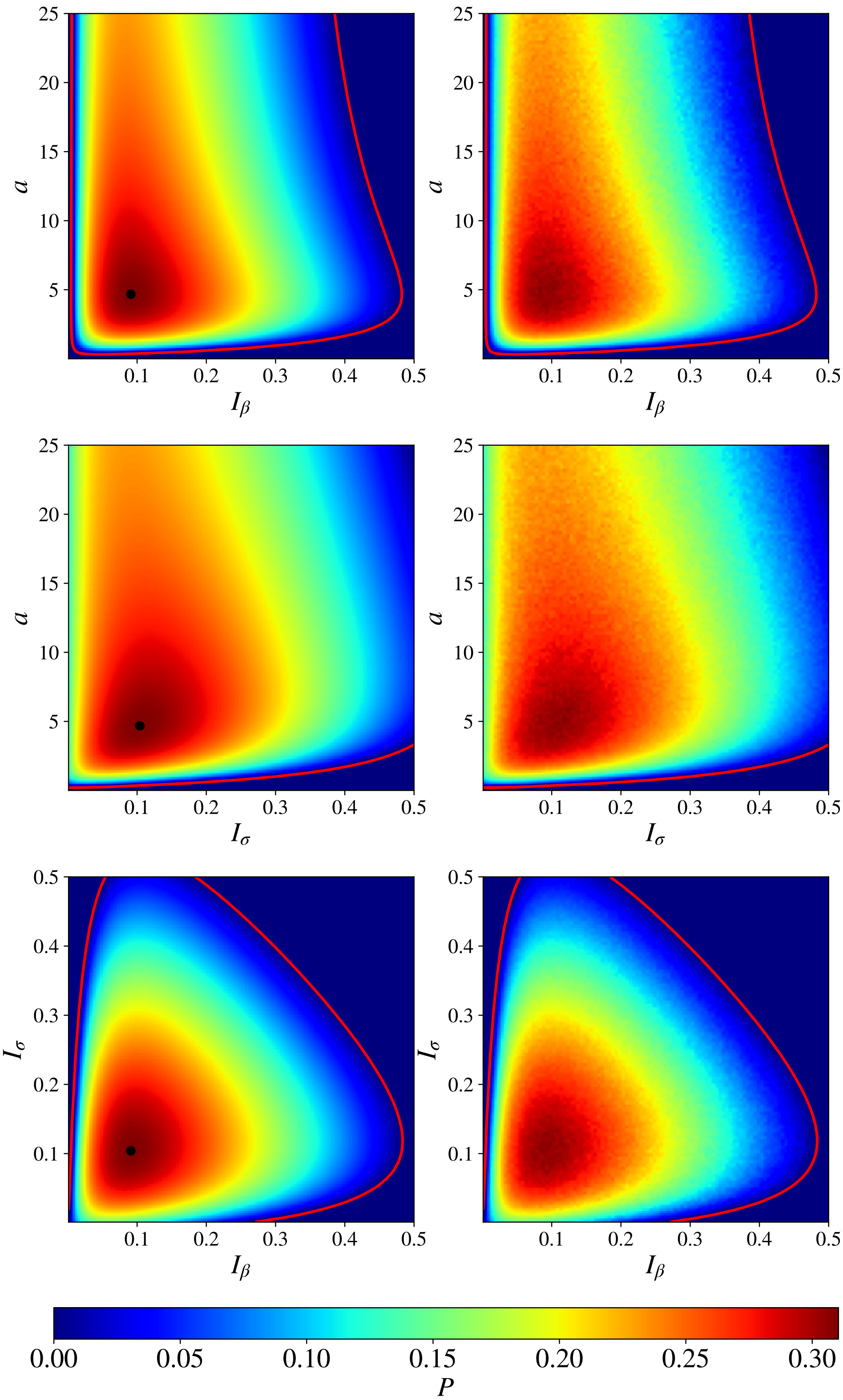

Figure 3.6: Player A's profit for varying strategy parameters (a, I_β, I_σ) when $I_{50} = 0.02$. The plots on the left are the result of numerical integration, while the plots on the right are the result of averaging over $n = 10^4$ simulations of the game with constant target data value $x = 1$. In each plot, the hidden parameter is set to its optimal value. The maximum mean profit achieved is marked by a black dot. The red curve marks where the mean profit is equal to 0.

integrating P_∞ to within an absolute error order $O\left(10^{-11}\right)$, as shown in Fig. 3.7(d). Therefore, precisely estimating the attacker's expected net profit using numerical integration is far more efficient than using agent-based simulation, which justifies our efforts to reduce the problem to numerical integration.

3.6 Viability of targeted ransomware under varying economic conditions

Thus far, we have neglected I_{50}, the parameter that reflects economic conditions for the attacker, setting the scale for the amount of investment needed by the attacker to achieve a certain level of precision in decryptor reliability and target data value estimation. Now that we have considered how the attacker's profit varies with their chosen strategy, we can consider how their profit under optimal strategy varies with I_{50}. Fig. 3.8 shows the attacker's optimal strategy and expected profit under that strategy for for varying I_{50} and fixed $x = 1$, so that as I_{50} increases, the cost of precision and estimation increases relative to the value of the target's data.

As I_{50} increases, the attacker's expected profit decreases towards 0, as expected. The optimal strategy on the other hand, undergoes an interesting transition. Firstly, for small $I_{50} \lesssim 0.029$, $I_\sigma > I_\beta$, indicating that when costs are extremely low relative to the value of the files, the attacker should focus on value estimation over reliability. $I_{50} = 0.02$, the value of I_{50} used throughout our analysis, lies in this domain. For $I_{50} \gtrsim 0.029$, $I_\sigma < I_\beta$, and the attacker should focus on reliability. For large $I_{50} \gtrsim 0.082$ however, we see I_σ drop to 0 in the optimal strategy; when the value of the target's data is small relative to the cost involved in estimating its value, the attacker should invest no effort in estimating its value! Furthermore, for $I_{50} \gtrsim 0.082$, a remains constant while I_β continues to decrease; once the attacker stops investing in estimation, there's no reason to change their negotiation tactics. This case compares well to the more classic style of ransomware that existed before targeted ransomware. The attacker made no investment in estimation, and accepted no negotiations, so that $I_\sigma = 0$ and $a \to \infty$. In our model, while a is constant, it does not go to ∞. This can be attributed to estimation being a built-in part of our game. As noted earlier in this chapter, cybercriminals operating classic ransomware simply set a fixed price for decryption based on the expected distribution of the value of their victims' files. Allowing for the attacker to adopt either a classic or targeted strategy based on economic conditions might inject a further note of realism; however, it would be a significant extension to the model, and so will not be considered here. It remains clear that beyond a certain economic threshold, the attacker may be forced to change their strategy fundamentally to continue to prosper.

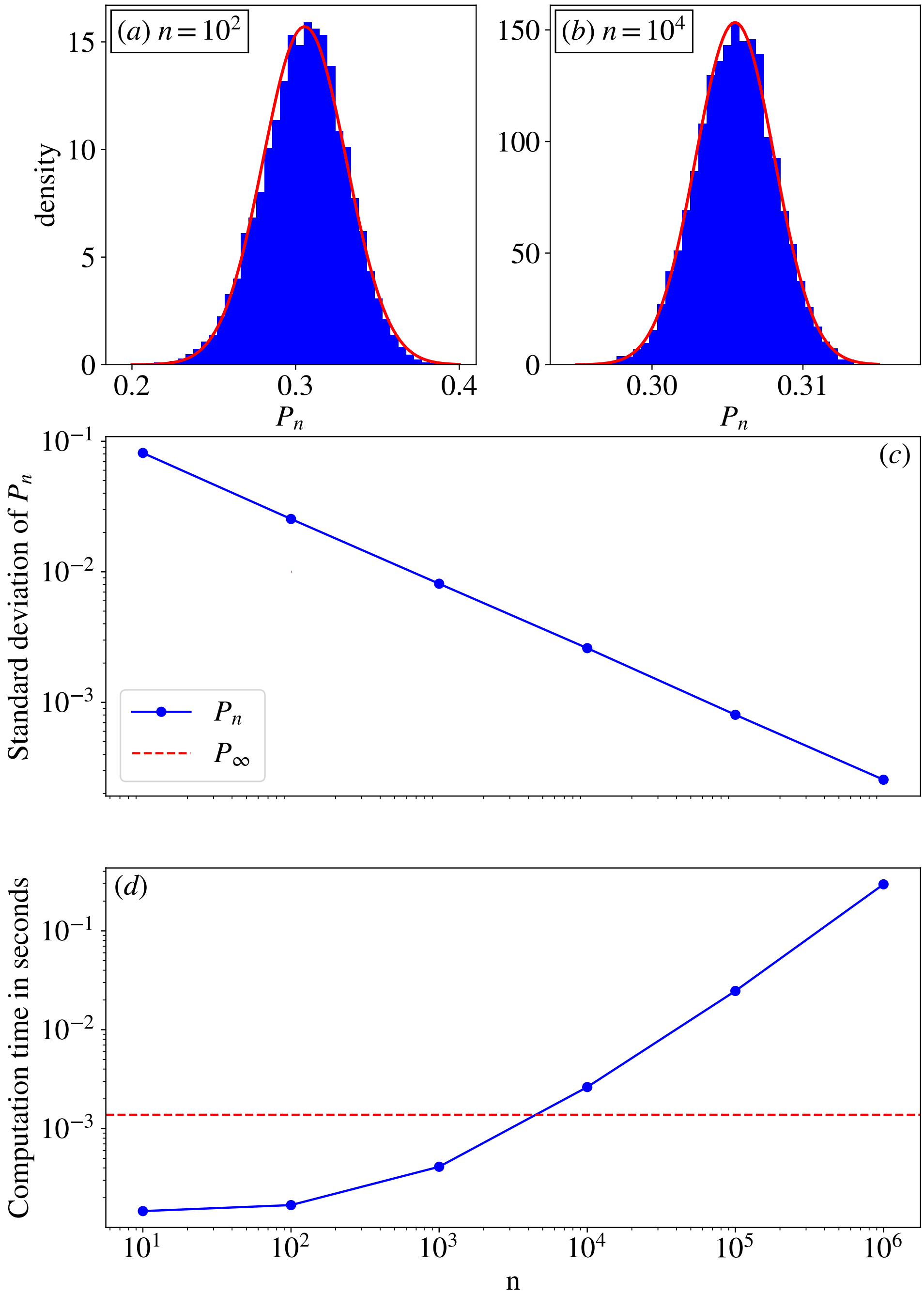

Figure 3.7: Comparison of error and computation time of the attacker's average net profit P between numerical integration (P_∞) and agent-based simulation (P_n) as the number of simulations n is varied. In (a) and (b), a Normal probability density function (red) is fitted to the distribution of P_n (blue) for two values of n. (c) shows the standard deviation of P for varying n. (d) shows the computation time of P_n for varying n (blue) and P_∞ (dotted red).

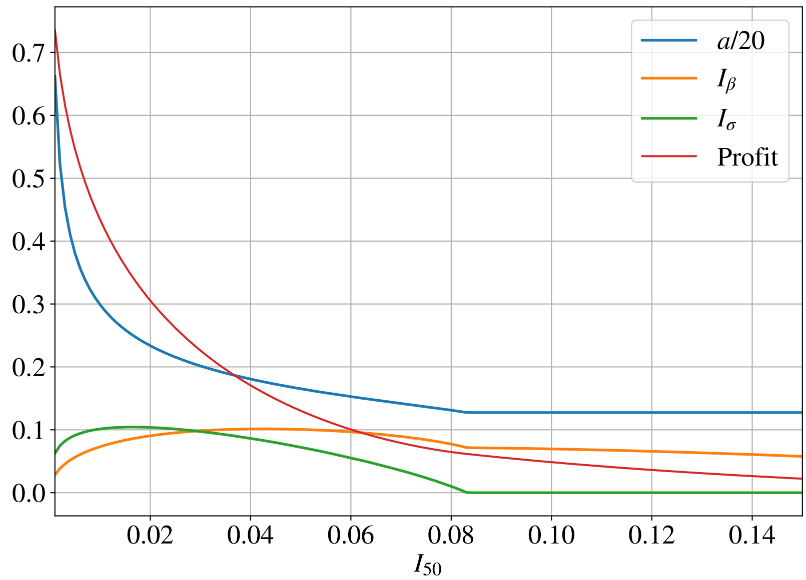

Figure 3.8: Attacker's optimal strategy and expected profit under that strategy for varying I_{50} and fixed $x = 1$.

3.7 Conclusion

We have constructed a simple game-theory model of targeted ransomware negotiations between two agents; an attacker, a cybercriminal operating targeted ransomware, and a defender, their target. Our model focuses on three key elements of the attacker's strategy which have a direct effect on the negotiations; aggression, investment in the reliability of their ransomware, and investment in their estimation of target data value. By thoroughly analysing our model, we demonstrate how the necessity of negotiation arises from optimal decision made by the agents under imperfect information. We show how the key elements of the attacker's strategy interact with each other, and demonstrate by numerical integration and by agent-based simulation that the attacker's profit depends on developing a balance of investment and aggression. Finally, we consider how economic conditions for the attacker set a limit on where targeted ransomware may no longer be an effective strategy. While ransomware strains do show a significant variety of behaviour, even within the subgroup of targeted ransomware, the features present in this model are quite generic, and so we expect the insights provided to be broadly applicable in the study of targeted ransomware.

In this analysis have focused solely on the negotiations of targeted ransomware that follow the infection of an organisation in order to construct a simple model that can be analysed in detail. In this scenario, the attacker has greater agency to determine the outcome, and so our analysis is focused on the decisions of the attacker. As noted in the Modelling section, this model neglects the attacker's investment in circumventing security, as it has no direct effect on the negotiations. By extension, the defender's investment in security is omitted. One could grant greater agency to the defender by considering how the defender's investment in security acts as a deterrent to a potential attacker. Every additional effort, such as redundant backups, data exfiltration countermeasures, and employee training makes an organisation harder to hold to ransom, and so less attractive as a target. Such deterrents, and other solutions to the growing problem of ransomware, are of vital importance to strengthening our shared internet security. Our model contributes to this goal by improving our understanding of various negotiation strategies through the use of game theory. As cybercrime continues to threaten our increasingly technology-dependent lives, we hope that the application of game theory to targeted ransomware will be of interest to a wide audience.

Chapter 4

Stochastic models of Time-To-Live caching systems

This chapter presents currently unpublished material which will require further development before publication. The first two sections form the basis for an invention disclosure made by P. Ryan and N. Fitzgerald. My contribution to this work consisted of developing and analysing the models of caching systems.

4.1 Introduction

Many online services operate by providing data to customers through a *data network* [135, 136, 137]. This is typically achieved by storing data on a high capacity server which sends data items to a customer's endpoint computer system over the Internet in response to a query from that endpoint, as shown in Fig. 4.1. The success of the service relies on the data network's ability to provide endpoints with data at low cost, to reduce the cost of service, and with low latency, to improve the quality of service. As a result, the study of such systems has been of great interest [138, 139, 140, 141].

Of all the strategies used to make data network and similar systems more efficient, few are as ubiquitous as caching. Caching is the process of temporarily storing recently used data so that it can be accessed more quickly when needed again later. There are many variations of caching, each suited to a different application. For example, the various caches in a computer's Central Processing Unit (CPU) typically operate according to a Least Recently Used (LRU) strategy, where space for a new data item is freed up in the cache by removing the data item that was least recently used [142]. As the cache is on the CPU itself, the cached data can be accessed far more rapidly than data that is stored on a hard drive.

However, LRU caching is not appropriate in applications where data items are volatile and may change without warning. For example, most cybersecurity services operate by sharing reputations (data items) for various files, websites and IP ad-

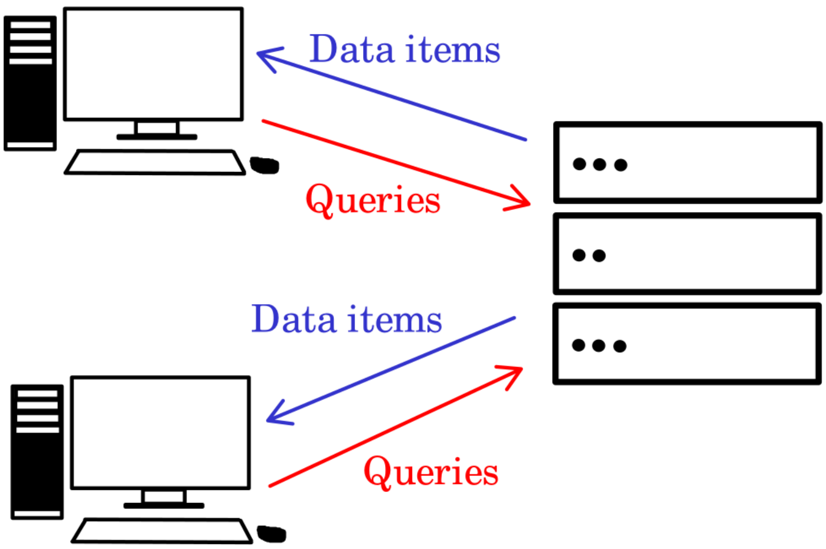

Figure 4.1: Sketch of interaction between server and endpoints.

dresses with customers worldwide. These reputations can change over time as new information becomes available. In such applications, *Time-To-Live* (TTL) caching is more appropriate [143]. When an item enters the cache, it is assigned a TTL in seconds, which counts down to zero, regardless of how often it is used. When the TTL reaches 0, the item is removed from the cache. This ensures that data cannot persist indefinitely in a cache while losing consistency with the server over time [144, 145]. Increasing the Time-To-Live increases the probability that an item will still be cached when it is needed again. This decreases the latency associated with loading data, improving quality of service, and decreases the volume of queries received by the server, reducing operating costs. However, increasing the Time-To-Live also increases the probability that a volatile data item changes while cached. The use of outdated data by endpoints has a large negative impact on the quality of service; this raises the problem of determining the optimal TTL based on the volatility of the data items and the activity of the endpoint.

We further note that caching systems have a weakness in that the server is less able to observe the activity of the endpoint. If the endpoint uses a data item when the item is not cached, this results in a query that the server can record. If the endpoint used a data item that is cached, the server is unaware. Without information detailing when the caching system prevented a query, it is difficult to estimate how effective that caching mechanism is. This makes it difficult for the service provider running the server to optimise their caching system, as they lack the data with which to judge the efficacy of the system.

One method of acquiring this data is to configure endpoints to send telemetry; data describing the local operation of an endpoint. Telemetry is used in a variety of applications, including healthcare, aerospace testing and industrial process management [146, 147, 148]. However, telemetry systems can be very costly to operate.

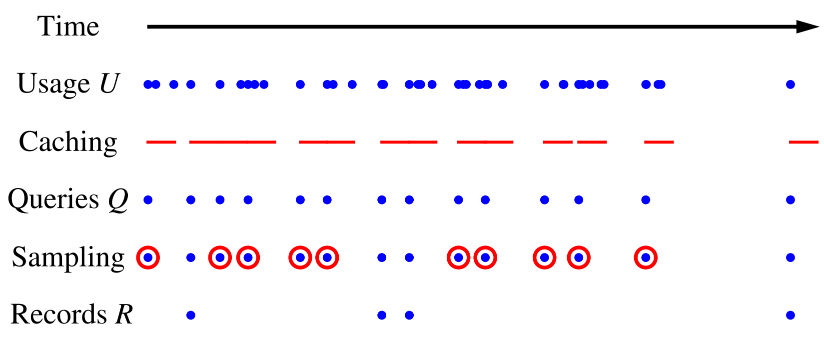

Figure 4.2: Visualisation of model of data usage, caching and sampling.

In the examples given, this cost is justifiable, as telemetry is practically essential to prevent catastrophic failure of the system. In other applications, such as data networks, telemetry might enable more efficient operation of the system, but it is not critical to its operation, and so the value of telemetry is marginal. Instead, we note that it is common for a server operating a data network to maintain records of queried data items. This *query data* is often randomly sampled to reduce the volume of data that must be stored on the server, and hence, the cost of storing that data. We propose to develop a method for estimating the activity of an endpoint based on randomly sampled query data. This enables the optimisation of caching systems in applications where telemetry is of marginal value.

In this chapter, we develop models of TTL caching of volatile data which allow us to consider the optimisation of data networks. First, we propose a simple stochastic model of endpoint/server interaction which simulates the generation of query data from on data item usage on an endpoint. We use this model to estimate the rate at which an endpoint uses data items based on observed query data, and hence estimate the reduction in query volume as a result of implementing a TTL caching system. We expand this model to consider data items that may change randomly to calculate an optimal TTL based on data volatility and data usage rates. Finally, we note interesting dynamical behaviour of the model and analyse this behaviour in the deterministic limit, where the model is comparable to the circle map [149, 150, 151].

4.2 Stochastic model of query data generation

Consider the following model of server/endpoint interaction, where a server and an endpoint are connected via the Internet. The server ensures that the endpoint has access to data items on demand. The model is visualised in Fig. 4.2.

1. The endpoint uses a particular data item according to some *random process*, which we will define shortly. Let $U = \{u_i\}_{i=1}^{N_U}$ be the set of times at which

the data item is used in a time interval $t \in [0, T]$. The random process has a rate $\lambda_U > 0$; as λ_U increases, so does the expected number of times the item is used in time T.

2. When the endpoint uses the item, it sends a query to the server, unless it has already queried the item in the preceding time interval of length $\tau \geq 0$, in which case the information is already cached on the endpoint and there is no need to query the server. Let $Q = \{q_i\}_{i=1}^{N_Q}$ be the set of times at which the endpoint queries the server; then $Q \subseteq U$ and $N_Q \leq N_U$.

3. The server observes Q, but it applies *simple random sampling* (defined shortly) to avoid storing excessive amounts of data. Each query is recorded with probability $\pi \in [0, 1]$. Let $R = \{r_i\}_{i=1}^{N_R}$ be the set of times at which a query is recorded; then $R \subseteq Q$ and $N_R \leq N_Q$.

These sets and set cardinalities evolve stochastically with time t. Formally, $U(t)$, $Q(t)$ and $R(t)$ are random processes, and their final states $U = U(T)$, $Q = Q(T)$ and $R = R(T)$ are *random variables*; the same treatment applies to $N_U(t)$, $N_Q(t)$ and $N_R(t)$.

Simple random sampling The type of sampling employed here is referred to as simple random sampling [134] with probability π. A simple explanation is that each query in Q is individually added to R with probability π. More formally, for $r \in R$ and $q \in Q$,

1. The probability that a specific query q_i is recorded is

$$P[q_i \in R] = \pi \tag{4.1}$$

2. Each query q_i can only be recorded once;

$$P[r_j = q_i | r_k = q_i] = 0 \tag{4.2}$$

   if $j \neq k$.

3. The probability that a query q_i is in R is independent of whether any other query q_l is in R;

$$P[r_j = q_i | r_k = q_l] = \frac{1}{N_Q - 1} \tag{4.3}$$

   if $j \neq k$ and $i \neq l$. If query q_l is already known to be in record r_k, then record r_j can be any of the other $N_Q - 1$ queries independently.

The model has three parameters; data item usage rate λ_U, caching time τ and sampling rate π. From the perspective of the server, the system produces observable data Q. The server chooses π to reduce the cost of storing query data, and chooses τ as part of its strategy to reduce query volume through caching. To estimate how effective this strategy is, the server would like to estimate how many times an item is used while it is cached. To enable this, we would like to determine a method for estimating λ_U by defining an estimator $\hat{\lambda}_U$ as a function of Q, τ and π [134, 152].

4.2.1 Poisson process of item usage

We assume that $U(t)$ is generated according to a Poisson process [153, 154, 134] with constant rate $\lambda_U > 0$, such that $N_U(t)$ is a *counting process;* a non-decreasing random walk on the non-negative integers $\{0\} \cup \mathbb{Z}^+$. Let $P_k(t) = P[N_U(t) = k]$, then the random walk has transition probabilities

$$\frac{d}{dt} P_0(t) = -\lambda_U P_0(t) \tag{4.4}$$

$$\frac{d}{dt} P_{k+1}(t) = \lambda_U P_k(t) - \lambda_U P_{k+1}(t) \quad k > 0 \tag{4.5}$$

$$P_k(0) = \begin{cases} 1 & k = 0 \\ 0 & k > 0 \end{cases} \tag{4.6}$$

The Poisson process is one of the simplest stochastic processes for modelling a counting process, and has been widely used across diverse disciplines including astronomy [155, 156, 157, 158, 159, 160], biology [161, 162, 163] and communication networks [164, 165, 166]. If $U(t)$ evolves according to Eq. (A.73), then $N_U(T)$ follows a Poisson distribution [153] with rate $\lambda_U T$

$$N_U(T) \sim \mathrm{Poi}(\lambda_U T) \tag{4.7}$$

such that for constant $t = T$, the *probability mass function* is

$$P[N_U(T) = k] = \frac{(\lambda_U T)^k e^{-\lambda_U T}}{k!} \tag{4.8}$$

(Derived in Appendix A.3.1). The expected value of $N_U(T)$ is

$$\langle N_U(T) \rangle = \lambda_U T \tag{4.9}$$

(Derived in Appendix A.3.2) which is the rate per unit time multiplied by elapsed time.

An important property of the Poisson process which will used throughout this chapter is the distribution of the *inter-event times.* If $u_i \in U(t)$ are item usage times such that the u_i are ordered by the index i; that is,

$$i < j < k \Rightarrow u_i < u_j < u_k \tag{4.10}$$

then the *inter-usage times* $\overline{u}_i = u_{i+1} - u_i$ are independent, identically distributed Exponential(λ_U) random variables [167], such that if

$$\overline{u}_i \sim \text{Exp}\left(\lambda_U\right) \tag{4.11}$$

then $\overline{u}_i \in (0, \infty)$ and

$$P\left[\overline{u}_i < c\right] = \int\limits_0^c \lambda_U e^{-\lambda_U \overline{u}_i} d\overline{u}_i = 1 - e^{-\lambda_U c} \tag{4.12}$$

(see Appendix A.3.3) with probability density function

$$f\left(\overline{u}_i, \lambda\right) \lambda_U e^{-\lambda_U \overline{u}_i} \tag{4.13}$$

The expected value of $\overline{u}_i$ [154] is given by

$$\langle \overline{u}_i \rangle = \frac{1}{\lambda_U} \tag{4.14}$$

Hence, another interpretation of Eq. (4.9) is

$$\langle N_U\left(T\right) \rangle = \frac{T}{\langle \overline{u}_i \rangle} \tag{4.15}$$

so that the average number of events is the total time divided by the average time between events.

4.2.2 Query and record processes under Poisson item usage

Now consider the query process $Q\left(t\right)$. $Q\left(t\right)$ is a *inhibitory point process* [168, 169, 170, 171, 172, 173] based on a Poisson process, so that there is a limit to how close together events can occur. If $q_i \in Q\left(t\right)$ are query times such that the q_i are ordered by the index i; that is,

$$i < j < k \Rightarrow q_i < q_j < q_k \tag{4.16}$$

then the *inter-query times* are $\overline{q}_i = q_{i+1} - q_i$. Each query occurs when the item is used while not in the cache. The item is then cached for period τ, so the minimum inter-query time is τ. The next query occurs at the first time the item is used after leaving the cache. The elapsed time between the item leaving the cache and being used again is simply $\overline{q}_i - \tau \sim \text{Exp}\left(\lambda_U\right)$ due to the memoryless property of the Exponential distribution [134] (see Appendix A.3.4); the time until the next query after the item leaves the cache is independent of the time since the last query before the item left the cache. Hence, the interquery times are *shifted* Exponential(λ_U, τ)

random variables such that if

$$\overline{q}_i \sim \mathrm{Exp}\left(\lambda_U, \tau\right) \tag{4.17}$$

then $\overline{q}_i \in (\tau, \infty)$ and

$$P\left[\overline{q}_i < c\right] = \int_{\tau}^{c} \lambda_U e^{-\lambda_U\left(\overline{q}_i - \tau\right)} d\overline{q}_i = \begin{cases} 0 & c \leq \tau \\ 1 - e^{-\lambda_U c} & c > \tau \end{cases} \tag{4.18}$$

The expected value of $\overline{q}_i$ is

$$\left\langle \overline{q}_i \right\rangle = \frac{1}{\lambda_U} + \tau \tag{4.19}$$

(see Appendix A.3.5) and hence, the expected number of queries is the total time divided by the expected inter-query time

$$\left\langle N_Q\left(T\right)\right\rangle = \frac{T}{\left\langle \overline{q}_i \right\rangle} = \frac{T}{\frac{1}{\lambda_U} + \tau} \tag{4.20}$$

Finally, the expected number of records is given by

$$\left\langle N_R\left(T\right)\right\rangle = \pi \left\langle N_Q\left(T\right)\right\rangle \tag{4.21}$$

We define the rate of accumulation of records as

$$\lambda_R = \lim_{T \to \infty} \frac{\left\langle N_R\left(T\right)\right\rangle}{T} \tag{4.22}$$

which is estimated by

$$\hat{\lambda}_R = \frac{\left\langle N_R\left(T\right)\right\rangle}{T} \tag{4.23}$$

Combining Eq. (4.20), Eq. (4.21) and Eq. (4.23) we derive an estimator $\hat{\lambda}_U$

$$\hat{\lambda}_R = \frac{\left\langle N_R\left(T\right)\right\rangle}{T} = \frac{\pi \left\langle N_Q\left(T\right)\right\rangle}{T} = \frac{\pi}{\frac{1}{\lambda_U} + \tau} \tag{4.24}$$

Rearranging, we derive an estimator $\hat{\lambda}_U$

$$\hat{\lambda}_U = \frac{1}{\frac{\pi}{\hat{\lambda}_R} - \tau} = \frac{1}{\frac{\pi T}{N_R(T)} - \tau} \tag{4.25}$$

In Fig. 4.3, we simulate the system for $T = 1000$, and plot $\hat{\lambda}_R$ for varying parameters λ_U, τ and π, showing that our analytic result matches our simulation.

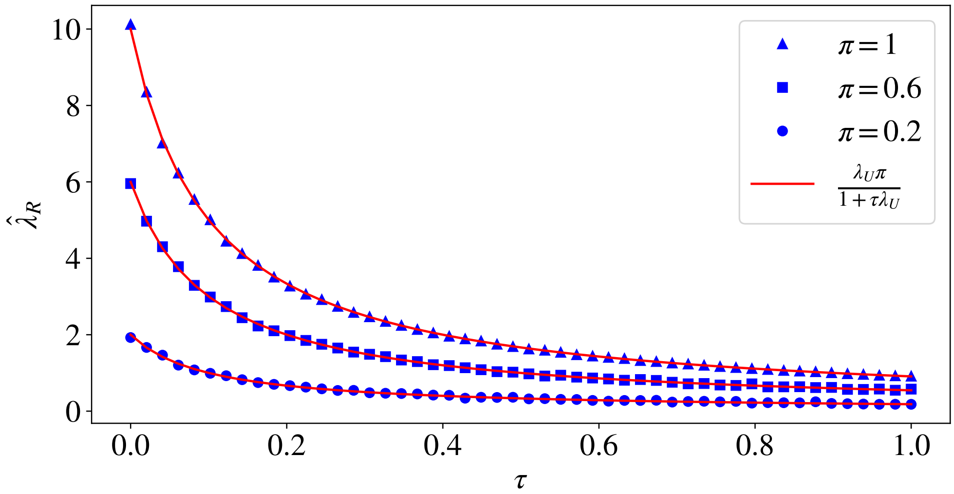

Figure 4.3: We show experimentally that the observed query rate is $\lambda_R = \frac{\lambda_U \pi}{1+\tau\lambda_U}$. Here $\lambda_U = 10$ and $T = 1000$.

4.2.3 Estimation of query volume reduction due to caching

We now apply our estimate for $\hat{\lambda}$ derived above to estimate the reduction in query volume due to caching for a single endpoint and data item, and hence across the entire data network. Let $E = \{e_i\}_{i=1}^{N_E}$ be the set of endpoints, and let $D = \{d_j\}_{j=1}^{N_D}$ be the set of data items. There are $N_R^{ij}(T)$ records of endpoint e_i querying data item d_j after time T. Hence, the item usage rate is estimated as

$$\hat{\lambda}_{ij} = \frac{1}{\frac{\pi T}{N_R^{ij}(T)} - \tau} \tag{4.26}$$

The expected number of times that e_i queries d_j is

$$\left\langle N_Q^{ij}(T) \right\rangle = \frac{1}{\pi} N_R^{ij}(T) \tag{4.27}$$

Therefore, the expected aggregate time during which d_j was cached on e_i is $\frac{\tau}{\pi} N_R^{ij}(T)$, during which it was still being used with estimated rate $\hat{\lambda}_{ij}$. Using the additive property of Poisson random variables [154] (see Appendix A.3.6) the number of times the item was used while cached is

$$\sum_{k=1}^{\frac{1}{\pi}N_R^{ij}(T)} \mathrm{Poi}\left(\hat{\lambda}_{ij}\tau\right) = \mathrm{Poi}\left(\hat{\lambda}_{ij}\frac{\tau}{\pi}N_R^{ij}(T)\right) \tag{4.28}$$

so that the expected number of queries prevented over time interval T is

$$P_{ij}(T) = \hat{\lambda}_{ij}\frac{\tau}{\pi}N_R^{ij}(T) \tag{4.29}$$

By summing over all endpoints and items, we estimate the expected total number of queries prevented by caching as

$$P(T) = \frac{\tau}{\pi} \sum_{i,j} \hat{\lambda}_{ij} N_R^{ij}(T) \tag{4.30}$$

This estimate provides a useful metric for how effective a caching system is at preventing queries. However, it does not take into account whether there is a cost associated in reducing the number of queries through caching. In section 4.3, we consider a common scenario where reducing queries has its own cost.

4.3 Optimal TTL for caching volatile data

In section 4.2 we consider some of the complexities involved in estimating the rate of data item usage on an endpoint. Now we consider how such an estimate may be put to use in estimating the optimal TTL.

As before, consider a single data item. The item is used on the endpoint according to a Poisson process with rate λ_U; if the item is used while it is not cached, the endpoint queries the server for the data item and then caches the item on the endpoint for time τ. However, this data item is also *volatile*; it changes on the server according to some random process (presumed to be a Poisson process) with rate μ. If the item changes on the server while cached on an endpoint, it does not change on the endpoint and so becomes *outdated*. An endpoint using outdated data is a highly undesirable drop in service quality, and so there is a cost associated with an outdated item being used which is greater than the cost of the server responding to a query for a data item every time it is needed. Hence, the cost of a data item being used varies dynamically over time, as visualised in Fig. 4.4. When a data item is queried (blue circle), the usage cost for later queries drops for an interval τ while the item is cached, but if the data changes (red triangle) in that interval, the usage cost is greatly increased until the item leaves the cache. Hence, the usage cost is piecewise-constant and varies dynamically based on the history of the item in the last τ time.

As the cost of using outdated data items is relatively high, we will focus on scenarios where it is also relatively uncommon. Hence, we assume that $\mu\tau \ll \lambda_U\tau$; so that the expected number of times that a cached data item becomes outdated is much less than the expected number of times that a cached data item is used. Unlike estimating λ_U, it is relatively straightforward for the server to estimate μ, as data item usage takes place at the endpoint, while data item changes takes place on the server. If a data item changes according to a Poisson process, then the server need only keep track of the number of times that the item has changed, and the interval of time the item has existed on the server, to estimate rate at which changes

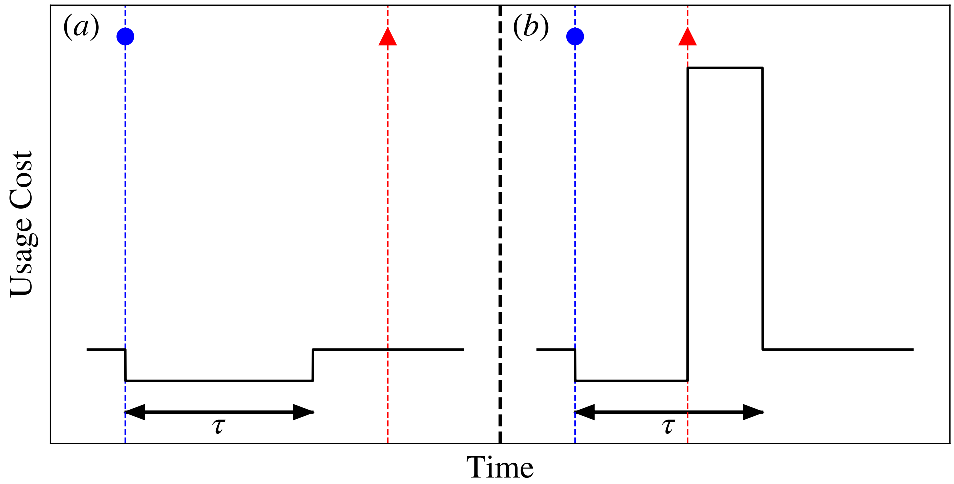

Figure 4.4: Sketch of dynamically varying data usage cost. When a data item is queried (blue circle), the usage cost for later queries drops for an interval τ while the item is cached, but if the data changes (red triangle) in that interval, the usage cost is greatly increased until the item leaves the cache.

occur.

For given μ and λ_U, there is an optimal choice of τ which minimises the combined cost of queries and outdated data usage. In this section, we consider a model of item usage and item volatility which enables us to calculate the optimal choice of τ when the rate of data item change μ is small.

4.3.1 Rate of outdated data usage

We assume that data items are used according to a Poisson process with rate λ_U as in the previous section, and we assume that data items change according to a Poisson process with constant rate μ. Furthermore, we assume that the probability that a data item changes, and then changes back to its previous state, with a time interval of length τ is negligible. This is the case when μ is small, or potentially when the state space of the data item is large. We assume that μ is small; if μ is large, and using outdated data items is undesirable, then caching would not be an appropriate data management method. If a data item changes while cached, we say that the cached data item is now outdated.

Consider a data item which is cached on the endpoint at time t_0, becomes outdated at time $t_0 + T$, and exits the cache at time $t_0 + \tau$. The time T until the data item becomes outdated is an Exponential(μ) random variable such that

$$T \sim \text{Exp}\left(\mu\right) \tag{4.31}$$

$$p\left(T, \mu\right) = \mu e^{-\mu T} \tag{4.32}$$

$$P\left[T < c\right] = \int_0^c \mu e^{-\mu T} dT = 1 - e^{-\mu c} \tag{4.33}$$

Let N_O be the expected number of times that the item is used while outdated in the interval $[0, \tau]$. If $T \geq \tau$, then the number of times the item is used while outdated must be zero; by the time the item has changed, it has left the cache. However, if $T < \tau$, then the expected number of times that the item is used in time interval $\tau - T$ is $\lambda_U (\tau - T)$. Then

$$N_O (T, \tau, \lambda_U) = \begin{cases} \lambda_U (\tau - T) & T < \tau \\ 0 & T \geq \tau \end{cases}$$

Then for each caching of an item, the expected number of times that an outdated item is used while cached is given by the integral

$$F (\mu, \lambda_U, \tau) = \int_0^\infty \left(\mu e^{-\mu T}\right) N_O (T, \tau, \lambda_U) \, dT \tag{4.34}$$

$$= \int_0^\tau \left(\mu e^{-\mu T}\right) \lambda_U (\tau - T) \, dT + \int_\tau^\infty \left(\mu e^{-\mu T}\right) (0) \, dT \tag{4.35}$$

$$= \lambda_U \tau \int_0^\tau \mu e^{-\mu T} dT - \lambda_U \int_0^\tau \mu T e^{-\mu T} dT \tag{4.36}$$

$$= \lambda_U \left[\tau - \frac{1}{\mu} \left(1 - e^{-\mu \tau}\right)\right] \tag{4.37}$$

by integration by parts. F increases linearly with λ_U, and the relationships between F and the other two parameters are plotted in blue in Fig. 4.5. The red circles in Fig. 4.5 show the result of simulating the cache for given parameters, average over 1000 simulations. The simulations agree with our analytic result.

4.3.2 Defining a cost function

Consider a single item and a single endpoint. Each time the item is queried by the endpoint, there is a cost C_Q of the server responding to the query. Each time an outdated data item is used while cached, there is a cost C_O due to company reputation/customer satisfaction. Let γ be the cost of the usage of an outdated data item relative to the cost of a query; if $\gamma = 1000$, then cost of the usage of an outdated data item is 1000 times higher the cost of a query, so that

$$C_O = \gamma C_Q \tag{4.38}$$

Then if there are N_O outdated usages and N_Q queries in a time period T, the total cost is

$$C (T) = N_O \gamma C_Q + N_Q C_Q \tag{4.39}$$

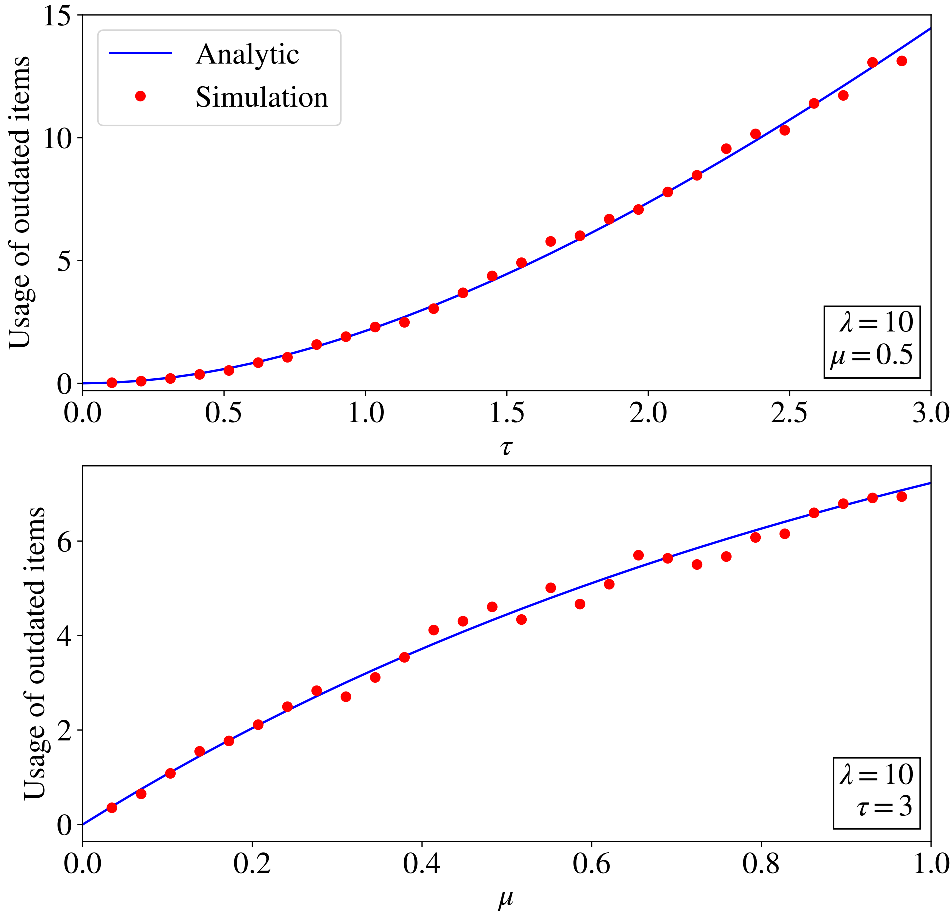

Figure 4.5: For non-zero μ, the expected number of times that an outdated item is used while cached increases monotonically with all parameters μ, λ_U, τ. The analytically derived expected value is shown in blue, while the red dots show the average value from simulations.

in units of the cost of a single query. The expected number of queries to the server in period T is given by Eq. (4.20) as

$$\langle N_Q(T)\rangle = \frac{T}{\frac{1}{\lambda_U}+\tau} \tag{4.40}$$

Each time the item is queried and added to the cache, the probability that it is used while outdated is

$$F(\mu,\lambda,\tau) = \lambda_U\tau - \frac{\lambda_U}{\mu}\left(1-e^{-\mu\tau}\right) \tag{4.41}$$

Hence, the expected number of usages while outdated in period T is

$$\langle N_O(T)\rangle = \langle N_Q(T)\rangle F(\mu,\lambda_U,\tau) \tag{4.42}$$

$$= \frac{T}{\frac{1}{\lambda_U}+\tau}\left(\lambda_U\tau - \frac{\lambda_U}{\mu}\left(1-e^{-\mu\tau}\right)\right) \tag{4.43}$$

Then the expected cost is

$$\langle C(T)\rangle = \frac{T}{\frac{1}{\lambda_U}+\tau}\left(\lambda_U\tau - \frac{\lambda_U}{\mu}\left(1-e^{-\mu\tau}\right)\right)\gamma C_Q + \frac{T}{\frac{1}{\lambda_U}+\tau}C_Q \tag{4.44}$$

$$= TC_Q\frac{1}{\frac{1}{\lambda}+\tau}\left[\left(\lambda\tau - \frac{\lambda}{\mu}\left(1-e^{-\mu\tau}\right)\right)\gamma+1\right] \tag{4.45}$$

By normalising $C = \frac{\langle C(T)\rangle}{TC_Q}$, we get unitless expected cost per unit time of a single item on a single endpoint

$$C = \frac{1}{\frac{1}{\lambda_U}+\tau}\left[\left(\lambda_U\tau - \frac{\lambda_U}{\mu}\left(1-e^{-\mu\tau}\right)\right)\gamma+1\right] \tag{4.46}$$

The server cannot control μ or λ_U, as μ characterises the data item and λ_U characterises the endpoint. The server may have limited control over γ, as γ is determined by the relative cost of incorrect usage relative to queries, and this can perhaps be altered through higher level business decisions. However, the server has complete control over τ. Hence, we try to maximise C w.r.t to τ. Normally, we would do so by simply setting $\frac{dC}{d\tau} = 0$. However, due to the form of C, this does not yield a explicit optimal τ. We can approach an analytic solution by making use of our assumption that μ is small.

4.3.3 Optimisation under small μ approximation

We wish to maximise

$$C = \frac{1}{\frac{1}{\lambda_U}+\tau}\left[\left(\lambda_U\tau - \frac{\lambda_U}{\mu}\left(1-e^{-\mu\tau}\right)\right)\gamma+1\right] \tag{4.47}$$

with respect to τ. The first minus sign suggests that there is a non-trivial optimal value of τ, but due to the exponential term, we cannot obtain this optimal value by differentiation analytically. The $e^{-\mu\tau}$ term has a series expansion

$$e^{-\mu\tau} = 1 - \mu\tau + \frac{(\mu\tau)^2}{2} - \frac{(\mu\tau)^3}{6} + ... \tag{4.48}$$

If we take $\mu \ll \lambda_U$ and $\mu \ll \tau$, then for small $\mu\tau$, we can approximate the exponential term by by truncating the series. A first order approximation of $e^{-\mu\tau}$

$$e^{-\mu\tau} = 1 - \mu\tau + O\left((\mu\tau)^2\right) \tag{4.49}$$

yields

$$\begin{align} C &= \frac{1}{\frac{1}{\lambda_U} + \tau}\left[\left(\lambda\tau - \frac{\lambda_U}{\mu}(1-(1-\mu\tau))\right)\gamma + 1\right] \tag{4.50} \\ &= \frac{1}{\frac{1}{\lambda_U} + \tau} \tag{4.51} \end{align}$$

which is minimised by $\tau \to \infty$. This is trivial, and contradicts both our approximation that $\mu\tau$ is small, and our expectation that the TTL should expire at some point. If we instead take a second order approximation of $e^{-\mu\tau}$

$$e^{-\mu\tau} = 1 - \mu\tau + \frac{(\mu\tau)^2}{2} + O\left((\mu\tau)^3\right) \tag{4.52}$$

then

$$\begin{align} C &= \frac{1}{\frac{1}{\lambda_U} + \tau}\left[\left(\lambda_U\tau - \frac{\lambda_U}{\mu}\left(1 - \left(1 - \mu\tau + \frac{(\mu\tau)^2}{2}\right)\right)\right)\gamma + 1\right] \tag{4.53} \\ &= \left(\frac{1}{\frac{1}{\lambda_U} + \tau}\right)\left[\left(\frac{\lambda_U\mu\tau^2}{2}\right)\gamma + 1\right] \tag{4.54} \end{align}$$

Differentiating, we get

$$\frac{dC}{d\tau} = \left[\left(\frac{\lambda_U\mu\tau^2}{2}\right)\gamma + 1\right]\frac{d}{d\tau}\left(\frac{1}{\frac{1}{\lambda_U} + \tau}\right) + \left(\frac{1}{\frac{1}{\lambda_U} + \tau}\right)\frac{d}{d\tau}\left[\left(\frac{\lambda_U\mu\tau^2}{2}\right)\gamma + 1\right] \tag{4.55}$$

$$\begin{align} 0 &= \left[\left(\frac{\lambda_U\mu\tau^2}{2}\right)\gamma + 1\right]\frac{-1}{\left(\frac{1}{\lambda_U} + \tau\right)^2} + \left(\frac{1}{\frac{1}{\lambda_U} + \tau}\right)(\lambda_U\mu\tau)\gamma \tag{4.56} \\ &= \frac{\lambda_U\mu\gamma}{2}\tau^2 + \mu\gamma\tau - 1 \tag{4.57} \end{align}$$

which is quadratic in τ, hence

$$\tau = \frac{-\mu\gamma \pm \sqrt{(\mu\gamma)^2 - 4\left(\frac{\lambda_U \mu\gamma}{2}\right)(-1)}}{2\frac{\lambda_U\mu\gamma}{2}} \tag{4.58}$$

$$= -\frac{1}{\lambda_U} \pm \frac{1}{\lambda_U}\sqrt{1 + \frac{2\lambda_U}{\mu\gamma}} \tag{4.59}$$

$\tau > 0$, so that the optimal TTL, $\hat{\tau}$ is given by

$$\hat{\tau} = \frac{1}{\lambda_U}\left(-1 + \sqrt{1 + \frac{2\lambda_U}{\mu\gamma}}\right) \tag{4.60}$$

Fig. 4.6 shows C for varying parameters; our small μ approximation appears to provide a reasonable estimate for the optimal value of τ.

By applying our estimated item usage rate from Eq. (4.25), and estimating the rate at which the data item changes using server logs, Eq. (4.60) can provide an optimal choice of τ for a single data item, and a single endpoint. However, all of this rests on the assumption that data items are used according to a Poisson process with constant (homogeneous) rate for a period of time $T \gg \tau$. In the next section, we consider whether this is accurate, and consider some dynamical behaviour that arises from modelling item usage as a non-homogeneous Poisson process.

4.4 Symbolic dynamics of TTL caching with piecewise-constant periodic rates of data usage

4.4.1 Non-homogeneous Poisson processes

So far, we have assumed that data item usage follows a Poisson process with constant rate; this is also referred to as a homogeneous Poisson process. As noted earlier in this chapter, that assumption has been widely used because it is very simple and often adequate for the purpose of constructing a model of a natural counting process. However, it has long been recognised that there are natural counting processes which deviate significantly from the Poisson process. A famous early example of this Ugo Fano's work on the ionization yield of radiation in the 1940s [174, 175]. Fano found that the variation in the number of ionizations in a gas was three times less than what would be predicted by a Poisson process with the same rate of ionization. This work led to the *Fano factor* $F(t)$ [176], defined for a counting process $N(t)$ as

$$F(t) = \frac{\left\langle N(t) - \langle N(t)\rangle^2 \right\rangle}{\langle N(t)\rangle} = \frac{V[N(t)]}{E[N(t)]} \tag{4.61}$$

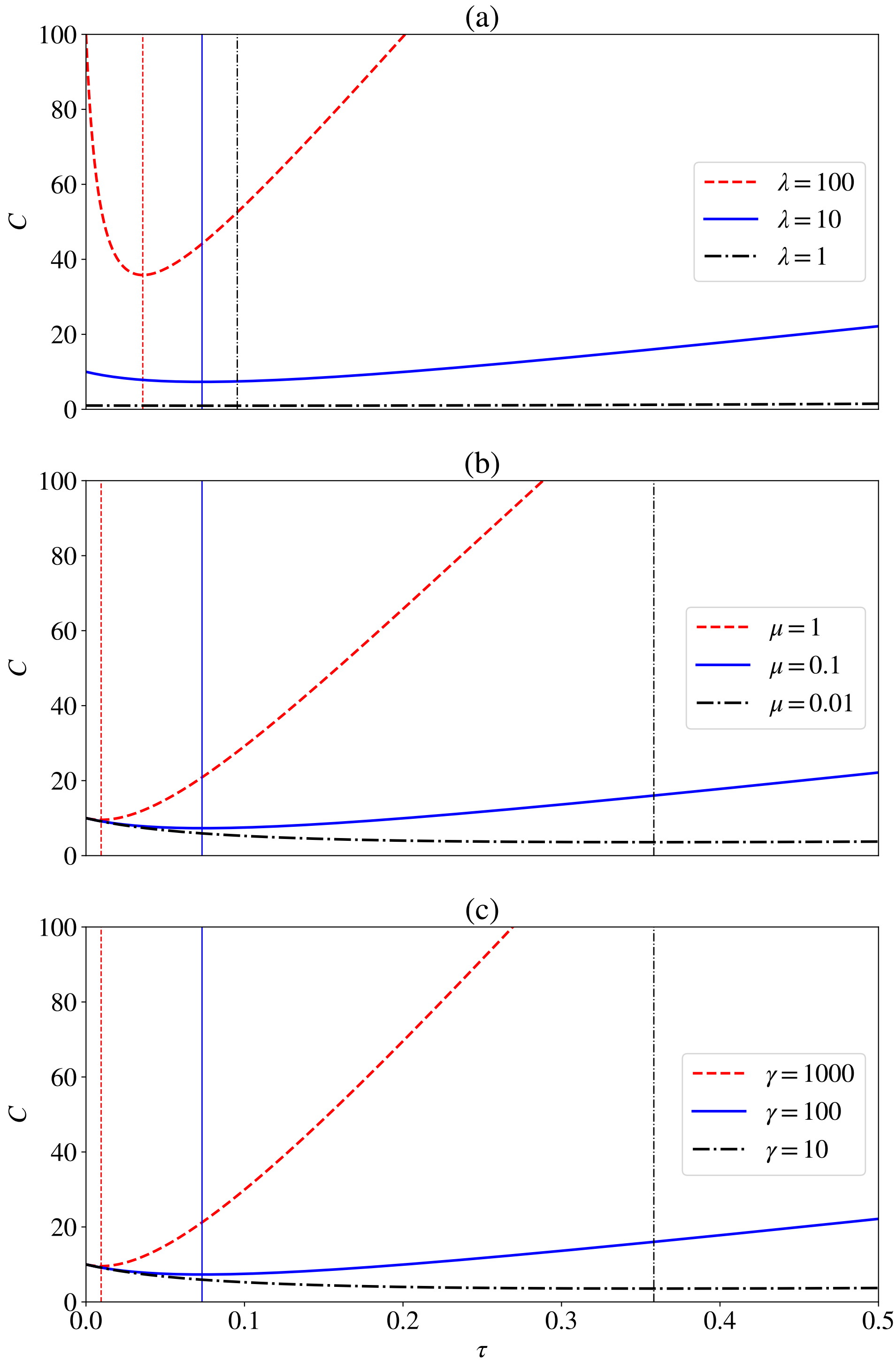

Figure 4.6: The expected cost per unit time C of serving a data item for varying τ. The default hidden parameters are $(\lambda, \mu, \gamma) = (10, 0.01, 100)$. The thin vertical lines indicate the optimal TTL $\hat{\tau}$.

or the variance of the process divided by its mean. A variation of the Fano factor is the *index of dispersion* D [177] which takes the form of the Fano factor in the limit $t \to \infty$

$$D = \lim_{t\to\infty} \frac{\left\langle N(t) - \langle N(t)\rangle^2 \right\rangle}{\langle N(t)\rangle} = \lim_{t\to\infty} \frac{V[N(t)]}{E[N(t)]} \tag{4.62}$$

For example, $N_U(t)$ has index of dispersion

$$D_U = \frac{\lambda_U}{\lambda_U} = 1 \tag{4.63}$$

which is characteristic of a standard Poisson process. $N_Q(t)$ on the other hand, has mean and variance given by

$$\lim_{t\to\infty} \frac{\langle N(t)\rangle}{t} = \frac{1}{\frac{1}{\lambda_U} + \tau} \tag{4.64}$$

$$\lim_{t\to\infty} \frac{\left\langle N(t) - \langle N(t)\rangle^2 \right\rangle}{t} \simeq \frac{1}{\lambda_U^2 \left(\frac{1}{\lambda_U} + \tau\right)^3} \tag{4.65}$$

(see Appendix A.3.8) Hence, $N_Q(t)$ has index of dispersion

$$D_Q = \frac{\frac{1}{\lambda_U^2\left(\frac{1}{\lambda_U}+\tau\right)^3}}{\frac{1}{\frac{1}{\lambda_U}+\tau}} = \frac{1}{\lambda_U^2 \left(\frac{1}{\lambda_U} + \tau\right)^2} = \frac{1}{(1+\lambda_U\tau)^2} \tag{4.66}$$

As λ_U increases, the mean and variance of the stochastic component of the inter-query time decrease. As τ increases, the deterministic component of the inter-query time increases. As $\lambda_U\tau \to \infty$, $D_Q \to 0$, and the query process becomes deterministic.

Once we move beyond the strict definition of the homogeneous Poisson process, there is a wide range of more complex stochastic processes which have been found to be of use in modelling diverse phenomena such as earthquakes [178], neuron spiking [179], and spreading dynamics in social networks [180]. Some real-world processes are found to be 'bursty', with autocorrelation between event times [181, 182, 183], with particularly relevant examples appearing in the context of Internet communications [184, 185, 186] and data networks [187, 188, 189, 190]. Others are modelled using non-homogeneous variations of the Poisson process that have a rate which varies in time [191, 192, 193, 194].

To motivate this point in our own context of cybersecurity, we consider query data from a data network in the cybersecurity industry. In this context, endpoints send queries about objects such as files, URLs, and IP addresses, and data items are the current reputations of the queried object based on current information about that object. Our dataset consists of queries made by endpoints in the Republic of Ireland using the McAfee product WSS-LAM, which provides reputations for files. We restrict our attention to a class of files with caching period $\tau = 30$ minutes over

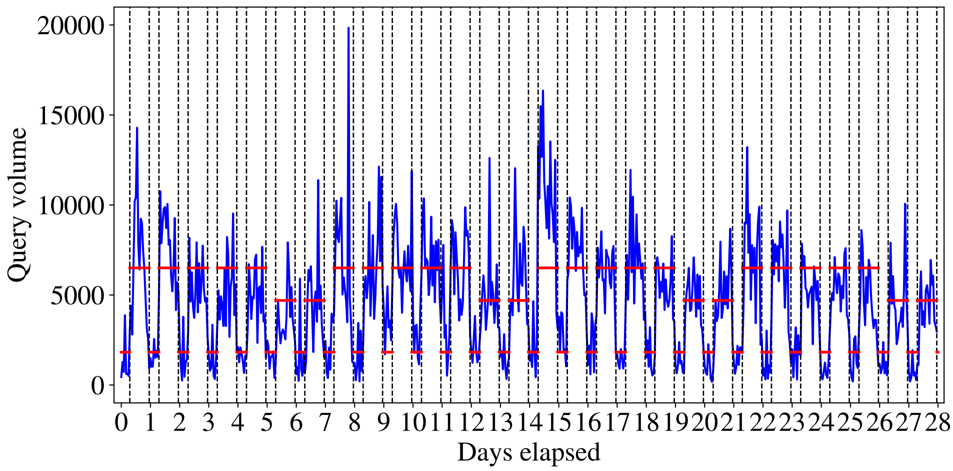

Figure 4.7: Total query volume per hour from endpoints using McAfee cybersecurity software in the Republic of Ireland plotted in blue. The average total query volume for daytime on weekdays, average total query volume for daytime on weekends, and average total query volume for nighttime are plotted in red.

a four week period between midnight on Monday 1st August and midnight on 29th August 2022. The data query data is sampled at a rate $\pi = 0.5$, so that every query has a 50% chance of being included in the dataset. In Fig. 4.7, we show the recorded query volume per hour over 28 days plotted in blue.

The query volume varies significantly over time, and is prone to large spikes. However, we can note that total query volume per hour is generally lower in daytime on weekends than in daytime during the week. It is also significantly lower during the night than during the day. The average total query volume for weekday days, weekend days, and nights are plotted in red. The day/night cycle was chosen such that nighttime is the eight hour period starting on a whole hour such that the difference in query volume between night and day is maximised. Even using very simple metrics, we see that in aggregate, the query volume, and hence, the rate of querying, varies periodically and significantly between day and night. This does not exclude the possibility that the process is also a bursty process; that will require further study at the level of individual endpoints. For now, we would like to consider the dynamical effects of periodic piecewise-constant query rate.

4.4.2 Stochastic model

Having seen that query rates vary periodically over time, we now consider a variation of our model where $\pi = 1$ and λ_U varies according to periodic step function, so that item usage is a piecewise-constant Poisson process [195, 196, 197]. Let

$$\lambda_U(t) = \begin{cases} \lambda_1 & \text{if } t \mod 1 \leq \rho \\ \lambda_2 & \text{if } t \mod 1 > \rho \end{cases} \tag{4.67}$$

where ρ determines the proportion of the time $\lambda_U = \lambda_1$. $\lambda_U(t)$ is periodic with period 1. If $\lambda_1 = \lambda_2$, the rate is constant, so we assume $\lambda_1 > \lambda_2$. The system now has parameters λ_1, λ_2, ρ and τ. With the introduction of the step function, the dynamics of the system now take place on a circle with circumference 1, and so we observe the dynamics with respect to $\theta = t \mod 1$. Fig. 4.8 shows the result of simulating the system for varying ρ and τ with $\lambda_1 = 100$ and $\lambda_2 = 0$. Both plots show four peaks where queries take place, but the overall shape is different. Consider Fig. 4.8(a), where $\tau = 0.2$. The system starts at $\theta = 0$ with nothing cached. The time until the data item is queried is an $\text{Exp}(\lambda_1)$ random variable; hence, the shape of the first peak is given by an $\text{Exp}(\lambda_1)$ probability density function which starts at $\theta = 0$. The time interval between the first and second query is $\tau+\text{Exp}(\lambda_1)$, so that the second query occurs at time $\text{Exp}(\lambda_1) + \tau + \text{Exp}(\lambda_1)$,where the random variables are independent. The sum of k independent $\text{Exp}(\lambda)$ random variables is a $\text{Gamma}(k, \lambda^{-1})$ random variable [154, 167] (see Appendix A.3.7), with probability density function

$$f(x, k, \lambda) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \tag{4.68}$$

Hence, the second peak has the shape of a $\text{Gamma}(2, \lambda^{-1})$ probability density function, the third has $\text{Gamma}(3, \lambda^{-1})$ shape, and the fourth has a $\text{Gamma}(4, \lambda^{-1})$ shape. There is no fifth peak, as the data item leaves the cache at $\theta > \rho$; instead, the system arrives at $\theta = 0$ with nothing in the cache to repeat the cycle. This creates an orbit on the circle where each successive peak is closer to $[\rho, 1]$, eventually entering the interval where $\lambda = 0$ and returning to $\theta = 0$. Fig. 4.8(c) shows an analytic reconstruction of the query distribution in Fig. 4.8(a) using the $\text{Gamma}(k, \lambda^{-1})$ probability density function, which is identical to that found as a histogram.

If we compare this to Fig. 4.8(b), we note that the shapes of the peak are the same, but out of order. This is due to τ now being large enough that it is possible for the orbit jumps over $[\rho, 1]$ from an appropriate starting point. We can further explore the dynamics of this model by observing that the two orbits have a different maximum value for θ. In Fig. 4.9, we plot the maximum θ achieved by iterating the model for two different values of λ. We note that as we increase λ, a pattern of triangles begins to emerge. To better understand this pattern, we note that as $\lambda \to \infty$, $\text{Exp}(\lambda) \to 0$, so that the next query happens as soon as the data item leaves the cache. Hence, in the limit $\lambda_1 \to \infty$ and $\lambda_2 = 0$, the system simplifies to a deterministic map.

4.4.3 Deterministic limit as a circle map

We construct a map

$$\theta_{n+1} = f(\theta_n) \tag{4.69}$$

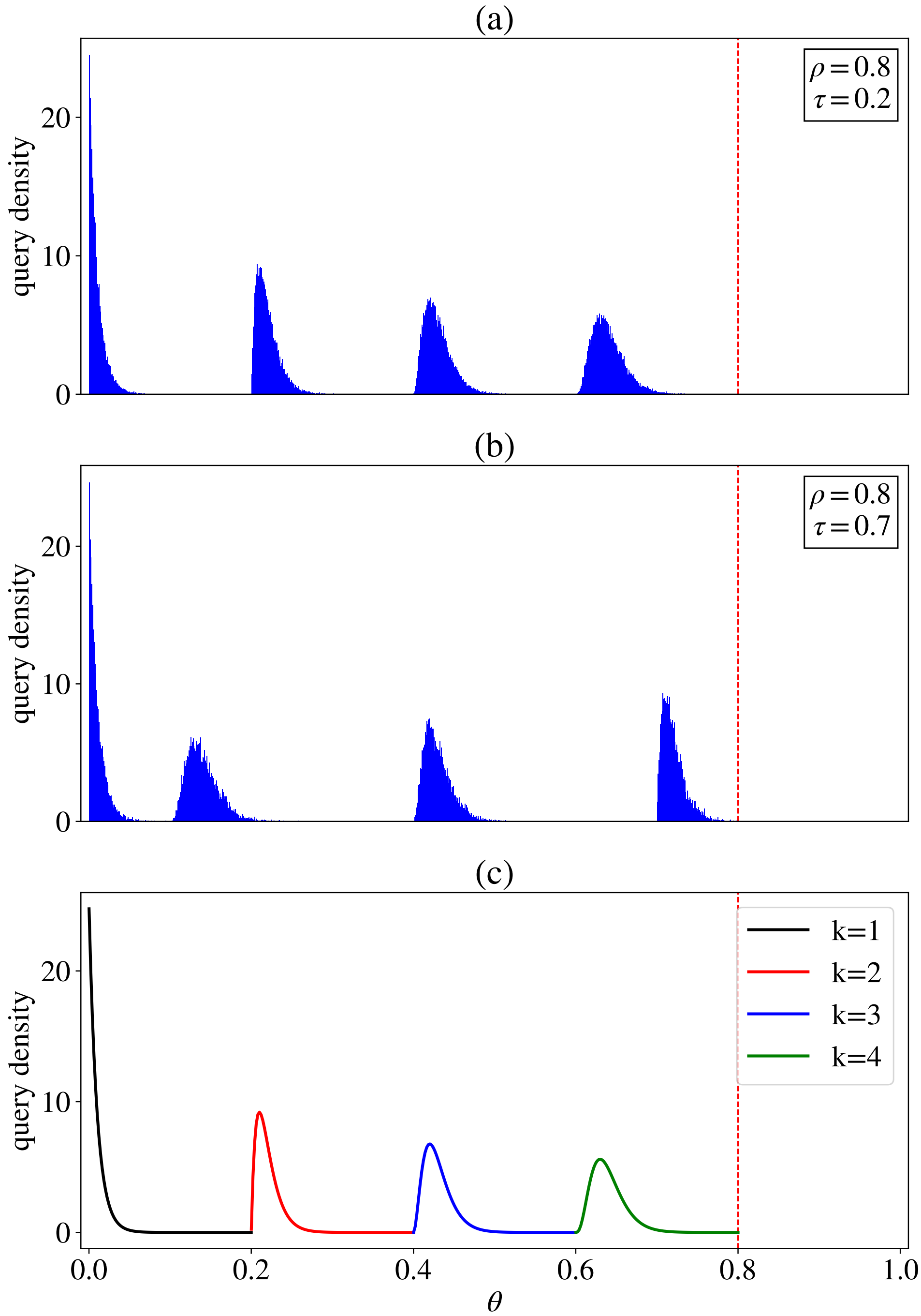

Figure 4.8: Two comparable caching time trajectories for $\lambda_1 = 100$, $\lambda_2 = 0$ and $\rho = 0.8$. $\tau = 0.1$ in (a) and $\tau = 0.7$ in (b). (c) shows the distribution of queries for (a) analytically derived from the Gamma(k, λ) distribution.

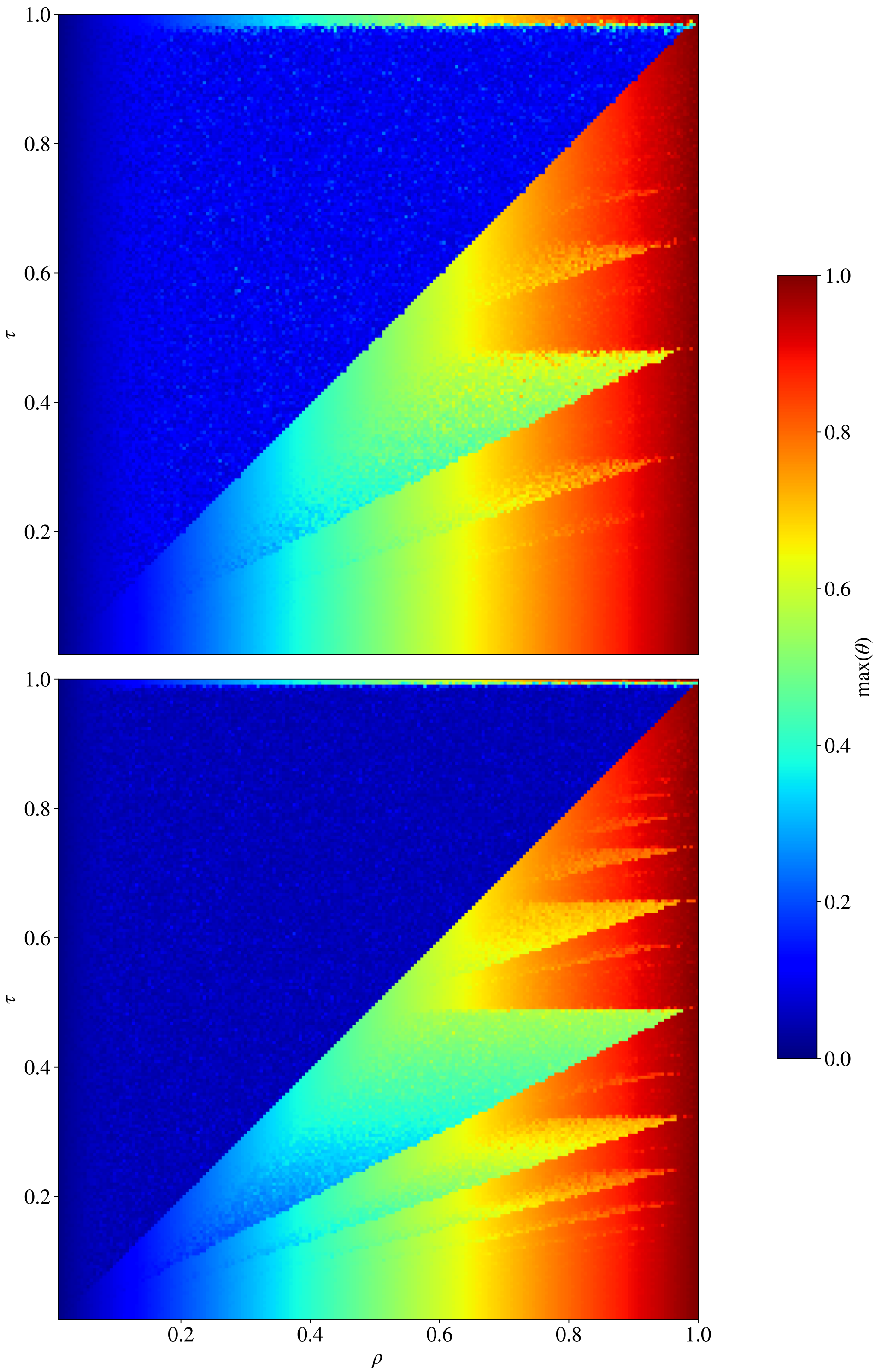

Figure 4.9: Colour maps showing the maximum θ after iterating for varying parameters (ρ, τ). $\max(\theta)$ ranges from 0 (dark blue) to 1 (dark red). $\lambda_1 = 50$ in the first colour map, and $\lambda_2 = 100$ in the second. A distinctive pattern emerges as λ_1 increases.

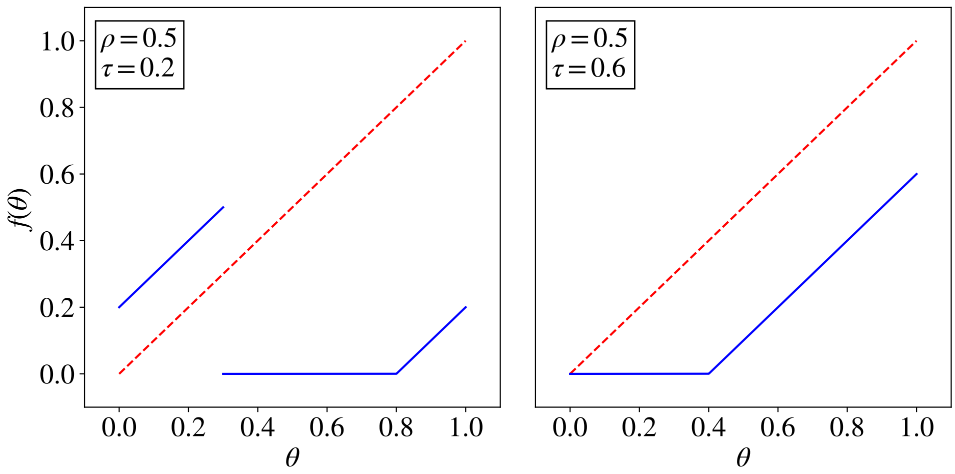

Figure 4.10: Plot of the analytic map $f(\theta)$.

where if the item is queried at time θ_n, θ_{n+1} is the next time at which the item is queried, such that

$$f(\theta;\rho,\tau) = \begin{cases} \theta+\tau & \text{if } \theta+\tau<\rho \\ 0 & \text{if } \rho\leq\theta+\tau<1 \\ (\theta+\tau) \mod 1 & \text{if } 1\leq\theta+\tau \end{cases} \tag{4.70}$$

as shown in Fig. 4.10. This map is comparable to the circle map [149]; we can write Eq. (4.70) in the form

$$\theta_{n+1} = \tau + g(\theta) \tag{4.71}$$

where $g(t)$ is a nonlinear function

$$g(\theta) = \begin{cases} 0 & \theta+\tau<\rho \\ 1-(\theta+\tau) & \theta+\tau\geq\rho \end{cases} \tag{4.72}$$

The circle map has been studied for a wide variety of nonlinear functions g [149, 150, 151, 198, 199, 200, 201], including a number of discontinuous functions [202, 203, 204, 205, 206, 207, 208]. The unique feature of our map is that it is linear within $[0,\rho]$, but contracts $[\rho,1]$ to a point. We iterate the map for varying (ρ,τ), which can be seen in Fig. 4.11. We observe that the distinctive pattern of triangles seen in Fig. 4.8 have resolved clearly. As τ increases, $\max(\theta)$ increases linearly within each triangle. However, at the boundary of each triangle, $\max(\theta)$ changes discontinuously. To understand why, we consider the system through a symbolic representation.

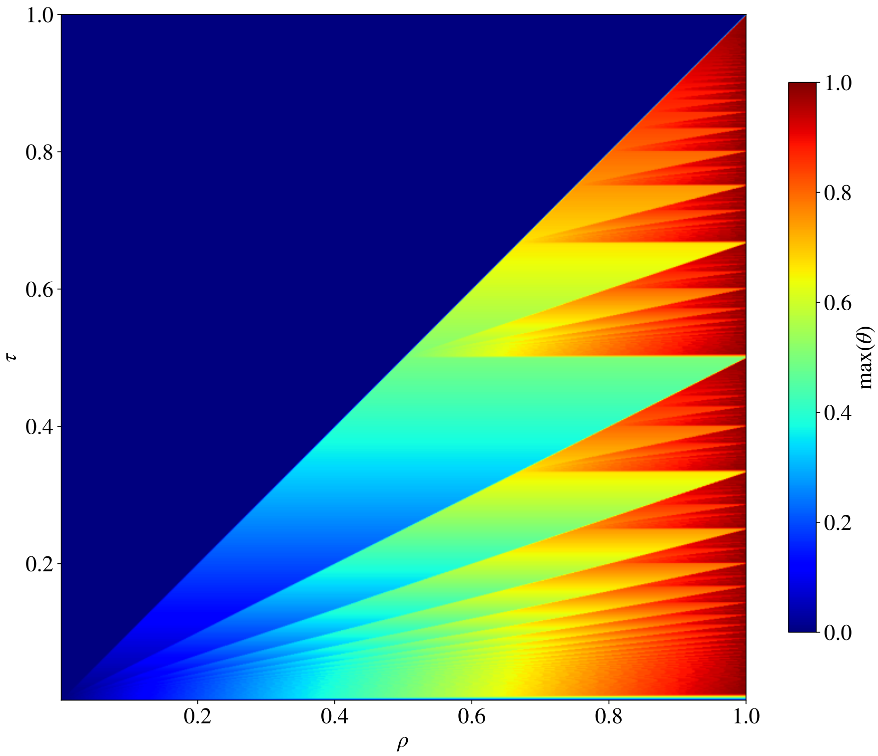

Figure 4.11: Maximum observed value of θ for orbits of the unit circle under the map $\theta_{n+1} = f(\theta_n)$ which begin at $\theta_0 = 0$.

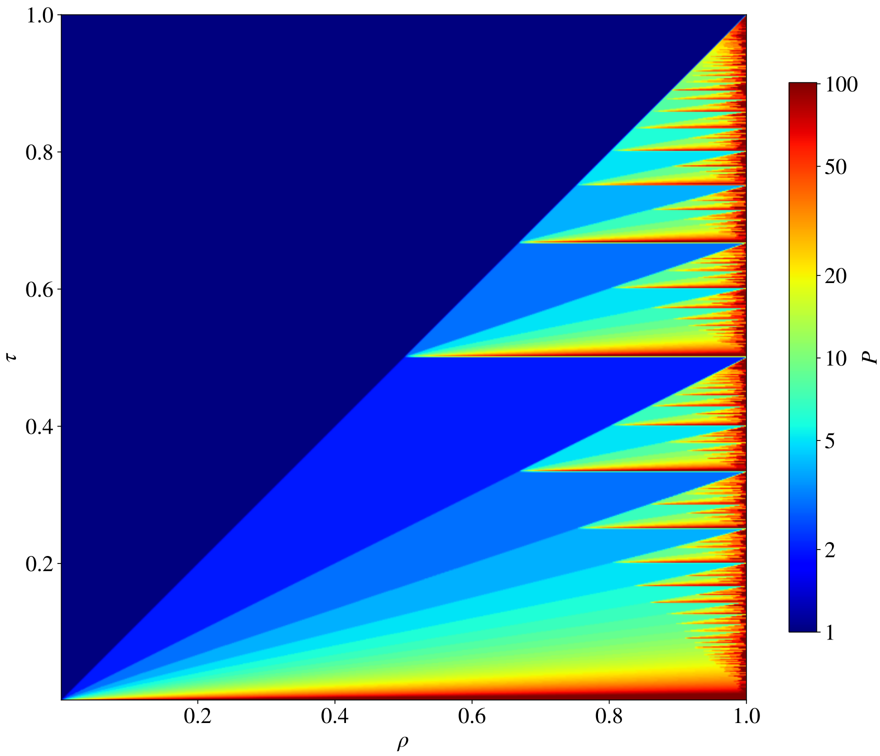

Figure 4.12: Period P for orbits of the unit circle under the map $\theta_{n+1} = f(\theta_n)$ which begin at $\theta_0 = 0$.

4.4.4 Symbolic dynamics

We recall the interesting ordering effect in Fig. 4.8, where the shape of a peak was determined by its order in the orbit. Similarly, $f(\theta;\tau,\rho)$ generates an orbit $\left\{f^k(\theta;\tau,\rho)\right\}_{k=0}^{\infty}$ which can be represented symbolically in terms of where a particular point maps to; left of ρ, between ρ and 1, or right of 1. We can define this by

$$s(\theta;\rho,\tau) = \begin{cases} L & \text{if } \theta+\tau<\rho \\ C & \text{if } \rho \le \theta+\tau < 1 \\ R & \text{if } 1 \le \theta+\tau \end{cases} \tag{4.73}$$

A periodic orbit can be defined by a periodic sub-sequence of symbols of length P. In Fig. 4.12, P is plotted for varying parameters.

We observe that the pattern of triangles observed before is more clear, and conclude that the pattern is associated with changing sequences. Fig. 4.13 shows the sequences for some of the larger tongues.

There are multiple tongues of the same period, but each tongue has a unique sequence representation in terms of the symbols L, R, and C. However, a period

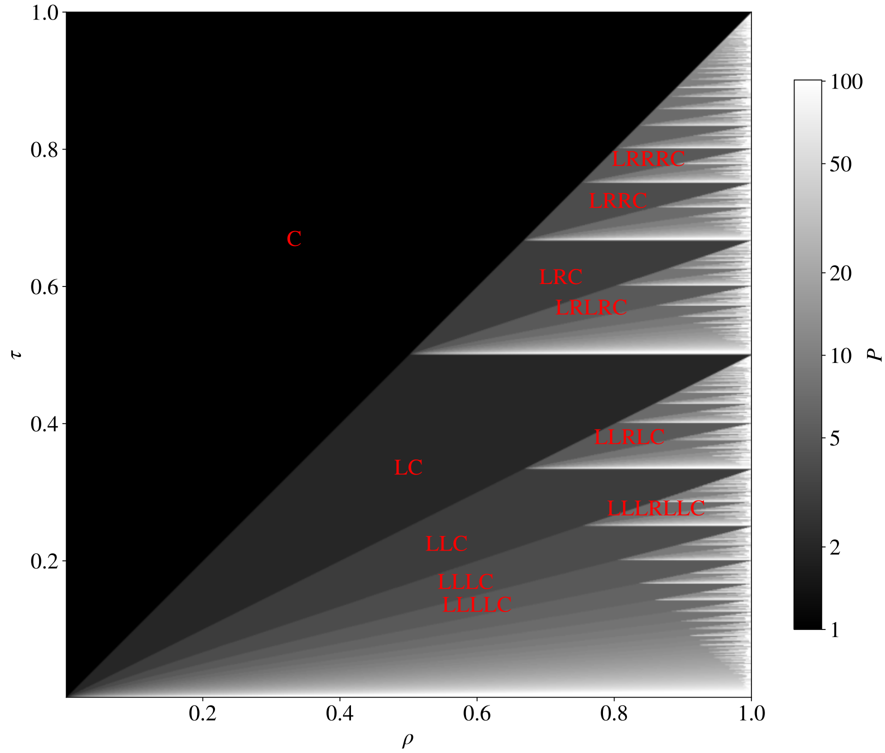

Figure 4.13: Period P for orbits of the unit circle under the map $\theta_{n+1} = f(\theta_n)$ which begin at $\theta_0 = 0$ (identical to Fig. 4.12). Some of the larger tongues have been labelled with the sequence of their orbit.

(ρ, τ)	Sequence	Orbit in θ
$(0.33, 0.67)$	C	$\{0\}$
$(0.50, 0.33)$	LC	$\{0.33, 0\}$
$(0.56, 0.22)$	LLC	$\{0.22, 0.44, 0\}$
$(0.72, 0.61)$	LRC	$\{0.61, 0.22, 0\}$
$(0.58, 0.17)$	$LLLC$	$\{0.17, 0.34, 0.51, 0\}$
$(0.81, 0.72)$	$LRRC$	$\{0.72, 0.44, 0.16, 0\}$
$(0.60, 0.13)$	$LLLLC$	$\{0.13, 0.26, 0.39, 0.52, 0\}$
$(0.85, 0.78)$	$LRRRC$	$\{0.78, 0.56, 0.34, 0.12, 0\}$
$(0.77, 0.57)$	$LRLRC$	$\{0.57, 0.14, 0.71, 0.28, 0\}$
$(0.82, 0.38)$	$LLRLC$	$\{0.38, 0.76, 0.14, 0.52, 0\}$

Table 4.1: Sample orbits for unique legal sequences up to period 5 which end at C.

$n+1$ sequence has $n+1$ symbols, which allows for a total of 2^n sequence permutations (the C at the end is constant). There are only two observed period 6 orbits, but $2^6 = 32$ possible sequences. For sequences of length $n+1$, we observe up to n different tongues; this is only observed when $n+1$ is prime. This means that for a sequence of a given length, only a small fraction of them are *legal*, representing a orbit that can exist. Table 4.1 shows all legal sequences which end at C up to period 5 with orbits for given (ρ, τ). Given a sequence, we can generate a set of inequalities using $s(\theta; \rho, \tau)$, which indicate a region of the plane where the orbit exists. For example, the sequence LRC yields inequalities

$$f(\theta) = \tau \mod 1 < \rho \tag{4.74}$$

$$f^2(\theta) = (2\tau) \mod 1 \geq 1 \tag{4.75}$$

$$f^3(\theta) = (3\tau) \mod 1 < 1 \tag{4.76}$$

$$f^3(\theta) = (3\tau) \mod 1 \geq \rho \tag{4.77}$$

However, the low ratio of legal sequences to illegal sequences makes this an inefficient method of searching for legal sequences. Fortunately, we can determine rules for how legal sequences are generated based on observations from Fig. 4.13.

1. For $\tau \geq \rho$, the only orbit is the C orbit. This occurs because for $\tau \geq \rho$, any orbit that starts at $\theta = 0$ will map to the C region, and get sent back to 0.

2. For $\tau < 1 - \rho$, there are no orbits containing R symbols, as there is no $\theta < \rho$ such that $\theta + \tau \geq 1$.

3. All tongues without an R in their name connect to the origin and follow a common rule. As τ decreases from $\tau = \rho$, the integer k such that $k\tau < \rho \leq (k+1)\tau$ increases, and hence, the period of the orbit increases. This generates a set of tongues fanning out from the origin. Starting from the C tongue and moving clockwise, add an L in front of the C to find the sequence for the next tongue, generating C, LC, LLC, $LLLC$, ... L^kC.

4. Tongues with an R in their sequence are born from adjacent lower period tongues. A higher period 'child' tongue has two adjacent lower period 'parent' tongues. The left parent meets the child along its left edge. The lower parent meets the child tongue at its lower left vertex. In order to derive the sequence for the child tongue, start with the sequence of the lower parent. Change the C at the end of the sequence to an R, then append the sequence of the left parent. For example, the tongue containing the sequence LRC has C as left parent and LC as lower parent.

Using these rules, we can find orbits that exist, and then use $s(\theta; \rho, \tau)$ to find regions of the (ρ, τ) plane where they exist.

Theorem: If $\tau, \rho \in (0, 1)$ then every orbit that begins at $\theta = 0$ must pass through $\theta = 0$ again.

Proof: Consider the circle map

$$h(\theta; \tau) = (\theta + \tau) \mod 1 \tag{4.78}$$

and note that

$$h(\theta; \tau) = f(\theta; \tau, 1) \tag{4.79}$$

$h(\theta; \tau)$ defines an orbit $\left\{h^k(\theta; \tau)\right\}_{k=1}^{\infty}$ on the unit circle. An n-periodic orbit is an orbit such that such that

$$h^n(\theta; \tau) = \theta \tag{4.80}$$

Then an orbit beginning at $\theta = 0$ is defined by

$$h^n(0) = (n\tau) \mod 1 \tag{4.81}$$

From here, we consider two cases.

(a) First, consider rational $\tau \in \mathbb{Q}$. If τ is rational, then

$$\tau = \frac{m}{n} \tag{4.82}$$

for some integers $m, n \in \mathbb{N}$. Then

$$\begin{aligned} h^n(0) &= \left(n\frac{m}{n}\right) \mod 1 & (4.83)\\ &= (m) \mod 1 & (4.84)\\ &= 0 & (4.85) \end{aligned}$$

Hence, $h^n(0)$ is n-periodic. Let

$$n_f = \begin{cases} n & \text{if } g^k(0) < \rho \text{ for } k = 1, 2, ..., n \\ \min\left\{k : \rho \leq h^k(0) < 1\right\} & \text{o.w.} \end{cases} \tag{4.86}$$

Then the orbit $\left\{f^k(\theta;\tau,\rho)\right\}_{k=1}^{\infty}$ is n_f-periodic. $\left\{f^k(\theta;\tau,\rho)\right\}_{k=1}^{\infty}$ is at most n-periodic, but the orbit can return to 0 earlier if a point on the orbit enters $[\rho, 1]$.

(*b*) Next, consider irrational $\tau \in \mathbb{R}\backslash\mathbb{Q}$. Then $\left\{h^k(\theta;\tau)\right\}_{k=1}^{\infty}$ is dense in $[0,1]$ [209, 149], and therefore the orbit passes arbitrarily close to every point in $[\rho, 1]$. Hence, there are points in $\left\{h^k(\theta;\tau)\right\}_{k=1}^{\infty}$ which enter $[\rho, 1]$, and so $\left\{f^k(\theta;\tau,\rho)\right\}_{k=1}^{\infty}$ must also include points in $[\rho, 1]$, which map to $\theta = 0$.

Therefore, if $\tau, \rho \in (0,1)$, then every orbit that begins at $\theta = 0$ must pass through $\theta = 0$ again. ■

4.4.5 Orbits that do not pass through $\theta = 0$

Above, we consider orbits which start at $\theta = 0$ and end at $\theta = 0$, which must therefore contain C in their sequence. However, these are not the only possible orbits. Consider $\tau = \frac{1}{3}$ and $\rho > \frac{5}{6}$. Let $\theta = \frac{1}{2}$, then

$$\theta = \frac{1}{2}, \quad f(\theta) = \frac{5}{6}, \quad f^2(\theta) = \frac{1}{6}, \quad f^3(\theta) = \frac{1}{2} \tag{4.87}$$

This yields a 3-periodic orbit with sequence LRR which does not contain C. Such orbits are only possible for rational $\tau > 1 - \rho$; otherwise, the orbit would be dense in $[\rho, 1]$, or wouldn't be able to jump $[\rho, 1]$. Considering initial $\theta > 0$ extends the original problem into an additional dimension, and so it would be be interesting to study such orbits to see how they compare to orbits that start at $\theta = 0$.

4.5 Conclusion

In this chapter, we construct a model for generating query data from a data network which features caching to reduce query volume and random sampling to reduce record volume. By modelling data usage as a Poisson process, we estimate the rate of data usage based on the rate of data recording. We expand our model to estimate the optimal caching period for volatile data, and finally, we consider the dynamical effect of non-homogeneous data usage rates through the lens of symbolic dynamics in a deterministic limit of our model.

As noted at the beginning of this chapter, the material presented requires further development before it wilinvl be suitable for publication. This further development may be undertaken in several different directions. To take the first section as an example, the model of query data generation is somewhat simple due to our choice of a Poisson process for data usage. The Poisson process is attractive because it is simple and relatively tractable. One interesting avenue of enquiry would be to continue to study the presented version of the model and derive analytic results for distributions and moments. However, we have noted that there are applications where the Poisson process is not sufficiently accurate, and that this may be one of those applications. In that case, it would be of practical interest to extend the model to a more general data usage process. Even then, a full analysis of the Poisson version would make a valuable point of reference.

Our extension to the model for finding the optimal caching period would also be worth considering in the light of a more general process. For example, if data is used in bursts of activity, followed by periods of inactivity, then we would expect that the optimal caching period is correlated with the expected length of a burst. We might also consider expanding the model to multiple endpoints and multiple items. In such a case, assigning a unique caching period to each endpoint/item pair might be unnecessarily complex, and would complicate any analysis of query data generated by the data network. In practice, caching periods are determined by assigning items to different classes; for example, the data shown in Fig. 4.7 consisted of queries for data files which had not yet been positively identified as being either safe or unsafe. Such data is likely to change as more information is collected, so the caching period is short. As confidence in identification increases, so can the caching period. Identifying malicious files and URLs is an evolving battle between cybersecurity providers and malware creators, but the manner in which that information is used remains the same. Hence, further analysis of how to optimise the propagation of this information will always be of value.

Finally, our circle map model generates interesting figures which describe a hierarchical pattern that can be understand through symbolic dynamics. This model is a significant abstraction from the original stochastic model, and so it may lack the same level of applicability as the other sections within the context of cybersecurity. However, models like this can help us identify signals that are otherwise lost in the noise of of the query data. Outside of this context, the circle map is so ubiquitous that every opportunity to understand it is worth considering. As noted in the final subsection, there is a clear path to further research here by considering orbits that do not pass through the point $\theta = 0$, and how those orbits connect with the ones studied in this section. We also note that the striking hierarchical pattern of triangles visible in Fig. 4.12 may feature some degree of self-similarity, and it would be worth investigating whether a transformation can be found which would map the pattern to some subset of itself.

There is clearly much scope for further work on this topic, falling into a number of areas of mathematics, and as many internet systems rely on efficient and reliable dissemination of data, we expect that this work will be interest to a wide audience.

Conclusion & Outlook

The Internet is a highly complex system that has a significant impact on our world. It is, therefore, crucial that we understand its dynamical behaviour. In this thesis, we formalise small-scale problems derived from Internet systems mathematically so that their dynamics can be studied rigorously. We observe that stochasticity is a crucial feature in modelling the negotiations following a targeted ransomware attack, while modelling the server/endpoint interactions of a data network requires dynamical features such as time delay, periodic forcing, switching and stochasticity. In addition, we study two dynamical systems from climate science and signal processing which feature time delay, switching, and either periodic forcing or linear flow. These systems give rise to complex resonance phenomena and rich bifurcation structures.

In Chapter 1, we studied a dynamical system derived from a climate science model which featured switched time-delayed feedback and switched periodic forcing. We observed a rich structure of torus bifurcations spanned by Arnold tongues divided into 'strings of sausages' by zero-width shrinking points. Solutions within the same Arnold tongue had the same period P, but the number of zero crossings R per period varied along the string. Due to the switching, we were able to study the dynamics of this non-smooth system analytically using a symbolic representation. This enabled us to map out key features of the bifurcation structure analytically, something which could only be achieved numerically in earlier smooth systems [27]. One area of particular interest which could merit further study is the relationship between the Chenciner (generalised Neimark-Sacker) bifurcations and the curve which bounds the region inside which we observe Arnold tongues of stable solutions with characteristic ratios $P : R$ where $R < P$. It may be possible to calculate points along this curve analytically by finding the zero-width shrinking points of Arnold tongues where the symbolic representation changes. This may yield further insight to the dynamics of the earlier smooth systems.

In Chapter 2, we studied a dynamical system which featured linear flow and switched time-delayed feedback derived from a model of a band-pass filter with self-feedback. An intriguing feature of this system is that the dynamics depended significantly on whether the system was underdamped or overdamped. In the underdamped regime, we observed rich dynamical behaviour which included both classical and discontinuity-induced bifurcations, and periodic orbits became unstable if their frequencies were too much higher or lower than the centre frequency of the band-pass filter. This is intuitive as the band-pass filter attenuates signals with frequencies outside of a band centred on the centre frequency. In the overdamped regime, we observed only classical bifurcations, and periodic orbits became unstable only if their frequencies were much *higher* than the centre frequency. We also noted that some of the studied unstable solutions appeared to vanish in highly non-generic manner. This could be remedied by introducing a hysteresis band in our time-delayed feedback, which might also be more justified from a physical perspective.

In Chapter 3, we developed a game-theory model of the negotiations between a cybercriminal and their target following a successful targeted ransomware attack. A key feature in this model was the addition of stochasticity to the cybercriminal's estimate of the target's willingness to pay the ransom. Without stochasticity, the outcome of the negotiation was deterministic and cybercriminal's optimal strategy was trivial. Including stochasticity caused the model to better reflect real-world observations. It would be interesting to expand on our model by considering an iterative game, where the cybercriminal sequentially attacks different targets and dynamically adjusts their strategy depending on how close they are to bankruptcy. Such an extension to our model may improve our understanding of the long-term strategic behaviour of cybercriminals.

Finally, in Chapter 4, we developed a model of interactions between a human-operated endpoint computer and a server operating a data network which employs Time-To-Live caching. We considered three different variations of this model which featured some combination of time delay, periodic forcing, switching and stochasticity. These variations (i) included the generation of records kept by the server of queries from the endpoint, (ii) the optimal choice of Time-To-Live for volatile data that changes stochastically, and (iii) the effect of human sleep cycles on query patterns. The deterministic limit of the latter produced a circle map which gave rise to interesting dynamics that we studied through a symbolic representation. There is much scope for further study in this area. We might expand on research conducted here by considering volatile data which changes dynamically in response to being used; this would be representative of a Trojan horse-style malware [210]. We might also consider large-scale data network optimisation methods such as precaching [138, 139, 140, 141, 211, 212], which preemptively loads data onto endpoints predictively to reduce query volume.

In this thesis, we have formalised two small-scale Internet systems; the interaction between the perpetrator of targeted ransomware and their target, and the server/endpoint interactions of data networks. Both systems are of great importance as the Internet continues to grow in scale and impact. Ransomware threatens every person who has data stored on the Internet, while data networks form the backbones of many Internet systems. These systems also contain dynamical features which can induce complex dynamics. We therefore expect this research to be of wide interest, and for the Internet to provide a rich supply of dynamical systems research well into the future.

Appendix A

Appendix

A.1 Border-collision bifurcations in a driven time-delay system

A.1.1 Determining legality of a sequence

Each event in a sequence has a time and a position associated with it. The times can be used to generate a system of simultaneous equations:

- Each D occurs at a time $t_D = n$, $n \in \mathbb{Z}$.
- Each $\bar{D}$ occurs at a time $t_{\bar{D}} = n + \frac{1}{2}$, $n \in \mathbb{Z}$.
- Each H occurs at a time $t_H = t_{\bar{Z}} + \tau$, where $t_{\bar{Z}}$ is a time at which a $\bar{Z}$ occurs.
- Each $\bar{H}$ occurs at a time $t_{\bar{H}} = t_Z + \tau$, where t_Z is a time at which a Z occurs.
- Each Z is connected to the previous $\bar{Z}$ by the equation showing that the displacement of the trajectory between the two events is 0.
- Each $\bar{Z}$ is connected to the previous Z by the equation showing that the displacement of the trajectory between the two events is 0.

As there is one equation for each symbol, we can solve the system of equation to calculate the times. We then solve for the positions, using the events to calculate the slope between the calculated times. Finally, we generate a set of inequalities for each event:

- The time at which each event occurs must be consistent with the order of sequence.
- Each event occurring between consecutive Z and $\bar{Z}$ events occurs at $x > 0$.
- Each event occurring between consecutive $\bar{Z}$ and Z events occurs at $x < 0$.

If the solved system of equations satisfies the set of inequalities then the sequence is legal, for a given (b, τ). These techniques allow us to use symbolic sequences to study solutions systematically. Solving for the times and positions associated with events in a sequence, we can plot the solution represented by the sequence. Additionally, by changing the inequalities to equalities, we obtain the bifurcation curves shown in Fig. 1.3.

A.1.2 Calculation of the bifurcation curve of $\mathbb{P}$

For $\tau > \frac{1}{2}$, $\mathbb{P}$ can written as

$$\mathbb{P}(S_z) = AS_z + B \tag{A.1}$$

Due to the sparsity of A, the characteristic equation can be readily calculated as

$$(-1-\lambda)(-\lambda)^{n-1} + (-1)^{n-1}\frac{2(-1)^{n-1}}{b+(-1)^{n-1}} = 0 \tag{A.2}$$

This simplifies to

$$\lambda^n + \lambda^{n-1} - \frac{2(-1)^{n-1}}{b+(-1)^{n-1}} = 0 \tag{A.3}$$

By substituting $\lambda = e^{i\rho}$, we can solve for b to determine $b = b_{\text{bif}}$ for which the fixed point of $\mathbb{P}$ is bifurcating as

$$e^{i\rho n} + e^{i\rho(n-1)} - \frac{2(-1)^{n-1}}{b_{\text{bif}}+(-1)^{n-1}} = 0 \tag{A.4}$$

Note that as $\frac{2(-1)^{n-1}}{b_{\text{bif}}+(-1)^{n-1}} \in \mathbb{R}$,

$$e^{i\rho n} = \overline{e^{i\rho(n-1)}} \tag{A.5}$$

So, $e^{i\rho(2n-1)} = 1 = e^{i2\pi k}, k \in \mathbb{Z}$, and so $\rho = \frac{2\pi k}{2n-1}$. In order to satisfy Eq. (A.5), $k = n-1$. Substituting ρ back into Eq. (A.4),

$$\cos\left(\frac{2\pi n(n-1)}{2n-1}\right) + \cos\left(\frac{2\pi\,(n-1)^2}{2n-1}\right) - \frac{2(-1)^{n-1}}{b_{\text{bif}}+(-1)^{n-1}} = 0 \tag{A.6}$$

which simplifies under a sum-to-product cosine identity to

$$\cos\left(\frac{\pi\,(n-1)}{2n-1}\right) - \frac{1}{b_{\text{bif}}+(-1)^{n-1}} = 0 \tag{A.7}$$

Therefore,

$$b_{\text{bif}}(n) = \frac{1}{\cos\left(\frac{\pi(n-1)}{2n-1}\right)} - (-1)^{n-1} \tag{A.8}$$

where $n = \lceil 2\tau \rceil$.

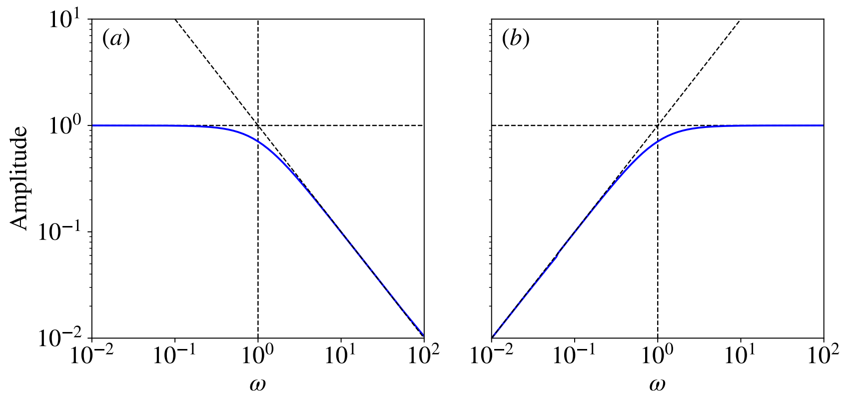

Figure A.1: Low-pass and high-pass filter output amplitude with input $f(t) = \cos(\omega t)$ for varying input frequency ω, where the corner frequency $\omega_L = \omega_H = 1$.

A.2 Dynamics of a band-pass filter system with switched time-delayed feedback

A.2.1 Derivation of the bandpass-filter system

A low-pass filter is a system or device that passes frequencies lower than the corner frequency ω_L and attenuates signals with frequency higher than ω_L [89, 213, 90]. A high-pass filter is a system or device that passes frequencies higher than the corner frequency ω_H and attenuates signals with frequency lower than ω_H. The dynamics of low-pass and high-pass filter systems are described in Eq. (A.9) and Eq.(A.9) respectively [89, 90, 213]

$$\tau_L \dot{x}_L + x_L = f(t) \tag{A.9}$$

$$\dot{x}_H + \frac{x_H}{\tau_H} = \frac{d}{dt}[g(t)] \tag{A.10}$$

where $\tau_L = \omega_L^{-1}$ and $\tau_H = \omega_H^{-1}$, and f and g are the input signals which drive the filters. When undriven, the filters relax to $x_L = x_H = 0$. When we drive Eq. (A.9) and Eq. (A.10) with input signal $f(t) = g(t) = \cos(\omega t)$, the outputs $x_L(t)$ and $x_H(t)$ are signals signal with frequency ω whose amplitude depends on ω relative to the corner frequency, as shown in Fig. A.1 (a) and (b) respectively. A band-pass filter can be constructed by coupling a low-pass filter with a high-pass filter [89, 213, 90]. This is achieved by letting $g(t) = x_L(t)$, yielding the equations

$$\tau_L \dot{x}_L + x_L = f(t) \tag{A.11}$$

$$\dot{x}_H + \frac{x_H}{\tau_H} = \dot{x}_L \tag{A.12}$$

The system can be simplified by first integrating Eq. (A.12) to yield

$$x_H + \frac{1}{\tau_H}\int\limits_0^t x_H\left(t'\right)dt' = x_L \tag{A.13}$$

with the integration constant being satisfied by the lower bound of the integral. Substituting Eq. (A.12) and Eq. (A.13) into Eq. (A.11) yields

$$\tau_L\left[\dot{x}_H + \frac{x_H}{\tau_H}\right] + \left[x_H + \frac{1}{\tau_H}\int\limits_0^t x_H\left(s\right)ds\right] = f\left(t\right)$$

which we rewrite as

$$x_H + \frac{\tau_H\tau_L}{\left(\tau_H+\tau_L\right)}\dot{x}_H + \frac{1}{\left(\tau_H+\tau_L\right)}\int\limits_0^t x_H\left(s\right)ds = \frac{\tau_H}{\left(\tau_H+\tau_L\right)}f\left(t\right)$$

where $\frac{\tau_H}{(\tau_H+\tau_L)}$ is the signal gain due to passing through the band-pass filter. We substitute centre frequency $\Omega = \frac{1}{\sqrt{\tau_H\tau_L}}$ and quality factor $Q = \frac{\sqrt{\tau_H\tau_L}}{(\tau_H+\tau_L)}$ to yield

$$x_H + \frac{Q}{\Omega}\frac{dx_H}{dt} + Q\Omega\int\limits_0^t x_H\left(s\right)ds = \frac{\tau_H}{\left(\tau_H+\tau_L\right)}f\left(t\right) \tag{A.14}$$

We rescale x_H to get dimensionless output $x = x_H\left(\frac{\tau_H}{(\tau_H+\tau_L)}\right)^{-1}$, yielding the equation

$$x + \frac{Q}{\Omega}\frac{dx}{dt} + Q\Omega\int\limits^t x\left(s\right)ds = f\left(t\right) \tag{A.15}$$

Finally, we introduce a variable $y = Q\Omega\int^t x\left(s\right)ds$ such that $\frac{dy}{dt} = Q\Omega x$, yielding the non-smooth delay differential equation

$$\begin{aligned} Q\Omega^{-1}\dot{x} &= -x - y + f\left(t\right) \\ \dot{y} &= Q\Omega x \end{aligned} \tag{A.16}$$

previously studied in [89, 90, 213] with different input signals $f(t)$.

A.2.2 Useful properties of Pauli matrices

The Pauli matrices, with the identify matrix I, form a basis for the vector space of 2×2 real matrices [214]. They are

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{A.17}$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{A.18}$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{A.19}$$

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.20}$$

Pauli matrices have some very useful properties. Firstly, the square of any Pauli matrix σ_j yields the identity matrix

$$(\sigma_i)^2 = I \tag{A.21}$$

Secondly, Pauli matrices anti-commute; for a pair of different Pauli matrices σ_j and σ_k, $j \neq k$,

$$\sigma_j \sigma_k = -\sigma_k \sigma_j \tag{A.22}$$

These two properties can be summarised in Einstein notation [215] as

$$\sigma_j \sigma_k = \delta_{jk} I + i\epsilon_{jkl}\sigma_l \tag{A.23}$$

As the Pauli matrices and the identify matrix form a basis for the vector space of 2×2 real matrices, a 2×2 real matrix A can be written as

$$A = a_0 I + \bar{a} \cdot \bar{\sigma} \tag{A.24}$$

$$= a_0 I + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z \tag{A.25}$$

where

$$\bar{a} = (a_x, a_y, a_z) \tag{A.26}$$

$$\bar{\sigma} = (\sigma_x, \sigma_y, \sigma_z) \tag{A.27}$$

A further useful property can be obtained from the Einstein notation

$$(\bar{a}\cdot\bar{\sigma})^2 = a_j\sigma_j a_k\sigma_k \tag{A.28}$$
$$= a_j a_k \sigma_j\sigma_k \tag{A.29}$$
$$= a_j a_k \left(\delta_{jk} I + i\epsilon_{jkl}\sigma_l\right) \tag{A.30}$$
$$= a_j a_k \delta_{jk} I + i a_j a_k \epsilon_{jkl}\sigma_l \tag{A.31}$$
$$= a_j a_j I + i a_j a_k \epsilon_{jkl}\sigma_l \tag{A.32}$$

The second term term collapses to 0 as

$$a_j a_k \epsilon_{jkl}\sigma_l = \epsilon_{jkl} a_j a_k \sigma_l \tag{A.33}$$
$$= \epsilon_{kjl} a_k a_j \sigma_l \text{ by swapping order of summation} \tag{A.34}$$
$$= -\epsilon_{jkl} a_k a_j \sigma_l \text{ by permutation of } \epsilon_{jkl} \tag{A.35}$$
$$= -\epsilon_{jkl} a_j a_k \sigma_l \text{ by commutativity of scalar multiplication} \tag{A.36}$$
$$\rightarrow \epsilon_{jkl} a_j a_k \sigma_l = -\epsilon_{jkl} a_j a_k \sigma_l = 0 \tag{A.37}$$

Thus,

$$(\bar{a}\cdot\bar{\sigma})^2 = a_j a_j I \tag{A.38}$$

A.2.3 Calculating the matrix exponential e^{At}

We can decompose A from Eq. (2.9) in terms of Pauli Matrices as

$$A = -\frac{\Omega}{2Q} I + \frac{\Omega}{2Q}\left(Q^2-1\right)\sigma_x - \frac{i\Omega}{2Q}\left(Q^2+1\right)\sigma_y - \frac{\Omega}{2Q}\sigma_z \tag{A.39}$$

or

$$A = a_0 I + \bar{a}\cdot\bar{\sigma} \tag{A.40}$$

where

$$\begin{aligned} \bar{a} &= \left(\frac{\Omega}{2Q}\left(Q^2-1\right), -\frac{i\Omega}{2Q}\left(Q^2+1\right), -\frac{\Omega}{2Q}\right) \\ a_0 &= -\frac{\Omega}{2Q} \end{aligned} \tag{A.41}$$

For our convenience later on, we rewrite $\bar{a}$ to get

$$A = a_0 I + \bar{a}\cdot\bar{\sigma} = a_0 I + c_a \hat{a}\cdot\bar{\sigma} \tag{A.42}$$

where we require $\hat{a}$ to be a unit vector such that

$$(\hat{a}\cdot\bar{\sigma})^2 = I \tag{A.43}$$

Using Eq. (A.38) and Eq. (A.42), we can calculate c_a as

$$
\begin{aligned}
(\bar{a}\cdot\bar{\sigma})^2 &= c_a^2\,(\hat{a}\cdot\bar{\sigma})^2 && \text{(A.44)}\\
a_j a_j I &= c_a^2 I && \text{(A.45)}\\
c_a &= \sqrt{a_j a_j} && \text{(A.46)}\\
&= \sqrt{\left[\frac{\Omega}{2Q}\left(Q^2-1\right)\right]^2 + \left[-\frac{i\Omega}{2Q}\left(Q^2+1\right)\right]^2 + \left(-\frac{\Omega}{2Q}\right)^2} && \text{(A.47)}\\
&= \frac{\Omega}{2Q}\sqrt{1-4Q^2} && \text{(A.48)}
\end{aligned}
$$

Then

$$
\begin{aligned}
\hat{a} &= \frac{1}{c_a}\bar{a} && \text{(A.49)}\\
&= \frac{1}{\frac{\Omega}{2Q}\sqrt{1-4Q^2}}\left(\frac{\Omega}{2Q}\left(Q^2-1\right), -\frac{i\Omega}{2Q}\left(Q^2+1\right), -\frac{\Omega}{2Q}\right) && \text{(A.50)}\\
&= \frac{1}{\sqrt{1-4Q^2}}\left(Q^2-1, -i\left(Q^2+1\right), -1\right) && \text{(A.51)}
\end{aligned}
$$

so that

$$
\begin{aligned}
\hat{a}\cdot\bar{\sigma} &= \frac{1}{\sqrt{1-4Q^2}}\left(Q^2-1, -i\left(Q^2+1\right), -1\right)\cdot\left(\sigma_x,\sigma_y,\sigma_z\right) && \text{(A.52)}\\
&= \frac{1}{\sqrt{1-4Q^2}}\left[\left(Q^2-1\right)\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix} - i\left(Q^2+1\right)\begin{pmatrix} 0 & -i\\ i & 0\end{pmatrix} - 1\begin{pmatrix} 1 & 0\\ 0 & -1\end{pmatrix}\right] && \text{(A.53)}\\
&= \frac{1}{\sqrt{1-4Q^2}}\begin{pmatrix} -1 & -2\\ 2Q^2 & 1\end{pmatrix} && \text{(A.54)}
\end{aligned}
$$

Using Eq. (A.42) and Eq. (A.54), we can now calculate the matrix exponential

$$e^{At} = \exp\left(ta_0 I + t\bar{a}\cdot\bar{\sigma}\right) \tag{A.55}$$
$$= \exp\left(ta_0\right)\exp\left(t\bar{a}\cdot\bar{\sigma}\right) \tag{A.56}$$
$$= \exp\left(ta_0\right)\exp\left(tc_a\hat{a}\cdot\bar{\sigma}\right) \tag{A.57}$$
$$= \exp\left(ta_0\right)\sum_{k=0}^{\infty}\frac{1}{k!}\left(tc_a\hat{a}\cdot\bar{\sigma}\right)^k \tag{A.58}$$
$$= \exp\left(ta_0\right)\left\{\sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(tc_a\hat{a}\cdot\bar{\sigma}\right)^{2n} + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(tc_a\hat{a}\cdot\bar{\sigma}\right)^{2n+1}\right\} \tag{A.59}$$
$$= \exp\left(ta_0\right)\left\{\sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(tc_a\right)^{2n}\left(\hat{a}\cdot\bar{\sigma}\right)^{2n} + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(tc_a\right)^{2n+1}\left(\hat{a}\cdot\bar{\sigma}\right)^{2n}\left(\hat{a}\cdot\bar{\sigma}\right)\right\} \tag{A.60}$$
$$= \exp\left(ta_0\right)\left\{\sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(tc_a\right)^{2n} I^n + \sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(tc_a\right)^{2n+1} I^n\left(\hat{a}\cdot\bar{\sigma}\right)\right\} \tag{A.61}$$
$$= \exp\left(ta_0\right)\left\{I\sum_{n=0}^{\infty}\frac{1}{(2n)!}\left(tc_a\right)^{2n} + \left(\hat{a}\cdot\bar{\sigma}\right)\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(tc_a\right)^{2n+1}\left(\hat{a}\cdot\bar{\sigma}\right)\right\} \tag{A.62}$$
$$= \exp\left(ta_0\right)\left\{I\cosh\left(tc_a\right) + \left(\hat{a}\cdot\bar{\sigma}\right)\sinh\left(tc_a\right)\right\} \tag{A.63}$$

We can substitute from A.54 and A.48 to get the matrix exponential in the overdamped regime

$$\begin{aligned} e^{At} &= e^{-\frac{\Omega}{2Q}t}\cosh\left(t\frac{\Omega}{2Q}\sqrt{1-4Q^2}\right)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \\ &+ e^{-\frac{\Omega}{2Q}t}\frac{1}{\sqrt{1-4Q^2}}\sinh\left(t\frac{\Omega}{2Q}\sqrt{1-4Q^2}\right)\begin{pmatrix}-1 & -2\\ 2Q^2 & 1\end{pmatrix}\end{aligned} \tag{A.64}$$

In the underdamped regime, the sinh and cosh terms switch over to sin and cos terms, as

$$\sinh x = -i\sin\left(ix\right) \tag{A.65}$$
$$\cosh x = \cos\left(ix\right) \tag{A.66}$$

yielding

$$\begin{aligned} e^{At} &= e^{-\frac{\Omega}{2Q}t}\cos\left(t\frac{\Omega}{2Q}\sqrt{4Q^2-1}\right)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \\ &+ e^{-\frac{\Omega}{2Q}t}\frac{1}{\sqrt{4Q^2-1}}\sin\left(t\frac{\Omega}{2Q}\sqrt{4Q^2-1}\right)\begin{pmatrix}-1 & -2\\ 2Q^2 & 1\end{pmatrix}\end{aligned} \tag{A.67}$$

A note on convergence At first glance, the second term in Eq. (A.64) and Eq. (A.67) respectively may have a convergence issue as $Q \to \frac{1}{2}$ and $\pm\left(1-4Q^2\right) \to 0$. We can doublecheck this by calculating the limit of $\frac{\sinh(rx)}{x}$ as $x \to \infty$ where $r = t\frac{\Omega}{2Q}$.

$$\lim_{x\to 0}\frac{\sinh\left(rx\right)}{x} = \lim_{x\to 0}\frac{1}{x}\sum_{n=0}^{\infty}\frac{1}{(2n+1)!}\left(rx\right)^{2n+1} \tag{A.68}$$

$$= \lim_{x\to 0}\sum_{n=0}^{\infty}\frac{r^{2n+1}x^{2n}}{(2n+1)!} \tag{A.69}$$

$$= \lim_{x\to 0}\left[r + \frac{r^3x^2}{6} + ...\right] \tag{A.70}$$

$$= r \tag{A.71}$$

and likewise for $\frac{\sin(rx)}{x}$. Thus, there is no issue with convergence, and the matrix exponential along the line of critical damping $Q = \frac{1}{2}$ is

$$e^{At} = e^{-\Omega t}\left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t\Omega\begin{pmatrix} -1 & -2 \\ \frac{1}{2} & 1 \end{pmatrix}\right\} \tag{A.72}$$

A.3 Stochastic models of Time-To-Live caching systems

A.3.1 Derivation of Poisson distribution from the Poisson process

The Poisson process is defined by the transition probabilities

$$\frac{d}{dt}P_0\left(t\right) = -\lambda_U P_0\left(t\right) \tag{A.73}$$

$$\frac{d}{dt}P_{k+1}\left(t\right) = \lambda_U P_k\left(t\right) - \lambda_U P_{k+1}\left(t\right) \;\; k > 0 \tag{A.74}$$

$$P_k\left(0\right) = \begin{cases} 1 & k = 0 \\ 0 & k > 0 \end{cases} \tag{A.75}$$

Eq. (A.73) has solution $P_0\left(t\right) = e^{-\lambda_U t}$; the probability $P\left[N_U\left(t\right) = 0\right]$ decays exponentially with constant rate λ_U as time evolves. More generally, if we assume

$$P_k\left(t\right) = f\left(t,k\right)e^{-\lambda_U t}$$

for some function f and substitute into Eq. (A.74), we get

$$\frac{d}{dt}\left[f(t,k+1)\,e^{-\lambda_U t}\right] = \lambda_U\left[f(t,k)\,e^{-\lambda_U t}\right] - \lambda_U\left[f(t,k+1)\,e^{-\lambda_U t}\right] \tag{A.76}$$

$$e^{-\lambda_U t}\left[f'(t,k+1) - \lambda_U f(t,k+1)\right] = e^{-\lambda_U t}\left[\lambda_U f(t,k) - \lambda_U f(t,k+1)\right] \tag{A.77}$$

$$f'(t,k+1) = \lambda_U f(t,k) \tag{A.78}$$

Hence, $f(t,k,\lambda_U) = C(k)(\lambda_U t)^k$ is polynomial in $\lambda_U t$ so that

$$\frac{d}{dt}\left[C(k+1)(\lambda_U t)^{k+1}\right] = \lambda_U C(k)(\lambda_U t)^k \tag{A.79}$$

$$(k+1)\,\lambda_U C(k+1)(\lambda_U t)^k = \lambda_U C(k)(\lambda_U t)^k \tag{A.80}$$

$$C(k+1) = \frac{C(k)}{(k+1)} \tag{A.81}$$

$$\Rightarrow C(k) = \frac{1}{k!} \tag{A.82}$$

yielding

$$P_k(t) = \frac{(\lambda_U t)^k\, e^{-\lambda_U t}}{k!} \tag{A.83}$$

Eq. (A.83) is the *probability mass function* for the Poisson distribution. For constant $t = T$, if

$$P\left[N_U(T) = k\right] = \frac{(\lambda_U T)^k\, e^{-\lambda_U T}}{k!} \tag{A.84}$$

then $N_U(T)$ follows a Poisson distribution [153] with rate $\lambda_U T$; that is,

$$N_U(T) \sim \text{Poi}(\lambda_U T) \tag{A.85}$$

A.3.2 Expected value of Poisson distribution

If

$$X \sim \text{Poi}(\lambda) \tag{A.86}$$

then the expected value of X is given by

$$\langle X\rangle = \sum_{k=0}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \tag{A.87}$$

$$= \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} \tag{A.88}$$

$$= \lambda e^{-\lambda} e^{\lambda} \tag{A.89}$$

$$= \lambda \tag{A.90}$$

A.3.3 Inter-event times of a Poisson process

The inter-event times of a Poisson process with rate λ are independent $\text{Exp}(\lambda)$ random variables.

Proof Consider the case where a Poisson process is in the state $N(T) = m$ at some time T. Let $t_0 > 0$ be the inter-event time, or the time until the next event occurs; that is, $N(T + t_0) = m + 1$ and $N(T + t) = m$ for $t < t_0$. Then

$$\frac{d}{dt}P_{k+1}(t) = \lambda P_k(t) - \lambda P_{k+1}(t) \tag{A.91}$$

as before, but now we have condition

$$P_k(T) = \begin{cases} 1 & k = m \\ 0 & k \neq m \end{cases} \tag{A.92}$$

Hence

$$P_m(T + t) = e^{-\lambda t} \tag{A.93}$$

If $N(T + t) = m$, then $t < t_0$, so that the probability that $N(T + t) = m$ is equal to the probability that $t < t_0$

$$\begin{aligned} P[t < t_0] &= P[N(T + t) = m] && \text{(A.94)} \\ &= P_m(T + t) && \text{(A.95)} \\ &= e^{-\lambda t} && \text{(A.96)} \end{aligned}$$

Hence,

$$\begin{aligned} P[t_0 < t] &= 1 - e^{-\lambda t} && \text{(A.97)} \\ &= \int_0^t \lambda e^{-\lambda t} dt && \text{(A.98)} \end{aligned}$$

If a random variable t_0 obeys Eq. (A.97), we say that t_0 follows an Exponential(λ) distribution

$$t_0 \sim \text{Exp}(\lambda) \tag{A.99}$$

with probability density function

$$f(t, \lambda) = \lambda e^{-\lambda t} \tag{A.100}$$

A.3.4 Memoryless property of the Exponential distribution

The Exponential process is described as memoryless because its increments are independent; if a random variable X follows an Exponential distribution $X \sim \text{Exp}(\lambda)$, then

$$P[X > x + y | X > y] = P[X > x] \tag{A.101}$$

for $x, y \geq 0$. This can be seen by considering

$$P[X > x + y | X > y] = \frac{P[X > x + y, X > y]}{P[X > y]} \tag{A.102}$$
$$= \frac{P[X > x + y]}{P[X > y]} \tag{A.103}$$
$$= \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} \tag{A.104}$$
$$= e^{-\lambda x} \tag{A.105}$$
$$= P[X > x] \tag{A.106}$$

A.3.5 Expected value of shifted Exponential distribution

If a random variable X follows a shifted Exponential distribution $X \sim \text{Exp}(\lambda, \tau)$, then $X \in (\tau, \infty)$ and

$$\langle X \rangle = \int\limits_{s=\tau}^{s=\infty} s\lambda e^{-\lambda(s-\tau)} ds \tag{A.107}$$

Substitute $r = s - \tau$. Then

$$\langle X \rangle = \int\limits_{s=\tau}^{s=\infty} s\lambda e^{-\lambda(s-\tau)} ds \tag{A.108}$$
$$\int\limits_{r=0}^{r=\infty} (r + \tau) \lambda e^{-\lambda r} dr \tag{A.109}$$
$$= \int\limits_{0}^{\infty} r\lambda e^{-\lambda r} dr + \int\limits_{0}^{\infty} \tau\lambda e^{-\lambda r} dr \tag{A.110}$$

Let $u = r$ and $dv = \lambda e^{-\lambda r} dr$, then by parts

$$\langle X \rangle = \left[-re^{-\lambda r}\right]_0^\infty - \int_0^\infty \left(-e^{-\lambda r}\right) dr + \int_0^\infty \tau\lambda e^{-\lambda r} dr \tag{A.111}$$

$$= [-0+0] - \left[\frac{1}{\lambda}e^{-\lambda r}\right]_0^\infty + \left[-\tau e^{-\lambda r}\right]_0^\infty \tag{A.112}$$

$$= -\left(0 - \frac{1}{\lambda}\right) + [0 - (-\tau)] \tag{A.113}$$

$$= \frac{1}{\lambda} + \tau \tag{A.114}$$

A.3.6 Sum of Poisson random variables

If $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$ are independent Poisson random variables, then their sum $X + Y \sim \text{Poi}(\lambda + \mu)$ [152].

Proof If $X \sim \text{Poi}(\lambda)$, then X has probability mass function

$$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

This can be proven by noting that if $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$ are independent Poisson random variables, then their joint probability mass function is

$$\begin{aligned} P[X = x, Y = y] &= P[X = x]\, P[Y = y] \\ &= \frac{\lambda^x e^{-\lambda}}{x!} \frac{\lambda^y e^{-\lambda}}{y!} \end{aligned}$$

Then the probability mass function of the sum $X + Y$ is

$$\begin{aligned} P[X + Y = k] &= \sum_{i=0}^{k} P[X = k - i, Y = i] \\ &= \sum_{i=0}^{k} \frac{\lambda^{k-i} e^{-\lambda}}{(k-i)!} \frac{\mu^i e^{-\mu}}{i!} \\ &= \frac{e^{-\lambda} e^{-\mu}}{k!} \sum_{i=0}^{k} \frac{k!}{(k-i)!i!} \lambda^{k-i} \mu^i \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} \lambda^{k-i} \mu^i \\ &= \frac{(\lambda + \mu)^k e^{-(\lambda+\mu)}}{k!} \end{aligned}$$

yielding the probability mass function for a $\text{Poi}(\lambda + \mu)$ random variable.

A.3.7 Sum of Exponential random variables

If $X_i \sim \text{Exp}(\lambda)$ for $i = 1, 2, ...n$ are independent random variables, then

$$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \lambda^{-1}\right) \tag{A.115}$$

Proof An Exponential(λ) random variable X has probability density function

$$f_X(x, \lambda) = \lambda e^{-\lambda x}$$

such that

$$P[X < c] = \int_0^c \lambda e^{-\lambda x} dx$$

The Exponential(λ) distribution is a special case of the Gamma(k, λ^{-1}) distribution, as the Gamma(k, λ^{-1}) has probability density function

$$f(x, k, \lambda) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \tag{A.116}$$

with rate parameter $\lambda > 0$ and scale parameter $k > 0$. If $k = 1$, then the Gamma distribution is identical to the Exponential distribution. If $X \sim \text{Gamma}(\alpha, \lambda^{-1})$ and $Y \sim \text{Gamma}(\beta, \lambda^{-1})$ are independent Gamma(k, λ^{-1}) random variables, then their joint probability density function is

$$f_{X,Y}(x, y, \alpha, \beta, \lambda) = f_X(x, \alpha, \lambda) f_Y(y, \beta, \lambda)$$

We can derive the probability density function for their sum $Z = X + Y$ by letting $y = z - x$ and integrating over x:

$$f_Z(z) = \int_0^z f_X(x, \alpha, \lambda) f_Y(z - x, \beta, \lambda) dx \tag{A.117}$$

$$= \int_0^z \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \frac{\lambda^\beta (z-x)^{\beta-1} e^{-\lambda(z-x)}}{\Gamma(\beta)} dx \tag{A.118}$$

$$= \lambda^\alpha \lambda^\beta \int_0^z \frac{x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} \frac{(z-x)^{\beta-1} e^{-\lambda z} e^{\lambda x}}{\Gamma(\beta)} dx \tag{A.119}$$

$$= \lambda^{\alpha+\beta} e^{-\lambda z} \int_{x=0}^{x=z} \frac{x^{\alpha-1} (z-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} dx \tag{A.120}$$

Substitute $x = zt$. Then $dx = zdt$ and

$$f_Z(z) = \lambda^{\alpha+\beta} e^{-\lambda z} \int_{t=0}^{t=1} \frac{(zt)^{\alpha-1}(z-zt)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} zdt \tag{A.121}$$

$$= \lambda^{\alpha+\beta} e^{-\lambda z} \int_0^1 \frac{z^{\alpha-1}t^{\beta-1}z^{\beta-1}(1-t)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} zdt \tag{A.122}$$

$$= \lambda^{\alpha+k_2} z^{\alpha+\beta-1} e^{-\lambda z} \int_0^1 \frac{t^{\alpha-1}(1-t)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} dt \tag{A.123}$$

$$= \frac{\lambda^{\alpha+\beta} z^{\alpha+\beta-1} e^{-\lambda z}}{\Gamma(\alpha+\beta)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \tag{A.124}$$

The integral in equation A.124 defines the Beta function [216]

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \tag{A.125}$$

By recursive integration by parts [217], we can show that

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{A.126}$$

Therefore,

$$f_Z(z) = \frac{\lambda^{\alpha+\beta} z^{\alpha+\beta-1} e^{-\lambda z}}{\Gamma(\alpha+\beta)} \tag{A.127}$$

which is the probability density function for a Gamma$(\alpha+\beta, \lambda^{-1})$ random variable. Hence, if $X_i \sim \text{Gamma}(k_i, \lambda^{-1})$ for $i = 1, 2, ...n$, then

$$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(\sum_{i=1}^{n} k_i, \lambda^{-1}\right) \tag{A.128}$$

If $k_i = 1$ for each X_i, then $X_i \sim \text{Exp}(\lambda)$, and

$$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \lambda^{-1}\right) \tag{A.129}$$

A.3.8 Variance of query process $Q(t)$

While the usage process $U(t)$ is well understood, being a straightforward Poisson process, we cannot say the same for $Q(t)$. The inhibitory effect of caching which prevents multiple queries in rapid succession has a significant effect on the process. The process is a temporal equivalent to the Matern type-3 process, which is an inhibitory *spacial* point process in $\mathbb{R}^n$, which was originally used to model the inhibitory effect

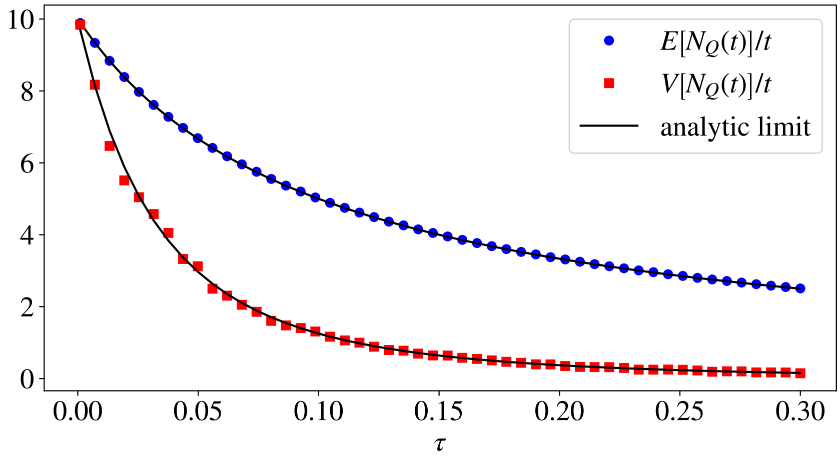

Figure A.2: The mean and variance of $N_Q(t)/t$ appear to be approximately equal to analytic closed form expression for $\lambda = 10$, $\tau = 0.1$ and $t = 100$.

that mature trees had on young trees growing nearby [218, 171]. These processes are known to quite intractable analytically, which makes calculating moments of $N_Q(t)$ quite difficult. In Subsection 4.2.2, we derived the first moment (expected value) of $N_Q(t)$

$$\frac{\langle N_Q(t)\rangle}{t} = \frac{1}{\frac{1}{\lambda} + \tau} \tag{A.130}$$

by considering the ratio of the total time over the interquery time; this shortcut doesn't work for finding the second moment, which we need to obtain the variance. The probability mass function is somewhat intractable; calculating probabilities analytically is possible by recursive integration by parts, but calculating moments analytically is problematic, and will be left for further research. However, if we calculate the variance of the process by simulation, we observe that

$$\frac{V[N_Q(t)]}{t} \simeq \frac{1}{\lambda^2\left(\frac{1}{\lambda} + \tau\right)^3} \tag{A.131}$$

as shown in Fig. A.2. This observation is by no means a proof of equality. However, the fact that both the mean and variance of $N_Q(t)$ appear to approach analytic closed form expression suggests that further analysis of this system would be worthwhile.

Bibliography

[1] J. M. Gillies, J. Gillies, R. Cailliau et al. *How the Web was born: The story of the World Wide Web*. Oxford University Press, USA, 2000.

[2] D. Shivalingaiah and U. Naik. Comparative study of Web 1.0, Web 2.0 and Web 3.0, 2008.

[3] C. Fuchs, W. Hofkirchner, M. Schafranek, C. Raffl, M. Sandoval and R. Bichler. Theoretical foundations of the web: cognition, communication, and co-operation. towards an understanding of Web 1.0, 2.0, 3.0. *Future internet*, 2(1):41–59, 2010.

[4] C. Beek and A. Furtak. Targeted ransomware no longer a future threat. February 2016. URL: `www.mcafee.com/enterprise/en-us/assets/reports/rp-targeted-ransomware.pdf` (visited on 07/10/2020).

[5] Coveware. Ransomware demands continue to rise as data exfiltration becomes common, and maze subdues. November 2020. URL: `www.coveware.com/blog/q3-2020-ransomware-marketplace-report` (visited on 25/11/2020).

[6] E. Kalaimannan, S. K. John, T. DuBose and A. Pinto. Influences on ransomware's evolution and predictions for the future challenges. *Journal of Cyber Security Technology*, 1:23–31, 2017.

[7] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442, 1998.

[8] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

[9] L. Larger, B. Penkovsky and Y. Maistrenko. Virtual chimera states for delayed-feedback systems. *Phys. Rev. Lett.*, 111(5):054103, 2013.

[10] L. P. Shayer and S. A. Campbell. Stability, bifurcation, and multistability in a system of two coupled neurons with multiple time delays. *SIAM Journal on Applied Mathematics*, 61(2):673–700, 2000.

[11] J. Foss, A. Longtin, B. Mensour and J. Milton. Multistability and delayed recurrent loops. *Physical Review Letters*, 76(4):708, 1996.

[12] S. Kim, S. H. Park and C. Ryu. Multistability in coupled oscillator systems with time delay. *Physical review letters*, 79(15):2911, 1997.

[13] A. Keane, B. Krauskopf and C. Postlethwaite. Delayed feedback versus seasonal forcing: resonance phenomena in an El Niño Southern Oscillation model. *SIAM J. Appl. Dyn. Syst.*, 14(3):1229, 2015.

[14] D. A. Barton, B. Krauskopf and R. E. Wilson. Explicit periodic solutions in a model of a relay controller with delay and forcing. *Nonlinearity*, 18(6):2637, 2005.

[15] S. Banerjee and C. Grebogi. Border collision bifurcations in two-dimensional piecewise smooth maps. *Phys. Rev. E*, 59(4):4052, 1999.

[16] M. di Bernardo, C. Budd, A. Champneys, P. Kowalczyk, A. Nordmark, G. Tost and P. Piiroinen. Bifurcations in nonsmooth dynamical systems. *SIAM Rev.*, 50(4):629, 2008.

[17] A. Colombo, M. di Bernardo, S. J. Hogan and M. R. Jeffrey. Bifurcations of piecewise smooth flows: Perspectives, methodologies and open problems. *Physica D*, 241(22):1845, 2012.

[18] D. A. W. Barton, B. Krauskopf and R. E. Wilson. Periodic solutions and their bifurcations in a non-smooth second-order delay differential equation. *Dynamical Systems: An International Journal*, 21(3):289–311, 2006.

[19] S. Boccaletti, A. N. Pisarchik, C. I. Del Genio and A. Amann. *Synchronization: from coupled systems to complex networks*. Cambridge University Press, 2018.

[20] A. Marchionne, P. Ditlevsen and S. Wieczorek. Synchronisation vs. resonance: Isolated resonances in damped nonlinear oscillators. *Physica D*, 380:8, 2018.

[21] S. Wieczorek, B. Krauskopf, T. B. Simpson and D. Lenstra. The dynamical complexity of optically injected semiconductor lasers. *Phys. Rep.*, 416:1, 2005.

[22] E. Tziperman, L. Stone, M. A. Cane and H. Jarosh. El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator. *Science*, 264:72, 1994.

[23] R. Selten. A simple game model of kidnapping. In *Models of strategic rationality*, pages 77–93. Springer, 1988.

[24] A. S. Pikovsky and J. Kurths. Coherence resonance in a noise-driven excitable system. *Physical Review Letters*, 78(5):775, 1997.

[25] O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241, 2001.

[26] P. Ryan, A. Keane and A. Amann. Border-collision bifurcations in a driven time-delay system. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(2):023121, 2020.

[27] M. Ghil, I. Zaliapin and S. Thompson. A delay differential model of ENSO variability: Parametric instability and the distribution of extremes. *Nonlin. Processes Geophys.*, 15, 2008.

[28] I. Belykh, R. Kuske, M. Porfiri and D. J. Simpson. Beyond the Bristol book: Advances and perspectives in non-smooth dynamics and applications. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(1):010402, 2023.

[29] A. Agliari, N. Pecora and A. Szuz. Appearance of closed invariant curves in a piecewise Cournot model with advertising. *Chaos, Solitons & Fractals*, 158:112013, 2022.

[30] P. Ryan, J. Fokker, S. Healy and A. Amann. Dynamics of targeted ransomware negotiation. *IEEE Access*, 10:32836–32844, 2022.

[31] T. V. Guy, J. Homolová and A. Gaj. Indirect dynamic negotiation in the Nash demand game. *IEEE Access*, 10:105008–105021, 2022.

[32] A. Parvaneh and F. Nasseri. Identify and neutralize negotiation tactics and tricks. *International Journal of Innovation in Marketing Elements*, 1(1):49–60, 2021.

[33] H. R. Skeoch. Modelling ransomware attacks using POMDPs. In *Workshop on the Economics of Information Security*. WEIS, 2022.

[34] P. Bajpai and R. Enbody. Know thy ransomware response: a detailed framework for devising effective ransomware response strategies. *Digital Threats: Research and Practice*, 2023.

[35] F. Teichmann, S. R. Boticiu and B. S. Sergi. The evolution of ransomware attacks in light of recent cyber threats. how can geopolitical conflicts influence the cyber climate? *International Cybersecurity Law Review*:1–22, 2023.

[36] S. P. Beeby, R. Torah, M. Tudor, P. Glynne-Jones, T. O'Donnell, C. Saha and S. Roy. A micro electromagnetic generator for vibration energy harvesting. *J. Micromech. Microeng.*, 17:1257, 2007.

[37] S. Daneshgar, O. De Feo and M. P. Kennedy. Observations concerning the locking range in a complementary differential LC injection-locked frequency divider–part I: qualitative analysis. *IEEE Trans. Circuits Syst. I*, 57:179, 2010.

[38] U. Parlitz and W. Lauterborn. Superstructure in the bifurcation set of the Duffing equation. *Phys. Lett. A*, 107(8):351, 1985.

[39] T. Heil, I. Fischer, W. Elsässer, J. Mulet and C. R. Mirasso. Chaos synchronization and spontaneous symmetry-breaking in symmetrically delay-coupled semiconductor lasers. *Phys. Rev. Lett.*, 86:795, 2001.

[40] S. Terrien, V. A. Pammi, B. Krauskopf, N. G. Broderick and S. Barbay. Pulse-timing symmetry breaking in an excitable optical system with delay. *Physical Review E*, 103(1):012210, 2021.

[41] E. Schöll, G. Hiller, P. Hövel and M. A. Dahlem. Time-delayed feedback in neurosystems. *Phil. Trans. Royal Soc. A*, 367:1079, 2009.

[42] J. Runge, V. Petoukhov and J. Kurths. Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models. *J. Climate*, 27:720, 2014.

[43] K. Pyragas. Control of chaos via extended delay feedback. *Phys. Lett. A*, 206:323, 1995.

[44] J. K. Hale and S. M. V. Lunel. *Introduction to functional differential equations.* Springer Science & Business Media, 1993.

[45] H. Hu, E. H. Dowell and L. N. Virgin. Resonances of a harmonically forced Duffing oscillator with time delay state feedback. *Nonlin. Dynamics*, 15:311, 1998.

[46] A. Maccari. Vibration control for the primary resonance of the van der Pol oscillator by a time delay state feedback. *Int. J. Non-Linear Mechanics*, 38(1):123, 2003.

[47] A. Keane, B. Krauskopf and C. Postlethwaite. Investigating irregular behavior in a model for the El Niño Southern Oscillation with positive and negative delayed feedback. *SIAM J. Appl. Dyn. Syst.*, 15(3):1656, 2016.

[48] R. D. Nussbaum. Uniqueness and nonuniqueness for periodic solutions of $\dot{x}(t) = -g(x(t-1))$. *J. Diff. Equ.*, 34(1):25, 1979.

[49] S.-N. Chow and H.-O. Walther. Characteristic multipliers and stability of symmetric periodic solutions of $\dot{x}(t) = -g(x(t-1)))$. *Trans. American Math. Soc.*, 307(1):127, 1988.

[50] Y. Wei-Ming and H. Bai-Lin. How the Arnold tongues become sausages in a piecewise linear circle map. *Comm. Theo. Phys.*, 8(1):1, 1987.

[51] D. J. W. Simpson and J. D. Meiss. Shrinking point bifurcations of resonance tongues for piecewise-smooth, continuous maps. *Nonlinearity*, 22(5):1123, 2009.

[52] D. J. W. Simpson. The structure of mode-locking regions of piecewise-linear continuous maps: I. Nearby mode-locking regions and shrinking points. *Nonlinearity*, 30(1):382, 2016.

[53] D. J. W. Simpson. The structure of mode-locking regions of piecewise-linear continuous maps: II. Skew sawtooth maps. *Nonlinearity*, 31(5):1905, 2018.

[54] N. Metropolis, M. Stein and P. Stein. On finite limit sets for transformations on the unit interval. *J. Combinat. Theory, Series A*, 15(1):25, 1973.

[55] G. Stépán and T. Insperger. Stability of time-periodic and delayed systems—a route to act-and-wait control. *Annual Reviews in control*, 30(2):159–168, 2006.

[56] T. Insperger and G. Stepan. On the dimension reduction of systems with feedback delay by act-and-wait control. *IMA Journal of Mathematical Control and Information*, 27(4):457–473, 2010.

[57] X.-S. Wang and J. Wu. Seasonal migration dynamics: periodicity, transition delay and finite-dimensional reduction. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2139):634–650, 2012.

[58] K. Nandakumar and M. Wiercigroch. Galerkin projections for state-dependent delay differential equations with applications to drilling. *Applied Mathematical Modelling*, 37(4):1705–1722, 2013.

[59] A. N. W. Hone, M. V. Irle and G. W. Thurura. On the Neimark-Sacker bifurcation in a discrete predator-prey system. *J. Biol. Dyn.*, 4(6):594, 2010.

[60] M. di Bernardo, F. Garofalo, L. Iannelli and F. Vasca. Bifurcations in piecewise-smooth feedback systems. *Int. J. of Control*, 75(16-17):1243, 2002.

[61] A. Granados, M. Krupa and F. Clément. Border collision bifurcations of stroboscopic maps in periodically driven spiking models. *SIAM J. Appl. Dyn. Syst.*, 13(4):1387, 2014.

[62] D. Meiss. Neimark-Sacker bifurcations in planar, piecewise-smooth, continuous maps. *SIAM J. Appl. Dyn. Syst.*, 7(3):795, 2008.

[63] H. E. Nusse, E. Ott and J. A. Yorke. Border-collision bifurcations: An explanation for observed bifurcation phenomena. *Phys. Rev. E*, 49(2):1073, 1994.

[64] U. Holmberg. *Relay feedback of simple systems.* PhD thesis, Lund Institute of Technology, 1991.

[65] A. Colombo, M. di Bernardo, S. J. Hogan and P. Kowalczyk. Complex dynamics in a hysteretic relay feedback system with delay. *Journal of Nonlinear Science*, 17(2):85–108, 2007.

[66] A. Keane and B. Krauskopf. Chenciner bubbles and torus break-up in a periodically forced delay differential equation. *Nonlinearity*, 31(6):R165, 2018.

[67] D. J. W. Simpson and J. D. Meiss. Resonance near border-collision bifurcations in piecewise-smooth, continuous maps. *Nonlinearity*, 23(12):3091, 2010.

[68] T. Sorrentino, C. Quintero-Quiroz, A. Aragoneses, M. C. Torrent and C. Masoller. Effects of periodic forcing on the temporally correlated spikes of a semiconductor laser with feedback. *Opt. Expr.*, 23(5):5571, 2015.

[69] V. Semenov, A. Zakharova, Y. Maistrenko and E. Schöll. Delayed-feedback chimera states: Forced multiclusters and stochastic resonance. *EPL*, 115(1):10005, 2016.

[70] K. E. Callan, L. Illing, Z. Gao, D. J. Gauthier and E. Schöll. Broadband chaos generated by an optoelectronic oscillator. *Phys. Rev. Lett.*, 104(11):113901, 2010.

[71] Y. K. Chembo, D. Brunner, M. Jacquot and L. Larger. Optoelectronic oscillators with time-delayed feedback. *Rev. Mod. Phys.*, 91:035006, 3, 2019.

[72] M. Peil, M. Jacquot, Y. K. Chembo, L. Larger and T. Erneux. Routes to chaos and multiple time scale dynamics in broadband bandpass nonlinear delay electro-optic oscillators. *Phys. Rev. E*, 79(2):026208, 2009.

[73] L. Larger, B. Penkovsky and Y. Maistrenko. Virtual chimera states for delayed-feedback systems. *Phys. Rev. Lett.*, 111:054103, 2013.

[74] L. Illing and D. J. Gauthier. Ultra-high-frequency chaos in a time-delay electronic device with band-limited feedback. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16:033119, 2006.

[75] E. Schöll, G. Hiller, P. Hövel and M. A. Dahlem. Time-delayed feedback in neurosystems. *Philos. Trans. R. Soc. A*, 367:1079, 2009.

[76] J. Hale, L. T. Magalhães and W. L. Oliva. *Dynamics in Infinite Dimensions*, volume 47 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2nd edition, 2002.

[77] T. Erneux. *Applied Delay Differential Equations*. S. S. Antman, J. E. Marsden and L. Sirovich, editors, volume 3 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, 2009.

[78] L. Larger and J. M. Dudley. Nonlinear dynamics: Optoelectronic chaos. *Nature*, 465(7294):41–42, 2010.

[79] K. Pyragas. Continuous control of chaos by self-controlling feedback. *Phys. Rev. Lett.*, 170(6):421–428, 1992.

[80] A. Amann, E. Schöll and W. Just. Some basic remarks on eigenmode expansions of time-delay dynamics. *Physica A*, 373:191–202, 2007.

[81] E. Schöll and H. G. Schuster, editors. *Handbook on Chaos Control*. Wiley-VCH, Weinheim, 2nd edition, 2008.

[82] E. Schöll, S. H. L. Klapp and P. Hövel, editors. *Control of Self-Organizing Nonlinear Systems*. Springer Int. Pub. AG, Switzerland, 2016.

[83] E. W. Hooton and A. Amann. Analytical limitation for time-delayed feedback control in autonomous systems. *Phys. Rev. Lett.*, 109(15):154101, 2012.

[84] L. Illing, D. J. Gauthier and R. Roy. Controlling optical chaos, spatio-temporal dynamics, and patterns. In P. R. Berman, E. Arimondo and C. Lin, editors, *Advances in Atomic, Molecular, and Optical Physics*. Volume 54, pages 615–695. Elsevier, Amsterdam, 2007.

[85] G. D. VanWiggeren and R. Roy. Chaotic communication using time-delayed optical systems. *Int. J. Bifurcation Chaos*, 09(11):2129–2156, 1999.

[86] J. D. Hart, L. Larger, T. E. Murphy and R. Roy. Delayed dynamical systems: networks, chimeras and reservoir computing. *Philos. Trans. R. Soc. A*, 377(2153):20180123, 2019.

[87] J. Sieber, P. Kowalczyk, S. Hogan and M. Di Bernardo. Dynamics of symmetric dynamical systems with delayed switching. *Journal of Vibration and Control*, 16(7-8):1111–1140, 2010.

[88] J. Sieber. Dynamics of delayed relay systems. *Nonlinearity*, 19(11):2489–2527, September 2006.

[89] V. S. Udaltsov, L. Larger, J.-P. Goedgebuer, M. W. Lee, E. Genin and W. T. Rhodes. Bandpass chaotic dynamics of electronic oscillator operating with delayed nonlinear feedback. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 49(7):1006–1009, 2002.

[90] L. Illing and D. J. Gauthier. Hopf bifurcations in time-delay systems with band-limited feedback. *Physica D: Nonlinear Phenomena*, 210(3-4):180–202, 2005.

[91] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM review*, 45(1):3–49, 2003.

[92] S. H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

[93] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 2013.

[94] A. J. Homburg and B. Sandstede. Homoclinic and heteroclinic bifurcations in vector fields. *Handbook of dynamical systems*, 3:379–524, 2010.

[95] A. Longtin, M. Mackey, J. Milton, R. Scholl et al. Oscillatory modes in a nonlinear second-order differential equation with delay. *Journal of Dynamics and Differential Equations*, 2(4):423–449, 1990.

[96] W. Bayer and U. an der Heiden. Oscillation types and bifurcations of a nonlinear second-order differential-difference equation. *Journal of Dynamics and Differential Equations*, 10(2):303–326, 1998.

[97] A. M. Maigida, S. M. Abdulhamid, M. Olalere et al. Systematic literature review and metadata analysis of ransomware attacks and detection mechanisms. *Journal of Reliable Intelligent Environments*, 5:67–89, 2019.

[98] P. Bajpai and R. Enbody. Dissecting. net ransomware: Key generation, encryption and operation. *Network Security*, 2020(2):8–14, 2020.

[99] A. Zimba and M. Chishimba. On the economic impact of crypto-ransomware attacks: The state of the art on enterprise systems. *European Journal for Security Research*, 4(1):3–31, 2019.

[100] R. B. Myerson. *Game theory*. Harvard University Press, 2013.

[101] J. S. Brown, J. W. Laundré and M. Gurung. The ecology of fear: Optimal foraging, game theory, and trophic interactions. *Journal of mammalogy*, 80:385–399, 1999.

[102] B. J. McGill and J. S. Brown. Evolutionary game theory and adaptive dynamics of continuous traits. *Annual Review of Ecology, Evolution, and Systematics*, 38:403–435, 2007.

[103] J. M. Smith and G. R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

[104] D. Friedman. On economic applications of evolutionary game theory. *Journal of evolutionary economics*, 8:15–43, 1998.

[105] R. Selten. Bounded rationality. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 146:649–658, 1990.

[106] E. Lukas, J. J. Reuer and A. Welling. Earnouts in mergers and acquisitions: A game-theoretic option pricing approach. *European Journal of Operational Research*, 223:256–263, 2012.

[107] D. A. Cerdeiro. Contagion exposure and protection technology. *Games and Economic Behavior*, 105:230–254, 2017.

[108] A. Kydd. Game theory and the spiral model. *World Politics*, 49:371–400, 1997.

[109] H. Ward. Game theory and the politics of the global commons. *Journal of Conflict Resolution*, 37:203–235, 1993.

[110] T. Caulfield and A. Fielder. Optimizing time allocation for network defence. *Journal of Cybersecurity*, 1:37–51, 2015.

[111] J. R. Lindsay. Tipping the scales: The attribution problem and the feasibility of deterrence against cyberattack. *Journal of Cybersecurity*, 1:53–67, 2015.

[112] A. Laszka, S. Farhang and J. Grossklags. On the economics of ransomware. In *International Conference on Decision and Game Theory for Security*, pages 397–417. Springer, 2017.

[113] E. Cartwright, J. Hernandez Castro and A. Cartwright. To pay or not: Game theoretic models of ransomware. *Journal of Cybersecurity*, 5:tyz009, 2019.

[114] H. Hu, Y. Liu, C. Chen, H. Zhang and Y. Liu. Optimal decision making approach for cyber security defense using evolutionary game. *IEEE Transactions on Network and Service Management*, 17(3):1683–1700, 2020.

[115] N. Caporusso, S. Chea and R. Abukhaled. A game-theoretical model of ransomware. In *International Conference on Applied Human Factors and Ergonomics*, pages 69–78. Springer, 2018.

[116] J. Hernandez-Castro, A. Cartwright and E. Cartwright. An economic analysis of ransomware and its welfare consequences. *Royal Society Open Science*, 7:190023, 2020.

[117] A. Cartwright and E. Cartwright. Ransomware and reputation. *Games*, 10:26, 2019.

[118] Z. Li and Q. Liao. Game theory of data-selling ransomware. *Journal of Cyber Security and Mobility*:65–96, 2021.

[119] Panda Security. Targeted ransomware attacks: The easy choice for cybercriminals. July 2019. URL: `www.pandasecurity.com/en/mediacenter/security/targeted-ransomware/` (visited on 04/02/2021).

[120] S. Stahie. Why is targeted ransomware so dangerous? March 2020. URL: `businessinsights.bitdefender.com/why-is-targeted-ransomware-so-dangerous` (visited on 14/01/2021).

[121] C. Gerritz. 2019 mid-market threat and incident response report. 2019. URL: `www.infocyte.com/resources/mid-market-threat-and-incident-response-report/` (visited on 14/01/2021).

[122] Coalition. Cyber insurance claims report. 2020. URL: `info.coalitioninc.com/download-2020-cyber-claims-report.html` (visited on 14/01/2021).

[123] A. Asokan. Experts warn: Targeted ransomware attacks to surge. February 2020. URL: `www.bankinfosecurity.com/experts-warn-targeted-ransomware-attacks-to-surge-a-13782` (visited on 14/01/2021).

[124] A. Young and M. Yung. Cryptovirology: Extortion-based security threats and countermeasures. In *Proceedings 1996 IEEE Symposium on Security and Privacy*, pages 129–140. IEEE, 1996.

[125] A. Ng. Malware now comes with customer service. July 2017. URL: `www.cnet.com/news/ransomware-goes-pro-customer-service-google-25-million-black-hat/` (visited on 14/01/2021).

[126] J. Rundle. Ransomware negotiators gain prominence as attacks increase. August 2020. URL: `www.wsj.com/articles/ransomware-negotiators-gain-prominence-as-attacks-increase-11598866201` (visited on 14/01/2021).

[127] E. L. Crow and K. Shimizu. *Lognormal distributions*. Marcel Dekker New York, 1987.

[128] S. Gualandi and G. Toscani. Human behavior and lognormal distribution. a kinetic description. *Mathematical Models and Methods in Applied Sciences*, 29:717–753, 2019.

[129] D. M. Kreps and R. Wilson. Reputation and imperfect information. *Journal of economic theory*, 27(2):253–279, 1982.

[130] A. Barrachina, Y. Tauman and A. Urbano. Entry and espionage with noisy signals. *Games and economic behavior*, 83:127–146, 2014.

[131] K. Durkota, V. Lisý, B. Bošanský and C. Kiekintveld. Approximate solutions for attack graph games with imperfect information. In *International Conference on Decision and Game Theory for Security*, pages 228–249. Springer, 2015.

[132] R. Piessens, E. de Doncker-Kapenga, C. W. Überhuber and D. K. Kahaner. *Quadpack: a subroutine package for automatic integration*, volume 1. Springer Science & Business Media, 2012.

[133] E. Jones, T. Oliphant, P. Peterson et al. SciPy: open source scientific tools for Python, 2001–. URL: `http://www.scipy.org/`.

[134] G. Casella and R. L. Berger. *Statistical inference*. Brooks/Cole Publishing Company, 1990.

[135] Y. Peng, L. Shangguan, Y. Hu, Y. Qian, X. Lin, X. Chen, D. Fang and K. Jamieson. PLoRa: A passive long-range data network from ambient lora transmissions. In *Proceedings of the 2018 conference of the ACM special interest group on data communication*, pages 147–160, 2018.

[136] C. U. Saraydar, N. B. Mandayam and D. J. Goodman. Pricing and power control in a multicell wireless data network. *IEEE Journal on selected areas in communications*, 19(10):1883–1892, 2001.

[137] C. Williamson, E. Halepovic, H. Sun and Y. Wu. Characterization of CDMA2000 cellular data network traffic. In *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05) l*, Z000–719. IEEE, 2005.

[138] P. Cao, E. W. Felten, A. R. Karlin and K. Li. A study of integrated prefetching and caching strategies. *ACM SIGMETRICS Performance Evaluation Review*, 23(1):188–197, 1995.

[139] W. Ali, S. M. Shamsuddin, A. S. Ismail et al. A survey of web caching and prefetching. *Int. J. Advance. Soft Comput. Appl*, 3(1):18–44, 2011.

[140] T. M. Kroeger, D. D. Long and J. C. Mogul. Exploring the bounds of web latency reduction from caching and prefetching. In *USENIX Symposium on Internet Technologies and Systems (USITS 97)*, 1997.

[141] T. Kimbrel and A. R. Karlin. Near-optimal parallel prefetching and caching. *SIAM Journal on computing*, 29(4):1051–1082, 2000.

[142] C. Fricker, P. Robert and J. Roberts. A versatile and accurate approximation for LRU cache performance. In *2012 24th international teletraffic congress (ITC 24)*, pages 1–8. IEEE, 2012.

[143] D. S. Berger, P. Gland, S. Singla and F. Ciucu. Exact analysis of TTL cache networks. *Performance Evaluation*, 79:2–23, 2014.

[144] R. H. Katz, S. J. Eggers, D. A. Wood, C. L. Perkins and R. G. Sheldon. Implementing a cache consistency protocol. *ACM SIGARCH Computer Architecture News*, 13(3):276–283, 1985.

[145] J. Gwertzman and M. I. Seltzer. World-wide web cache consistency. In *USENIX annual technical conference*, volume 141, page 152, 1996.

[146] M. J. Ward, M. H. Eckman, D. P. Schauer, A. S. Raja and S. Collins. Cost-effectiveness of telemetry for hospitalized patients with low-risk chest pain. *Academic Emergency Medicine*, 18(3):279–286, 2011.

[147] D. Nacapricha, N. Amornthammarong, K. Sereenonchai, P. Anujarawat and P. Wilairat. Low cost telemetry with PC sound card for chemical analysis applications. *Talanta*, 71(2):605–609, 2007.

[148] M. L. Don. A low-cost software-defined telemetry receiver. In International Foundation for Telemetering, 2015.

[149] E. Ott. *Chaos in dynamical systems*. Cambridge University Press, 2002.

[150] V. Arnold. Cardiac arrhythmias and circle mappingsa. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 1(1):20–24, 1991.

[151] L. Glass. Cardiac arrhythmias and circle maps- a classical problem. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 1(1):13–19, 1991.

[152] R. V. Hogg and A. T. Craig. Introduction to mathematical statistics.(5th edition). *Englewood Hills, New Jersey*, 1995.

[153] C. Gardiner. *Stochastic methods*, volume 4. Springer Berlin, 2009.

[154] P. J. Bickel and K. A. Doksum. *Mathematical statistics: Basic ideas and selected topics, vol. I Second Edition*. Chapman and Hall/CRC, 2001.

[155] G. J. Babu and E. D. Feigelson. Spatial point processes in astronomy. *Journal of statistical planning and inference*, 50(3):311–326, 1996.

[156] R. S. Stoica, E. Tempel, L. J. Liivamägi, G. Castellan and E. Saar. Spatial patterns analysis in cosmology based on marked point processes. *European Astronomical Society Publications Series*, 66:197–226, 2014.

[157] J. D. Scargle and G. J. Babu. Ch. 20. Point processes in astronomy: Exciting events in the universe. *Handbook of Statistics*, 21:795–825, 2003.

[158] R. Alfaro, C. Alvarez, J. Álvarez, R. Arceo, J. Arteaga-Velázquez, D. A. Rojas, H. A. Solares, A. Barber, N. Bautista-Elivar, A. Becerril et al. Search for very-high-energy emission from gamma-ray bursts using the first 18 months of data from the HAWC gamma-ray observatory. *The Astrophysical Journal*, 843(2):88, 2017.

[159] E. Tempel, R. S. Stoica, R. Kipper and E. Saar. Bisous model-detecting filamentary patterns in point processes. *Astronomy and Computing*, 16:17–25, 2016.

[160] S. Kočiščák, A. Kvammen, I. Mann, S. H. Sørbye, A. Theodorsen and A. Zaslavsky. Modeling solar orbiter dust detection rates in the inner heliosphere as a poisson process. *Astronomy & Astrophysics*, 670:A140, 2023.

[161] A. Gosztolai and M. Barahona. Cellular memory enhances bacterial chemotactic navigation in rugged environments. *Communications Physics*, 3(1):47, 2020.

[162] E. A. Codling, M. J. Plank and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.

[163] H. G. Othmer, S. R. Dunbar and W. Alt. Models of dispersal in biological systems. *Journal of mathematical biology*, 26(3):263–298, 1988.

[164] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse and M. Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE journal on selected areas in communications*, 27(7):1029–1046, 2009.

[165] R. W. Heath, M. Kountouris and T. Bai. Modeling heterogeneous network interference using poisson point processes. *IEEE Transactions on signal processing*, 61(16):4114–4126, 2013.

[166] Y. Hmamouche, M. Benjillali, S. Saoudi, H. Yanikomeroglu and M. Di Renzo. New trends in stochastic geometry for wireless networks: A tutorial and survey. *Proceedings of the IEEE*, 109(7):1200–1252, 2021.

[167] Y. Pawitan. *In all likelihood: Statistical modelling and inference using likelihood.* Oxford University Press, 2001.

[168] A. M. Ibrahim, T. ElBatt and A. El-Keyi. Coverage probability analysis for wireless networks using repulsive point processes. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1002–1007. IEEE, 2013.

[169] I. Flint, X. Lu, N. Privault, D. Niyato and P. Wang. Performance analysis of ambient RF energy harvesting with repulsive point process modeling. *IEEE Transactions on Wireless Communications*, 14(10):5402–5416, 2015.

[170] V. Rao, R. P. Adams and D. D. Dunson. Bayesian inference for Matérn repulsive processes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(3):877–897, 2017.

[171] B. Matérn. *Spatial variation*. Springer Science & Business Media, 2013.

[172] S. Chen, A. Shojaie, E. Shea-Brown and D. Witten. The multivariate Hawkes process in high dimensions: beyond mutual excitation. *arXiv preprint arXiv:1707.04928*, 2017.

[173] J. Feng and G. Wei. Increasing inhibitory input increases neuronal firing rate: Why and when? diffusion process cases. *Journal of Physics A: Mathematical and General*, 34(37):7493, 2001.

[174] U. Fano. On the theory of ionization yield of radiations in different substances. *Physical Review*, 70(1-2):44, 1946.

[175] U. Fano. Ionization yield of radiations. II. The fluctuations of the number of ions. *Physical Review*, 72(1):26, 1947.

[176] D. R. Cox. *Renewal theory*. Methuen, 1962.

[177] D. R. Cox and P. A. Lewis. The statistical analysis of series of events, 1966.

[178] Á. Corral. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letters*, 92(10):108501, 2004.

[179] G. Maimon and J. A. Assad. Beyond Poisson: Increased spike-time regularity across primate parietal cortex. *Neuron*, 62(3):426–440, 2009.

[180] H.-H. Jo, J. I. Perotti, K. Kaski and J. Kertész. Analytically solvable model of spreading dynamics with non-Poissonian processes. *Physical Review X*, 4(1):011041, 2014.

[181] A.-L. Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207–211, 2005.

[182] M. Livny, B. Melamed and A. K. Tsiolis. The impact of autocorrelation on queuing systems. *Management science*, 39(3):322–339, 1993.

[183] R. Gusella. Characterizing the variability of arrival processes with indexes of dispersion. *IEEE Journal on Selected Areas in Communications*, 9(2):203–211, 1991.

[184] Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman and J. Leskovec. Seismic: a self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1513–1522, 2015.

[185] V. Paxson and S. Floyd. Wide area traffic: the failure of Poisson modeling. *IEEE/ACM Transactions on networking*, 3(3):226–244, 1995.

[186] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on networking*, 2(1):1–15, 1994.

[187] T. Hiraoka, N. Masuda, A. Li and H.-H. Jo. Modeling temporal networks with bursty activity patterns of nodes and links. *Physical Review Research*, 2(2):023073, 2020.

[188] B. D'Auria and S. I. Resnick. Data network models of burstiness. *Advances in Applied Probability*, 38(2):373–404, 2006.

[189] H.-Q. Ye, J. Ou and X.-M. Yuan. Stability of data networks: stationary and bursty models. *Operations Research*, 53(1):107–125, 2005.

[190] A. Sun, D. D. Zeng and H. Chen. Burst detection from multiple data streams: A network-based approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(3):258–267, 2010.

[191] R. W. Klein and S. D. Roberts. A time-varying poisson arrival process generator. *Simulation*, 43(4):193–195, 1984.

[192] B. N. Dimitrov, V. V. Rykov and Z. L. Krougly. Periodic poisson processes and almost-lack-of-memory distributions. *Automation and Remote Control*, 65(10):1597–1610, 2004.

[193] A. Ihler, J. Hutchins and P. Smyth. Adaptive event detection with time-varying poisson processes. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 207–216, 2006.

[194] E. Lawrence, S. Vander Wiel, C. Law, S. B. Spolaor and G. C. Bower. The nonhomogeneous Poisson process for fast radio burst rates. *The Astronomical Journal*, 154(3):117, 2017.

[195] S. Rajaram, T. Graepel and R. Herbrich. Poisson-networks: A model for structured Poisson processes. In *International Workshop on Artificial Intelligence and Statistics*, pages 277–284. PMLR, 2005.

[196] S.-H. Kim and W. Whitt. Are call center and hospital arrivals well modeled by nonhomogeneous Poisson processes? *Manufacturing & Service Operations Management*, 16(3):464–480, 2014.

[197] A. De Santis, T. Giovannelli, S. Lucidi, M. Messedaglia and M. Roma. Determining the optimal piecewise constant approximation for the nonhomogeneous Poisson process rate of emergency department patient arrivals. *Flexible Services and Manufacturing Journal*, 34(4):979–1012, 2022.

[198] L. Glass. Synchronization and rhythmic processes in physiology. *Nature*, 410(6825):277–284, 2001.

[199] M. H. Jensen, P. Bak and T. Bohr. Complete devil's staircase, fractal dimension, and universality of mode-locking structure in the circle map. *Physical review letters*, 50(21):1637, 1983.

[200] M. McGuinness, Y. Hong, D. Galletly and P. Larsen. Arnold tongues in human cardiorespiratory systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 14(1):1–6, 2004.

[201] W. Briggs and W. L. Briggs. Anatomy of a circle map. *Mathematics magazine*, 72(2):116–125, 1999.

[202] G. Derks, P. A. Glendinning and A. C. Skeldon. Creation of discontinuities in circle maps. *Proceedings of the Royal Society A*, 477(2251):20200872, 2021.

[203] S. Coombes and S. Doole. Neuronal population dynamics with post inhibitory rebound: A reduction to piecewise linear discontinuous circle maps. *Dynamics and Stability of Systems*, 11(3):193–217, 1996.

[204] M. Bauer, S. Habip, D.-R. He and W. Martienssen. New type of intermittency in discontinuous maps. *Physical review letters*, 68(11):1625, 1992.

[205] S.-X. Qu, S. Wu and D.-R. He. A multiple devil's staircase in a discontinuous map. *Physics Letters A*, 231(3-4):152–158, 1997.

[206] M. Bailey, G. Derks and A. Skeldon. Circle maps with gaps: Understanding the dynamics of the two-process model for sleep–wake regulation. *European Journal of Applied Mathematics*, 29(5):845–868, 2018.

[207] A. Granados, L. Alsedà and M. Krupa. The period adding and incrementing bifurcations: from rotation theory to applications. *SIAM Review*, 59(2):225–292, 2017.

[208] G. Lajoie and E. Shea-Brown. Shared inputs, entrainment, and desynchrony in elliptic bursters: From slow passage to discontinuous circle maps. *SIAM Journal on Applied Dynamical Systems*, 10(4):1232–1271, 2011.

[209] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Courier Corporation, 2012.

[210] C. E. Landwehr, A. R. Bull, J. P. McDermott and W. S. Choi. A taxonomy of computer program security flaws, with examples. Technical report, Naval Research Lab Washington DC, 1993.

[211] R. H. Patterson, G. A. Gibson, E. Ginting, D. Stodolsky and J. Zelenka. Informed prefetching and caching. In *Proceedings of the fifteenth ACM symposium on Operating systems principles*, pages 79–95, 1995.

[212] B. D. Higgins, J. Flinn, T. J. Giuli, B. Noble, C. Peplin and D. Watson. Informed mobile prefetching. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, pages 155–168, 2012.

[213] J. N. Blakely, L. Illing and D. J. Gauthier. High-speed chaos in an optical feedback system with flexible timescales. *IEEE Journal of Quantum Electronics*, 40(3):299–305, 2004.

[214] E. U. Condon and P. M. Morse. *Quantum mechanics.* In McGraw Hill, New York, 1929, pages 55–72.

[215] A. Einstein. Die grundlage der allgemeinen relativitätstheorie. *Annalen der Physik*, 354(7):769–822, 1916.

[216] E. Artin. *The gamma function.* Courier Dover Publications, 2015.

[217] M. Aerin. Beta distribution - intuition, examples, and derivation. January 2020. URL: `towardsdatascience.com/beta-distribution-intuition-examples-and-derivation-cf00f4db57af` (visited on 03/03/2023).

[218] M. L. Huber and R. L. Wolpert. Likelihood-based inference for Matérn type-III repulsive point processes. *Advances in Applied Probability*, 41(4):958–977, 2009.