An Optimal Control Perspective on Externally Induced Tipping of Rigidly Shifting Systems

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Grace Zhao Zhang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Richard McGehee

February, 2024

Acknowledgements

I would like to express my gratitude to all those who have played a role in the completion of this PhD dissertation.

First and foremost, I am deeply indebted to my advisor, Richard McGehee, for his unwavering support and mentorship throughout this degree. His expertise, patience, trust, kindness, and broad encouragement to think expansively about the meaning of mathematics in the world have been instrumental in shaping the direction and trajectory of my research and my journey through graduate school.

I extend sincere thanks to my dissertation committee chair, Richard Moeckel, and the remaining members of my committee, Katherine Meyer, Arnd Scheel, and Tiffany Wolf, for their valuable insights, constructive criticism, and willingness to invest time and expertise into evaluating this work.

I gratefully acknowledge members of the Minnesota Mathematics of Climate group and the Mathematics and Climate Research Network for providing a stimulating and nourishing academic environment through numerous workshops, seminars, and thoughtful discussions that helped keep my mathematical motivations always grounded. I would particularly like to acknowledge the Minnesota Mathematics of Climate seminar, the 2020 Dynamics and Data workshop at the American Institute of Mathematics, and the 2023 Dynamical Systems and MCRN meeting.

I am further indebted to the UMN Institute on the Environment, the Ecosystem Health Graduate Scholars cohort program, the Ecosystem Health Seminar, and the UMN Grand Challenges Curriculum for deeply broadening my understanding of the interconnected nature of knowledge, enriching the scope of my academic engagements beyond the bounds of my home department, and teaching me about the rich interdisciplinary value of mathematics.

Dedication

This dissertation is dedicated to my family. To my parents Lulu and Yuanpeng whose unwavering love, support, and sacrifices have enabled me to pursue my dreams. To my brother Sebastian who always makes me laugh and think more than anyone else. With all my love, Grace.

Abstract

The standard setting for rate-induced tipping involves fixing a particular parameterized family of smooth forcing functions and identifying a critical value of the rate parameter. In contrast, we consider a broad collection of all possible forcing functions, continuous but not necessarily smooth, and seek a general property possessed by those which effect tipping behavior. We focus on rigidly shifting asymptotically autonomous scalar systems $\dot{x} = f(x + \lambda(t))$ and identify a nonsmooth choice of forcing function $\lambda(t)$ which is an optimal tipping strategy in the sense that it utilizes the least possible maximum speed. Under a co-moving change of coordinates, the problem of finding this optimal $\lambda(t)$ becomes dual to the problem of finding an additive control function that achieves basin escape with minimum fuel. We show the optimizer is a bang-bang control.

The outcome is a lower bound on the speed $|\dot{\lambda}(t)|$ that must be attained at least once in order to induce tipping. Its value depends on the total arclength $\int_{-\infty}^{\infty} |\dot{\lambda}(t)|\, dt$ of forcing, and may be interpreted as a safe threshold rate associated to each given arclength, such that if the speed of forcing remains everywhere slower than this, tipping cannot occur. The bound is tight in the sense that there exists a forcing function which induces tipping, possesses the required arclength, and never exceeds the threshold speed. Further, the threshold speed is a strictly decreasing function of arclength, thus capturing the abstract trade off between *how fast* and *how far* of a minimal disturbance characterizes tipping. While our results assume a scalar setting, the prospect of generalizing to n-dimensions is discussed and formulated as a conjecture.

The control-theoretic construction used in deriving the above inspires a new theory of resilience, which is a slight modification of the intensity kof attraction framework of McGehee and Meyer. This is a family of resilience values parameterized by a number representing the allowable L^1 norm of perturbations; in the limit as the integral-constraining parameter grows unbounded, these values approach the intensity of attraction. This integral-constrained intensity of attraction has the advantage of increased descriptiveness under scenarios where limited total resources are available for perturbing the system. We suggest it to be the natural choice for quantifying the resilience of a rigidly shifting system to externally-forced tipping.

Contents

Chapter 1

Introduction

The broad motivation for this dissertation is to explore two distinct but closely intertwined concepts in dynamical systems: tipping points and resilience. Specifically, we are concerned with rate-induced tipping of rigidly shifting systems and a control-theoretic framework for quantifying resilience known as intensity of attraction [1].

A tipping point or critical transition occurs in a dynamical system when a perturbation to the system state or to environmental conditions causes an abrupt overall shift in qualitative behavior. Tipping points have been studied in contexts as diverse as Earth's climate [2, 3], emerging infectious diseases [4], aquatic and land ecosystems [5, 6], the onset of medical health states [7, 8], financial markets [9], and more. Since tipping points often represent a shift into an undesirable or catastrophic regime, and since such transitions may not be easily or at all reversible, it has long been of interest to understand the dynamical mechanisms by which they are triggered. There are numerous formally defined behaviors that align with the intuitive notion of a tipping point, including traditional local and global bifurcations of autonomous dynamical systems, noise-induced tipping of stochastically driven systems, spatial pattern-formation mechanisms of instability, transitions between chaotic regimes, nonautonomous bifurcations of pullback or forward attractors, and rate-induced tipping.

Rate-induced tipping, or R-tipping is a relatively recently described phenomenon, first appearing in the mathematical literature in the previous decade [10]; very roughly speaking, while classical bifurcation-induced tipping is determined by *how much* environmental conditions change, rate-induced tipping is determined by *how quickly* environmental conditions change. Notably, the same total amount of change may or may not trigger tipping, depending on whether that change occurs slowly or rapidly. The framework of rate-induced tipping has gained attention partly for its clear relevance to modern day human and natural systems, such as climate and ecological systems that face the threat of rapid changes under the influence of anthropogenic forces [11, 12].

Chapter 3 begins with introductory examples of rate-induced tipping and gradually builds up to the central results of the dissertation. Along the way we construct most of the necessary theory for stating and proving the results; the exception is some preliminary background that is built in advance in Chapter 2 (the paragraph below here will describe what material is relegated to Chapter 2). Our core result can be summarized as the assertion that there exists a certain arclength-rate trade off for tipping in rigidly shifting externally driven systems; that is a trade off between *how far* and *how fast* of an external disturbance may be withstood safely without crossing a critical boundary. A unique aspect of our methods is the idea that under a co-moving change of coordinates, the problem of finding an external force that achieves tipping with minimal rate becomes dual to the problem of finding a control function that achieves tipping with minimal fuel; to our knowledge this is the first time that optimal control theory techniques have been utilized in order to derive a result about rate-induced tipping.

Prior to developing the results of Chapter 3, some preliminary background on the mathematics of quantifying resilience in dynamical systems is presented in Chapter 2. Loosely, resilience refers to the capacity for a dynamical system to retain its overall qualitative structure in the face of disturbances. Its precise definition depends highly on context (resilience *of what*, *to what*? [13]), and an abundance of approaches to formally quantifying resilience have been proposed. In Chapter 2 we focus on a set of three different measures of resilience that have been used by different authors. The first is asymptotic resilience, which is included because it is the most traditional mathematical definition of resilience and thus sets the context for the others. The other two are known as reactivity [14] and intensity of attraction [15, 1]. In brief, asymptotic resilience is good for measuring long-term recovery rates from small, isolated perturbations. In contrast, reactivity measures the capacity for a system to amplify rather than dampen the effect of a small perturbation in the short term, even if that system fully recovers from the perturbation in the long term.

Unlike asymptotic resilience and reactivity, which are both local measures of resilience, intensity of attraction is a control-theoretic approach to quantifying resilience whose primary advantage is that it considers large, continuous perturbations that might drive the state across the entire basin of attraction. This framework is clearly evocative of the kind of large continuous external forcing observed in rate-induced tipping. Indeed, the intensity of attraction of the attracting rest point of interest in the rigidly shifting system easily gives a lower bound on the rate of external forcing required to induce tipping; another way to state the purpose of our main results is to improve this loose lower bound to one that is tight. The method of obtaining this improvement inspires a refinement of intensity of attraction, which we call integral-constrained intensity of attraction, and the material on it is given in Chapter 4. This is a family of resilience values parameterized by a number representing the allowable L^1 norm of perturbations; in the limit as the integral-constraining parameter grows unbounded, the integral-constrained intensity of attraction approaches the intensity of attraction.

Chapter 2

Resilience Quantification Background

In this chapter, we present the definitions of asymptotic resilience, reactivity, and intensity of attraction. Firstly, asymptotic resilience, which is the most traditional mathematical definition of resilience, is a *local* measure of *long-term* resilience. In contrast, the remaining two metrics measure resilience in ways that consider transient dynamics. Reactivity measures a type of local, short-term resilience. Intensity of attraction measures resilience in a nonlocal way by taking into account the larger basin of attraction. For a review of other definitions of resilience from the point of view of mathematical ecology, one can refer to [16].

We begin by establishing some relevant dynamical systems preliminaries.

2.1 Dynamical Preliminaries

Let $U \subset \mathbb{R}^n$ be an open set, and assume that $f : U \to \mathbb{R}^n$ is a locally Lipschitz continuous function. Consider the ordinary differential equation (ODE)

$$\dot{x} = f(x) \tag{2.1}$$

The Lipschitz condition guarantees well-defined solutions, but only for sufficiently short intervals of time; hence we call the solutions local. We will use flow notation to collect all solution trajectories into one convenient object, called the **local flow** $\varphi : D \subset \mathbb{R} \times U \to U$, which is defined such that $\varphi(t, x_0) = x(t)$ is a solution to the initial value problem

$$\dot{x}(t) = f(x(t)), \qquad x(0) = x_0.$$

Depending on context, f may be globally Lipschitz continuous, in which case trajectories are defined for as long as they remain within the domain U. If no trajectories grow unbounded in U, then φ is a **global flow**, meaning it is defined for all time.

We will take the following notational conveniences. For a flow φ, we denote the time-t map as $\varphi_t : U \to U$, $x_0 \mapsto \varphi(t, x_0)$. We naturally extend this notation to allow set-valued inputs $S \subset U$:

$$\varphi_t(S) = \{x \in U \mid \varphi_t(x_0) = x \text{ for some } x_0 \in S\}.$$

In other words, the map φ_t outputs the location of any input point after it flows for t units of time. If the input is a set, then the output is also a set, consisting of all the locations reached at time t.

Two central objects of study in this paper are attractors and their associated basins of attraction. Attractors characterize the system's behavior as $t \to \infty$, by pulling trajectories toward them – at least, those trajectories which begin within their basin of attraction. Tipping behavior often comes down to either an abrupt shift in the nature of an attractor or an abrupt switch from one attractor to an alternative attractor. When we talk about resilience in this work, we are referring to the resilience *of an attractor*.

In order to define attractors and basins, we must first formalize some aspects of long-term behavior.

Definition 1. Consider a subset $S \subset U$. S is **invariant** under the flow φ if it contains all its own images in time: $\varphi_t(S) \subset S$ for all $t \in \mathbb{R}$.

Intuitively, an invariant set is one which is sealed off – nothing ever enters or exits it (although it can be approached asymptotically). The next definition collects the locations where an arbitrary set ends up, or at least approaches, in the long tun.

Definition 2. The **omega limit set** of $S \subset U$ is

$$\omega(S) = \bigcap_{T>0} \overline{\bigcup_{t>T} \varphi_t(S)}.$$

Now we have the vocabulary to formally define attractors and basins.

Definition 3. An **attractor** $A \subset U$ is a non-empty, compact, invariant set which is the omega limit set $\omega(N)$ of some neighborhood N of itself. Its **basin of attraction**, also called its **domain of attraction**, is

$$D(A) = \{x \in U \mid \omega(x) \subset A, \omega(x) \neq \emptyset\}.$$

Thus, attractors are the fixed structures in a system which are approached by nearby points in the long run. Each attractor has a certain dominion of rule – those trajectories beginning within its basin are the ones attracted toward it. While attractors may have interesting structures – periodic or chaotic, for instance – we will begin with the simplest type of attractor: an **attracting rest point**. Also referred to as asymptotically stable rest points or hyperbolic stable rest points, these points capture the intuitive idea of a "steady state."

The next definition says that a rest point is any unmoving point, while the subsequent proposition, which is standard theory, gives conditions under which a rest point is an attracting one.

Definition 4. x_* is a **rest point** or **equilibrium** of the ODE (2.1) if $f(x_*) = 0$.

Proposition 1. *If all eigenvalues of linearization at the rest point x_* have negative real part, that is,*

$$Re(\lambda) < 0 \textit{ for all } \lambda \in spec(\boldsymbol{D}f(x_*)),$$

where $\boldsymbol{D}f$ is the Jacobian of f, then x_* is an attractor.

Finally, we give standard terminology to classes of rest points which do not fall into the above category.

Definition 5. If all eigenvalues of linearization at the rest point x_* have non-zero real part, then x_* is called **hyperbolic**. Otherwise, at least one eigenvalue has zero real part, and we call x_* **non-hyperbolic**.

Definition 6. If x_* is hyperbolic, and at least one eigenvalue of linearization at the rest point x_* has positive real part, then x_* is called **unstable**.

Hyperbolic rest points can be thought of as "nice" rest points, ones near which the dynamics are predictable in some sense. Unstable rest points match the intuitive notion of unstable states – around them, nearly all trajectories are repelled away. At a non-hyperbolic point the behavior is not determined by its linear part.

This concludes our set up of the preliminary framework, and we continue next to quantifications of resilience.

2.2 Asymptotic Resilience

Throughout this subsection, we will assume that x_* is a hyperbolic attracting rest point of a continuously differentiable ODE. Probably the most commonly used and traditional mathematical definition of resilience, originating in theoretical ecology [17, 18, 19, 20], represents long-term return rates to x_*, and is measured by the real part of the dominant eigenvalue at linearization.

Definition 7. Let $\mathbf{A} = Df(x_*)$ denote the Jacobian, and recall that all eigenvalues of $\mathbf{A}$ have negative real part. Let $\lambda_1(\mathbf{A})$ be an eigenvalue with maximum (closest to 0) real part. The **asymptotic resilience** of the system at the attracting rest point is equal to the negative of that real part,

$$-Re(\lambda_1(\mathbf{A})).$$

We will refer to λ_1 as the **dominant eigenvalue** or the **slow eigenvalue** of $\mathbf{A}$.

For the linearized system $x' = \mathbf{A}x$, asymptotic resilience estimates the rate at which trajectories approach the equilibrium. The following theorem is standard theory for linear ODEs.

Theorem 1. *For an $n \times n$ matrix $\mathbf{A}$, if $Re(\lambda) < L < 0$ for all eigenvalues λ of $\mathbf{A}$, then there is some constant $C > 0$ such that for all $x \in \mathbb{R}^n$ and $t \geq 0$,*

$$|e^{t\mathbf{A}}x| \leq Ce^{Lt}|x|.$$

And in the long term C can be taken to equal 1. That is, there is some $T \geq 0$ such that

$$|e^{t\mathbf{A}}x| \leq e^{Lt}|x| \quad \text{for all } t \geq T.$$

Note the operator $e^{t\mathbf{A}}$ in the left hand side is exactly the flow φ_t for the linear system $x' = \mathbf{A}x$. So in the long term trajectories must decay to the origin at an exponential rate governed by the asymptotic resilience.

For nonlinear systems, similar results for decay rate are justified by the Stable Manifold Theorem, standard theory which says that, at sufficiently nice rest points, the linear approximation is a good approximation. A special case of the Stable Manifold Theorem is stated here, while a full version can be found in any standard text.

Theorem 2. *(Stable Manifold Theorem, for attracting rest points) Consider a non-linear system*

$$x' = \mathbf{A}(x) + h(x),$$

*where $\mathbf{A}, h : \mathbb{R}^n \to \mathbb{R}^n$ with $\mathbf{A}$ linear and $Dh(0) = 0$. Let φ_t be the local flow. Assume there is an attracting rest point at the origin. Let λ_1 be the dominant eigenvalue of $\mathbf{A}$. Then there exists a neighborhood $N \ni 0$ which is a **local stable manifold** of the origin. That is, for all $x \in N$, $\lim_{t\to\infty} \varphi_t(x) = 0$.*

Furthermore, for any $Re(\lambda_1) < L < 0$, there exists $C > 0$ such that for all $x \in N$, $t \geq 0$,

$$|\varphi_t(x)| \leq Ce^{Lt}|x|,$$

and for some $T \geq 0$, C can be taken to equal 1

$$|\varphi_t(x)| \leq e^{Lt}|x| \quad \text{for } t \geq T.$$

The Stable Manifold Theorem implies that any trajectory beginning sufficiently close to equilibrium decays toward equilibrium at an exponential rate, where that rate is determined in the long term by asymptotic resilience. Any point which is very close to, but not quite at, the equilibrium represents a state slightly perturbed away from steady state. Hence, the rate of decay can be thought of as the recovery rate from a small perturbation.

One major drawback of asymptotic resilience is the assumption that the system is well approximated by its asymptotic or long-term behavior; another is that it relates only to small disturbances which do not drive the system state very far from equilibrium. In particular, there is a neglect of transient dynamics, such as short term behavior and behavior far away from the attractor.

2.3 Reactivity

Note that trajectories need not decay monotonically in distance to the attracting rest point, not even for linear systems. In the short term, a perturbation can initially be amplified in magnitude before eventually decaying to the stable equilibrium – a phenomenon termed **reactivity** by Neubert and Caswell in [14] (Figure 2.1). A real world example of positive reactivity occurs in infectious disease systems, where a new introduction of a pathogen leads to a short term epidemic, even if the disease eventually dies out naturally.

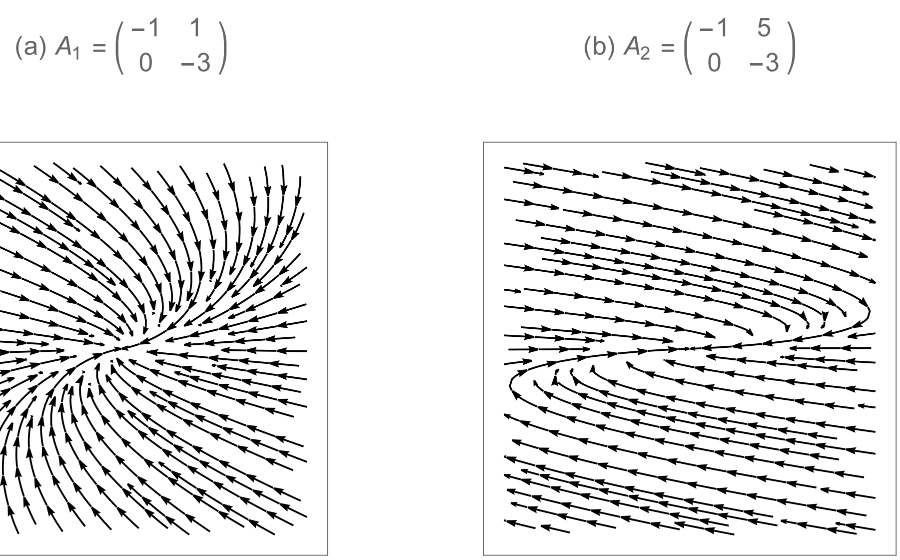


Figure 2.1: Phase portraits of two linear systems $\dot{x} = \mathbf{A}x$. (a) All trajectories decay monotonically in magnitude. (b) There are trajectories beginning arbitrarily close to the origin which initially increase in magnitude. Notice that both matrices have the same eigenvalues $\lambda = -1, -3$; hence asymptotic resilience cannot discern whether an attracting equilibrium is reactive.

Definition 8. Suppose that x_* is an attracting rest point of the ODE (2.1). Let $\mathbf{A} = Df(x_*)$ be the Jacobian, and let $\mathbf{H} = \frac{\mathbf{A}+\mathbf{A}^T}{2}$ be its symmetric part. Since $\mathbf{H}$ is a real symmetric matrix, it has real eigenvalues. Let $\lambda_1(\mathbf{H})$ be the maximum eigenvalue.

The **reactivity** of the system at the stable rest point is $\lambda_1(\mathbf{H})$.

If this number is positive, the attractor is called **reactive**.

Reactivity measures the maximum possible relative rate of initial amplification. The following proposition captures this assertion.

Proposition 2 (Neubert and Caswell)**.** *For the linear system* $\dot{x} = \mathbf{A}x$, $\lambda_1(\mathbf{H}) = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{1}{||x||} \frac{d}{dt} ||x||$.

Proof.

$$\begin{aligned} \frac{1}{||x||} \frac{d}{dt} ||x|| &= \frac{1}{||x||} \frac{d}{dt} \left(x^T x\right)^{1/2} \\ &= \frac{1}{2||x||^2} (x^T \dot{x} + \dot{x}^T x) \\ &= \frac{1}{2||x||^2} (x^T \mathbf{A} x + x^T \mathbf{A}^T x) \\ &= \frac{x^T \mathbf{H} x}{||x||^2} \end{aligned}$$

This expression is a scale invariant function of x, so to maximize it, we only need to consider unit vectors:

$$\max_x \frac{x^T \mathbf{H} x}{||x||^2} = \max_{||x||=1} x^T \mathbf{H} x.$$

$\mathbf{H}$ is real symmetric, hence diagonalizable with an orthogonal change of basis. Let $\{\lambda_1, \lambda_2, \ldots \lambda_n\} = spec(\mathbf{H})$, in order from largest to smallest.

$$\begin{aligned} x^T \mathbf{H} x &= x^T (\mathbf{B} \mathbf{D} \mathbf{B}^T) x \\ &= (x^T \mathbf{B}) \mathbf{D} (\mathbf{B}^T x) \\ &= y^T \mathbf{D} y, \quad \text{where } y = \mathbf{B}^T x \text{ is also unit.} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2. \end{aligned}$$

The maximum of this expression over all $||y|| = 1$ is clearly λ_1, when $y = (y_1, \ldots, y_n) = (1, 0, \ldots, 0)$.

□

Remark 1. Since $\mathbf{H}$ is a symmetric matrix, it defines a quadratic form $Q(x) = x^T\mathbf{H}x$. The maximum value of a quadratic form restricted to the unit sphere is the largest eigenvalue of its matrix representation H. Thus, to say that the attracting rest point has negative reactivity is equivalent to saying that $Q(x)$ is negative definite.

Remark 2. Non-negative reactivity can only occur in two or greater dimensions, since for scalar systems $\mathbf{A} = \mathbf{H}$.

2.4 Intensity of Attraction

Asymptotic resilience and reactivity notably rely on linearizing at a point attractor. In contrast, intensity of attraction [1], originally introduced by McGehee for discrete maps [21], and extended to the continuous case by Meyer [15], measures resilience not only for rest points but also for any other type of attractor. Even more importantly, it captures metric information across the entire basin of attraction rather than simplifying to a topologically equivalent approximation within a local neighborhood. While the local point of view may suffice for measuring resilience to a small, isolated perturbation, the larger basin of attraction becomes important when the system state is driven far away from equilibrium. Such a transient state can result from large, continual disturbances, as are common in many ecological and other real world settings, such as environmental forces or human-driven pressure on an ecosystem. We now review the necessary background in order to define intensity of attraction.

2.4.1 Bounded Control, Reachable Sets, and Intensity

First of all, the idea of perturbation will now be represented by a **control function** combined additively with an underlying vector field. We assume that the control function

$$u : \mathbb{R} \to \mathbb{R}^n$$

is taken from the space $L^\infty(\mathbb{R}, \mathbb{R}^n)$ of measurable and essentially bounded functions. In particular, the measurability of u implies that it is allowed some points of discontinuity, but is continuous almost everywhere (except possibly on a set of Lebesgue measure 0), while essential boundedness means bounded almost everywhere. Here the norm is

$$||u||_\infty = \inf\{C \geq 0 : ||u(x)|| \leq C \text{ for almost every } x \in \mathbb{R}\}.$$

Next, we formalize how the perturbation is added to the underlying system.

Definition 9. A **bounded control system** is a non-autonomous ODE

$$\dot{x} = f(x) + u(t) \tag{2.2}$$

where $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz, $u \in L^\infty(\mathbb{R}, \mathbb{R}^n)$.

Here, the underlying system is thought of as an ODE $\dot{x} = f(x)$; but it is altered by adding a perturbation $u(t)$ to the vector field $f(x)$ on the right hand side. The effect of $u(t)$ is to adjust, at every point in time, the path of solutions somewhat away from what would have been their original trajectory.

It remains to be justified whether this construction produces a well-defined system. Because the right hand side $f(x) + u(t)$ may be a discontinuous function, solutions $x(t)$ of the ODE must be considered in an extended sense. The right hand side of the ODE (3.1) satisfies the Carathéodory conditions [22]:

- For every fixed t, $f(x) + u(t)$ is clearly continuous in x.
- For every fixed y, $f(x) + u(t)$ is clearly measurable in t.
- $u(t)$ is essentially bounded on $\mathbb{R}$ and $f(x)$ is continuous hence bounded on every compact set $K \subset \mathbb{R}^n$, so $|f(x) + u(t)| \leq \sup_K |f| + \text{ ess } \sup_{\mathbb{R}} |u|$, which is a constant function and thus Lebesgue-integrable on $K \times \mathbb{R}$.

This is enough to guarantee local existence of an absolutely continuous solution $x(t)$

to any initial value problem $x(t_0) = x_0$ in the extended sense that

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(s)) + u(s)\ ds$$

and $\dot{x}(t) = f(x) + u(t)$ almost everywhere.

Furthermore, f is locally Lipschitz continuous, so fixing any t gives a locally Lipschitz continuous $f(y) + u(t)$, where the local Lipschitz constant is clearly not affected by the choice of t, hence is uniform in t. This additional condition guarantees uniqueness of the local solution.

So we have well-defined solutions, and can therefore extend the standard local flow notation to the bounded control setting. Fixing an underlying vector field f, we will denote as follows the flow obtained by applying a choice of perturbation u.

Definition 10. $\varphi_u(t, x_0) : D \subset \mathbb{R} \times U \to U$ is the local flow defined by

$$\varphi_u(t, x_0) = x(t)$$

where $x(t)$ solves in the extended sense the ODE (2.2), with initial condition $x(0) = x_0$.

Intensity of attraction considers not just one single control function, but entire families of control functions – specifically, those where every function is bounded by some maximum magnitude r. The next definition gives a notation for these families.

Definition 11. Denote by $B_r \subset L^\infty[\mathbb{R}, \mathbb{R}^n]$ the set of control functions bounded above by r:

$$B_r = \{u : ||u||_\infty < r\}$$

This leads, next, into the notion of all possible states reachable in forward time, under the family of all possible control functions bounded by r, and beginning from some arbitrary initial set of states.

Definition 12. Consider $S \subset U$. The **reachable set** of S under r-bounded control is the set

$$R_r(S) = \bigcup_{u \in B_r} \bigcup_{x_0 \in S} \bigcup_{t \geq 0} \varphi_u(t, x_0)$$

Finally, we are ready to define intensity of attraction, which captures the folllowing idea: what is the smallest magnitude of control necessary in order to escape from the basin of attraction?

Definition 13. If A is an attractor of $\dot{x} = f(x)$ and $D(A)$ denotes its basin of attraction, then the **intensity of attraction** of A is

$$\mu(A) = \sup\{r \geq 0 \mid R_r(A) \subset D(A)\}.$$

2.4.2 Intensity in Scalar Systems

For scalar systems, the intensity of attraction is simple to describe. There are up to two possible sides through which to escape the basin of attraction. Over each escapable side, consider the largest magnitude of the vector field. The intensity of attraction equals the minimum of these.

Proposition 3. *Consider the one-dimensional system $\dot{x} = f(x)$ with an attracting rest point at $x = a$ with basin of attraction $D = (\alpha, \beta)$, where $-\infty \leq \alpha < a < \beta \leq \infty$. Define*

$$\mu_- = \begin{cases} \sup\{f(x) \mid x \in [\alpha, a]\} \text{ if } \alpha \neq -\infty \\ \infty \text{ otherwise} \end{cases}$$

$$\mu_+ = \begin{cases} -\inf\{f(x) \mid x \in [a, \beta]\} \text{ if } \beta \neq \infty \\ \infty \text{ otherwise} \end{cases}$$

The intensity of attraction of the attractor $A = \{a\}$ is $\mu(A) = \min\{\mu_-, \mu_+\}$.

Proof. We omit this proof because it follows very similarly to Proposition 4.2 in [1], which describes all the reachable sets in a scalar system. □

Remark 3. Another way to interpret intensity of attraction relates to the idea of basin steepness. What is the steepest part of the basin that must be overcome in order to escape the influence of the attractor? For scalar systems, this intuition is precise: the vector field $f : \mathbb{R} \to \mathbb{R}$ is always integrable, producing a potential function, and the maximum steepness of that potential on the basin determines intensity of attraction [15].

Unfortunately, for two and higher dimensional systems, no potential function necessarily exists, complicating the landscape analogy. Still, the idea of basin steepness serves as a rough heuristic interpretation of the intensity of attraction.

2.4.3 A Lake Eutrophication Model

Example 1. A scalar example of intensity of attraction. A simple model of lake eutrophication [23] has two alternative stable states, corresponding to a low-nutrient oligotrophic lake versus a nutrient-overloaded eutrophic lake.

$$\dot{p} = \ell - sp + r(\frac{p^q}{m^q - p^q})$$

Here the state variable p represents the total amount of phosphorus in the lake. Its evolution is affected by the phosophrus inflow ℓ, loss rate s, and recycling rate r, while m is a parameter that modulates the recycling term, which has a sigmoidal shape.

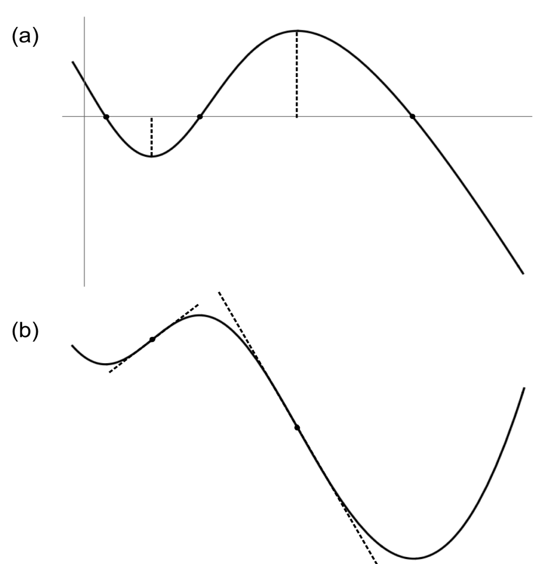


Figure 2.2: Intensity of attraction in a 1D lake eutrophication model, as illustrated on (a) the vector field and (b) its potential landscape.

Figure 2.2a shows the vector field $f(p)$. The left attractor is an oligotrophic state with low phosphorus content. The attractor on the right is a eutrophic state with high phosphorus content. The intensities of attraction are indicated by the heights of the two vertical lines occuring at two local extrema of f. Figure 2.2b shows a potential landscape $h(p)$ such that $f(p) = -\frac{d}{dp}h(p)$. The intensities of attraction are now equal to the two extreme slopes of h.

Under a plausible parameter regime, the eutrophic attractor has a higher intensity of attraction. Although the model is simplistic, this fact could be interpreted as a reflection of the reality that tipping an originally oligotrophic lake to eutrophy (for example, via increased nutrient run-off from agricultural and industrial activities) is typically much easier than attempting to restore an already eutrophic lake to oligotrophy (for example, via an intensive program to remove or inactivate nutrients).

2.4.4 Comparison to Meyer-McGehee Construction

Remark 4. We have given a slightly different definition of intensity of attraction compared to the original McGehee and Meyer construction, which is

$$\mu(A) = \sup\{r \geq 0 \mid R_r(A) \subset K \subset D(A) \text{ for some compact } K\}.$$

The difference lies in whether one wishes to consider a trajectory going to infinity without leaving the basin as a successful escape or not. For our purposes, it is more useful to say no. We require trajectories to escape through a finite boundary point of the basin. The next example illustrates the difference between the two definitions.

Example 2. An example illustrating the difference between the two approaches to defining the intensity of attraction. In Figure 2.3, the vector field has an attractor at $f(-1) = 0$, whose basin of attraction is the interval $(-\infty, 1)$. On the left side $(-\infty, -1)$ of the basin $f(x) < 1$ and $\lim\limits_{x \to -\infty} f(x) = 1$. On the right side $(-1, 1)$ of the basin, $\max |f(x)| = 2$.

According to the original McGehee and Meyer definition, the intensity of attraction of the attractor $A = \{-1\}$ is $\mu(A) = 1$, since the control $u(t) \equiv -1$ applied to the initial condition $x(0) = -1$ steers the solution $x(t) \to -\infty$ as $t \to \infty$, thereby escaping every compact subset of the basin.

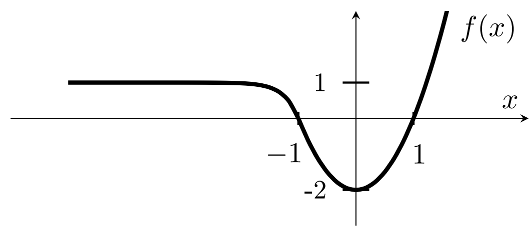


Figure 2.3: An example where the two definitions for intensity of attraction differ.

According to the definition used in the present work, the intensity of attraction of the attractor $A = \{-1\}$ is $\mu(A) = 2$, since true escape from the basin can only occur through the finite boundary point on the right side.

For a nontrivial example of intensity of attraction in two dimensions, refer to the McGehee and Meyer source [1]. Analytically computing the intensity of attraction in dimensions higher than one is generally very difficult. The numerical approximation of reachable sets also presents challenges and remains an active research area in control theory. Nevertheless, the abstract utility of intensity of attraction for our present purposes is significant.

Chapter 3

An Application of Optimal Control Theory to Rate-Induced Tipping

This chapter aims to establish for rigidly shifting forced systems $\dot{x} = f(x + \lambda(t))$ the existence of a minimum critical rate associated to any given amplitude of external forcing, such that if an arbitrary forcing function of the specified amplitude is to successfully induce tipping it exhibit a speed at or above the critical rate at least once. Equivalently, this critical rate may be interpreted as a safe threshold rate, such that if the speed of forcing remains everywhere slower than this tipping does not occur. Additionally, the threshold rate is a strictly decreasing function of the forcing amplitude, thereby yielding a generalized amplitude-rate trade off, that is a trade off between *how far* and *how fast* of an external force the system can safely withstand.

The key to our approach is to transform the rigidly shifting rate-induced tipping system into an additive control system using a co-moving change of coordinates. Then, the problem of finding a translational external force that achieves tipping with minimal rate is transformed into a dual problem of finding an additive control function that achieves tipping with minimal fuel.

We separate the presentation into a scalar special case where the forcing function is assumed to be monotone and the basin of attraction one-sided, followed by the general scalar case with these assumptions removed. Lastly, we present a sketch toward obtaining a possible n-dimensional generalization.

3.1 Introductory Examples

3.1.1 Smooth Prototype

Example 3. A prototypical example of rate-induced tipping, given in [24], involves a base vector field $\dot{x} = x^2 - 1$ which is nonautonomously shifted to the left by a smooth ramp function $\lambda(rt)$. A fixed constant $\lambda_\infty > 2$ defines the total amplitude of the shift, while a variable rate parameter $r > 0$ modulates its steepness, with smaller r corresponding to a slower shift and larger r to a faster shift.

$$\dot{x} = (x + \lambda)^2 - 1$$
$$\lambda(rt) = \frac{\lambda_\infty}{2}\left(1 + \tanh\left(\frac{\lambda_\infty rt}{2}\right)\right)$$

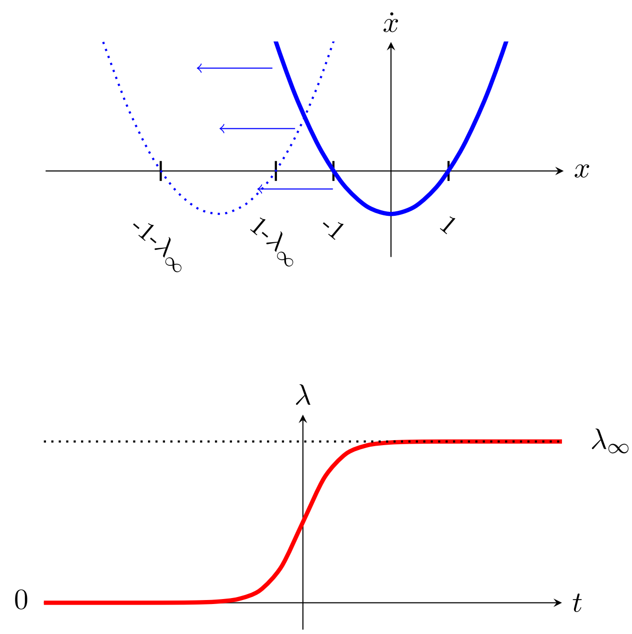


Figure 3.1: A prototypical example of rate-induced tipping.

As the vector field translates rigidly leftwards, the attracting equilibrium at $x = -1$ is displaced. Rate-induced tipping concerns itself with whether a trajectory beginning at the original steady state can adapt to this displacement or if it becomes destabilized.

Here it is known that for each parameter regime (λ_∞, r) there exists a unique solution $\hat{x}(t)$ to the ODE such that $\lim_{t\to-\infty} \hat{x}(t) = -1$. Further, fixing λ_∞, there exists a critical value $r = r_c(\lambda_\infty)$ such that

$$\begin{cases} \lim_{t\to\infty} \hat{x}(t) = -1 - \lambda_\infty & \text{for } r < r_c \\ \lim_{t\to\infty} \hat{x}(t) = 1 - \lambda_\infty & \text{for } r = r_c \\ \hat{x}(t) \to \infty \text{ (in finite time)} & \text{for } r > r_c \end{cases},$$

and an exact expression for this critical value [25] is known to be

$$r_c = \frac{4}{\lambda_\infty(\lambda_\infty - 2)}.$$

Intuitively, a sufficiently slow shift allows the trajectory to seamlessly "track" the moving attractor. But a too-fast shift destabilizes it onto the other side of the moving repeller. There is a critical rate in between where the trajectory ends up balanced precisely on the basin boundary.

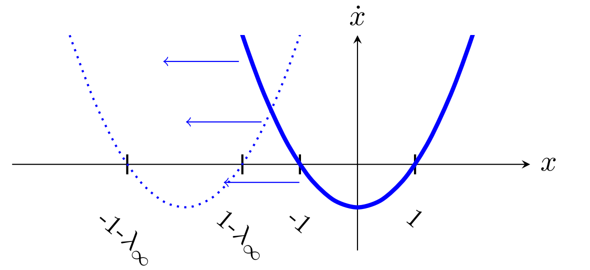
3.1.2 Piecewise Linear Prototype

Most literature on rate-induced tipping customarily assumes smoothness of the ramping function; however, the next instance of a nonsmooth, piecewise linear ramp will be of core importance to this chapter.

Example 4. Replace the smooth ramping function in Example 3 with the following piecewise linear ramping function, where the slope of the increasing portion is $m > 0$, and as before $\lambda_\infty > 2$.

$$\dot{x} = (x + \lambda)^2 - 1$$

$$\lambda(mt) = \begin{cases} 0 & \text{if } t < 0 \\ mt & \text{if } 0 \le t \le \lambda_\infty/m \\ 0 & \text{if } t > \lambda_\infty/m \end{cases}$$



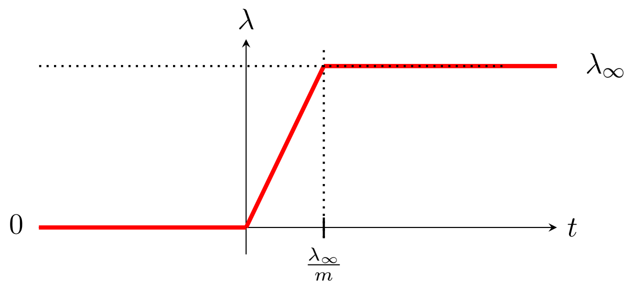


Figure 3.2: A piecewise linear ramping function.

The formal setting for this and related non-smooth ODEs will be reviewed later; however, it can be shown that tipping behavior analogous to the previous example occurs. For each parameter regime (λ_∞, m) the ODE possesses a unique solution $\hat{x}(t)$ such that $\lim_{t\to-\infty} \hat{x}(t) = -1$.* Fixing λ_∞, there exists a critical value $m = m_c(\lambda_\infty)$ with

$$\begin{cases} \lim_{t\to\infty} \hat{x}(t) = -1 - \lambda_\infty & \text{for } m < m_c \\ \lim_{t\to\infty} \hat{x}(t) = 1 - \lambda_\infty{}^\dagger & \text{for } m = m_c \\ \hat{x}(t) \to \infty \text{ (in finite time)} & \text{for } m > m_c \end{cases}.$$

For this example, it can be shown that m_c is given by the unique solution to the equation

$$\frac{2m_c}{\sqrt{m_c - 1}} \arctan\left(\frac{1}{\sqrt{m_c - 1}}\right) = \lambda_\infty.$$

We remark that m_c is a strictly decreasing function of λ_∞; that is, a larger amplitude

*In this case, $\hat{x}(t) = -1$ for all $t \in (-\infty, 0]$

†In this case $\hat{x}(t) = 1 - \lambda_\infty$ for all $t \in [\lambda_\infty/m, \infty)$.

shift allows for a more gentle critical rate of shift while a smaller amplitude shift requires a steeper critical rate of shift. A similar amplitude-rate trade off between r_c and λ_∞ may be observed in Example 3, for instance by focusing on the maximum slope of the critical ramp function, which occurs at $t = 0$.

Each of the two previous examples involves fixing a particular family of ramp functions, parameterized by a variable r or m that controls the overall steepness of the ramp, and then identifying a critical ramp from within the predetermined family. This is the standard point of view that prevails in the rate-induced tipping literature.

In contrast, we are motivated by a different approach: to instead consider the entire collection of all possible functions that interpolate from 0 to λ_∞ (subject to appropriate conditions) and to seek a general property about the steepness of those which effect tipping behavior.

Our core insight is that the critical piecewise linear ramp function of Example 4 is actually an optimal tipping strategy in an important sense. In particular, it utilizes the least possible maximum slope; in other words, we will show that any arbitrary scalar ramp function $\lambda(t)$ (monotone non-decreasing, for now) that induces tipping when applied to the same base vector field must attain a slope greater than or equal to $m_c(\lambda_\infty)$ at least once. The result is a necessary but not sufficient criterion for tipping, or equivalently a safe threshold for non-tipping.

3.2 Change to Co-Moving Coordinates

For $\dot{x} = f(x + \lambda(t))$, we consider a change of coordinates to co-moving coordinates,

$$\begin{aligned} y &= x + \lambda(t) \\ \implies \dot{y} &= f(y) + \dot{\lambda}(t) \end{aligned}$$

This transformation has the effect of converting the ramp function that originally translated the base vector field $\dot{x} = f(x)$ leftwards into a pulse function that now translates the same base vector field up and then back down.

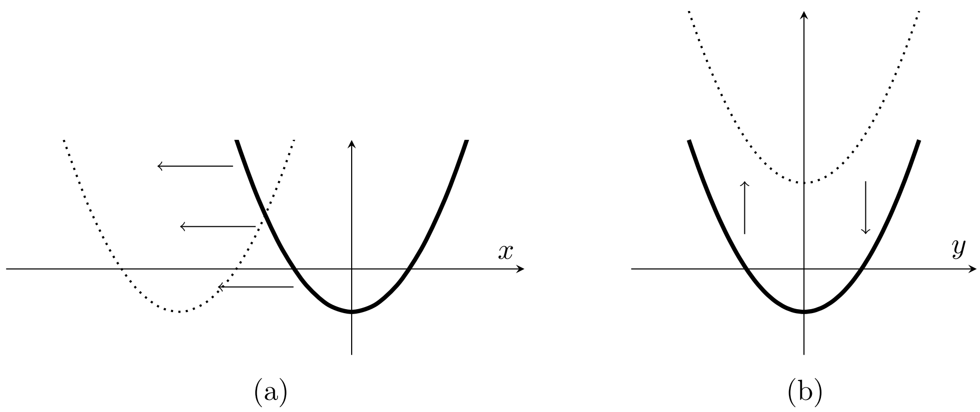


Figure 3.3: A change to co-moving coordinates converts the leftward translating ramp function to an up-and-down translating pulse function.

3.2.1 Smooth Prototype

Example 5. For $\lambda(rt)$ as in Example 3 the pulse function is

$$\dot{\lambda}(r,t) = \left(\frac{\lambda_\infty}{2}\right)^2 r \,\mathrm{sech}^2\left(\frac{\lambda_\infty rt}{2}\right)$$

where, fixing λ_∞, a smaller value of the parameter r corresponds to a shorter and wider peak, while a larger value of the parameter r corresponds to a taller and narrower peak.

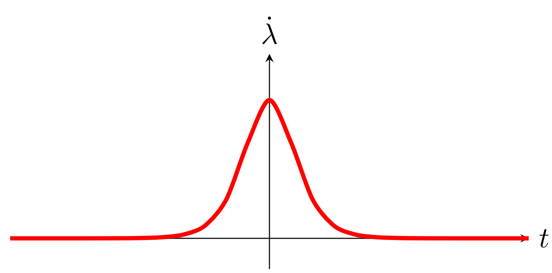


Figure 3.4: The pulse function that results from the smooth prototype ramp.

Regardless of the choice of r, note that the total area under the pulse is unaffected, since it is always equal to λ_∞.

3.2.2 Piecewise Linear Prototype

Example 6. For $\lambda(m,t)$ as in Example 4 we obtain a discontinuous step pulse

$$\dot{\lambda}(mt) = The \begin{cases} 0 & \text{if } t < 0 \\ m & \text{if } 0 \le t \le \lambda_\infty/m \\ 0 & \text{if } t > \lambda_\infty/m \end{cases}.$$

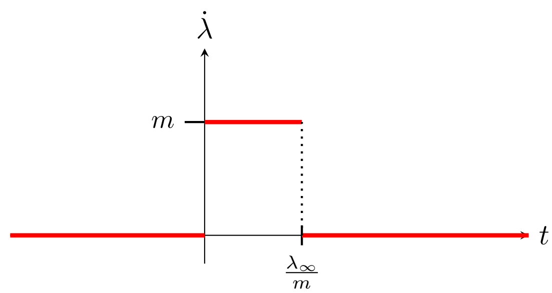


Figure 3.5: The pulse function that results from the piecewise linear ramp.

The formal setting for this non-smooth change of coordinates between a non-smooth vector field and a discontinuous vector field will be discussed soon, in section 3.3. But here we can see that smaller m corresponds to a shorter and wider step while larger m results in a taller and narrower step, again while always preserving the area λ_∞ underneath. Here, the critical step has a height of $m_c = m_c(\lambda_\infty)$ as defined in Example 4.

In the co-moving frame of reference, our desired assertion becomes the statement that any arbitrary tipping pulse whose total area equals λ_∞ must at some point reach or surpass the height $m_c(\lambda_\infty)$ of the critical step from Example 6.

We adopt the point of view that an "arbitrary pulse" is a measurable and essentially bounded (and non-negative, for now) control function $u(t)$ added to the base vector field to obtain the nonautonomous ODE

$$\dot{y} = f(y) + u(t)$$

Then, fixing the restriction $\int_{-\infty}^{\infty} u(t)\ dt = \lambda_\infty$, we would wish to demonstrate that the essential supremum of any control u that induces tipping is at least the height of the critical step pulse from Example 6. Actually, in order to cast this problem into an optimization problem with a more amenable cost function and constraint, we will instead prove a near-contrapositive before subsequently recovering the full result.

Remark 5. A control function which simply toggles between a minimum and maximum value, such as the critical step function we have described, is commonly known as a **bang-bang control**. Bang-bang control arises as an optimal control in several contexts [26].

3.3 Formal Setting

Although we at first restrict ourselves to the scalar case $f : \mathbb{R} \to \mathbb{R}$, we will subsequently wish to discuss the possibility of generalizing into n-dimensions. Hence, it is convenient to present a portion of this section with the generality of n-dimensional dynamics.

We set our attention on control systems of the form

$$\dot{y} = f(y) + u(t). \tag{3.1}$$

Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^2 (hence locally Lipschitz). Assume $u(t) : \mathbb{R} \to \mathbb{R}^n$ is Lebesgue measurable and essentially bounded. As discussed earlier in Section 2.4, the right hand side of the ODE (3.1) satisfies the Carathéodory conditions for existence and uniqueness of solutions on $\mathbb{R}^n \times \mathbb{R}$ in the extended sense that:

$$y(t) = y(t_0) + \int_{t_0}^{t} f(y(s)) + u(s)\ ds$$

and $\dot{y}(t) = f(y) + u(t)$ almost everywhere.

3.3.1 Tipping Induced by Additive Control

Now add the assumption that the ODE $\dot{y} = f(y)$ has a hyperbolic attracting rest point at $y = a$.

Proposition 4. *Assuming f, u, a as above with the additional condition that* $\lim_{t\to-\infty} u(t) = 0$*, there exists a unique solution* $\hat{y}(t)$ *to the ODE* $\dot{y} = f(y) + u(t)$ *such that* $\lim_{t\to-\infty} \hat{y}(t) = a$.

Proof. This claim is closely related to a fundamental result in R-tipping where the external force is assumed to be smooth and is not necessarily applied rigidly (see Theorem 2.2 in [24]). In our context, an essentially identical argument still carries through, and the co-moving coordinate change makes it slightly less burdensome.

First, assume without loss of generality that $a = 0$; otherwise, translate the base vector field by replacing y with $y - a$. Define

$$\omega(\epsilon) = \sup\{|Df(y) - Df(0)| : |y| < \epsilon\}$$
$$\delta(T) = \operatorname*{ess\,sup}_{t<-T} |u(t)|$$

Because $f \in C^2$ we have $\omega(\epsilon) \to 0$ as $\epsilon \to 0$. Because $\lim_{t\to-\infty} u(t) = 0$ we have $\delta(T) \to 0$ as $T \to \infty$. Hyperbolic stability of the equilibrium at 0 means that there exist $K > 0, \alpha > 0$ such that

$$|e^{At}| \leq Ke^{-\alpha t} \text{ for } t \leq 0$$

where $A = Df(0)$. Now define $h(y,t) = f(y) + u(t) - Ay$ and rewrite the ODE as

$$\dot{y} = Ay + h(y,t).$$

Notice $D_y h = Df(y) - A = Df(y) - Df(0)$, so for all $t < -T$ we have

$$|d_y h(y,t)| \leq \omega(|y|)$$

$$\text{and } |h(0,t)| = |u(t)| \leq \delta(T) \text{ almost everywhere.}$$

Now choose any $\Delta > 0, T_0 > 0$ such that

$$K\alpha^{-1}\omega(\Delta) \leq \frac{1}{2} \text{ and } K\alpha^{-1}\delta(T_0) \leq \frac{\Delta}{2}$$

and consider the space of continuous functions

$$S = \{y(t) \in C^0((\infty, -T_0]) : |y(t)| \leq \Delta \text{ for } t < -T_0\}.$$

We define an operator on S

$$\Phi(y) = \int_{-\infty}^{t} e^{A(t-s)} h(y(s), s)\ ds$$

and verify first that it is well defined, second that it is a contraction mapping.

For the former,

$$\begin{aligned} |\Phi(y)(t)| &\leq \int_{-\infty}^{t} Ke^{-\alpha(t-s)}[\delta(T_0) + |y(s)|\omega(|y(s)|)]\ ds \\ &\leq \int_{-\infty}^{t} \\ &\leq K\alpha^{-1}\delta(T_0) + K\alpha^{-1}\Delta\omega(\Delta) \\ &\leq \Delta \end{aligned}$$

and for the latter, since $h(y,t)$ is Lipschitz continuous in y with Lipschitz constant $\omega(\Delta)$,

$$\begin{aligned} ||\Phi(y_1) - \Phi(y_2)|| &= \sup_{t \leq -T_0} \left| \int_{-\infty}^{t} e^{A(t-s)} \left(h(y_1(s), s) - h(y_2(s), s)\right)\ ds \right| \\ &\leq \sup_{t \leq -T_0} \int_{-\infty}^{t} Ke^{-\alpha(t-s)} \left|h(y_1(s), s) - h(y_2(s), s)\right|\ ds \\ &\leq K\alpha^{-1} \sup_{t \leq -T_0} |h(y_1(t), t) - h(y_2(t), t)| \\ &\leq K\alpha^{-1}\omega(\Delta)||y_1 - y_2|| \\ &\leq \frac{1}{2}||y_1 - y_2|| \end{aligned}$$

So Φ has a unique fixed point $\hat{y}$. By the nonlinear variation of parameters formula $y(t) = y_0 e^{A(t-t_0)} + \int_{t_0}^{t} e^{A(t-s)} h(y(s), s)\, ds$, and by taking the limit as $t_0 \to -\infty$, it follows that the fixed point $\hat{y}(t)$ is also the unique solution of the ODE that satisfies $|y(t)| \leq \Delta$ for all $t \leq -T_0$. Since Δ can be chosen arbitrarily small, $\lim_{t \to -\infty} \hat{y}(t) = 0$. □

Next, we define tipping based on the forward time behavior of the solution $\hat{y}(t)$ from the previous proposition. Specifically, we would like tipping to depend on the relationship between the future trajectory and the boundary of the basin of attraction.

Letting $D \subset \mathbb{R}^n$ denote the basin of attraction of the attracting rest point at $y = a$, assume that its boundary ∂D is nonempty. Since an arbitrary basin of attraction in higher than one dimension may have a quite complicated boundary, for instance featuring a fractal structure, we should require some regularity constraints on ∂D. We delay this thought until Section 3.5, and restrict attention now to the scalar case $n = 1$. Here, the basin is simply an interval, and if the boundary is nonempty then it consists either of one or two isolated points. We will assume only the generically true property that the boundary points are hyperbolic.

Additionally, we let u decay to 0 in forward time, and desire sufficiently fast decay such that solutions to the nonautonomous ODE limit nicely in forward time to solutions of the base autonomous ODE. For simplicity we assume that u is eventually C^1 smooth with exponential decay; the reason is that with these conditions we may call upon a compactification technique developed in [27] to obtain the forward limiting behavior. Though it should be possible to relax the smoothness assumption, we leave this prospect as an open direction for future work.

Proposition 5. *Assume $n = 1$ and let $\hat{y}(t)$ be the solution with $\lim_{t \to -\infty} \hat{y}(t) = a$, whose existence and uniqueness were shown in Proposition 4. Assume that ∂D consists of either one or two hyperbolic unstable rest points. Assume that there exists a time T such that u restricted to (T, ∞) is C^1 smooth. Also assume that $\lim_{t \to \infty} u(t) = 0$ with exponential decay, meaning there exists a number ρ such that $\lim_{t \to \infty} \frac{\dot{u}(t)}{e^{-\rho t}}$ exists. Then $\hat{y}$ must exhibit exactly one of three long-term behaviors in forward time:*

- $\lim_{t \to \infty} \hat{y}(t) = a$
- $\lim_{t \to \infty} \hat{y}(t) \in \partial D$

- *$\hat{y}(t)$ escapes the closure $\overline{D}$ of the basin. That is, there exists a T within the maximal interval of existence of the solution $\hat{y}(t)$ such that for all $t > T$ where $\hat{y}(t)$ is defined, $\hat{y}(t) \notin \overline{D}$.*

Proof. We leave this proof as a brief sketch, and direct the reader toward the sources [27, 28] for full details on the compactification procedure. In the future limiting autonomous system $\dot{y} = f(y)$, the listed behaviors comprise the only 3 possible behaviors for any solution. The C^1-smooth exponential decay of u to zero allows the use of a compactification trick in forward time by "gluing on" the forward limiting autonomous system. This results in a smooth $(n+1)$-dimensional autonomous ODE where the hyperbolic basin boundary gains one stable time dimension but remains hyperbolic. (The exponential decay eliminates any pathological behaviors that might arise from compactifying.) This provides a correspondence between forward behaviors in the glued-on cross section and forward behaviors of the original nonautonomous solutions.

□

Definition 14. Assume $n = 1$ and let $\hat{y}(t)$ be the unique solution with $\lim_{t\to-\infty} \hat{y}(t) = a$. Out of the three possible forward behaviors from Proposition 5, if it is not the case that $\lim_{t\to\infty} \hat{y}(t) = a$, then we say $u(t)$ **induces tipping** in Equation (3.1). If $\lim_{t\to\infty} \hat{y}(t) \in \partial D$ we say that $u(t)$ is **critical**.

3.3.2 Tipping Induced by Translational External Force

Next, consider again the original rigidly shifting ODE

$$\dot{x} = f(x + \lambda(t)). \tag{3.2}$$

where we assume $\lambda(t) : \mathbb{R} \to \mathbb{R}^n$ is globally Lipschitz continuous and satisfies the asymptotic conditions

- $\lim_{t\to-\infty} \lambda(t) = 0.$
- $\lim_{t\to\infty} \lambda(t) = \lambda_\infty$ for a finite constant λ_∞.

Lipschitz continuity of λ implies absolute continuity of λ, which guarantees its almost-everywhere differentiability. The resulting measurable derivative $u = \dot{\lambda}$ is essentially bounded (actually, bounded) by the global Lipschitz constant of λ.

The transformation

$$y = x + \lambda(t) \tag{3.3}$$

is absolutely continuous with absolutely continuous inverse, establishing a one-to-one, absolutely continuous correspondence between solutions of Equation (3.2) and solutions of

$$\dot{y} = f(y) + \dot{\lambda}(t). \tag{3.4}$$

Hence we obtain existence and uniqueness of local solutions to Equation (3.2) in the same extended sense as before.

From the asymptotic conditions on λ it follows that $\lim_{t\to\pm\infty} u(t) = 0$. Add also now the assumption that there exists a T such that λ restricted to (T, ∞) is C^2 smooth and the derivative $u = \dot{\lambda}$ decays exponentially as $t \to \infty$, so that the conditions for Propositions 4 and 5 are satisfied. Then it is straightforward to check the following proposition:

Proposition 6. *The solution $\hat{y}$ from Proposition 4 of Equation* (3.1) *corresponds to a unique solution $\hat{x}$ of* (3.2) *such that $\lim_{t\to-\infty} \hat{x}(t) = a$, and the 3 cases from Proposition 5 correspond respectively to 3 long term behaviors for $\hat{x}$ in forward time:*

- $\lim_{t\to\infty} \hat{x}(t) = a - \lambda_\infty$
- $\lim_{t\to\infty} \hat{x}(t) \in \partial D - \lambda_\infty$
- $\hat{x}(t)$ *escapes* $\overline{D} - \lambda_\infty$.

where $S - \lambda_\infty$ denotes the set $\{s - \lambda_\infty \mid s \in S \subset \mathbb{R}^n\}$

Definition 15. We say λ **induces tipping** in Equation (3.2) if and only if $u = \dot{\lambda}$ induces tipping in Equation (3.1). And similarly we say λ is **critical** if and only if u is critical.

Remark 6. For an r-parameterized family of smooth scalar ramp functions λ, rate-induced tipping is typically defined via end-point tracking/non-tracking of a quasi-static equilibrium [24], or as a nonautonomous bifurcation of a pullback attractor that loses its forward attraction [29]. In our present application, we require no more than the simply stated definition above. It is equivalent to the quasi-static equilibrium and pullback attractor definitions found in the literature for the parameterized smooth case.

3.4 The Scalar Case

3.4.1 Monotone Ramp, One-Sided Basin Version

Let us initially restrict our attention to the case where D is half infinite. Without loss of generality, assume

$$D = (-\infty, \beta) \text{ and } \beta < \infty.$$

Additionally, let us initially assume that $\lambda(t)$ is monotone non-decreasing, hence

$$u(t) = \dot{\lambda}(t) \geq 0 \text{ wherever it is defined.}$$

Both these restrictions, that D be half infinite and that λ be monotone, are primarily for the sake of ease in the initial exposition. Afterward we explain how remove both these assumptions, which requires slight adjustment to the statement of the result.

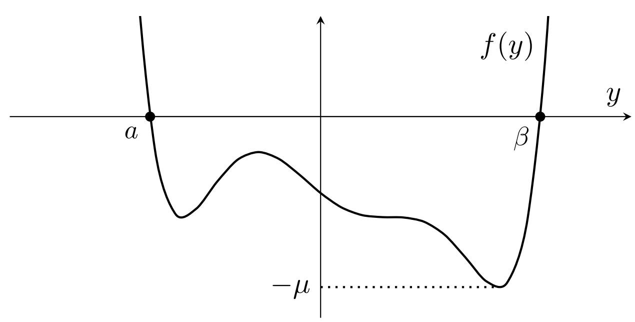


Figure 3.6: An arbitrary scalar vector field $f(y) \in C^2$ with an attracting rest point at $y = a$ whose basin of attraction is the half open interval $(-\infty, \beta)$. The minimum value of $f(y)$ on $[a, \beta]$ is denoted $-\mu$. In the vocabulary of Section 2.4, μ equals the intensity of attraction of the attractor at $y = a$.

Let $-\mu$ be the minimum value of f on $[a, \beta]$. For any constant $M > \mu$ consider the initial value problem

$$\dot{y} = f(y) + M$$
$$y(0) = a$$

Since the right hand side of the ODE is positive for all $y \in [a, \beta]$, the solution $y(t)$ of the initial value problem is strictly increasing there. Clearly a unique time $T_M > 0$ exists such that $y(T_M) = \beta$. Define the following bang-bang control function

$$B_M(t) = \begin{cases} 0 & \text{if } t < 0 \\ M & \text{if } 0 \leq t \leq T_M \\ 0 & \text{if } T_M < t \end{cases}. \tag{3.5}$$

By construction $B_M(t)$ is critical in the sense of Definition 14. That is, it steers the initial condition $y = a$ to an end state exactly balanced on the boundary $y = \beta$ of the basin of attraction.

We now introduce the optimization problem for which we will claim that $B_M(t)$ is an optimal solution.

Problem 1. Fix a constant $M > \mu$. Call a pair $(y(t), u(t))$ an **admissible pair** if y is absolutely continuous on $\mathbb{R}$, u is measurable on $\mathbb{R}$, and they solve the ODE $\dot{y} = f(y)+u(t)$ subject to the constraints:

- $\lim_{t\to-\infty} y(t) = a$,
- $\lim_{t\to\infty} y(t) = \beta$,
- $u(t) \in [0, M]$ for almost every t.

In this case, we also say u is an **admissible control** and y is the corresponding **admissible trajectory**. For an admissible pair (y, u), if it achieves a global minimum value of the integral

$$\int_{-\infty}^{\infty} u(t)dt$$

among all admissible pairs, then (y, u) is called an **optimal pair**. In this case, we also say u is an **optimal control** and y is the corresponding **optimal trajectory**.

Lemma 1. *The bang-bang function B_M defined in* (3.5) *is an optimal control for Problem 1, and every optimal control is equal to B_M almost everywhere and up to time translation.*

Proof. Note if an optimal control exists, the corresponding optimal state trajectory $y(t)$ must be strictly increasing during all times t such that $y(t) \in (a, \beta)$. Otherwise, there would exist $t_1 < t_2$ with $y(t_1) = y(t_2)$, $y(t) \in (a, \beta)$ for all $t \in [t_1, t_2]$, and $\dot{y}(t) \geq 0$ on a subset of positive measure of $[t_1, t_2]$. Because $f(y(t)) < 0$ for all $t \in [t_1, t_2]$ we must have $u(t) > 0$ on that same subset of positive measure of $[t_1, t_2]$. By definition $u(t) \geq 0$ everywhere. So by entirely excising the interval $[t_1, t_2)$ we produce a strictly lower cost admissible control.

Restricting to t such that $y(t) \in (a, \beta)$, we have that $y(t)$ is invertible and $\dot{y}(t) > 0$ for almost all t. We can now directly compute a lower bound on the integral

$$\begin{aligned}
&\int_{-\infty}^{\infty} u(t)\ dt \\
=&\int_{-\infty}^{\infty} \dot{y}(t) - f(y(t))\ dt \\
=&\int_{-\infty}^{\infty} \dot{y}(t)\ dt - \int_{-\infty}^{\infty} f(y(t))\ dt \\
=&(\beta - a) - \int_{-\infty}^{\infty} f(y(t))\ dt \\
=&(\beta - a) - \int_{a}^{\beta} \frac{f(y)}{\dot{y}}\ dy \\
=&(\beta - a) - \int_{a}^{\beta} \frac{f(y)}{f(y) + u(t(y))}\ dy \\
\geq&(\beta - a) - \int_{a}^{\beta} \frac{f(y)}{f(y) + M}\ dy
\end{aligned}$$

The latter inequality follows from the facts that $\dot{y} = f(y) + u > 0$ and $f(y) < 0$, thus

$$\begin{aligned} 0 &< f(y) + u \le f(y) + M \\ \implies \frac{1}{f(y)+u} &\ge \frac{1}{f(y)+M} \\ \implies \frac{-f(y)}{f(y)+u} &\ge \frac{-f(y)}{f(y)+M} \end{aligned}$$

This lower bound on the value of the cost integral is achieved exactly by the bang-bang control B_M, hence we conclude B_M is an optimal control. Any other optimal control must be equal to M at almost all times when $y \in (a, \beta)$, or its cost integral would not achieve the lower bound. Clearly the optimal strategy outside of this is to set $u = 0$ almost everywhere. Thus any optimal control is equal to B_M almost everywhere and up to time-translation. □

Lemma 2. *The integral of the bang-bang control function* (3.5), *which is*

$$\int_{-\infty}^{\infty} B_M(t)dt = M \cdot T_M,$$

is a strictly decreasing continuous function of M. Additionally, its limiting behavior satisfies

- $\lim\limits_{M\to\infty} M \cdot T_M = \beta - a,$
- $\lim\limits_{M\to\mu^+} M \cdot T_M = \infty.$

Proof. Take two different values $\mu < M_1 < M_2$ and compare the respectively associated bang-bang functions B_{M_1} and B_{M_2}. Set $M = M_2$ in the optimization problem (Problem 1), so that by Lemma 1 B_{M_2} is an optimal control and B_{M_1} is a strictly suboptimal control. This yields the strict decreasing order $\int_{-\infty}^{\infty} B_{M_1}(t)dt > \int_{-\infty}^{\infty} B_{M_2}(t)dt$.

Continuity of $M \cdot T_M$ follows if the switching time T_M when y arrives at the terminal state β is continuous in M. This follows from a well known property of globally continuous dependence on parameters for solutions to initial value problems with a globally Lipschitz vector field. Here, $\dot{y} = f(y) + M$ is locally Lipschitz on $\mathbb{R}$ thus globally Lipschitz on the compact set of interest $y \in [a, \beta]$.

For the first limit, recall T_M is defined so that $\beta - a = \int_0^{T_M} f(y(s)) + M\ ds$. Since $-\mu \leq f(y(s)) \leq 0$ we have

$$T_M(M-\mu) \leq \beta - a \leq MT_M$$
$$\implies -\mu T_M \leq (\beta - a) - MT_M \leq 0$$

and taking the limit as $M \to \infty$ gives $T_M \to 0$ and $(\beta - a) - MT_M \to 0$.

For the second limit, it suffices to show that $\lim_{M \to \mu^+} T_M = \infty$. Let

$$y_\mu = \min\{y \in [a, \beta] \ : \ f(y) = -\mu\}$$

be the first point at which $f(y)$ achieves its minimum value in $[a, \beta]$. Then y_μ is a rest point of the ODE $\dot{y} = f(y) + \mu$ such that the solution to the initial value problem $y(0) = a$ approaches y_μ as $t \to \infty$. By globally continuous dependence of solutions on parameters, we may choose $M \approx \mu$ such that the solution of $\dot{y} = f(y) + M,\ y(0) = a$ takes an arbitrarily large time to reach y_μ.

□

The final ingredient is the next lemma, which is modeled on the piecewise linear ramping example (Example 4) from the beginning of this chapter, except that it replaces the example base vector field with our arbitrary one $\dot{x} = f(x)$. Exactly the same tipping behavior still occurs, though of course the value of the critical slope m_c depends on the choice of f.

Lemma 3. *Fixing an f as before with attracting rest point at a and basin boundary β, and a constant $\lambda_\infty > \beta - a$, consider the following parameterized family of piecewise linear ramp functions with parameter $m > 0$:*

$$\lambda(mt) = \begin{cases} 0 & \text{if } t < 0 \\ mt & \text{if } 0 \leq t \leq \lambda_\infty/m \\ 0 & \text{if } t > \lambda_\infty/m \end{cases}.$$

There exists a unique solution $\hat{x}(t)$ of the ODE $\dot{x} = f(x + \lambda(mt))$ such that $\lim_{t \to -\infty} \hat{x}(t) =$

a, and there exists a unique critical parameter value $m = m_c$ *such that*

$$\begin{cases} \lim_{t\to\infty} \hat{x}(t) = a - \lambda_\infty & \textit{for } m < m_c \\ \lim_{t\to\infty} \hat{x}(t) = \beta - \lambda_\infty & \textit{for } m = m_c \\ \hat{x}(t) \textit{ escapes } \overline{D} - \lambda_\infty & \textit{for } m > m_c \end{cases}.$$

Furthermore, when considered as a function of λ_∞*,* m_c *is continuous and strictly decreasing and* $\lim_{\lambda_\infty \to \beta - a} m_c = \infty$*,* $\lim_{\lambda_\infty\to\infty} m_c = \mu$ *where* $-\mu$ *is the minimum value of* f *on* $[a, \beta]$*.*

Proof. Proposition 6 established the existence of the unique solution $\hat{x}(t)$ and fact that the three forward behaviors mentioned are the only possible forward behaviors for $\hat{x}(t)$. Under the co-moving change of coordinates, the ramp $\lambda(mt)$ is transformed into a control function $u(mt) = \dot{\lambda}(mt)$ which is a bang-bang style step function where the step has height m and width λ_∞/m. First of all, if $m \leq \mu$ then u certainly cannot induce tipping, thus the first option out of the three forward behaviors occurs when $m \leq \mu$. Now assuming $m > \mu$, compare $u(mt)$ to the critical bang-bang function $B_m(t)$ (3.5), whose step also has height m but has a possibly different width T_m. By definition of T_m, we see that $u(mt)$ induces tipping if and only if $T_m \leq \lambda_\infty/m$, with equality giving criticality. Rewriting slightly, $u(mt)$ induces tipping if and only if $mT_m \leq \lambda_\infty$.

By Lemma 2, mT_m is a continuous decreasing function of m with range $(\beta - a, \infty)$; thus it intersects the constant $\lambda_\infty > \beta - a$ exactly once, which gives $m = m_c$ with the desired behaviors on either side of m_c, as well as the strict decreasing property, the continuity, and the limiting behaviors.

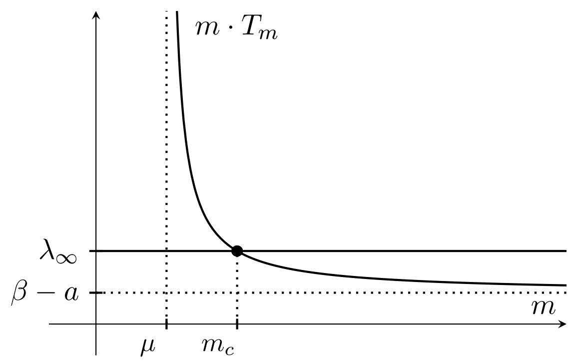


Figure 3.7: The continuous curve mT_m is strictly decreasing in m and approaches $\infty, \beta - a$ in the limits as $m \to \mu, \infty$, respectively. The curve crosses the constant $\lambda_\infty > \beta - a$ exactly once, giving the desired critical slope m_c. As $\lambda_\infty \to \beta - a$ we have $m_c \to \infty$; as $\lambda_\infty \to \infty$ we have $m_c \to \mu$.

□

Theorem 3 (Scalar, Monotone Ramp, One-Sided Basin Version)**.** *Assume that* $f : \mathbb{R} \to \mathbb{R}$ *is* C^2 *and the ODE* $\dot{x} = f(x)$ *has an attracting rest point at* $x = a$ *whose basin of attraction* D *is the half-infinite interval* $D = (-\infty, \beta)$ *where* $\beta > a$ *is a finite number, and* $x = \beta$ *is a hyperbolic unstable rest point of* $\dot{x} = f(x)$*. Fix a finite constant* $\lambda_\infty > \beta - a$*, and assume* $\lambda : \mathbb{R} \to \mathbb{R}$ *is globally Lipschitz continuous and monotone non-decreasing with* $\lim_{t \to -\infty} \lambda(t) = 0$*,* $\lim_{t \to \infty} \lambda(t) = \lambda_\infty$*. Assume there exists a* T *such that* λ *is* C^2 *when restricted to* (T, ∞)*, and there exists a number* ρ *such that* $\lim_{t \to \infty} \frac{\ddot{\lambda}(t)}{e^{-\rho t}}$ *exists. Let* $-\mu < 0$ *equal the minimum value of* f *on* $[a, \beta]$*. Then there exists a number* $m_c > \mu$ *such that if* $\lambda(t)$ *induces tipping in the ODE* $\dot{x} = f(x + \lambda(t))$ *then* $\dot{\lambda}(t) \geq m_c$ *at least once. Moreover, there exists a choice of* $\lambda(t)$ *satisfying the given conditions which does induce tipping with* $\max_t \dot{\lambda}(t) = m_c$*. Finally,* m_c *is continuous and strictly decreasing when viewed as a function of* λ_∞ *and satisfies* $\lim_{\lambda_\infty \to \beta - a} m_c = \infty$*,* $\lim_{\lambda_\infty \to \infty} m_c = \mu$*.*

Proof. Choose m_c as defined in Lemma 3. The critical piecewise linear ramp function that it corresponds to in Lemma 3 induces tipping with maximum slope m_c. To show that $\dot{\lambda}(t) \geq m_c$ at least once, a slightly different argument is used depending on whether $\dot{\lambda}(t)$ attains its supremum or not.

Case 1. $\dot{\lambda}(t)$ attains its supremum.

Suppose for contradiction that $\lambda(t)$ induces tipping but $\max_t \dot{\lambda}(t) = N < m_c$. Let $u = \dot{\lambda}$ and assume u is critical – otherwise, truncate it (and set equal to 0) on the right end while decreasing its integral; thus, $\int_{-\infty}^{\infty} u(t)\ dt \le \lambda_\infty$.

Note u is now an admissible control for the optimization problem (Problem 1) when the control constraint is $u \in [0, N]$. So by optimality of the bang-bang function B_N in Lemma 1, its integral satisfies

$$\begin{aligned}\int_{-\infty}^{\infty} B_N(t)\ dt &\le \int_{-\infty}^{\infty} u(t)\ dt \\ &\le \lambda_\infty\end{aligned}$$

But also

$$\begin{aligned}\int_{-\infty}^{\infty} B_N(t)\ dt &> \int_{-\infty}^{\infty} B_{m_c}(t)\ dt \text{ by Lemma 2} \\ &= \lambda_\infty \text{ by definition}\end{aligned}$$

so we have reached a contradiction.

Case 2. $\dot{\lambda}(t)$ does not attain its supremum.

Suppose for contradiction that $\lambda(t)$ induces tipping but $\sup_t \dot{\lambda}(t) \le m_c$. Let $u = \dot{\lambda}$ and assume u is critical – otherwise, truncate it (and set equal to 0) on the right end while decreasing its integral; thus, $\int_{-\infty}^{\infty} u(t)\ dt \le \lambda_\infty$.

Note u is now an admissible, but strictly suboptimal, control for the optimization problem (Problem 1) when the control constraint is $u \in [0, m_c]$. So by optimality of the bang-bang function B_{m_c} from Lemma 1, its integral satisfies

$$\begin{aligned}\int_{-\infty}^{\infty} B_{m_c}(t)\ dt &< \int_{-\infty}^{\infty} u(t)\ dt \\ &\le \lambda_\infty\end{aligned}$$

But also

$$\int_{-\infty}^{\infty} B_{m_c}(t)\ dt = \lambda_\infty \text{ by definition}$$

so we have reached a contradiction.

□

3.4.2 General Scalar Version

At this point, we discuss how to expand from the above special case to a general scalar version where assumptions of the monotonicity of the ramp and the one-sidedness of the basin are removed. In summary,

- Allowing a two-sided boundary for the basin of attraction. An optimal escape trajectory would traverse through only one side of the basin; it would not cross back over the attractor.
- Removing monotonicity. This requires a notable revision to our interpretation of the amplitude of the perturbation. By amplitude we no longer mean the forward limiting constant λ_∞ but instead

$$L = \int_{-\infty}^{\infty} |\dot{\lambda}(t)|\ dt,$$

  the total **arclength** of the perturbation, which is at least as large as λ_∞, with equality if λ is monotone.

Remark 7. Since a non-monotone λ may no longer bear a resemblance to a ramp, we refer to it as an **external forcing function** rather than a ramp function.

Theorem 4 (General Scalar Version)**.** *Assume that $f : \mathbb{R} \to \mathbb{R}$ is C^2 and the ODE $\dot{x} = f(x)$ has an attracting rest point at $x = a$ whose basin of attraction D has a boundary consisting of either one or two hyperbolic unstable rest points. Write $D = (\alpha, \beta)$, where $-\infty \leq \alpha < a < \beta \leq \infty$ and at least one of α, β is finite. Let R be the radius of the* *basin, that is $R = \min\{a - \alpha, \beta - a\}$. Let $\mu > 0$ equal the intensity of attraction of the attractor. Fix a constant $L > R$, and assume the external forcing function $\lambda : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous with $\lim\limits_{t\to-\infty} \lambda(t) = 0$, $\lim\limits_{t\to\infty} \lambda(t)$ finite, and $\int\limits_{-\infty}^{\infty} |\dot{\lambda}(t)|\, dt = L$. Assume there exists a T such that λ is C^2 when restricted to (T, ∞), and there exists a number ρ such that $\lim\limits_{t\to\infty} \frac{\ddot{\lambda}(t)}{e^{-\rho t}}$ exists. Then there exists a number $m_c > \mu$ such that if $\lambda(t)$ induces tipping in the ODE $\dot{x} = f(x + \lambda(t))$ then $|\dot{\lambda}(t)| \geq m_c$ at least once. Moreover, there exists a choice of $\lambda(t)$ satisfying the given conditions which does induce tipping with $\max_t \dot{\lambda}(t) = m_c$. Finally, m_c is continuous and strictly decreasing when viewed as a function of L and satisfies $\lim_{L\to R} m_c = \infty$, $\lim_{L\to\infty} m_c = \mu$.*

Remark 8. In this case a threshold speed can be defined similarly to before, but possibly one on each side of the basin, and m_c will be their minimum. Note that a minimum of two continuous functions is also continuous.

Proof. Instead of proving this directly here, we leave it as a corollary of the exposition in the next section. There, we present a conjectural generalization to n-dimensions (Conjecture 6), but prove the truth of the conjecture in 1 dimension. □

3.5 Toward N-Dimensions

In this section, we present a sketch toward obtaining a generalization in n-dimensions. We highlight the primary areas of difficulty that remain to be overcome, and leave them as avenues for future research. We present a conjecture for an n-dimensional result, but prove the truth of the conjecture in 1 dimension. The latter is equivalent to Theorem 4, whose proof we delayed, from the previous section.

3.5.1 Basin Regularity

Basins of attraction in n dimensions generally may have a complicated boundary structure, hence we expect to require some regularity conditions. The first task we leave for future work is that of identifying the appropriate conditions. A likely useful resource is [28], which contains a detailed discussion of the compactification procedure and its effect on objects called regular thresholds. These regular thresholds are codimension one, embedded, orientable manifolds which are normally hyperbolic and repelling, and are closely related to but not necessarily equal to the boundary of the basin of attraction if $n > 1$. What we desire is a sufficiently reasonable basin such that three possible forward behaviors analogous to Proposition 5 follow through. One difference to accommodate is that the basin boundary in one dimension consists purely of rest points so we could speak of a forward time limit toward it; here, we speak instead of an omega limit set.

Definition 16. The basin of attraction D is called **regular** if for every control u that is eventually smooth and decays exponentially to zero in the sense of Proposition 5 the unique solution $\hat{y}(t)$ with $\lim_{t \to -\infty} \hat{y}(t) = a$ as defined in Proposition 4 exhibits one of these three behaviors in forward time:

- $\lim_{t \to \infty} \hat{y}(t) = a$.
- $\omega(\{\hat{y}(t)\}) \subset K \subset \partial D$, where K is an attractor of the flow of $\dot{x} = f(x)$ restricted to ∂D.
- $\hat{y}(t)$ escapes the closure $\overline{D}$ of the basin.

3.5.2 Control Optimizability

Next we introduce an updated control optimization problem, where magnitude bars are inserted in the cost functional.

Problem 2. Let μ be the intensity of attraction (see Definition 13) of the attractor $y = a$. Fix a constant $M > \mu$. Call $(y(t), u(t))$ an **admissible pair** if y is absolutely continuous on $\mathbb{R}$, u is measurable on $\mathbb{R}$ and they solve the ODE $\dot{y} = f(y) + u(t)$ subject to the constraints:

- $\lim_{t \to -\infty} y(t) = a$,
- $\omega(\{\hat{y}(t)\}) \subset K \subset \partial D$, where K is an attractor of the flow of $\dot{x} = f(x)$ restricted to ∂D.
- $|u(t)| \leq M$ for almost every t.

In this case we also say u is an **admissible control** and y is the corresponding **admissible trajectory**. For an admissible pair (y, u) if it achieves a global minimum value of the integral

$$\int_{-\infty}^{\infty} |u(t)| dt$$

among all admissible pairs, then (y, u) is called an **optimal pair**. In this case we also say u is an **optimal control** and y is the corresponding **optimal trajectory**.

Definition 17. We say Problem 2 is **optimizable** if there exists an optimal control.

We leave for future work the identification of the most general conditions under which Problem 2 is optimizable. To this end, it should be possible to apply a Filippov-style theorem for the existence of optimal control. There is an extension of the Filippov Existence Theorem (first proven in the case of a finite time domain here [30]) to the infinite time domain, due to Cesari and Baum [31]. In the following lemma, we show how to use the Cesari and Baum result under strong assumptions.

Lemma 4. *If the basin D is bounded and every trajectory on ∂D limits to a hyperbolic rest point, then Problem 2 is optimizable. If $n = 1$ (with D not necessarily bounded) then Problem 2 is optimizable.*

Proof. In order to satisfy the conditions of the infinite time domain Filippov Existence Theorem (as stated in Appendix A), we desire a simplification in the statement of Problem 2. Specifically, we wish to speak of the future limit of an admissible trajectory, rather than its omega limit set. By assuming that every trajectory on ∂D limits to a hyperbolic rest point, we may replace the second constraint in the optimization problem with $\lim_{t\to\infty} y(t) \in \partial D$, instead of the original constraint $\omega(\{\hat{y}(t)\}) \subset K \subset \partial D$.

We also desire a compact space $S \subset \mathbb{R}^n$ and the assumption that $y(t) \in S$ for all t. The simplest case occurs when the basin of attraction D is bounded. If an optimal control exists, the corresponding state trajectory $y(t)$ must remain in the closure $\overline{D}$ of the basin of attraction, otherwise, it can be improved by truncation upon first arrival at the boundary; so restrict the trajectory space to the compact set $S = \overline{D}$. However, if the basin D is unbounded, then we need a way to further restrict to a compact subset of $\overline{D}$. When $n = 1$ it is clear that we can restrict to the compact interval between the attracting rest point a and the single boundary point of D, since an optimal trajectory would not make an excursion into the opposite of the basin.

Now we may verify that the compactness, closure, and convexity conditions in Filippov's Existence Theorem are satisfied. In particular, we check:

- The trajectory space S is compact.
- The control space $U = \{u \in \mathbb{R}^n : |u| \leq M\}$ is compact.
- The integrand $|u|$ in the cost functional is bounded below (by 0).
- The space $\{a\} \times \omega(\partial D)$ of limiting states for the trajectory is compact.
- For every $y \in S$, the set $Q(y) = \{(w, z) \mid w \geq |u|, z = f(y) + u, u \in U\}$ is closed and convex, and $Q(y)$ is upper semi-continuous when considered as a set-valued function of y.

For the latter, it can be straightforwardly verified that for any $(w_1, z_1), (w_2, z_2) \in Q(y)$ the line segment $(1 - t)(w_1, z_1) + t(w_2, z_2)$ is in $Q(y)$ for $t \in [0, 1]$. Upper semi-continuity follows from the continuity of $f(y)$. Additionally, we note that there exists at least one admissible control, because we chose M to be larger than the intensity of attraction. This suffices to conclude there exists an optimal control.

□

While the version of Cesari and Baum's Filippov Existence Theorem used here and stated in the Appendix requires that an admissible trajectory limits to a point, the original source [31] actually depends on a more general closure property of admissible controls. It is possible that the closure property stated there directly accommodates an omega limit set instead of limit condition in the optimization problem, so that the assumption of hyperbolic rest points being the only attractors of the flow restricted to the boundary can be removed, but this possibility is left for future inspection.

The other remaining difficulty lies in the case of an unbounded basin, where we must argue for each choice of M that the trajectory space may be restricted to a compact subset of $\overline{D}$. It should be possible to achieve this restriction via the idea that an optimal control would not take an extremely long route out of the basin; it would always escape through a fairly nearby portion of the basin boundary. A challenge to consider here is that solutions within the basin do not necessarily decay monotonically toward the attractor in n dimensions, so some additional requirement of a global decay bound may be in order.

3.5.3 Optimal Control Behaviors via Pontryagin's Maximum Principle

Next, recall that in the scalar case the simple topology of the basin allowed us to easily construct all optimal controls explicitly. In n dimensions we are not afforded such a luxury; however, assuming that optimal controls exist, it is not necessary to explicitly construct them in order to gain a usable descriptive understanding of what they must look like. The next lemma states that optimal controls are not necessarily bang-bang in magnitude, but their possible behaviors are limited. In particular, the only times during which an optimal control u is allowed a magnitude not equal to M or 0 is when u points perpendicular to the underlying vector field f.

Lemma 5. *Any optimal control is only allowed the following three possible behaviors at different times (with exceptions only on a set of measure 0).*

1. $|u(t)| = M$ *and* $u(t) \cdot f(y(t)) < 0$*, that is, the control has maximum allowable magnitude and points roughly opposite to the vector field.*

2. $u(t) = 0$.

3. $u(t) \neq 0$ *and* $u(t) \cdot f(y(t)) = 0$*, that is, the control points orthogonal to the underlying vector field. Here u might, in general, have any magnitude in* $(0, M]$.

Here it is convenient to state an additional property that must be satisfied whenever $u(t)$ *falls into category 3. In this circumstance, we have*

$$u^T \left(Df(y)\right) u = 0,$$

where $Df(y)$ *is the Jacobian matrix of* f.

Remark 9. The above is a necessary, not sufficient, description of optimal control.

Proof. Pontryagin's Maximum Principle (Appendix B) is a tool used for identifying all possible candidates for optimal control, though its standard setting involves a finite time interval. Taking an optimal trajectory $y(t)$ and any two finite times $t_1 < t_2$, fix the points $c = y(t_1)$, $d = y(t_2)$. We temporarily modify our optimization problem to the following: Minimize the cost functional

$$\int_0^T C(y, u, t)\ dt = \int_0^T |u(t)|\ dt \tag{3.6}$$

subject to the constraints

$$\begin{aligned} &|u(t)| \in [0, M] \\ &y(0) = c \\ &y(T) = d \\ &\dot{y} = f(y(t)) + u(t) \end{aligned} \tag{3.7}$$

where the terminal time $T > 0$ is finite but not fixed. For this finite-time subproblem, we call a choice of (y, u, T) with y absolutely continuous on $[0, T]$ and u measurable on $[0, T]$ admissible or optimal in an analogous way to before. Here we remark that a finite-time version of Filippov's Existence Theorem with weaker conditions guarantees existence of an optimal control for the subproblem (Appendix A).

Now to apply Pontryagin's Maximum Principle, introduce the constant variable p_0 and a so-called costate variable $p(t)$ and formulate the expression known as the control Hamiltonian,

$$\begin{aligned} H(y(t), u(t), p(t), t, p_0) &= -p_0 C(y, u, t) + p^T \dot{y} \\ &= -p_0 |u| + p^T (f(y) + u) \\ &= -p_0 |u| + p^T f(y) + p^T u \end{aligned} \tag{3.8}$$

Pontryagin's Maximum Principle states that if $(y(t), u(t), T)$ is optimal, then there also exist a p_0 and a continuous $p(t) : [0, T] \to \mathbb{R}^n$ such that, for almost all t, the

Hamiltonian satisfies the conservation property

$$H(y, u, p, t) = 0$$

and the maximum property

$$u \text{ maximizes } H \text{ when all other inputs are held constant.}$$

Additionally, either $p_0 = 0$ or $p_0 = 1$, and $(p_0, p(t)) \neq 0$ for all $t \in [0, T]$. Finally, the Euler-Lagrange equation or adjoint equation

$$-\dot{p}^T = p^T Df(y)$$

is satisfied almost everywhere. Since the magnitude and direction of u may be chosen independently, maximizing H means choosing u parallel to p, that is $u = k(t)p$ for a time-varying scalar $k(t) \geq 0$. Then we may write

$$\begin{aligned} H(y(t), u(t), p(t), t) &= -p_0|u| + p^T f(y) + |p||u| \\ &= |u|(|p| - p_0) + p^T f(y) \end{aligned} \tag{3.9}$$

Beginning with the case $p_0 = 0$, we have $p \neq 0$ and $H = |u||p| + p^T f(y) = 0$, which is linear in $|u|$ and maximimize by setting $|u| = M$. Then $p^T f(y) = -M|p| < 0 \implies u^T f(y) < 0$ almost everywhere. This falls into category 1 of the 3 behaviors.

Next, in the case $p_0 = 1$, we have

$$H = |u|(|p| - 1) + p^T f(y) = 0.$$

Since H is linear in the magnitude $|u|$, it is maximized by

$$|u| = \begin{cases} M & \text{if } |p| > 1 \\ 0 & \text{if } |p| < 1 \end{cases}.$$

If $p > 1$ then $p^T f(y) = -M(|p| - 1) < 0 \implies u^T f(y) < 0$ almost everywhere. This falls into category 1 of the three behaviors. The case $|p| < 1$ gives category 2.

It remains to investigate what happens if $|p| = 1$ on an interval. In this case we have $0 = H = p^T f(y) \implies 0 = u^T f(y)$ which gives category 3.

For the additional property of category 3, note p has constant magnitude so its velocity $\dot{p}$ satisfies $\dot{p}^T p = 0$. Taking the Euler-Lagrange equation $-\dot{p}^T = p^T(Df(y))$ and performing the inner product with p yields $-\dot{p}^T p = 0 = p^T(Df(y))p \implies 0 = u^T(Df(y))u$.

So far we have proven for the finite-time subproblem that only the 3 behaviors outlined in the lemma statement are possible for an optimal control. Now changing scope back to the infinite time problem, an optimal control in this context must be optimal on every finite time subproblem. Otherwise, it may be improved by switching a finite subtrajectory with a lower cost replacement. Hence it is also only allowed the same set of three possible behaviors. □

3.5.4 Regularity of Optimal Control

A key component of the argument used in the scalar case was the fact that every optimal control exhibits the maximum allowable control magnitude M. This yielded the strict decreasing order of optimal cost with respect to M: what was an optimal control becomes strictly suboptimal whenever M increases.

From the previous lemma we know that optimal controls have magnitude M whenever $u \cdot f < 0$. Intuitively, one would expect that to escape from the basin of an attractor, one must push roughly opposite to the underlying vector field for some time. However, in n dimensions it is possible to ascend Lyapunov level sets without using $u \cdot f < 0$.

We propose this as a regularity property of the optimization problem in the next definition, and leave open the question of how to prove it along with what additional conditions, if any, are required. Perhaps one may consider reachable sets from the attractor under controls bounded by fixed L^1 norm $I = \int_{-\infty}^{\infty} |u(t)| dt$ for every $0 \leq I$ (see Definition 22) and try to characterize optimal ascent up the nested boundaries of reachable sets, but reachable sets in n dimensions are generally difficult to describe.

We offer in the subsequent lemma one special case which we show is definitely sufficient to guarantee this regularity: when the attractor has negative reactivity, such as when the vector field is a gradient.

Definition 18. If Problem 2 is optimizable, and every optimal control exhibits the maximum magnitude $|u(t)| = M$ on a set of positive measure, then we say Problem 2 is **regular**.

Lemma 6. *Assume the attractor has negative reactivity, meaning that the symmetric part $A + A^T$ of the linearization $A = Df(y)(a)$ is negative definite (see Section 2.3). Note this assumption automatically holds when f is a gradient vector field, including when $n = 1$. Then out of the three possible behaviors stated in Lemma 5, every optimal control must exhibit behavior 1 on a set of positive measure. Thus Problem 2 is regular.*

Proof. The intuition is that we cannot even leave the attractor if we do not work against the vector field. Since its symmetric part $A + A^T$ is negative definite, the linearization A is also negative definite in the sense that

$$u^T A u < 0 \text{ for every } u \neq 0, \ u \in \mathbb{R}^n.$$

Since $Df(y)$ is continuous, there is a neighborhood N of the attractor such that

$$u^T Df(y) u < 0 \text{ for every } u \neq 0, \ u \in \mathbb{R}^n, \ y \in N.$$

Thus by the additional property stated at the end of Lemma 5, behavior 3 is ruled out for almost all t such that $y(t) \in N$. If $u = 0$ for almost all such t, then the trajectory clearly does not leave the attractor. So u must use behavior 1 on a set of positive measure.

□

Lemma 7. *Assuming Problem 2 is optimizable and regular, let u_M denote an optimal control. The integral $\int_{-\infty}^{\infty} |u_M(t)| dt$ is a strictly decreasing function of M.*

Proof. Take any two choices $M_1 < M_2$, and let u_{M_1}, u_{M_2} be respectively associated optimal controls. Now set $M = M_2$ so that u_{M_2} is an optimal control. By Definition 18 every optimal control has magnitude M_2 on a set of positive measure. Thus u_{M_1} is strictly suboptimal because its maximum magnitude is M_1. So $\int_{-\infty}^{\infty} |u_{M_1}(t)| dt > \int_{-\infty}^{\infty} |u_{M_2}(t)| dt$. □

3.5.5 Upper Semicontinuity of Optimal Cost

Next, in two or more dimensions we do not expect the optimal value of the cost integral to be continuous in M, because increasing the value of M can cause the optimal trajectory to jump suddenly to a new path. Rather, we conjecture that it is upper semicontinuous. For $n = 1$ we prove that it is continuous. We also conjecture limits analogous to those in Lemma 2 and prove them for $n = 1$.

Conjecture 5. *Denote by μ the intensity of attraction of the attractor a. Denote by $R = \inf\{|d-a| : d \in \partial D\}$ the radius of the basin of attraction. Assuming Problem 2 is optimizable and regular, let u_M denote an optimal control. The integral $J(M) = \int_{-\infty}^{\infty} |u_M(t)| dt$ is upper semicontinuous in M and satisfies $\lim_{M\to\infty} J(M) = R$ and $\lim_{M\to\mu^+} J(M) = \infty$.*

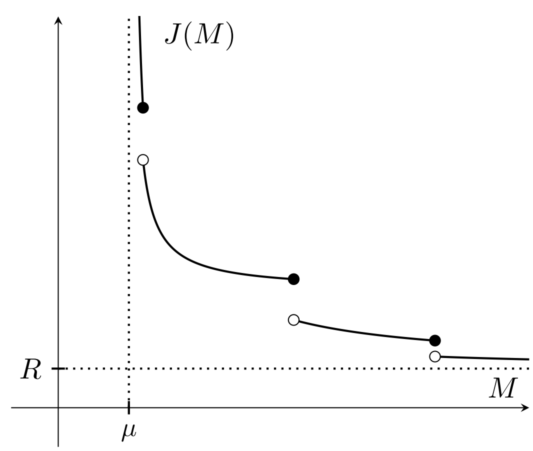


Figure 3.8: $J(M) = \int_{-\infty}^{\infty} |u_M(t)|\, dt$ is strictly decreasing. It is conjectured to be upper semicontinuous and to approach ∞, R as $M \to \mu, \infty$, respectively.

We briefly sketch a rough proof idea for Conjecture 5, but leave open the completion of details. $J(M)$ is the infimum of the optimal costs on every possible fixed simple path from the attractor to the boundary. Show that the optimal cost on each fixed path is continuous; then, taking the pointwise infimum of an arbitrary collection of continuous functions results in an upper semicontinuous function. Next, show that the optimal cost on a fixed path limits to the length of the path as $M \to \infty$. Then $\lim_{M\to\infty} J(M)$ equals the infimum of the lengths of the paths, which is radius R of the basin. For the other limit, show that the optimal cost on each fixed path limits to infinity as $M \to \gamma^+$ for some constant $\gamma(P) \geq \mu$ that depends on the choice of path P, and then show that $\mu = \inf_P\{\gamma(P)\}$.

Lemma 8. *If $n = 1$ then Conjecture 5 holds, but with the stronger claim that $J(M)$ is continuous in M.*

Proof. For $n = 1$ there are up to two possible escape paths and $J(M)$ is the pointwise minimum of two continuous functions, which is continuous. The argument for continuity and the limits over each of the up to two paths follows the same way as in Lemma 2. In particular, as $M \to \infty$ the optimal cost on each relevant path limits to the length of the path, so $\lim_{M\to\infty} J(M) = R$, the minimum path length. On the other hand, the optimal cost on each relevant path limits to infinity as M approaches the maximum magnitude μ_P of $|f|$ on that path, so $\lim_{M\to\min \mu_P} J(M) = \infty$. By Proposition 3, which describes the intensity of attraction for scalar systems, $\min \mu_P = \mu$. □

Remark 10. When $n = 1$, an optimal control is bang-bang, hence compactly supported. Therefore its integral is a piecewise linear function that belongs in the category of forcing functions $\lambda(t)$ with asymptotic constancy, eventual smoothness, and exponential approach to forward asymptotic constancy.

For $n > 1$ we know so far only that our optimal control is measurable, bounded, and L^1. A question to pursue here is whether the limits $\lim_{t\to\pm\infty} u(t)$ even exist, or what conditions would ensure they exist. We know for instance that when the attracting rest point is negatively reactive, an optimal control has compact support on the left. If the basin boundary is normally hyperbolic and repelling, does that imply compact support on the right?

3.5.6 Statement of Conjecture

Finally, we state our conjectural result that depends on the truth of Conjecture 5. Like Conjecture 5, it is sufficient to hold in the case of $n = 1$.

Conjecture 6 (N-Dimensional)**.** *Assume that $f : U \subset \mathbb{R}^n \to \mathbb{R}^n$ is C^2 and the ODE $\dot{x} = f(x)$ has an attracting rest point at $y = a$ whose basin of attraction D has nonempty* *boundary ∂D and is regular (Definition 16). Assume that Problem 2 is optimizable (Definition 17) and regular (Definition 18). Assume the external forcing function $\lambda : \mathbb{R} \to \mathbb{R}^n$ is globally Lipschitz continuous with $\lim_{t\to-\infty} \lambda(t) = 0$ and $\lim_{t\to\infty} \lambda(t)$ finite. Assume there exists a T such that λ is C^2 when restricted to (T, ∞), and there exists a number ρ such that $\lim_{t\to\infty} \frac{\ddot{\lambda}(t)}{e^{-\rho t}}$ exists. Let $R = \inf\{|d - a| : d \in \partial D\}$ be the radius of the basin of attraction. Fix a constant $L > R$ and assume $\int_{-\infty}^{\infty} |\dot{\lambda}(t)|\, dt = L$. Then there exists a threshold number m_c such that if $\lambda(t)$ induces tipping in the ODE $\dot{x} = f(x + \lambda(t))$ then $\dot{\lambda}(t)| \geq m_c$ at least once. m_c is upper semicontinuous (continuous if $n = 1$) and strictly decreasing when viewed as a function of L and satisfies $\lim_{M\to\infty} J(M) = R$ and $\lim_{M\to\mu} J(M) = \infty$. If $n = 1$ there exists a choice of $\lambda(t)$ satisfying the given conditions which does induce tipping with $\max_t \dot{\lambda}(t) = m_c$.*

Proof. We prove this under the assumption that Conjecture 5 is true.

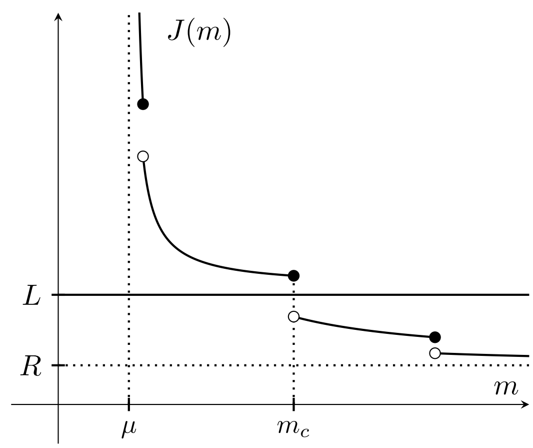


Figure 3.9: m_c is the maximum value of m such that $J(m) = \int_{-\infty}^{\infty} |u_m(t)|\, dt \geq L$.

Unlike in the scalar case, we do not give a constructive definition of m_c. Instead, let S be the set of numbers m such that an optimal control $u_m(t)$ for the optimization problem (Problem 2) when the control constraint is $u \in [0, m]$ satisfies $\int_{-\infty}^{\infty} |u_m(t)|\, dt \geq L$. By Conjecture 5 the set S is nonempty and bounded above. It has a maximum, due to the upper semicontinuity of the cost integral. Let $m_c = \max\{m : m \in S\}$.

To show that $\dot{\lambda}(t) \geq m_c$ at least once, a slightly different argument is used depending on whether $\dot{\lambda}(t)$ attains its supremum or not.

Case 1. $\dot{\lambda}(t)$ attains its supremum.

Suppose for contradiction that $\lambda(t)$ induces tipping but $\max_t \dot{\lambda}(t) = N < m_c$. Let $u = \dot{\lambda}$ and assume u is critical – otherwise, truncate it (and set equal to 0) on the right end while decreasing its integral; thus, $\int_{-\infty}^{\infty} |u(t)|\, dt \leq L$.

Note u is now an admissible control for the optimization problem (Problem 2) when the control constraint is $u \in [0, N]$. Also take an optimal control u_N for this problem. Comparing their L^1 norms,

$$\begin{aligned} \int_{-\infty}^{\infty} |u_N(t)|\, dt &\leq \int_{-\infty}^{\infty} |u(t)|\, dt \\ &\leq L \end{aligned}$$

But also

$$\begin{aligned} \int_{-\infty}^{\infty} |u_N(t)|\, dt &> \int_{-\infty}^{\infty} |u_{m_c}(t)|\, dt \text{ by Lemma 7} \\ &\geq L \text{ by definition} \end{aligned}$$

so we have reached a contradiction.

Case 2. $\dot{\lambda}(t)$ does not attain its supremum.

Suppose for contradiction that $\lambda(t)$ induces tipping but $\sup_t \dot{\lambda}(t) \leq m_c$. Let $u = \dot{\lambda}$ and assume u is critical – otherwise, truncate it (and set equal to 0) on the right end while decreasing its integral; thus, $\int_{-\infty}^{\infty} |u(t)|\, dt \leq L$.

Note u is now an admissible, but strictly suboptimal, control for the optimization problem (Problem 2) when the control constraint is $u \in [0, m_c]$. Comparing L^1

norms,

$$\int_{-\infty}^{\infty} |u_{m_c}(t)|\ dt < \int_{-\infty}^{\infty} |u(t)|\ dt$$
$$\leq L$$

But also

$$\int_{-\infty}^{\infty} |u_{m_c}(t)|\ dt = L \text{ by definition}$$

so we have reached a contradiction.

□

Chapter 4

Integral-Constrained Intensity of Attraction

The optimal control application to rate-induced tipping in the previous chapter inspires a new method of quantifying resilience, which we present in this brief chapter; it is a modification of intensity of attraction (Section 2.4), and we term it **integral-constrained intensity of attraction**. Unlike the original form of intensity of attraction, in which the total integral of the perturbation has no bound, this modification measures the resilience of the dynamical system under perturbations which are bounded in L^1 norm by a specific threshold number.

4.1 Definition via Bounded Integral-Constrained Control

As before, we consider the control system

$$\dot{x} = f(x) + u(t) \tag{4.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally Lipschitz. Now assume $u : \mathbb{R} \to \mathbb{R}^n$ with $u \in L^\infty \cap L^1$, fix a number $L > 0$ and assume that u satisfies

$$\int_{-\infty}^{\infty} |u(t)|\, dt \leq L$$

We refer to such controls u as **L-integral-constrained**. All the following definitions are exactly the same as those for intensity of attraction, with the only difference being this new constraint on the integral of the magnitude of u.

Definition 19. Denote by $\varphi_u(t, x_0) : D \subset \mathbb{R} \times U \to U$ the local flow defined by

$$\varphi_u(t, x_0) = x(t)$$

where $x(t)$ solves in the extended sense the ODE (4.1), with initial condition $x(0) = x_0$.

Families of r-magnitude and L-integral bounded control are defined as before:

Definition 20. Denote by $B_r^L \subset L^\infty[I, \mathbb{R}^n]$ the set of M-integral-constrained control functions whose essential supremum is bounded above by r:

$$B_r^L = \{u : ||u||_\infty < r\}$$

Next, we define reachable sets.

Definition 21. Consider $S \subset U$. The I-integral-constrained **reachable set** of S under r-bounded control is the set

$$R_r^L(S) = \bigcup_{u \in B_r^L} \bigcup_{x_0 \in S} \bigcup_{t \geq 0} \varphi_u(t, x_0)$$

And lastly we define L-integral constrained intensity of attraction.

Definition 22. If A is an attractor of $\dot{x} = f(x)$ and $D(A)$ the basin of attraction, then its **L-integral-constrained intensity of attraction** is

$$\mu^L(A) = \sup\{r \geq 0 \mid R_r^L(A) \subset D(A)\}$$

Remark 11. L-integral constrained intensity of attraction is a decreasing function of L and approaches the intensity of attraction as $L \to \infty$.

4.2 Relation to Externally Forced Tipping

Now we relate integral-constrained intensity of attraction to the core claims from the previous chapter.

Corollary 7 (Scalar)**.** *Take the same assumptions as in Theorem 4. Let* μ^L *equal the L-integral constrained intensity of attraction of the attractor. If* $\lambda(t)$ *induces tipping in the ODE* $\dot{x} = f(x + \lambda(t))$ *then* $|\dot{\lambda}(t)| \geq \mu^L$ *at least once.*

Conjecture 8 (N Dimensions)**.** *Take the same assumptions as in Conjecture 6. Let* μ^L *equal the L-integral constrained intensity of attraction of the attractor. If* $\lambda(t)$ *induces tipping in the ODE* $\dot{x} = f(x + \lambda(t))$ *then* $|\dot{\lambda}(t)| \geq \mu^L$ *at least once.*

Thus we posit integral-constrained intensity of attraction as the natural choice for quantifying resilience of a dynamical system to rigidly shifting externally-forced tipping, when the arclength of allowable external forcing functions is bounded by a fixed maximum.

4.3 An Ocean Circulation Box Model

We conclude this chapter by connecting integral-constrained intensity of attraction to a simple conceptual model for ocean circulation that originates from Cessi in [32].

The dynamics of global ocean circulation are driven by currents of mass exchange due to differences in water temperature and salinity (thermohaline circulation) between low-latitude and high-latitude regions. A classic Stommel two-box model [33] codifies these dynamics in a minimal way and gives rise to two alternate stable equilibria, which represent a strong and weak regime of ocean circulation. Indeed, recent research suggests that the Atlantic Meridional Ocean Circulation (AMOC) is weakening in response to anthropogenically driven climate change [34, 35] and warns of a possible future collapse of the AMOC [36].

4.3.1 Stommel Equations

We review the Stommel model (Figure 4.1), following the presentation in Cessi [32]. The model divides the ocean into a low-latitude warm water box and a high-latitude cold

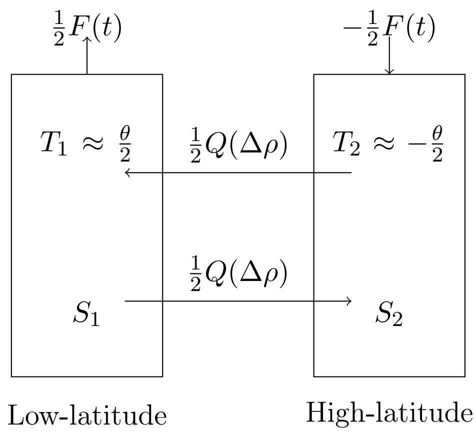


Figure 4.1: The Stommel two-box model of thermohaline circulation.

water box. T_i is the temperature of water, and similarly S_i is the salinity of water, in box i, for $i = 1, 2$. The density ρ_i of water is a function of T_i, S_i, and to a first order approximation may be written $\rho_i = \rho_0 + \alpha_S(S_i - S_0) - \alpha_T(T_i - T_0)$, where (ρ_0, T_0, S_0) is the point about which the linearization is taken. Then the density difference $\Delta\rho = \rho_1 - \rho_2$ between the two boxes is

$$\begin{aligned} \Delta\rho &= \alpha_S(S_1 - S_2) - \alpha_T(T_1 - T_2) \\ &= \alpha_S \Delta S - \alpha_T \Delta T \end{aligned}$$

We assume that the exchange of mass between the two boxes depends on the density difference and is given by some positive definite function $Q(\Delta\rho)$. Additionally, we assume that the temperatures relax linearly, at a fixed rate $1/t_r$, to prescribed values determined by a positive constant θ, while the salinities are forced by an external freshwater flux $F(t)$, and H is the height of each box. The evolution of temperature and salinity are governed by the equations

$$\begin{aligned} \dot{T}_1 &= -\frac{1}{t_r}\left(T_1 - \frac{\theta}{2}\right) - \frac{1}{2}Q(\Delta\rho)(T_1 - T_2) \\ \dot{T}_2 &= -\frac{1}{t_r}\left(T_2 + \frac{\theta}{2}\right) - \frac{1}{2}Q(\Delta\rho)(T_2 - T_1) \end{aligned}$$

$$\dot{S}_1 = \frac{F(t)}{2H}S_0 - \frac{1}{2}Q(\Delta\rho)(S_1 - S_2)$$
$$\dot{S}_2 = -\frac{F(t)}{2H}S_0 - \frac{1}{2}Q(\Delta\rho)(S_2 - S_1)$$

Subtracting pairs of equations gives the Stommel evolution equations for the salinity and temperature differences,

$$\frac{d}{dt}\Delta T = -\frac{1}{t_r}(\Delta T - \theta) - Q(\Delta\rho)\Delta T$$
$$\frac{d}{dt}\Delta S = \frac{F(t)}{H}S_0 - Q(\Delta\rho)\Delta S$$

4.3.2 Cessi Equation

Cessi proceeds from the Stommel equations by taking $Q(\Delta\rho) = \frac{1}{t_d} + \frac{1}{V}q(\Delta\rho)^2$ where V is the volume of each box, $1/t_d$ is a diffusion rate, and q is related to the transport of Poiseuille flow, then non-dimensionalizing by setting $x = \frac{\Delta T}{\theta}$, $y = \frac{\alpha_S \Delta S}{\alpha_T \theta}$, $t = t_d t$, thereby obtaining the system

$$\dot{x} = -\alpha(x-1) - x(1 + \beta(x-y)^2)$$
$$\dot{y} = p(t) - y(1 + \beta(x-y)^2)$$

where $\alpha = \frac{t_d}{t_r}$, $\beta = \frac{q t_d \alpha_T^2 \theta^2}{V}$, and $p(t) = \frac{\alpha_S S_0 t_d}{\alpha_T \theta H}F(t)$. Arguing that temperature relaxation occurs on a much faster timescale than any of the other processes, Cessi makes the simplifying assumption that the temperatures are fixed at the constants $\pm\theta/2$. This yields the one-dimensional equation

$$\dot{y} = -y(1 + \beta(1-y)^2) + p(t).$$

Writing $p(t) = \overline{p} + \tilde{p}(t)$ and $f(y) = -y(1 + \beta(1-y)^2) + \overline{p}$ puts this into the form

$$\dot{y} = f(y) + \tilde{p}(t). \tag{4.2}$$

Here $p(t)$ should be thought of as a nondimensionalized freshwater flux, and it is decomposed into a baseline constant value $\overline{p}$ plus a time-variable anomalous perturbation $\tilde{p}(t)$.

Cessi estimates $\bar{p} \approx 1.1$ and $\beta \approx 6.2$ as reasonable values for the North Atlantic.

The system (4.2) with no anomalous perturbation, $\tilde{p}(t) \equiv 0$, possesses two attracting states (Figure 4.2). The left attractor a corresponds to a large density difference and thus strong mass transport, meaning a pattern of strong ocean circulation. The right attractor c corresponds to a small density difference or pattern of weak ocean circulation.

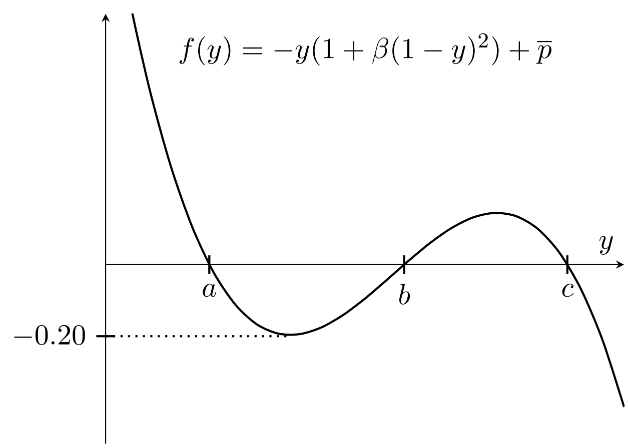


Figure 4.2: Cessi's ocean model $\dot{y} = f(y)$ with no flux perturbation, that is $\tilde{p}(t) \equiv 0$. Parameter values used are $\beta = 6.2$, $\overline{p} = 1.1$. Locations of equilibria are $a \approx 0.24$, $b \approx 0.69$, $c \approx 1.07$. The intensity of attraction of the left attractor $\{a\}$ is $\mu \approx 0.20$. The basin radius is $b - a \approx 0.45$.

4.3.3 Resilience to Freshwater Flux Perturbation

There exists a major debate over what mechanism of climate change is driving the currently observed weakening of the AMOC, but one explanation that has been offered is that of significantly increased freshwater flux, such as from the melting of Arctic sea ice [37].* In the Cessi model, let us consider the function $\tilde{p}(t)$ to represent the anomalous increase in freshwater flux arising from Arctic sea ice melt. We ask about the resilience of the strong circulation AMOC steady state $y = a$ to choices of $\tilde{p}(t)$.

*A competing explanation that we do not consider here is that of increased heat flux directly into the ocean surface, from overall increase in greenhouse gas concentration and global mean surface temperature. Additionally, we are neglecting ways other than polar ice melt in which hydrological responses to climate change may increase or affect freshwater flux.

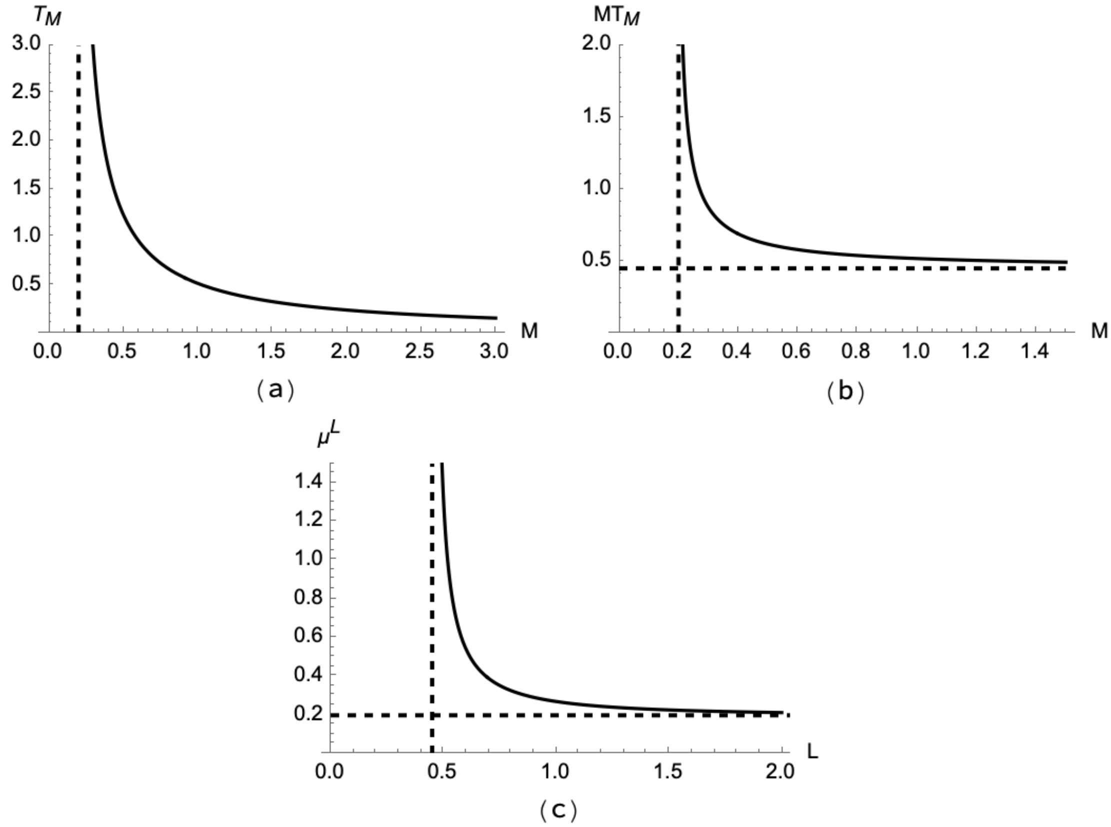


Figure 4.3: (a) The switching time T_M of a critical bang-bang perturbation $\tilde{p}_M(t)$ as a function of the magnitude $M > \mu \approx 0.2$ of control. (b) The L^1 norm $\int_{-\infty}^{\infty} \tilde{p}_M(t)\, dt = MT_M$ of a critical bang-bang perturbation as a function of the magnitude M of control. MT_M approaches the basin radius $R = b - a \approx 0.45$. (c) The integral-constrained intensity of attraction $\mu^L(\{a\})$ as a function of the constraint parameter $L \geq \int_{-\infty}^{\infty} \tilde{p}(t)\, dt$. The function plotted in (c) is the inverse of that in (b).

The L-integral constrained intensity of attraction of the attractor $\{a\}$ may be computed by finding all critical bang-bang perturbations

$$\tilde{p}_M(t) = \begin{cases} 0 & \text{if } t < 0 \\ M & \text{if } 0 \leq t \leq T_M \\ 0 & \text{if } T_M < t \end{cases} \tag{4.3}$$

where the switching time T_M is such that the initial condition $y(0) = a$ is steered exactly to $y(T_M) = b$.

We may solve for the value of the switching time T_M by integrating the ODE $\dot{y} = f(y) + M$ (Figure 4.3a). Next, multiplying by M gives the L^1 norm $M \cdot T_M$ of control (Figure 4.3b). Subsequently reflecting the curve gives the integral-constrained intensity $\mu^L(L)$ (Figure 4.3c).

4.3.4 Takeaway

To interpret the meaning of the integral-constrained intensity of attraction in this example, consider a situation where the total anomalous addition $\int_{-\infty}^{\infty} \tilde{p}(t)\, dt$ of freshwater is limited by a maximum L. For instance, L may equal the total (nondimensionalized) volume of meltwater trapped in the Arctic sea ice. Or, L may equal the volume of meltwater that would be released given a fixed number of hypothetical degrees of global warming.

Then μ^L prescribes a safe flux $\tilde{p}(t)$: if the sea ice melts at all times slowly enough that the anomalous freshwater flux remains below μ^L, the model ocean will not tip into the alternate basin. **The larger the total addition, the smaller the safe flux.** If ice melt exceeds the safety rate at some point, the model ocean could possibly, but may not necessarily, tip to the alternate regime of circulation. But if ice melt at all times exceeds the safety rate until the prescribed total melt is depleted, the model ocean will certainly tip to the alternate regime.

It should be emphasized that the model here is highly simplistic. Its significance lies not in any realistic predictive accuracy, but rather in simple illustrative power. The moral takeaway is akin to that of rate-induced tipping: **it is not only the cumulative total amount of a disturbance over time that matters but also the intensity with which that disturbance plays out at each point in time.**

Chapter 5

Conclusion

Motivated by the interplay between rate-induced tipping of rigidly shifting systems and the intensity of attraction framework for measuring resilience, we proved the existence of a critical lower bound on the speed $|\dot{\lambda}(t)|$ that must be reached at least once by any arbitrary forcing function $\lambda(t)$ inducing tipping in the nonautonomous scalar ODE $\dot{x} = f(x+\lambda(t))$. This threshold speed exhibits an inverse relationship with the arclength of forcing, illustrating the inherent trade-off between the total cumulative extent of a optimal forcing function and its maximal rate. Our approach is distinguished by a co-moving change of coordinates that transforms the optimal translational forcing function into a dual optimal additive control function. Inspired by the latter coordinate setting, we then modified the intensity of attraction framework to obtain a parameterized family of resilience values we called integral-constrained intensity of attraction. This integral-constrained construction codifies the idea that it is both the cumulative total norm (L^1) of control over time and the maximal control magnitude (L^∞) across pointwise times that matter in measuring the capacity for control to induce escape from the basin of attraction.

The types of externally-forced tipping considered here hold significance for contemporary physical, ecological, and social systems on all scales experiencing complex transient changes such as from modern anthropogenic practices. It is crucial to comprehend the resilience of systems in the face of disturbances that are simultaneously large and swift (as in R-tipping), or simultaneously large and intense (as in integral-constrained intensity). This work provides theoretical insight into a duality between the R-tipping and integral-constrained control settings, and highlights, in both cases, the interplay between two distinct axes of resilience. Although a dynamical system might individually endure considerable environmental disturbances in either *total cumulative extent* or in *pointwise severity*, there exists a trade-off where simultaneous exposure to excessive perturbation of both types surpasses the capacity of the system to retain its original regime.

References

[1] Katherine J. Meyer and Richard P. McGehee. Intensity—A Metric Approach to Quantifying Attractor Robustness in ODEs. *SIAM Journal on Applied Dynamical Systems*, 21(2):960–981, June 2022.

[2] Timothy M. Lenton, Hermann Held, Elmar Kriegler, Jim W. Hall, Wolfgang Lucht, Stefan Rahmstorf, and Hans Joachim Schellnhuber. Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences*, 105(6):1786–1793, February 2008.

[3] Vasilis Dakos, Marten Scheffer, Egbert H. van Nes, Victor Brovkin, Vladimir Petoukhov, and Hermann Held. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105(38):14308–14312, September 2008.

[4] Tobias S. Brett and Pejman Rohani. Dynamical footprints enable detection of disease emergence. *PLOS Biology*, 18(5):e3000697, May 2020.

[5] Marten Scheffer, Steve Carpenter, Jonathan A. Foley, Carl Folke, and Brian Walker. Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596, October 2001.

[6] S. R. Carpenter and W. A. Brock. Rising variance: A leading indicator of ecological transition. *Ecology Letters*, 9(3):311–318, 2006.

[7] Patrick E. McSharry, Leonard A. Smith, and Lionel Tarassenko. Prediction of epileptic seizures: Are nonlinear methods relevant? *Nature Medicine*, 9(3):241–242, March 2003.

[8] Jose G. Venegas, Tilo Winkler, Guido Musch, Marcos F. Vidal Melo, Dominick Layfield, Nora Tgavalekos, Alan J. Fischman, Ronald J. Callahan, Giacomo Bellani, and R. Scott Harris. Self-organized patchiness in asthma as a prelude to catastrophic shifts. *Nature*, 434(7034):777–782, April 2005.

[9] Hayette Gatfaoui and Philippe de Peretti. Flickering in Information Spreading Precedes Critical Transitions in Financial Markets. *Scientific Reports*, 9(1):5671, April 2019.

[10] Peter Ashwin, Sebastian Wieczorek, Renato Vitolo, and Peter Cox. Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962):1166–1184, March 2012.

[11] Paul D. L. Ritchie, Hassan Alkhayuon, Peter M. Cox, and Sebastian Wieczorek. Rate-induced tipping in natural and human systems. *Earth System Dynamics*, 14(3):669–683, June 2023.

[12] S. Wieczorek, P. Ashwin, C. M. Luke, and P. M. Cox. Excitability in ramped systems: The compost-bomb instability. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2129):1243–1269, November 2010.

[13] Steve Carpenter, Brian Walker, J. Marty Anderies, and Nick Abel. From Metaphor to Measurement: Resilience of What to What? *Ecosystems*, 4(8):765–781, 2001, 3659056.

[14] Michael G. Neubert and Hal Caswell. Alternatives to Resilience for Measuring the Responses of Ecological Systems to Perturbations. *Ecology*, 78(3):653–665, 1997.

[15] Katherine Meyer. *Metric Properties of Attractors for Vector Fields via Bounded, Nonautonomous Control*. PhD thesis, University of Minnesota, Twin Cities, May 2019.

[16] Katherine Meyer. A Mathematical Review of Resilience in Ecology. *Natural Resource Modeling*, 29(3):339–352, 2016.

[17] Stuart L. Pimm. The complexity and stability of ecosystems. *Nature*, 307(5949):321–326, January 1984.

[18] Robert May. *Stability and Complexity in Model Ecosystems*. Princeton University Press, 1974.

[19] C S Holling. Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, 4(1):1–23, 1973.

[20] S.L. Pimm and S.L. Pimm. *The Balance of Nature?: Ecological Issues in the Conservation of Species and Communities*. University of Chicago Press, 1991.

[21] Richard P McGehee. Some Metric Properties of Attractors with Applications to Computer Simulations of Dynamical Systems. 1988.

[22] Jack K. Hale. *Ordinary Differential Equations*. Number 21 in Pure and Applied Mathematics. Krieger, Malabar, Fla, 2. ed edition, 1980.

[23] S. R. Carpenter, D. Ludwig, and W. A. Brock. Management of Eutrophication for Lakes Subject to Potentially Irreversible Change. *Ecological Applications*, 9(3):751–771, 1999, 2641327.

[24] Peter Ashwin, Clare Perryman, and Sebastian Wieczorek. Parameter shifts for nonautonomous systems in low dimension: Bifurcation- and rate-induced tipping. *Nonlinearity*, 30(6):2185, April 2017.

[25] Clare G Perryman. How Fast is Too Fast? Rate-induced Bifurcations in Multiple Time-scale Systems. 2015.

[26] Zvi Artstein. Discrete and Continuous Bang-Bang and Facial Spaces Or: Look for the Extreme Points. *SIAM Review*, 22(2):172–185, April 1980.

[27] Sebastian Wieczorek, Chun Xie, and Chris K. R. T. Jones. Compactification for asymptotically autonomous dynamical systems: Theory, applications and invariant manifolds. *Nonlinearity*, 34(5):2970, May 2021.

[28] Sebastian Wieczorek, Chun Xie, and Peter Ashwin. Rate-induced tipping: Thresholds, edge states and connecting orbits. *Nonlinearity*, 36(6):3238, May 2023.

[29] Alanna Hoyer-Leitzel and Alice N. Nadeau. Rethinking the definition of rate-induced tipping. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(5):053133, May 2021.

[30] A. F. Filippov. On Certain Questions in the Theory of Optimal Control. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, 1(1):76–84, January 1962.

[31] R. F. Baum. Existence theorems for lagrange control problems with unbounded time domain. *Journal of Optimization Theory and Applications*, 19(1):89–116, May 1976.

[32] Paola Cessi. A Simple Box Model of Stochastically Forced Thermohaline Flow. *Journal of Physical Oceanography*, 24(9):1911–1920, September 1994.

[33] Henry Stommel. Thermohaline Convection with Two Stable Regimes of Flow. *Tellus*, 13(2):224–230, 1961.

[34] Stefan Rahmstorf, Jason E. Box, Georg Feulner, Michael E. Mann, Alexander Robinson, Scott Rutherford, and Erik J. Schaffernicht. Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, 5(5):475–480, May 2015.

[35] Chenyu Zhu, Zhengyu Liu, Shaoqing Zhang, and Lixin Wu. Likely accelerated weakening of Atlantic overturning circulation emerges in optimal salinity fingerprint. *Nature Communications*, 14(1):1245, March 2023.

[36] Peter Ditlevsen and Susanne Ditlevsen. Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nature Communications*, 14(1):4254, July 2023.

[37] Qin Wen, Chenyu Zhu, Deliang Chen, Mengyu Liu, Liang Ning, Mi Yan, Jian Liu, and Zhengyu Liu. Separating Direct Heat Flux Forcing and Freshwater Feedback on AMOC Change Under Global Warming. *Geophysical Research Letters*, 50(22):e2023GL105478, 2023.

[38] Lamberto Cesari. *Optimization—Theory and Applications.* Springer, New York, NY, 1983.

[39] Richard B. Vinter. Optimal Control and Pontryagin's Maximum Principle. In John Baillieul and Tariq Samad, editors, *Encyclopedia of Systems and Control*, pages 1–9. Springer, London, 2013.

[40] L. S. Pontryagin. *The Mathematical Theory of Optimal Processes*. Joint Publication Research Service ; JPRS-15089. Joint Publications Research Service, Arlington, Virginia, 1962.

Appendix A

Filippov's Existence Theorems

A.1 For Unbounded Time Domains

We reproduce here a somewhat simplified statement of Filippov's Existence Theorem for Unbounded Time Domains. For the original version in full generality, due to Baum and Cesari, and an explanation of why the simplification given here follows, see [31].

Theorem 9 (Filippov's Existence Theorem for Unbounded Time Domains)**.** *Consider the minimization of the cost functional*

$$J(t,x,u) = \int_{-\infty}^{\infty} C(t,x,u)\ dt,$$

subject to the constraints $\dot{x} = f(t,x,u)$, $(t,x,u) \in M$, $\lim_{t\to-\infty} x(t) \in B_1$, $\lim_{t\to\infty} x(t) \in B_2$.

Here, assume $A \subseteq \mathbb{R} \times \mathbb{R}^n$ *is closed, such that for any compact interval* $G \subset \mathbb{R}$ *the set* $(G \times \mathbb{R}^n) \cap A$ *is compact. Assume for every* $(t,x) \in A$ *there is a closed set* $U(t,x) \subseteq \mathbb{R}^m$ *such that when viewed as a set-valued function on* $\mathbb{R} \times \mathbb{R}^n$, $U(t,x)$ *is upper semicontinuous. Let* $M = \{(t,x,u) : (t,x) \in A, u \in U(t,x)\}$ *and assume* $C(t,x,u), f(t,x,u)$ *are continuous functions on* M*. Assume that for every* $(t,x) \in A$ *the set* $Q(t,x) = \{(w,z) : w \geq C(t,x,u), z = f(t,x,u), u \in U(t,x)\}$ *is convex and closed and* $Q(t,x)$ *is an upper-semicontinuous set-valued function. Also assume there exist Lebesgue-integrable functions* $\Phi(t) \geq 0$, $\psi(t) \geq 0$ *on* $\mathbb{R}$ *such that* $|f(t,x,u)| \leq \Phi(t)$ *and* $C(t,x,u) \geq \psi(t)$ *for all* $(t,x,u) \in M$*. Finally assume the set* $B = B_1 \times B_2 \subset \mathbb{R}^n \times \mathbb{R}^n$ *is compact.*

Any $x(t) : \mathbb{R} \to \mathbb{R}^n$ absolutely continuous with $u(t) : \mathbb{R} \to \mathbb{R}^m$ measurable satisfying the constraints is called an admissible pair. Then as long as there exists at least one admissible pair, the cost functional achieves a minimum value among admissible pairs.

A.2 For Finite Time Domains

We follow Cesari [38] (section 5.1) in stating the finite time domain version of Filippov's Existence Theorem. This theorem is originally attributed to Filippov [30].

Theorem 10 (Filippov's Existence Theorem for Finite Time Domains)**.** *Consider the minimization of the cost functional*

$$J(t,x,u) = \int_{t_1}^{t_2} C(t,x,u)\ dt,$$

subject to the constraints $\dot{x} = f(t,x,u)$, $(t,x,u) \in M$, $(t_1, x(t_1), t_2, x(t_2)) \in B \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$.

Here, assume $A \subseteq \mathbb{R} \times \mathbb{R}^n$ is compact and $U \subset \mathbb{R}^m$ is compact. Let $M = A \times U$ and assume $C(t,x,u), f(t,x,u)$ are continuous functions on M. Assume that for every $(t,x) \in A$ the set $Q(t,x) = \{(w,z) : w \geq C(t,x,u), z = f(t,x,u), u \in U\}$ is convex. Finally assume the set B is closed.

Any $x(t) : [t_1, t_2] \to \mathbb{R}^n$ absolutely continuous with $u(t) : [t_1, t_2] \to \mathbb{R}^m$ measurable satisfying the constraints is called an admissible pair. Then as long as there exists at least one admissible pair, the cost functional achieves a minimum value among admissible pairs.

Appendix B

Pontryagin's Maximum Principle

Following [39] (or see also the original source [40]) we state Pontryagin's Maximum Principle with free terminal time and fixed endpoints for minimizing the cost functional

$$J(T, x(t), u(t)) = \int_0^T C(t, x(t), u(t))\ dt + g(T, x(0), x(T))$$

A triple $(T, x(t), u(t))$ with $T \in \mathbb{R}$. measurable $u : [0, T] \to \mathbb{R}^m$ and absolutely continuous $x : [0, T] \to \mathbb{R}^n$ satisfying $\dot{x}(t) = f(t, x(t), u(t))$ almost everywhere, $u(t) \in U \subset \mathbb{R}^m$ almost everywhere, $x(0) = c \in \mathbb{R}^n$, $x(T) = d \in \mathbb{R}^n$ is called **admissible**. An admissible triple that achieves the minimum of $J(x(t), u(t))$ over all admissible triples is called **optimal**. Define the Hamiltonian function

$$H(t, x, p, u, p_0) = -p_0 C(t, x, u) + p^T f(t, x, u)$$

Theorem 11. *Let $(\bar{T}, \bar{x}(t), \bar{u}(t))$ be optimal. Assume the following hypotheses are satisfied:*

- *With $\tilde{f}(t, x, u) = (C(t, x, u), f(t, x, u))$, $\tilde{f}$ is continuously differentiable in x for every (t, u) and there exist $\epsilon > 0$, $k(t) \in L^1$ such that*

$$|\tilde{f}(t, x, u) - \tilde{f}(t, x', u)| \leq k(t)|x - x'|$$

for all $x, x' \in \mathbb{R}^n$ such that $|x - \bar{x}(t)| \leq \epsilon$ and $|x' - \bar{x}(t)| \leq \epsilon$, and $u \in U$ almost *everywhere,* $t \in [0, \bar{T}]$.

- U *is a Borel set.*

Then there exist a number $p_0 = 0$ *or* $p_0 = 1$ *and an absolutely continuous function* $p(t) : [0, \bar{T}] \to \mathbb{R}^n$ *with* $(p_0, p(t)) \neq 0$ *satisfying the conditions*

- $H(t, \bar{x}(t), p(t), \bar{u}(t)) = \max_{u \in U} H(t, \bar{x}(t), p(t), u(t))$ *almost everywhere (Maximization of the Hamiltonian)*

- $H(t, \bar{x}(t), p(t), \bar{u}(t)) = c$ *for some constant* $c \in \mathbb{R}$ *almost everywhere (Constancy of the Hamiltonian)*

- $c = p_0 \dfrac{\partial}{\partial T} g(\bar{T}, \bar{x}(0), \bar{x}(T))$ *(Transversality Condition)*

- $-\dot{p}^T = -p_0 \dfrac{\partial}{\partial x} C^T(t, \bar{x}(t), \bar{u}(t)) + p^T(t) \dfrac{\partial}{\partial x} f(t, \bar{x}(t), \bar{u}(t))$ *(Adjoint Equation)*