FIXED POINTS IN ANALOG NETWORK MODELS

Existence, Continuity, and Computability of Unique Fixed Points in Analog Network Models

By Nick D. James, B.Sc., M.A., M.Sc.

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements for the Degree
Doctor of Philosophy

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE (2012)
McMaster University, Hamilton, Ontario

TITLE: Existence, Continuity, and Computability of Unique Fixed Points in Analog Network Models

AUTHOR: Nick D. James, B.Sc. (University of Guelph), M.A. (University of Western Ontario), M.Sc. (University of Western Ontario)

SUPERVISOR: Jeffery I. Zucker

NUMBER OF PAGES: xii, 138

Abstract

The thesis consists of three research projects concerning mathematical models for analog computers, originally developed by John Tucker and Jeff Zucker. The models are capable of representing systems that essentially "diverge," exhibiting no valid behaviour—much the way that digital computers are capable of running programs that never halt. While there is no solution to the general Halting Problem, there are certainly theorems that identify large collections of instances that are guaranteed to halt. For example, if we use a simplified language featuring only assignment, branching, algebraic operations, and loops whose bounds must be fixed in advance (i.e. at "compile time"), we know that all instances expressible in this language will halt.

In this spirit, one of the major objectives of all three thesis projects is identify a large class of instances of analog computation (analog computer + input) that are guaranteed to "converge." In our semantic models, this convergence is assured if a certain operator (representing the computer and its input) has a unique fixed point. The first project is based on an original fixed point construction, while the second and third projects are based on Tucker and Zucker's construction. The second project narrows the scope of the model to a special case in order to concretely identify a class of operators with well-behaved fixed points, and considers some applications. The third project goes the opposite way: widening the scope of the model in order to generalize it.

Acknowledgements

Dedicated[1] *to my lovely, compassionate, and brilliant wife, Rebecca.*

I owe a great deal to my committee and to my family for giving me the tools I needed to do this research.

My exemplary supervisor, Prof. Jeff Zucker, invited me to harvest some low-hanging fruit in a grove that he and his colleague Prof. John Tucker had been cultivating for years. I am fortunate indeed to have enjoyed the privilege of being his student. He always gave me the strongest encouragement, guidance, and support possible. He had a particular knack for showing me that I had done good work and made significant progress even when I couldn't see it at all and was feeling that I had accomplished nothing.

I am also extremely lucky to have had Prof. Jacques Carette on my advisory committee. Prof. Carette went so far above and beyond the role of a committee member that he might have been more accurately described as a co–supervisor. As long as it took me to complete this research, I fear I might never have finished at all if it hadn't been for his input. We engaged in several rather lengthy discussions about my research, he suggested papers and lines of inquiry, warned me when he felt I was heading in a perilous direction, and he furnished me with some fascinating examples and proofs, specifically tailored to my questions. I am disappointed that I was unable to include more of his contributions and that I was unable to implement more of his creative ideas.

Rebecca, I simply cannot thank emphatically enough for her love and support, and for her implacable faith in my ability. I cannot imagine life without her.

[1] *…because nothing says "I love you" like a hundred and thirty-eight pages of tedious and coldly impersonal mathematics.*

Contents

List of Figures

List of Symbols

$\mathbb{N}$	The set of natural numbers $\{0, 1, 2, \ldots\}$
$\mathbb{Z}$	The set of integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{Z}^+$	The set of positive integers $\{1, 2, \ldots\}$
$\mathbb{R}$	The set of real numbers
$\mathbb{R}^+$	The set of positive real numbers
$\mathbb{R}^{\geq 0}$	The set of nonnegative real numbers
$\mathbb{T}$	A metric space representing time (an alias for $\mathbb{R}^{\geq 0}$)
$\mathcal{A}$	A metric space representing measurable data to be used in computation
$\mathcal{C}[\mathbb{T}, \mathcal{A}]$	A space of data streams: continuous functions from $\mathbb{T}$ into $\mathcal{A}$
$\mathbb{X}$	A σ-compact space representing an abstract "data index" space (e.g. spacetime)
$\mathcal{C}[\mathbb{X}, \mathcal{A}]$	A space of data "smoothies": continuous functions from $\mathbb{X}$ into $\mathcal{A}$

Declaration of Academic Achievement

Sections 1.3 and Appendix B are mostly from other sources

Almost everything in Sections 1.3 on page 10 and Appendix B on page 133 is taken from other sources but written in my own words. Most of it is from [TZ11], some of it from my old textbooks, and some of it is technically "mine," except that it's so obvious and elementary that I'm certain it exists somewhere else and I just abandoned the search for a source. There are only a few exceptions:

- All the proofs are mine (although, again, it's quite possible they have been proved elsewhere the same way I did them).
- All the figures are mine.
- Remark 1.3.19 on page 17 is my own observation.

The core ideas in Section 4.7 are from [TZ12], but the exposition is radically different and it includes a few adaptions of my own. In some respects, it is a simplified summary of half of that paper, generalized for a larger context.

Other sections are mostly original

Everything in the rest of the thesis is mine (to the best of my knowledge) unless otherwise noted, with the following exceptions:

- Jeff Zucker was responsible for numerous improvements in the grammar and exposition throughout the thesis.
- The comment in the in Section 1.1 about Hadamard's Principle (starting from "the significance of the latter question" to the end of the paragraph) was written by Jeff Zucker.

- Figure on page 67 was provided by Jeff Zucker.
- The discussion preceding Example 3.3.2 was written by me, but requested by an anonymous reviewer.
- Definition 4.7.2 was taken almost directly from [TZ12].
- I am assuming the following lemmas must surely exist in some other source, but I abandoned the search for references in these cases, despite the conviction that I was reinventing the wheel:
  - Lemma 1.4.1 on page 21.
  - Lemma 1.4.13 on page 27.
  - The Equicontinuity Lemma on page 63.

All other non-original content (e.g. the ODEs for the pendulum and the mass-spring-damper system, Banach's Fixed Point Theorem, the definition of a σ-compact space, etc.) has been attributed explicitly, or was felt to be common enough to not require attribution.

Finally, there are some things for which I deserve only partial credit:

- The Cauchy sequence in Example 4.1.5 is mine, but the sequence of nested sets upon which it is based was taken from [TZ11].
- The definition of $\boldsymbol{Contr}(\lambda, \tau)$ is definitely not mine (although I wish it were!), but it was invented by Jeff Zucker to solve a problem I spotted in an early draft of [TZ11]. So I pride myself on the tiny modicum of credit I deserve for its existence.
- The condition that $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ be locally bounded in Theorem TZ2 on page 19 was my suggestion. They were originally required to be continuous.
- The proofs of my two major theorems (Generalized TZ1 on page 94, and Generalized TZJ2 on page 102) were heavily inspired by the proofs of the two major theorems in [TZ11]. They've been thoroughly bent out of shape, but a close reading will reveal many common steps and tricks.

The originality of Chapter 2 is somewhat unknown

There is also one outstanding uncertainty: the originality of the vanishing delay construction upon which Chapter 2 is based. That was an idea which occurred to me independent of any other sources, but I have had several such "original" ideas over the course of my academic career. They seem exciting and wonderful until I discover that someone else already thought of and published the idea (and in some cases, quite a while ago).

To ensure this wasn't the case for the vanishing delay construction, I conducted literature searches and met with two professors in the Faculty of Engineering at the University of McMaster whom my committee felt would be most likely to know about it: Dr. Mark Lawford, from the Department of Software and Engineering and Dr. Shahin Sirouspour from the Department of Electrical and Computer Engineering. Neither of these researchers were remotely familiar with the idea.

As for the literature searches, they revealed that the term "vanishing delay" has often been applied to delay differential equations which involve a varying delay that may reach zero or approach it arbitrarily closely—which appears to be quite unrelated to the content of Chapter 2. Consequently, my supervisor and I deemed it safe for me to continue, although I still suspect someone has surely attempted this sort of thing before and I simply haven't found documentation of the attempt.

Chapter 1

Introduction

1.1 Analog Computation and Analog Networks

Analog computation concerns computation on continua rather than on discrete spaces. Where digital computation uses an abstract, symbolic encoding of data and explicitly written algorithms to operate upon them, analog computation uses—as its name would suggest—an analogy or transduction of measured data and a corresponding physical system which serves as a model of the original system, i.e. the system about which we wish to reason or make predictions.. The input data can be any sort of measurement (e.g. voltage, pressure, temperature, etc.) from the world outside the model, and it can be represented by any measurable quantity that is within the model. The model is set up to mimic the initial conditions of the original system, and then set in motion and observed. The "language" of analog computation comes directly from the laws of physics rather than from the minds of instruction set engineers and programming language designers.

Admittedly, digital computation often involves analogies as well. An array of bits in a digital computer might be used, for example, to directly represent the status of a series of locks in a canal. Metaphors for data structures, algorithms, and programming language constructs like binary "trees," "simulated annealing," and "inheritance" permeate the literature on digital computation. Hence, we might alternatively dichotomize computation into "algorithmic" and "non-algorithmic" paradigms, but the term, "analog computation" is already well-established and the notion of analogy is *inherent* in it (both in its representation of data and in its actual mechanisms of computation), while it appears only incidentally in digital computation, and often for only didactic purposes.

Putting aside such devices as the Antikythera Mechanism [F^+06], slide rules, planimeters, and similar devices used to compute individual values (we might call them analog "calculators" rather than "computers"), likely the first recorded account of analog computation was written in 1836 by Gaspard-Gustave Coriolis [Cor36], in which he described using gears and cylinders to integrate first-order differential equations. These ideas were further developed (or perhaps reinvented) in 1876 to tackle differential equations of arbitrary order by Lord Kelvin and his brother, James Thomson [Tho76]. While Kelvin and Thomson's ideas were implemented to some extent in the "Argo" fire control system used by the Royal Navy [Pol80], it was Vannevar Bush who designed and built what is likely the most advanced mechanical analog computer and one of the most famous and practical computers of its day: the differential analyzer.

Claude Shannon, working as a research assistant in Bush's lab, defined a mathematical model of the differential analyzer and named it the "General Purpose Analog Computer" (or "GPAC") in [Sha41]. The GPAC is an example of what could more generally be called an *analog network*, which may be visualized as a circuit: a directed graph in which the nodes are processing elements known as "modules" and the edges (known as "channels") act as wires or tubes to convey data streams (which are functions of time).

The network is merely a conceptual model, however, and is not intended to describe the actual appearance of the system. An electronic or hydraulic implementation of an analog network might physically resemble the directed graph itself, while a mechanical implementation often wouldn't. A module to perform scalar multiplication, for example, could be implemented as a step-up transformer or a transistor amplifier in an electronic circuit (both of which commonly appear in schematics), whereas the same module could be implemented mechanically as the physical interface between the teeth of two cogs of differing diameters (which does not so neatly suggest a node in a schematic). Hence, a physical system that bears no apparent resemblance to a network at all, may still qualify in our vernacular as an "analog network."

One of the main purposes of defining such a model is to determine the set of functions it is capable of generating, for if some physical device can reliably generate a particular function, it follows that this function is "computable" in the plainest and most intuitive sense of the word. Shannon proved[1] that the GPAC is capable of generating all and only the differentially algebraic functions. This is a very large class of functions, including polynomials of one real variable along with sinusoids, exponential functions, and solutions of ordinary differential equations consisting of these functions. It is not, however, without some disappointing limitations—Shannon's poster child being the well-known gamma function, which is not differentially algebraic.

[1] There were some problems with his proof which Marian Pour-El addressed and attempted to rectify in [PE74] using an alternative GPAC model. Unfortunately, there were also problems in her own approach which were spotted and corrected by Daniel S. Graça and José Félix Costa in [GC03] using a third GPAC model.

Partially inspired by these limitations and partially by the assumption that the brain is a type of analog computer which is known to perform spatial as well as temporal integration, Lee Rubel defined the "Extended Analog Computer" (or EAC) in [Rub93]. Rubel's EAC is theoretically capable of solving boundary value problems for partial differential equations, whereas the GPAC is limited (according to Shannon's definition) to initial value problems of ordinary differential equations. Jonathan Mills ran with Rubel's model, creating fully-functional analog computers inspired by the EAC from foam sheets typically used as packaging material and even blocks of salted gelatin [MPH+06]. There have been other implementations of analog computation that represent an even more profound departure from the GPAC model. Slime mold [YMTK95] and bees [LCR10] have been used to solve small instances of the Travelling Salesperson Problem and generate near-optimal solutions to larger instances.

While models of analog computation offer one approach for investigating the computability of functions involving continua, there has been a parallel research effort focused on extending classical computability theory (as defined by Turing, Church, Kleene, etc.) into this realm: *computable analysis*. Pioneered primarily by Andrzej Grzegorczyk [Grz55, Grz57] and Daniel Lacombe [Lac55], computable analysis puts real (and complex) analysis, functional analysis, and numerical analysis under the microscope of classical computability theory and asks the question central to most research on analog computation: which functions are computable? We already have a clear answer to that question in the domain of classical computability theory (i.e. for functions of the form $f : \mathbb{N} \to \mathbb{N}$), as all of the models of digital computation we've discovered so far are in agreement. This is, of course, the foundation for the famous Church-Turing Thesis.

Computability theory on continua has not yet reached the same degree of consensus, but much progress is being made. Olivier Bournez et al. showed in [BCGH06] that the GPAC is equivalent to Ker-I Ko's model of computability [Ko91] as long as the GPAC is permitted to *approximate* functions (to arbitrary precision) rather than produce them in real time. Viggo Stoltenberg-Hansen and John Tucker used domain representability in [SHT99] to prove that five different models of computation on topological algebras are equivalent (under some modest conditions). Further equivalence results (and exceptions) can be found in [Wei00]. The matter is still not entirely settled, so the question of computability pertaining to functions with uncountable domains or codomains remains open for now.

In [TZ07, TZ11], John Tucker and Jeff Zucker turn this question around and ask instead, given a particular analog network, under what conditions does it produce meaningful output, and under what conditions does this output vary continuously with the network's parameters? They argue[2] that the significance of the latter question is grounded in the imperative of experimental physics known as "Hadamard's Principle," first articulated by Hadamard [Had52] and later refined by Courant and Hilbert [CH53]. Its fundamental tenet is that for the solution to a problem in physics to be practically applicable, it must vary continuously with the parameters of the system so that small discrepancies or inaccuracies in the input produce only small variations in the output. The stability of measurements in the presence of noise is an essential feature for a physical system to qualify as an analog computer.

Like the GPAC, the data streams carried by the analog networks in [TZ07, TZ11] are functions of time. There are, of course, various ways of modelling time. The debate over whether spacetime is continuous, discrete, or even both simultaneously (see [Kem10]) is ongoing, but regardless of the outcome of that debate, the majority of our physical laws and theories treat measurable quantities (including time) as real numbers. This may suggest using the whole real line as a model of time, but regardless of the duration a computer is allowed to run while solving a problem it must at some point actually be built, initialized, and started. For this reason, the authors chose to represent time using the only the nonnegative real numbers (as we do here, up until Chapter 4, at which point several possible representations of time become merely special cases in a broader theory).

1.2 Chapter Summary

Chapter 1: Preliminary Concepts

The three research projects share a common foundation, rooted in [TZ11]. Briefly, we take $\mathbb{T}$ to be the nonnegative reals, which will represent time, and $\mathcal{A}$ to be a metric space which represents a physically measurable quantity (e.g. voltage, position, pressure, etc.) or a collection of physically measurable quantities. Our fundamental "object space" is $\mathcal{C}[\mathbb{T}, \mathcal{A}]$, which is the space of total continuous functions from $\mathbb{T}$ into $\mathcal{A}$, equipped with a metric topology. We call this "stream space," and the elements within it, "streams."

The model of computation in the first two projects concerns operators on $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ which represent physical systems to be used as computers. The semantics of the model are given by the existence of unique[3] fixed points for

[2]Note that I don't fully agree with Hadamard's Principle, as I explain in Appendix A.

[3]Technically, the fixed points need only be *distinguishable* to provide such semantics. That is, the model would still work even with a whole set of fixed points, provided there exists a selection function (e.g. *least* fixed point) with nice properties to provide the uniqueness. This would represent a generalization of the theory suitable for future work.

these operators. Two theorems from [TZ11] are presented here: the first provides a set of sufficient conditions for the existence of a unique fixed point, while the second provides conditions to ensure this fixed point varies continuously with the parameters and input streams. The first theorem is proved constructively in [TZ11], and the construction is imperative for most of the subsequent results. So I reproduce it in Chapter 1 for reference, albeit using different notation and slightly different methods, but keeping the spirit of the construction the same.

There are two operator properties of particular importance to the theorems: causality and contraction. Loosely speaking, a "causal" operator does not depend on the future and a "contracting" operator brings streams closer together (but only locally; this is somewhat different from the usual sense of contraction, as used in analysis). Causality is a basic requirement of the theory, without which we couldn't get off the ground at all, while contraction does most of the heavy lifting. This is not contraction in the usual sense, but rather a domain-restricted, conditional version of contraction. The properties are presented, along with a third which is essential to [TZ11], but less important here: shift invariance.

Up until Section 1.4, I stick very closely the original source material in order to better set it apart from my own work. After that section, I introduce some modest generalizations of the theory and some further preliminary results I need to use later.

Chapter 2: Constructing Fixed Points of Stream Operators Using Vanishing Delays

This chapter covers the work I did on my original project, which I felt I had to abandon because I had gone several months without making any progress. Recall that the model of analog computation upon which this thesis is based concerns fixed points of stream operators (aside from Chapter 4, in which I depart from streams). Some have fixed points, some don't.

If we compose any such operator with a delay, however (creating a delayed version of the original operator), this new operator is guaranteed to have a unique fixed point, and one that can even be constructed quite mechanically. So the idea explored in this chapter is to see what happens when we compose an operator with a delay, find the fixed point of the delayed operator (as a function of the delay duration), and then let that delay approach zero.

Intuitively, we expect that the fixed point of the delayed operator will converge to the fixed point of the original operator, if one exists, and that it will diverge otherwise.

Chapter Highlights:

1. The Delayed Operator Theorem on page 45: a delayed operator that satisfies a certain causality condition always has a unique fixed point.

2. The Vanishing Delay Theorem on page 47: if that (parametrized) fixed point converges to a stream as the delay approaches zero, the stream to which it converges is a fixed point of the original, non-delayed operator.

3. Theorem TZJ1 for Vanishing Delays on page 48: if a continuous operator satisfies Tucker and Zucker's sufficiency conditions for having a unique fixed point, then my technique of vanishing delays will converge to it (loosely speaking, if their construction works, so will mine—at least in the case of continuous operators).

Chapter 3: *A Class of Contracting Stream Operators*

In 2011, I co-authored a paper entitled "*A Class of Contracting Stream Operators*," which has just been published by *The Computer Journal* [JZ12]. Since that paper and the rest of this thesis share a common foundation of theory and since the paper was written to be self-contained, I disassembled it somewhat and spread the contents between Chapter 1 and Chapter 3. There are two, fairly distinct parts to this chapter. The operators discussed in [TZ11] are identified only indirectly by the properties they possess. In the first part of this chapter (the first two sections), I explicitly develop a class of operators whose members satisfy those properties. In the second part (the third section), I show how the case studies in [TZ07] and [TZ11] (mass-spring-damper systems) can be reorganized according to Part 1 to cover a broader range of systems, as well as including a new system (the simple pendulum), which yields only partially to the analysis in Part 1.

The abstract from the paper reads as follows:

> In [TZ07] and [TZ11], Tucker and Zucker present a model for the semantics of analog networks operating on streams from topological algebras. Central to their model is a parametrized stream operator representing the network along with a theory that concerns the existence, uniqueness, continuity, and computability of a fixed point of that stream operator. We narrow the scope of this paper from general topological algebras to algebras of streams that assume values only from a Banach space. This restriction facilitates the definition of a fairly broad class of stream operators to which the theory described in the above two papers applies.
>
> As a demonstration in their original work, the authors provide two case studies: analog networks which model the behaviour of simple mass-spring-damper systems. The case studies showcase the

theory well, but they seem to require the imposition of somewhat peculiar conditions on the parameters (the masses, the spring constants, and the damping coefficients). The extra conditions—while not catastrophic to the case studies—make them somewhat unsatisfying. We show here that while their original mass-spring-damper models do not fall within our new class, they can be easily reconfigured into equivalent models that do. This modification obviates the extra conditions on the parameters.

Chapter Highlights:

1. If we take $\mathcal{A}$ to be a Banach space, it is natural to define two corresponding stream spaces: one for scalar-valued streams and one for vector-valued streams. These work together as expected, using pointwise versions of the algebraic operations on $\mathcal{A}$. In fact, we can even generalize the former to scalar-matrix-valued streams.

2. The Building Block Lemma on page 55: an investigation into the way the two essential properties ***Lip*** and ***Caus*** are affected by integration and the pointwise stream operations on $\mathcal{A}$.

3. The Continuity Lemma on page 59: integration and the pointwise stream operations preserve the continuity of stream operators.

4. The General Form Theorem on page 65: this theorem identifies the titular class of contracting stream operators by pushing the two lemmas above as far as they can go without using any "foreign" operators.

5. The mass-spring-damper system from [TZ07, TZ11, TZ12] is reformulated in a way that requires no special conditions to be imposed on the parameters. Incidentally, my presentation of the mass-spring-damper system in Section 3.3.1.2 is the only proof of which I am aware that the ODE corresponding to the mass-spring-damper system has a solution for any continuous forcing function. The versions I've seen presented in textbooks always use a sinusoidal forcing function. This result may very well be proven elsewhere, of course, but I have never seen it.

6. The simplest form of pendulum is examined using this theory, but it requires the use of a function outside the class identified by the General Form Theorem on page 65.

Chapter 4: Generalizing the Theory Beyond Time Streams

The previous chapters along with most of [TZ07, TZ11, TZ12] concern stream operators, and as noted earlier, streams are functions of time. We model time (primarily) using $\mathbb{R}^{\geq 0}$, but the theory depends on very few of the special properties of $\mathbb{R}$. Tucker and Zucker do start out with a more general framework, using an arbitrary σ-compact space $\mathbb{X}$ instead of $\mathbb{T} = \mathbb{R}^{\geq 0}$, but they drop down to the special case of streams ($\mathbb{X} = \mathbb{T}$) as soon as causality is involved—since the concept of causality is inherently temporal.

I was able to generalize their two main properties (causality and contraction) to σ-compact spaces, alter the construction somewhat to be compatible in the more general framework, and prove variants of the two main theorems in [TZ11]. While I don't use shift invariance in my own theorems, I haven't been able to prove it is completely superfluous, so to help inspire future work, I suggest a way to generalize the shift operator as well. I also present some alternatives to the contraction property which may lead to other interesting results.

In the final section, I give a somewhat cursory treatment of the preceding material from the perspective of computability, and prove the final major theorem of the thesis.

Chapter Highlights:

1. Definitions of $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Lip}(\lambda, \mathfrak{X})$ (Definitions 4.2.1 and 4.2.6) form the basis of the generalization beyond time streams.

2. The definition of a retractable exhaustion (Definition 4.3.2 on page 87) is used to generalize the actual fixed point construction.

3. The Generalized TZ1 Theorem on page 94 shows that an operator which satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$ has a unique fixed point. This is one of the main thesis highlights. In addition to being a more general result than Theorem TZ1 on page 19, its proof *invokes* Banach's Fixed Point Theorem rather than cannibalizing key steps in the proof of Banach's theorem. So I believe it is both more general and more elegant than the theorem it supplants (it is just, admittedly, much less original, given that it is supplanting something in the first place).

4. The Generalized TZJ2 Theorem on page 102 is a strict generalization of Theorem TZJ2 on page 27. Loosely speaking, it shows that the fixed point of an operator $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ varies continuously with the parameters. This is perhaps the other main thesis highlight.

5. The Concrete Computability Theorem 4.7.11 on page 115 represents the first significant step toward generalizing Tucker and Zucker's follow-up paper to [TZ11]: [TZ12], in which the authors provide an analysis of the computability of the operators in [TZ11] using two different approaches to computability for stream operators (concrete and abstract). In this

theorem, I provide a set of conditions which are sufficient to ensure that the fixed point from the previous two theorems is concretely computable.

Chapter 5: Conclusion and Discussion

Given that I'm already covering the thesis highlights in this chapter summary, I use Chapter 5 to assess a few of the problems I encountered and review some of the ideas for further research.

Appendices

A. Hadamard's Principle and Supplementary Lemmas

Hadamard's Principle is a philosophical statement about the properties a mathematical model should possess if it is meant to correspond to a physical system. It was first expressed by Jacques Hadamard in [Had52], and explicated further by Richard Courant and David Hilbert in [CH53]. Its most weighty requirement is that the solution to such a problem should vary continuously with the parameters of the problem (or the input to the system). This is one of the reasons continuity is so heavily emphasized in [TZ11].

Hadamard's Principle seemed quite reasonable when it was first introduced to me, but something about it just didn't sit right. Despite the fact that Hadamard, Courant, and Hilbert were all far better mathematicians than I could ever hope to be (although I suppose that doesn't necessarily make them better philosophers), and despite the fact that I have yet to encounter any criticism of Hadamard's Principle from anyone else, I'm going to risk appearing impudent and voice my concerns with it in Appendix A, along with offering a suggestion about what I think might more aptly replace it. I relegate this discussion to the appendices, since it is more of opinion piece than a research topic.

B. Supplementary Propositions

Appendix B is small collection of assorted lemmas that are needed elsewhere, but which cluttered the exposition when inserted near the points in which they are invoked.

1.3 Preliminaries From Tucker and Zucker's Work

This thesis builds upon the work in [TZ11, TZ07, TZ12]. In order to make this document relatively self-contained, some of that foundational research must be reviewed, along with a few definitions and results from elementary topology and analysis. That is the purpose of this section.

1.3.1 The Space of Streams

Let $(\mathcal{A}, d_{\mathcal{A}})$ be a complete, separable metric space. We use the symbol $\mathbb{T}$ to represent time, taking[4] $\mathbb{T} = \mathbb{R}^+ \cup \{0\}$. We adopt $\mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ (for some $m \in \mathbb{Z}^+$) as our fundamental *stream space:* the space of m-tuples of continuous functions from $\mathbb{T}$ into $\mathcal{A}$.

Definition 1.3.1 (Pseudometrics on $\mathcal{C}[\mathbb{T}, \mathcal{A}]^m$)**.** For $m = 1$ we define a family of pseudometrics[5] $\{d_{a,b} : a, b \in \mathbb{T} \text{ and } a \leq b\}$ where $\forall u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$,

$$d_{a,b}(u, v) = \sup_{a \leq t \leq b} d_{\mathcal{A}}\left(u(t), v(t)\right) \tag{1.3.1}$$

Observe that if our stream space were instead $\mathcal{C}[[a, b], \mathcal{A}]$, then $d_{a,b}$ would be a metric. It is a pseudometric only because it "ignores" any differences between its arguments outside the interval $[a, b]$. For $m \in \mathbb{Z}^+$ and $\mathbf{u} = (u_1, u_{2,}, \ldots, u_m)$, $\mathbf{v} = (v_1, v_2, \ldots, v_m) \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ we define,

$$d^m_{a,b}(\mathbf{u}, \mathbf{v}) = \max_{1 \leq k \leq m} d_{a,b}\left(u_k, v_k\right)$$

In practice, however, we will drop the superscript since no ambiguity is introduced by overloading the symbol $d_{a,b}$. Furthermore, it is so often the case that we set $a = 0$ that typically we just write $d_b(\mathbf{u}, \mathbf{v})$ to mean $d^m_{0,b}(\mathbf{u}, \mathbf{v})$.

Remark 1.3.2. We will often form a product space of some metric space (X, d_X) and $\mathcal{C}[\mathbb{T}, \mathcal{A}]^m$. An equivalent family of pseudometrics ("equivalent" in the sense that they collectively generate the same topology as the metric) on this product space can be defined as,

$$d_T^{(X \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m)}\left((x, \mathbf{u}), (y, \mathbf{v})\right) = \max\left\{d_X(x, y), d_T(\mathbf{u}, \mathbf{v})\right\}$$

Again, without loss of specificity, we will drop the superscript and use simply d_T.

[4] Tucker and Zucker also develop their theory to address the case in which $\mathbb{T} = \mathbb{N}$, but here we'll be using only the continuum of nonnegative reals.

[5] A pseudometric is like a metric except that it is permitted to be zero even for distinct points. That is, if $d : X^2 \to Y$ is a pseudometric, then d is also a metric iff $\forall x, y \in X$ $[d(x, y) = 0 \Rightarrow x = y]$.

Remark 1.3.3. In [TZ11] it is shown that $\mathcal{C}[\mathbb{T},\mathcal{A}]^m$ is homeomorphic to $\mathcal{C}[\mathbb{T},\mathcal{A}^m]$, so the theory could be presented equivalently using either $\mathcal{C}[\mathbb{T},\mathcal{A}]$ or $\mathcal{C}[\mathbb{T},\mathcal{A}]^m$ as the fundamental stream space. If we adopt the former, we can always take $\mathcal{A} = \mathcal{B}^m$ (where $\mathcal{B}$ is some other space) whenever m-tuples are required, and if we adopt the latter, we can always take $m = 1$ when tuples are not wanted. We choose $\mathcal{C}[\mathbb{T},\mathcal{A}]$ for the sake of a cleaner exposition wherever possible, but sometimes we do need tuples (in Chapter 3, especially), so we will alternate between them according to convenience.

Definition 1.3.4 (Local Uniform Topology)**.** The family of pseudometrics in Definition 1.3.1 induces the *local uniform topology* on $\mathcal{C}[\mathbb{T},\mathcal{A}]^m$. A basis for this topology is given by open balls of the form,

$$B_{T,\varepsilon}(\mathbf{u}) = \{\mathbf{v} \in \mathcal{C}[\mathbb{T},\mathcal{A}]^m \,:\, d_T(\mathbf{u},\mathbf{v}) < \varepsilon\}$$

for $\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{A}]^m$, $T \in \mathbb{T}$, and $\varepsilon > 0$. See [TZ11] for a discussion of its equivalence to the compact-open topology and the inverse limit topology in this context. In fact, it is not even necessary to include every $T \in \mathbb{T}$. We can generate the topology using only countably many, equally spaced[6] values of $T \in \mathbb{T}$.

Definition 1.3.5 (Metric on $\mathcal{C}[\mathbb{T},\mathcal{A}]$)**.** There is actually a class of metrics that can be defined[7] on $\mathcal{C}[\mathbb{T},\mathcal{A}]$ (and hence on $\mathcal{C}[\mathbb{T},\mathcal{A}]^m$ as well) using the family of pseudometrics, given any $\tau > 0$:

$$d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(u,v) = \sum_{k=0}^{\infty} \min\left\{2^{-k}, d_{k\tau}(u,v)\right\}$$

These metrics are rather unwieldy, however. While they are important for showing that $\mathcal{C}[\mathbb{T},\mathcal{A}]$ (with the local uniform topology) is indeed metrizable, we prefer to use the pseudometrics when actually reasoning about the space. Of course, metrics are more widely known than pseudometrics, so I owe the reader some explanation of this last comment. Recall the following definitions for continuity from elementary topology.

Definition 1.3.6 (Continuity on Topological Spaces)**.** Let X, Y be topological spaces, let $f : X \to Y$, and let $x \in X$. Then,

1. f is continuous at x if for every open neighbourhood $U \subseteq Y$ of $f(x)$, there is an open neighbourhood $V \subseteq X$ of x such that $f(V) \subseteq U$.

[6] Even that is overly demanding, but we don't require anything more general at the moment. See Section 4.1.1 on page 82 for a more general treatment.

[7] Courtesy of Edwin Beggs.

2. f is continuous if it is continuous at every point $x \in X$. Equivalently, f is continuous if for every open $U \subseteq Y$, $f^{-1}(U)$ is open in X.

Since we can define the same topology on $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ using either the metric, $d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}$, or the family of pseudometrics $\{d_T\}_{T\in\mathbb{T}}$ (or, indeed, any subfamily $\{d_{n\tau}\}_{n\in\mathbb{N}}$, where $\tau \in \mathbb{R}^+$, as mentioned in Definition 1.3.4), we get the following lemma (along with Lemma 4.6.5), which is rather convenient for proving the continuity of stream functions.

The proof is routine, and therefore omitted.

Lemma 1.3.7. *A function* $f : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ *is continuous iff* $\forall \varepsilon > 0$ $\forall T \in \mathbb{T}$ $\forall \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ $\exists \delta > 0$ $\exists T' \in \mathbb{T}$ $\forall \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$,

$$d_{T'}(\mathbf{u}, \mathbf{v}) < \delta \Rightarrow d_T\left(f(\mathbf{u}), f(\mathbf{v})\right) < \varepsilon$$

That is (loosely speaking), f *is continuous if and only if the images of* $\mathbf{u}$ *and* $\mathbf{v}$ *under* f *can be made arbitrarily close on any closed interval* $[0, T]$*, as long as* $\mathbf{u}$ *and* $\mathbf{v}$ *are taken to be sufficiently close on some other closed interval* $[0, T']$.

Definition 1.3.8 (Stream Operations)**.** We'll often make use of the following three, time-based stream operations: ***shift***, ***hold***, ***delay***. Given $T, t \in \mathbb{T}$, each operation is of the form

$$f_T : \bigcup_{k=1}^{\infty} \mathcal{C}[\mathbb{T}, \mathcal{A}]^k \to \bigcup_{k=1}^{\infty} \mathcal{C}[\mathbb{T}, \mathcal{A}]^k$$

For a stream (or a portion of a stream), $\mathbf{u}$, they are defined as follows (also see Figure 1.3.1):

$$\begin{aligned}
\boldsymbol{shift}_T(\mathbf{u})(t) &= \mathbf{u}(t + T) \\
\boldsymbol{hold}_T(\mathbf{u})(t) &= \begin{cases} \mathbf{u}(t) & \text{if } t \leq T \\ \mathbf{u}(T) & \text{otherwise} \end{cases} \\
\boldsymbol{delay}_T(\mathbf{u})(t) &= \begin{cases} \mathbf{u}(0) & \text{if } t \leq T \\ \mathbf{u}(t - T) & \text{otherwise} \end{cases}
\end{aligned}$$

In some situations we'll need to treat them as functions of two variables: $\boldsymbol{shift}(T, \mathbf{u})$, $\boldsymbol{hold}(T, \mathbf{u})$, $\boldsymbol{delay}(T, \mathbf{u})$.

Remark 1.3.9. In [TZ11], the authors use an operation $\boldsymbol{ext}_T$, which is defined the same as $\boldsymbol{hold}_T$ except that its domain is $\mathcal{C}[[0, T], \mathcal{A}]$. I'm using ***hold*** so I can present a slightly different, but equivalent construction in Section 1.3.5 on page 19.

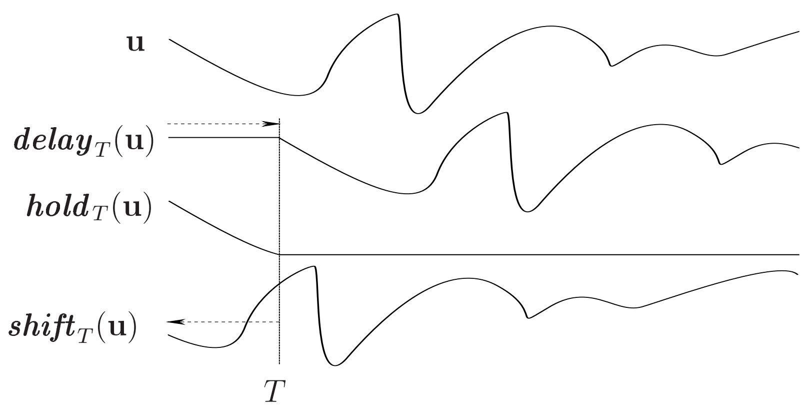

Figure 1.3.1: Time-based Stream Operations

1.3.2 The Analog Network Model

The streams represent data flowing through a network of channels and modules over time (which is considered a single, global property of the network). Each module has stream inputs, parameter inputs, and stream outputs, and thus, can be represented by a function of the form,

$$f : \mathcal{A}^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^r$$

We refer to the stream inputs and stream outputs as *channels*.

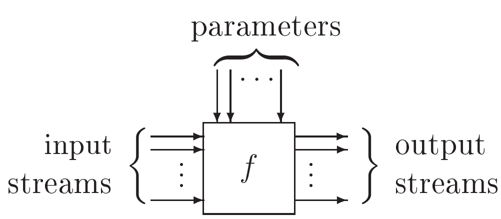

Figure 1.3.2: A Module in an Analog Network

Remark 1.3.10. The use of $\mathcal{A}^p$ as the parameter space is a feature of the original model and in this section I am striving to hew as closely as possible to the source material. In Section 1.4, this model will be generalized, allowing for the use of an arbitrary parameter space.

If all our networks were exclusively feed-forward (as in the following example), there would be no reason for any of this theory, and we could directly calculate the network output as a function of its input streams and parameters. We would simply compose all the module functions, working from the network input, all the way to the network output channels. That is, we could represent the network output as a straight-line program (see Chapter 4 of [BCS97]).

Example 1.3.11. Suppose $f_1 : \mathcal{A} \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$, $f_2 : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$, and $f_3 : \mathcal{A}^2 \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^2 \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$, and they are connected as shown in Figure 1.3.3. Then $\forall \mathbf{c} = (c_1, c_2, c_3) \in \mathcal{A}^3$ $\forall \mathbf{x} = (x_1, x_2) \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^2$, the network's output is well-defined and given by the function $f : \mathcal{A}^3 \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^2 \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ defined as

$$f(\mathbf{c}, \mathbf{x}) = f_3(c_2, c_3, f_1(c_1, x_1), f_2(x_2))$$

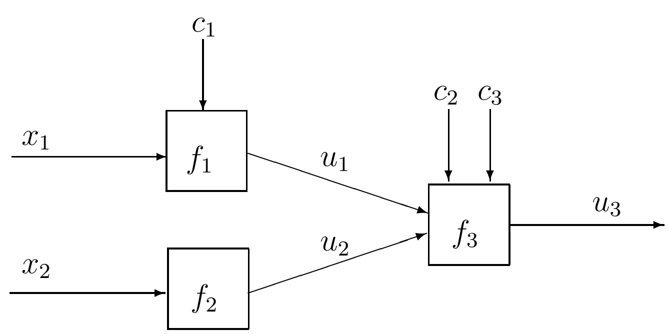

Figure 1.3.3: A simple, feedforward network

With feedback, however, such an attempt would lead to infinite regress (see Example 2.2.1). So, Tucker and Zucker adopt an alternative approach. Rather than looking at f itself, and trying to express the whole network's output as a function of its input (and parameters), they create a system of equations, one for every output channel. Each equation's left-hand side consists of a single stream variable representing the output of a module. If that channel is connected to the input of another module, the stream variable will appear within the expression on the right-hand side of another equation.

Example 1.3.11 would be written like this:

$$\begin{aligned} u_1 &= f_1(c_1, x_1) \\ u_2 &= f_2(x_2) \\ u_3 &= f_3(c_2, c_3, u_1, u_2) \end{aligned} \tag{1.3.2}$$

It is convenient to express this system as a single equation involving tuples. For a given $\mathbf{c}$, $\mathbf{x}$, define $F_{\mathbf{c},\mathbf{x}} : \mathcal{C}[\mathbb{T}, \mathcal{A}]^3 \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^3$ as follows:

$$F_{\mathbf{c},\mathbf{x}}\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\right) = \begin{bmatrix} f_1(c_1, x_1) \\ f_2(x_2) \\ f_3(c_2, c_3, u_1, u_2) \end{bmatrix}$$

The semantics of the network are then given by the fixed point for $F_{\mathbf{c},\mathbf{x}}$ (or a solution for Equation 1.3.2), if a unique one exists. Since the parameters and the input streams are meant to be adjustable, we often consider the function

$$F : \mathcal{A}^r \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \to (\mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m)$$

where $F(\mathbf{c}, \mathbf{x}, \cdot) = F_{\mathbf{c},\mathbf{x}}$. This is the real heart of the model, and along with it we define a *fixed-point function*:

$$\Phi : U \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \tag{1.3.3}$$

where $U \subseteq \mathcal{A}^r \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p$ and $\forall(\mathbf{c}, \mathbf{x}) \in U \ F_{\mathbf{c},\mathbf{x}}(\Phi(\mathbf{c}, \mathbf{x})) = \Phi(\mathbf{c}, \mathbf{x})$.

While this is consistent with the concept of fixed points, I always refer to such operators F in their uncurried form:

$$F : \mathcal{A}^q \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \tag{1.3.4}$$

Using this form, F can't really be said to have a "fixed point," per se, but it is isomorphic to an operator that can, hence the concept of fixed points is equally relevant, regardless of the form. So, in what might be considered an abuse of the vernacular, I will still refer to "fixed points" and the "fixed-point function," even when reasoning about an uncurried F.

This sort of operator F together with its fixed point function Φ is a slightly simplified version of the model of analog computation introduced by [TZ11, TZ07]. There are, of course, some properties to be imposed on F, which will be covered next. There is also one extra component to be added to the domains of the two functions which will be done when we turn to the property of shift invariance in Section 1.3.3.2 on page 17. After that, some of this structure will be undone when I present my own contributions to theory, but despite this undoing, it is important to see the intent behind the original model (which becomes somewhat less apparent as the model is generalized).

1.3.3 Properties of Stream Operators

As stated in the previous section, our objective is to find fixed points for a stream operator F. One of the distinguishing features of the theory is that these fixed points can be constructed, analyzed, or shown to exist in pieces rather than all at once. The following definition is helpful in this respect.

Definition 1.3.12 (T-approximate Fixed Points)**.** Let $f : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$, $T \in \mathbb{T}$, and $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$. Then we say $\mathbf{u}$ is a *T-approximate fixed point* of f if $d_T(\mathbf{u}, F(\mathbf{u})) = 0$.

There are two properties we must impose on a stream operator in order to facilitate this piecewise construction of the fixed point: causality and contraction.

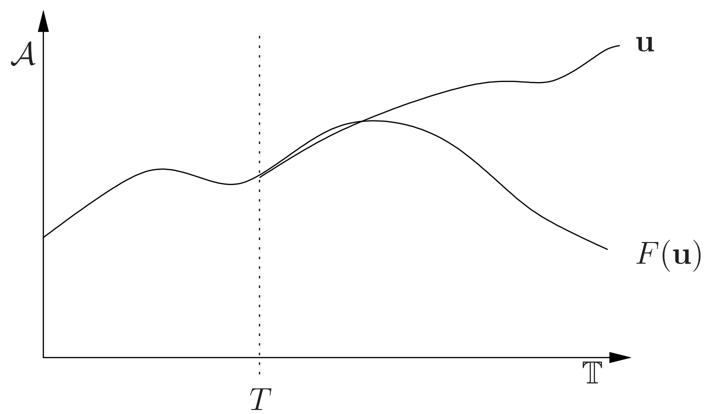

Figure 1.3.4: A T-approximate fixed point of a stream function, F

1.3.3.1 Causality and Contraction

Definition 1.3.13 (***Caus*** and ***WCaus***)**.** Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$. If $\forall T \in \mathbb{T}\ \forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$,

$$\mathbf{u} \restriction_{[0,T)} = \mathbf{v} \restriction_{[0,T)} \Rightarrow\ F(\mathbf{u})(T) = F(\mathbf{v})(T)$$

then we say that F satisfies ***Caus*** or $F \in$ ***Caus***. It is named as such since the property represents a form of causality. At each point in time, the value of $F(\mathbf{u})$ can be determined without any knowledge of future or present values of $\mathbf{u}$.

If instead,

$$\mathbf{u} \restriction_{[0,T]} = \mathbf{v} \restriction_{[0,T]} \Rightarrow\ F(\mathbf{u})(T) = F(\mathbf{v})(T)$$

then we say that F satisfies ***WCaus*** ("weak causality").

Remark 1.3.14. Causality conditions appear throughout control theory and signal processing (see [Son90] for example), and in several other contexts as well. Conditions almost identical to the two versions we define above (differing only in the domains and codomains of the operators involved), ***WCaus*** and ***Caus***, are identified in [Tra99] and [Rab03] as "retrospective" and "strongly retrospective," respectively.

Fact 1.3.15. *Since streams are continuous, it follows that*

$$F \in \textbf{\textit{Caus}} \text{ iff } F \in \textbf{\textit{WCaus}} \text{ and } \forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m\ F(\mathbf{u})(0) = F(\mathbf{v})(0)$$

Example 1.3.16. The pointwise addition of a constant to a real-valued stream is an example of an operator that satisfies ***WCaus*** but not ***Caus***. Define $F : \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]$ as

$$F(u)(t) = u(t) + 1$$

Let $T \in \mathbb{T}$ and $u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$. If $u \restriction_{[0,T]} = v \restriction_{[0,T]}$, then $u(T) = v(T)$. Hence, $F(u)(T) = u(T) + 1 = v(T) + 1 = F(v)(T)$. Thus, F satisfies $\boldsymbol{WCaus}$. But consider the streams $u(t) = 1$ and $v(t) = 0$. At no point $t \in \mathbb{T}$ is $F(u)(t) = F(v)(t)$ (the former is the constant stream 2, while the latter is the constant stream 1), but the interval $[0, T)$ is simply the empty set when $T = 0$. Thus, $u \restriction_{[0,0)} = v \restriction_{[0,0)}$ holds trivially, and yet $F(u)(0) = 2 \neq 1 = F(v)(0)$.

Remark 1.3.17. In light of Fact 1.3.15, the reader might wonder why we would bother with $\boldsymbol{Caus}$ when we have $\boldsymbol{WCaus}$. The latter is, indeed, sufficient for some purposes, but the former is essential for the most important theorems in which we prove that a unique fixed point stream of an operator, F, exists (and construct it). This fixed point stream is constructed one portion at a time, each successive portion created from the previous one. For this to work, the initial portion must already be in place and this is what $\boldsymbol{Caus}$ provides. If $F \in \boldsymbol{Caus}$, then every stream in the range of F is the same at time $t = 0$. Thus, the image of any stream in the range of F is a 0-approximate fixed point. From this, we can build a τ-approximate fixed point (where τ is some positive real number), and from that, a 2τ-approximate fixed point, and so on. If F satisfies only $\boldsymbol{WCaus}$, a starting place—let alone a whole fixed point stream—may not even exist! Consider the operator in Example 1.3.16, by inspection, it is clear that it has no T-approximate fixed points for any value of T, yet it satisfies $\boldsymbol{WCaus}$.

Definition 1.3.18 ($\boldsymbol{Contr}(\lambda, \tau)$)**.** Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$, and $\lambda, \tau \in \mathbb{R}^+$. If $\lambda < 1$ and $\forall T \in \mathbb{T}\ \forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$,

$$d_T(\mathbf{u}, \mathbf{v}) = 0 \ \Rightarrow\ d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v})$$

then we say that F satisfies $\boldsymbol{Contr}(\lambda, \tau)$ or $F \in \boldsymbol{Contr}(\lambda, \tau)$, named for the similarity this property shares with the notion of contraction[8] on a metric space. We refer to λ as the *modulus of contraction* of F (some authors use *contraction ratio*), and to τ as the *contraction increment* of F.

1.3.3.2 Shift Invariance

Somewhat central to [TZ11, TZ12] is the concept of shift invariance; major theorems in each of the two papers relies on it.

Remark 1.3.19 (Not the usual sort of shift invariance). In signal processing and control theory a shift invariant operator F is one that simply commutes with the shift operator: $\boldsymbol{shift}_T \circ F = F \circ \boldsymbol{shift}_T$ (see [Son90] for example). This won't work for an $F \in \boldsymbol{Caus}$, however, because for any such F there

[8] If (X, d_X) and (Y, d_Y) are metric spaces, and $f : X \to Y$, then f is *contracting* if $\exists \lambda > 0$ such that $\lambda < 1$ and $\forall x, y \in X\ d_Y(f(x), f(y)) \leq \lambda d_X(x, y)$.

is a constant $\mathbf{b} \in \mathcal{A}$ such that $\forall \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ $F(\mathbf{u})(0) = \mathbf{b}$. Thus, if F were shift invariant in the usual sense, then $\forall \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ $\forall t \in \mathbb{T}$ $F(\mathbf{u})(t) = \boldsymbol{shift}_t(F(\mathbf{u}))(0) = F(\boldsymbol{shift}_t(\mathbf{u}))(0) = \mathbf{b}$. In other words, the range of F would be the singleton set consisting of the stream with the constant value $\mathbf{b}$.

Shifting the output results in a glimpse of the future, while shifting the input effectively *erases* some of the past upon which that future output depends. So Tucker and Zucker's formulation of shift invariance avoids this problem by introducing a tuple of initial values which encodes the entire history of the input before T in a single snapshot, thus preserving all the essential information about the past input.

The space of parameters $\mathcal{A}^q$ from 1.3.4 on page 15 is factorized as $\mathcal{A}^q = \mathcal{A}^r \times \mathcal{A}^s$, where $\mathbf{c} \in \mathcal{A}^r$ is a tuple of system parameters (essentially these are freely configurable module settings), and $\mathbf{a} \in \mathcal{A}^s$ is a tuple of initial values, which comprises the aforementioned snapshot. The number s is chosen to be less than or equal to m and represents the number of components of $\mathbf{u}$ which must be "initialized" to reconstruct the past portions of $\mathbf{u}$ which are lost in the shift. The symbol $\mathbf{u}^s$ is used in this limited context to represent a tuple consisting of the first s components of $\mathbf{u}$ (i.e. a projection of $\mathbf{u}$ onto $\mathcal{C}[\mathbb{T}, \mathcal{A}]^s$).

Definition 1.3.20 (***Invar***)**.** Let

$$F : \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$$

Suppose that $\forall T \in \mathbb{T}$, $\forall (\mathbf{c}, \mathbf{a}, \mathbf{x}, \mathbf{u}) \in \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$, whenever

$$F(\mathbf{c}, \mathbf{a}, \mathbf{x}, \mathbf{u})\restriction_T = \mathbf{u}\restriction_T$$

the following two conditions also hold:

$$\begin{aligned} \mathbf{u}^s(0) &= \mathbf{a} \\ F(\mathbf{c}, \mathbf{u}^s(T), \boldsymbol{shift}_T(\mathbf{x}), \boldsymbol{shift}_T(\mathbf{u})) &= \boldsymbol{shift}(F(\mathbf{c}, \mathbf{a}, \mathbf{x}, \mathbf{u})) \end{aligned}$$

Then we say F satisfies ***Invar*** (or $F \in \boldsymbol{Invar}$).

Definition 1.3.21 (Closure of a domain under shifts)**.** Let

$$F : \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$$

and suppose that the fixed point function Φ for F is defined on a set $U \subseteq \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p$ (i.e. $\forall (\mathbf{c}, \mathbf{a}, \mathbf{x}) \in U$ $\exists! \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ such that $F(\mathbf{c}, \mathbf{a}, \mathbf{x}, \mathbf{u}) = \mathbf{u} = \Phi(\mathbf{c}, \mathbf{a}, \mathbf{x})$). Then U is *closed under shifts* with respect to Φ if $\forall T \in \mathbb{T}$ $\forall (\mathbf{c}, \mathbf{a}, \mathbf{x}) \in U$

$$(\mathbf{c}, \Phi(\mathbf{c}, \mathbf{a}, \mathbf{x})^s(T), \boldsymbol{shift}_T(\mathbf{x})) \in U$$

1.3.4 The Main Theorems From [TZ11]

Theorem 1.3.22 (Theorem TZ1)**.** *If $F : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ satisfies* ***Caus*** *and* ***Contr***(λ, τ) *for some $\tau > 0$ and $0 \leq \lambda < 1$, then F has a unique fixed point.*

Proof. See Theorem 1 from [TZ11]. □

Theorem 1.3.23 (Theorem TZ2)**.** *Let*

$$F : \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$$

and use the notation $F_{c,a,x}$ to represent the function $F(c, a, x, \cdot) : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$. Let $U \subseteq \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^p$ be an open set. Let $\boldsymbol{\lambda} = \{\lambda_{c,a,x} : (c, a, x) \in U\}$ be a family of contraction moduli and $\boldsymbol{\tau} = \{\tau_{c,a,x} : (c, a, x) \in U\}$ be a family of increments. Suppose the following conditions hold:

1. *$F_{c,a,x} \in$* ***Contr***$(\lambda_{c,a,x}, \tau_{c,a,x})$ *for all $(c, a, x) \in U$*
2. *$F \in$* ***Caus***
3. *$F \in$* ***Invar***
4. *F is continuous on U*
5. *$\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ are locally bounded on U (i.e. every point of U has a neighbourhood within which $\boldsymbol{\lambda}$ has an upper bound strictly less than 1, and $\boldsymbol{\tau}$ has a positive lower bound)*
6. *U is closed under shifts with respect to Φ (where Φ is the fixed point function defined in (1.3.3) on page 15)*

Then Φ is continuous on U.

Proof. See Theorem 2 from [TZ11]. □

1.3.5 The Mathematical Construction of the Fixed Point

As Theorem TZ1 on this page assures us, if $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ satisfies ***Caus*** and ***Contr***(λ, τ) (for some $\lambda, \tau > 0$ with $\lambda < 1$)[9], then it has a unique fixed point. The proof is constructive and while it is not necessary to include the whole thing in this thesis, we often need to refer to the construction it uses. The construction below is nearly identical to the one used in that proof, but

[9] Using the vernacular from Definition 1.4.2 on page 22, F satisfies ***Caus*** and ***Contr***.

since we're not reproducing the whole proof, we can simplify the exposition a bit. I'm also using rather different notation here, which (I think) greatly improves the clarity of some of my subsequent proofs. In my notation, given $F \in \boldsymbol{Caus} \cap \boldsymbol{Contr}(\lambda, \tau)$, we define a function $\Psi : \mathbb{N} \times \mathbb{N} \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ inductively as follows:

Construction 1.3.24.

1. Let $\Psi(0,0)$ be the constant stream, $\Psi(0,0)(t) = c \ \forall t \in \mathbb{T}$, where $c \in \mathcal{A}$ is the initial value constant associated with $F \in \boldsymbol{Caus}$. That is, $\forall u \in \mathcal{C}[\mathbb{T}, \mathcal{A}] \ F(u)(0) = c$.

2. For $n, k \in \mathbb{N}$, $\Psi(n, k+1) = \boldsymbol{hold}_{n\tau}(F(\Psi(n,k)))$, where τ is a number such that $F \in \boldsymbol{Contr}(\lambda, \tau)$.

3. Given $n \in \mathbb{N}$ define[10] $\Psi(n+1, 0) = \lim_{k \to \infty} \Psi(n, k)$

See Figure 1.3.5 for an overview. The central feature of the construction is that for any $n, k \in \mathbb{N}$, $\Psi(n,k)$ is an $n\tau$-approximate fixed point. That is, $\forall t \in [0, n\tau] \ \Psi(n,k)(t) = F(\Psi(n,k))(t)$.

$$\begin{array}{rrrr}
 & & c = & \Psi(0,0) \\
\Psi(0,1) = \boldsymbol{hold}_{\tau}(F(\Psi(0,0))) & \Psi(0,2) = \boldsymbol{hold}_{\tau}(F(\Psi(0,1))) & \cdots \to & \Psi(1,0) \\
\Psi(1,1) = \boldsymbol{hold}_{2\tau}(F(\Psi(1,0))) & \Psi(1,2) = \boldsymbol{hold}_{2\tau}(F(\Psi(1,1))) & \cdots \to & \Psi(2,0) \\
\Psi(2,1) = \boldsymbol{hold}_{3\tau}(F(\Psi(2,0))) & \Psi(2,2) = \boldsymbol{hold}_{3\tau}(F(\Psi(2,1))) & \cdots \to & \Psi(3,0) \\
\vdots & \vdots & & \vdots \\
 & & & \downarrow \\
 & & & v
\end{array}$$

Figure 1.3.5: Construction of the fixed point $v = F(v)$

Remark 1.3.25. Ordinarily it would be more natural to use a double sequence, but I've opted for a function on $\mathbb{N}^2$ to make it easier to talk about stages of the construction when F is augmented with parameters. When F is of the form $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ instead of merely $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$, we can easily (in terms of notational consistency) define

$$\Psi : P \times \mathbb{N} \times \mathbb{N} \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$$

Hence, for any $r \in P$, $\Phi(r) = \lim_{n \to \infty} \Psi(r, n, 0)$. The operator F and its interval of contraction τ are obviously central aspects of the construction, but unlike the parameter, they are always implicitly specified by the context.

[10] Theorem TZ1 proves that this limit exists using—as the reader might well guess from the invocation of contraction—mechanisms shared by the proof of Banach's Fixed Point Theorem on page 92.

Remark 1.3.26. At this point, the reader could hardly be blamed for wondering just how such a construction can possibly be related to computability. After all, we must perform infinitely many applications of F before we can even *begin* to approximate its fixed point at values of $t \in [\tau, 2\tau)$, and then infinitely many again before we can go beyond 2τ. The important thing to realize is that the purpose of this construction is to serve as a framework in which the fixed point can be analyzed (and shown to exist); it is clearly not suited to serve as a viable approximation algorithm.

This concludes the bulk of the prerequisite material from other sources. The remainder of this chapter will be used to cover a few of my own contributions to these rudiments which apply to at least two of the research projects (and hence belong in the neutral territory of the introductory thesis chapter rather than in any of the three project–specific chapters).

1.4 Observations and Addenda to the Core Preliminaries

There are a few more definitions and results to cover that apply to the whole thesis, but they are (for the most part) my own and not part of Tucker and Zucker's research. Because of that and because I feel they encumber the exposition somewhat if they are included in the section above, I've put them in their own section here.

1.4.1 Replace sup with max in Definition 1.3.1

This is admittedly somewhat pedantic, but if we wish to refer to $d_{a,b}$ from Definition 1.3.1 on page 10 as a pseudometric, and we define it as the supremum of a set of reals, it is incumbent on us to show that the set is always bounded. By the definition of a pseudometric, its codomain is the set of nonnegative real numbers (or just $\mathbb{R}$ in some texts), while the codomain of sup is the two-point compactification of the real numbers ($\mathbb{R} \cup \{-\infty, \infty\}$). Not only is it possible to show that the set is bounded, however, but it is also possible to show that it is closed. Hence, its supremum is not only finite, but actually contained within the set itself. Thus, it makes more sense to simply use max instead of sup. While it is fairly straightforward to show that this is possible, it is surprisingly nontrivial. First, we need a lemma.

Lemma 1.4.1 (Metrics are continuous)**.** *Let (X, d) be a metric space*[11]*. Then*

[11] In fact, this lemma holds even if d is only a pseudometric, but stating it this way would only lead to unnecessary confusion here since we need this lemma only for $d_{\mathcal{A}}$, which is a metric.

d is continuous with respect to the topology it induces on X^2.

Proof. Since the product is finite, we can work with the box topology on X^2, which consists of basic open sets

$$B_\varepsilon(x,y) = \left\{(x',y') \in X^2 \,:\, \max\{d(x,x'), d(y,y')\} < \varepsilon\right\}$$

Let $(x,y) \in X^2$ and let $U \subseteq \mathbb{R}^{\geq 0}$ be an open set that contains $d(x,y)$. Then there is an open interval $I \subseteq \mathbb{R}^{\geq 0}$ (open with respect to the subspace topology on $\mathbb{R}^{\geq 0}$) such that $d(x,y) \subseteq I \subseteq U$. Let $r > 0$ be the length of that interval. Let $V = B_r(x,y)$. Then $d(x,y) \in d\left(B_r(x,y)\right)$ (since $d(x,x) = d(y,y) = 0 < r$). Since $\forall (x',y') \in B_r(x,y)\ d(x,x') < r$ and $d(y,y') < r$, it follows that $d(B_r(x,y)) \subseteq I$. By Definition 1.3.6 on page 11, it follows that d is continuous. $\square$

Returning to the issue hand (replacing sup with max), since u and v are continuous on $\mathbb{T}$ and since $d_{\mathcal{A}}$ is continuous on $\mathcal{A}^2$ (by Lemma 1.4.1), it follows that $d_{\mathcal{A}}\left(u(t), v(t)\right)$ is continuous on $[a,b]$, which is compact with respect to the subspace topology on $\mathbb{T} \subseteq \mathbb{R}$. The continuous image of a compact set is compact, and a compact subset of $\mathbb{R}$ is closed and bounded. Thus, it contains its supremum, which is finite.

1.4.2 Generalize $\boldsymbol{Contr}(\lambda, \tau)$

Remark 1.4.2. There are times at which we need to refer to an operator F that satisfies $\boldsymbol{Contr}(\lambda,\tau)$ for some $\tau, \lambda > 0$ and $\lambda < 1$, but we don't care about the values of λ and τ. In such cases, it seems especially cumbersome to be obligated to specify that $\lambda, \tau > 0$ and $\lambda < 1$ since all three inequalities must hold just to satisfy Definition 1.3.18 on page 17. In these situations, it makes sense to write simply, "$F \in \boldsymbol{Contr}$" or "F satisfies $\boldsymbol{Contr}$."

While writing [JZ12], I found it necessary to be able to identify operators that *would* satisfy $\boldsymbol{Contr}(\lambda,\tau)$, but for values of λ that may be greater than or equal to one. Although such operators don't offer contraction per se, they are uniquely positioned to be composed with other operators to produce such contraction, so it is quite useful to be able to refer to this property. This is the subject of the Building Block Lemma (Lemma 3.2.1 on page 55).

Definition 1.4.3 (***Lip***)**.** Let $F : \mathcal{C}[\mathbb{T},\mathcal{A}]^m \to \mathcal{C}[\mathbb{T},\mathcal{A}]^m$. If $\exists \tau, \lambda \in \mathbb{R}^+ \cup \{0\}$ such that $\forall T \in \mathbb{T}\ \forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T},\mathcal{A}]^m$,

$$d_T(\mathbf{u},\mathbf{v}) = 0 \ \Rightarrow\ d_{T+\tau}\left(F(\mathbf{u}), F(\mathbf{v})\right) \leq \lambda d_{T+\tau}(\mathbf{u},\mathbf{v})$$

then we say that F satisfies $\boldsymbol{Lip}(\lambda,\tau)$ or $F \in \boldsymbol{Lip}(\lambda,\tau)$. The name is due to the similarity this property shares with the well-known Lipschitz continuity property from analysis (although traditionally α is in place of our λ).

Remark 1.4.4. It may seem as though $F \in \boldsymbol{Lip}(\lambda, \tau) \Rightarrow F \in \boldsymbol{WCaus}$. After all, if we take any $T \geq \tau$ and a pair of streams $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ such that $d_T(\mathbf{u}, \mathbf{v}) = 0$, then certainly $d_{T-\tau}(\mathbf{u}, \mathbf{v}) = 0$. Hence, $d_{\mathcal{A}}(F(\mathbf{u})(T), F(\mathbf{v})(T)) \leq d_T(F(\mathbf{u}), F(\mathbf{v})) = d_{(T-\tau)+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{(T-\tau)+\tau}(\mathbf{u}, \mathbf{v}) = \lambda d_T(\mathbf{u}, \mathbf{v}) = 0$. And therefore, $F(\mathbf{u})(T) = F(\mathbf{v})(T)$. Hence, any $F \in \boldsymbol{Lip}(\lambda, \tau)$ could be said to satisfy $\boldsymbol{WCaus}$ on $[\tau, \infty) \subseteq \mathbb{T}$.

There is, however, no way to establish the causality (weak or otherwise) of such an F on $[0, \tau)$, as the following example demonstrates.

Example 1.4.5 (***Lip*** $\not\Rightarrow$ ***WCaus***)**.** Take $\mathcal{A} = \mathbb{R}$ with the usual metric, let $\tau \in \mathbb{R}^+$, and choose $m = 1$. Define $F : \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]$ as follows:

$$F(u)(t) = \begin{cases} \frac{1}{2}u(\tau) & \text{if } 0 \leq t \leq \tau \\ \frac{1}{2}u(t) & \text{if } t > \tau \end{cases}$$

Then $F \in \boldsymbol{Lip}(1/2, \tau)$ (and it's even continuous), but it does not satisfy $\boldsymbol{WCaus}$. To see this, consider $u(t) = t$ and $v(t) = -t$. Taking $T = 0$, we see that $\forall t \leq T\ u(T) = v(T) = 0$, but $F(u)(0) = \tau/2 \neq -\tau/2 = F(v)(0)$. Note that such an example would not be possible if we were to take $\mathbb{T} = \mathbb{R}$ ($\boldsymbol{Lip}(\lambda, \tau)$ would give us $\boldsymbol{WCaus}$ "for free" on such a stream space), but adapting the rest of the theory to work on $\mathcal{C}[\mathbb{R}, \mathcal{A}]$ would not be trivial and nor would it necessarily be an improvement overall (see Section 3.1 on page 50 for an explanation).

Lemma 1.4.6. *If $F \in \boldsymbol{Lip}(\lambda, \tau)$ and $F \in \boldsymbol{WCaus}$ then $\forall \tau' \leq \tau$, $\forall \lambda' \geq \lambda$, $F \in \boldsymbol{Lip}(\lambda', \tau')$.*

Proof. Let $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$, $T \in \mathbb{T}$ and suppose $d_T(\mathbf{u}_1, \mathbf{u}_2) = 0$. For $\lambda' \geq \lambda$, it is obvious that $F \in \boldsymbol{Lip}(\lambda', \tau)$:

$$d_{T+\tau}(F(\mathbf{u}_1), F(\mathbf{u}_2)) \leq \lambda d_{T+\tau}(\mathbf{u}_1, \mathbf{u}_2) \leq \lambda' d_{T+\tau}(\mathbf{u}_1, \mathbf{u}_2)$$

The τ' assertion is less trivial. For $i = 1, 2$, define $\mathbf{u}_i^* \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ as follows:

$$\mathbf{u}_i^*(t) = \begin{cases} \mathbf{u}_i(t) & \text{if } t < T + \tau' \\ \mathbf{u}_i(T + \tau') & \text{if } t \geq T + \tau' \end{cases}$$

Then for $0 < \tau' < \tau$,

$$\begin{aligned} d_{T+\tau'}(F(\mathbf{u}_1), F(\mathbf{u}_2)) &= d_{T+\tau'}(F(\mathbf{u}_1^*), F(\mathbf{u}_2^*)) \quad \text{since } F \in \boldsymbol{WCaus} \text{ and } d_{T+\tau'}(\mathbf{u}_i^*, \mathbf{u}_i) = 0 \\ &\leq d_{T+\tau}(F(\mathbf{u}_1^*), F(\mathbf{u}_2^*)) \quad \text{since } t < r \Rightarrow d_t(\mathbf{v}, \mathbf{w}) \leq d_r(\mathbf{v}, \mathbf{w}) \\ &\leq \lambda' d_{T+\tau}(\mathbf{u}_1^*, \mathbf{u}_2^*) \quad \text{since } F \in \boldsymbol{Lip}(\lambda', \tau) \\ &= \lambda' d_{T+\tau'}(\mathbf{u}_1^*, \mathbf{u}_2^*) \quad \text{since } \mathbf{u}_i \text{are constant beyond } T + \tau' \\ &= \lambda' d_{T+\tau'}(\mathbf{u}_1, \mathbf{u}_2) \quad \text{since } d_{T+\tau'}(\mathbf{u}_i^*, \mathbf{u}_i) = 0 \end{aligned}$$

$\square$

Remark 1.4.7. As in Remark 1.4.4 on the preceding page, the only reason we must require F to satisfy $\boldsymbol{WCaus}$ in the proof of Lemma 1.4.6 is to establish the inequality for $T < \tau - \tau'$. For if $T \geq \tau - \tau'$ then,

$$\begin{aligned} d_T(\mathbf{u}_1, \mathbf{u}_2) = 0 \quad &\Rightarrow \quad d_{T-\tau+\tau'}(\mathbf{u}_1, \mathbf{u}_2) = 0 \\ &\Rightarrow \quad d_{(T-\tau+\tau')+\tau}(F\mathbf{u}_1, F\mathbf{u}_2) \leq \lambda' d_{(T-\tau+\tau')+\tau}(\mathbf{u}_1, \mathbf{u}_2) \\ &\quad\quad = \lambda' d_{T+\tau'}(\mathbf{u}_1, \mathbf{u}_2) \end{aligned}$$

This argument doesn't rely on $\boldsymbol{WCaus}$ at all, but it does require $T - \tau + \tau' \geq 0$ (so it isn't quite sufficient to show $F \in \boldsymbol{Lip}(\lambda', \tau')$).

Remark 1.4.8. Note that for any $\lambda \geq 0$, $\boldsymbol{WCaus}$ is actually equivalent to $\boldsymbol{Lip}(\lambda, 0)$. Putting this observation together with Lemma 1.4.6 yields the following result:

$$F \in Lip(\lambda, \tau) \cap \boldsymbol{WCaus} \Longleftrightarrow (\forall \tau' \leq \tau) \;\; F \in \boldsymbol{Lip}(\lambda, \tau')$$

Remark 1.4.9. In order to be more consistent with [TZ11] and to get the most general results possible, it would seem preferable to define $\boldsymbol{Lip}(\lambda, \tau)$ using the apparently weaker condition,

$$d_T(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v})$$

Call this condition $\boldsymbol{Lip}'(\lambda, \tau)$. One could not be faulted for thinking this definition is strictly more inclusive than $\boldsymbol{Lip}(\lambda, \tau)$, and it matches the definition of $\boldsymbol{Contr}(\lambda, \tau)$ in [TZ11] much more closely. In fact, it turns out that the two definitions are equivalent (so we stand by Definition 1.4.3).

Proposition 1.4.10 (Equivalence of $\boldsymbol{Lip}$ definitions). *Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$, $\lambda \in \mathbb{R}^+$, $\tau \in \mathbb{T}$. Then $F \in \boldsymbol{Lip}(\lambda, \tau)$ if and only if $F \in \boldsymbol{Lip}'(\lambda, \tau)$.*

Proof. Let $T \in \mathbb{T}$ and $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ such that $d_T(\mathbf{u}, \mathbf{v}) = 0$.
($\Rightarrow$) Suppose $F \in \boldsymbol{Lip}(\lambda, \tau)$. Then

$$\begin{aligned} d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v})) &\leq d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \\ &\leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

The first inequality holds because

$$d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) = \max\{d_T(F(\mathbf{u}), F(\mathbf{v})), d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v}))\}$$

and the second because $F \in \boldsymbol{Lip}(\lambda, \tau)$.
Now,

$$\begin{aligned} \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v}) &= \lambda \max\{d_T(\mathbf{u}, \mathbf{v}),\, d_{T,T+\tau}(\mathbf{u}, \mathbf{v})\} \\ &= \lambda \max\{0, d_{T,T+\tau}(\mathbf{u}, \mathbf{v})\} \\ &= \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

Therefore, $d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v})$.
($\Leftarrow$) Suppose $F \in \boldsymbol{Lip'}(\lambda, \tau)$. We must show that $d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v})$. As before, note that $\lambda d_{T+\tau}(\mathbf{u}, \mathbf{v}) = \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v})$. Similarly, $d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) = \max\{d_T(F(\mathbf{u}), F(\mathbf{v})), d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v}))\}$. Hence, we need to establish the following inequalities:

$$\begin{aligned} d_T(F(\mathbf{u}), F(\mathbf{v})) &\leq \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v}) \\ d_{T,T+\tau}(F(\mathbf{u}), F(\mathbf{v})) &\leq \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

The latter follows directly from the hypothesis, but the former requires a bit of work. We'll use an inductive approach for this. For the base case, suppose $0 \leq T < \tau$. Since $d_T(\mathbf{u}, \mathbf{v}) = 0$, it follows that $d_0(\mathbf{u}, \mathbf{v}) = 0$. Since $F \in \boldsymbol{Lip'}(\lambda, \tau)$,

$$\begin{aligned} d_\tau(F(\mathbf{u}), F(\mathbf{v})) &= d_{0,0+\tau}(F(\mathbf{u}), F(\mathbf{v})) \\ &\leq \lambda d_{0,0+\tau}(\mathbf{u}, \mathbf{v}) \\ &= \lambda d_\tau(\mathbf{u}, \mathbf{v}) \end{aligned}$$

Since $T < \tau$, $d_T(F(\mathbf{u}), F(\mathbf{v})) \leq d_\tau(F(\mathbf{u}), F(\mathbf{v}))$.
Since $T + \tau \geq \tau$, $\lambda d_\tau(\mathbf{u}, \mathbf{v}) \leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v})$.
Putting these last three results together we get,

$$\begin{aligned} d_T(F(\mathbf{u}), F(\mathbf{v})) &\leq d_\tau(F(\mathbf{u}), F(\mathbf{v})) \\ &\leq \lambda d_\tau(\mathbf{u}, \mathbf{v}) \\ &\leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v}) \\ &= \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

Now, for the inductive step, let $n \in \mathbb{Z}^+$ and assume that $\forall t < n\tau\ \forall u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$,

$$d_t(\mathbf{u}, \mathbf{v}) = 0 \ \Rightarrow\ d_{t+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{t+\tau}(\mathbf{u}, \mathbf{v})$$

Suppose $n\tau \leq T < (n+1)\tau$. We must show that

$$d_T(\mathbf{u}, \mathbf{v}) = 0 \ \Rightarrow\ d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \leq \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v})$$

Since $d_T(\mathbf{u}, \mathbf{v}) = 0$ and $0 \leq T - \tau \leq T$, it follows that,

$$d_{T-\tau}(\mathbf{u}, \mathbf{v}) = 0$$

So, by the inductive hypothesis and the fact that $T - \tau < n\tau$,

$$\begin{aligned} d_{(T-\tau)+\tau}(F(\mathbf{u}), F(\mathbf{v})) &\leq \lambda d_{(T-\tau)+\tau}(\mathbf{u}, \mathbf{v}) \\ &= \lambda d_T(\mathbf{u}, \mathbf{v}) \\ &= 0 \\ &\leq \lambda d_{T,T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

$\square$

1.4.3 Doesn't Continuity Follow from *Caus* and *Contr*?

Perhaps it's just my own flawed intuition, but it seemed to me that if an operator $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ satisfied ***Caus*** and ***Contr***, surely it *must* be continuous. I was particularly motivated to consider this assertion after having written Theorem TZJ1 for Vanishing Delays on page 48 (as it would have allowed me to omit one of the antecedents). After trying to prove it unsuccessfully for a while, a counterexample almost immediately occurred to me when I abandoned the proof attempt and tried to think of one.

Example 1.4.11. Define $F : \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]$ as follows:

$$F(u)(t) = \begin{cases} t & \text{if } u(0) \text{ is rational} \\ -t & \text{if } u(0) \text{ is irrational} \end{cases}$$

Let $u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$ and let $T \in \mathbb{T}$. Then $F(u)(0) = F(v)(0) = 0$, and if $d_T(u, v) = 0$ then $u(0) = v(0)$. Hence $F(u) = F(v)$. That means $F(u)(T) = F(v)(T)$ and therefore, $F \in \boldsymbol{Caus}$. It also means that for any $\lambda, \tau > 0$, $d_{T+\tau}(F(u), F(v)) = 0 \leq \lambda d_{T+\tau}(u, v)$. Hence, $F \in \boldsymbol{Contr}$.

As for continuity, let $\varepsilon = T_\varepsilon = 1$, let $\delta, T_\delta > 0$, and let $u \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$. Now choose a number $a \in (0, \delta)$ such that if $u(0) \in \mathbb{Q}$ then $u(0) + a \notin \mathbb{Q}$, and if $u(0) \notin \mathbb{Q}$ then $u(0) + a \in \mathbb{Q}$. Let $v(t) = u(t) + a$. Then $d_{T_\delta}(u, v) = a < \delta$, but $d_{T_\varepsilon}(F(u), F(v)) = 2 \geq \varepsilon$. Therefore, F is not continuous (by Lemma 1.3.7 on page 12).

1.4.4 Parameter-Relaying Tilde Functions

There are a few places in which I need to transmit a parameter value through a function that does not otherwise include the parameter space in its codomain. Up until my pre-defence revisions, I was unfamiliar with any notational convention for doing so. In the absence of such a convention, I began adorning my function names with a tilde when I needed to do this. It was only during these late-hour revisions that my attention was directed to the concept of "arrows" in functional programming (thanks to Prof. Jacques Carette!). Unfortunately, at this point I had tildes liberally sprinkled throughout my thesis, and more importantly, the notation for arrows does not appear to be well–suited for the use to which I would need to put them here. Consequently, I have left the tildes untouched.

Notation 1.4.12 (Tilde functions)**.** Let X and Y be sets and let $f : X \times Y \to Y$. Then we define $\widetilde{f} : X \times Y \to X \times Y$ as $\widetilde{f}(x, y) = (x, f(x, y))$ for $(x, y) \in X \times Y$.

Lemma 1.4.13. *Let X and $Y_1, Y_2, \ldots, Y_n$ be topological spaces. Let $x_0 \in X$. For $i = 1, 2, \ldots, n$, let $f_i : X \to Y_i$ be a function which is continuous at x_0. Let $f : X \to \prod_{i=1}^n Y_i$ be defined as $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$. Then f is continuous at x_0.*

Proof. Let $Y = \prod_{i=1}^n Y_i$ (with the product topology) and for each i, let $\pi_i : Y \to Y_i$ be the projection of Y on Y_i. Let $V \subseteq Y$ be an open neighbourhood of $f(x_0)$. Then, by definition of the product topology, there is a basic open set $B = B_1 \times B_2 \times \cdots \times B_n \subseteq V$, where each $B_i \subseteq \pi_i(V)$ and $f(x_0) \in B$. Since each f_i is continuous at x_0, there is an open neighbourhood $U_i \subseteq X$ of x_0 such that $f_i(U_i) \subseteq B_i$. Let $U = \bigcap_{i=1}^n U_i$. Then U is an open neighbourhood of x_0 (since it's only a finite union of open sets, each of which contains x_0), and $f(U) \subseteq B \subseteq V$. □

Corollary 1.4.14. *Let X and Y be topological spaces and suppose a function $f : X \times Y \to Y$ is continuous at a point $(x_0, y_0) \in X \times Y$. Then $\widetilde{f}$ (as defined in Notation 1.4.12) is also continuous at (x_0, y_0).*

Proof. $\tilde{f}$ can be rewritten as $\tilde{f}(x, y) = (\pi_X(x, y), f(x, y))$ (where $\pi_X : X \times Y \to X$ is the projection of $X \times Y$ on X). Both component functions are continuous at (x_0, y_0), so the result follows from Lemma 1.4.13. □

1.4.5 My Version of Theorem TZ2

While working on my original research project (Chapter 2), I found myself in need of something like Theorem TZ2 on page 19, but much to my chagrin, the function to which I needed to apply this theorem was *not* shift invariant and could not be made so by simply augmenting it with the extra initial value parameters. After many failed attempts using other theorems and constructions to get around this, I decided to dive into the proof to see whether I could substitute some other property for ***Invar***.

Much to my surprise, it initially appeared I didn't need to substitute *anything* for ***Invar***! My proof went through by apparently just omitting it. Upon later inspection, my supervisor and I together realized that indeed I had substituted something to replace ***Invar***: continuity on the entire domain rather than on only the parameter space. This leads to the following modified version of Theorem TZ2:

Theorem 1.4.15 (*Theorem TZJ2*)**.** *Let (P, d_P) be a metric space and let $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$. Let $p \in P$ and let $V \subseteq P$ be a neighbourhood of p. Let $\tau, \lambda \in \mathbb{R}^+$ with $\lambda < 1$. Using the notation $F_r(u) = F(r, u)$, suppose that for all $r \in V$ F_r satisfies **Caus** and **Lip**(λ, τ), and that for all $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$, F is continuous at (p, u). Then $\Phi : V \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ (as described in (1.3.3) on 15), whose existence is assured by Theorem TZ1 on page 19 is continuous at p.*

Proof. Deferred to Chapter 4, since this is now[12] merely a special case of the generalized version on page 102. □

Remark 1.4.16 (How do Theorems TZ2 and TZJ2 compare?). One obvious advantage to my requirement that F be continuous at (p, u) for every $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ (instead of being continuous on merely a subset of P, without regard to its behaviour on $P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]$) over Tucker and Zucker's requirement that $F \in$ ***Invar*** is that the latter places much heavier demands on the domain of F. In their version, P must be of the form $\mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q$, where $r, s, q \in \mathbb{N}$ and it must contain a subset U which has a nonempty interior and which is closed under shifts with respect to Φ. In my version, P is just an arbitrary metric space (which could be of the form $\mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q$, or of some other form).

Another (possible) advantage is that my version is pointwise rather than set-wise. They require F to be continuous on $U \subseteq P$ (which, as mentioned above, has a nonempty interior and is closed under shifts with respect to Φ), instead of at a single point $p \in P$. Of course, they do establish continuity on all of U—not at just a single point—so if one has no need for a pointwise version of the theorem, mine would offer no particular advantage in this respect.

Finally, the most obvious question to ask is how the two conditions overlap. That is, if we put aside the two advantages above (assume P is of the form $\mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q$ and F is continuous on U), are there any stream operators $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ that would satisfy one version and not the other? Are they, perhaps, the same under these conditions on the domain? After all, in many cases[13] it would be little more than a matter of bookkeeping (possibility rather elaborate and arduous bookkeeping, but bookkeeping nonetheless) to start with an operator $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ where P is a metric space that doesn't conform to the structure demanded by ***Invar*** and create an equivalent operator $F' : \mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ which is at least eligible to satisfy ***Invar***.

Unfortunately, I don't have the complete answer for this question, but do I have half the answer: there are some stream operators (in which P is of the correct form for ***Invar***) that satisfy the antecedents of my Theorem TZJ2, but not the antecedents of Tucker and Zucker's Theorem TZ2. Hence, their theorem may be a special case of mine, but the converse is not a possibility (even when the domain has the right form), as the following counterexample shows.

[12]Originally I proved this theorem directly, and that was long before the Generalized Theorem TZJ2 on page 102 even occurred to me. To prove the more general theorem required only a few adjustments in the proof of this theorem.

[13]In particular, I'm thinking of cases in which P can be embedded in a space of the form $\mathcal{A}^r \times \mathcal{A}^s \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^q$, for some $r, s, q \in \mathbb{N}$.

Example 1.4.17 (A non-***Invar*** stream operator)**.** Take $\mathcal{A} = \mathbb{R}^+$ and $P = \mathcal{A} \times \mathcal{A} \times \mathcal{C}[\mathbb{T}, \mathcal{A}]$ with the metric,

$$d_P\left((c_1, a_1, x_1), (c_2, a_2, x_2)\right) = \max\left\{|c_1 - c_2|, |a_1 - a_2|, d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(x_1, x_2)\right\}$$

Define $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ as follows, for $(c, a, x, u) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]$:

$$F(c, a, x, u) = \boldsymbol{delay}_c x$$

- $\forall (c, a, x) \in P \; F(c, a, x, \cdot) \in \boldsymbol{Caus}$. This is obvious since the value of $F(c, a, x, u)(t)$ doesn't depend on *any* values of u, let alone future or present values.
- F is continuous (on $P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]$). See Corollary 2.5.8 on page 45.
- $F(c, a, x, \cdot) \in \boldsymbol{Contr}(\lambda, |c|)$ for $\lambda = 1/2$ (in fact, for *any* $\lambda < 1$). See Lemma 2.5.1 on page 41.

Thus, F satisfies all the antecedents of Theorem TZJ2 on page 27, and it is clear by inspection that its fixed point is $u = \boldsymbol{delay}_c x$.

Now, take x to be a monophonic recording of somebody shouting[14], "Echo!", starting at time $t = 0$ and falling silent at $t = 1$ (and obviously transposed with direct current to ensure the recording stays in $\mathcal{C}[\mathbb{T}, \mathbb{R}^+]$ and never ventures below the $\mathbb{T}$-axis into $\mathcal{C}[\mathbb{T}, \mathbb{R}]$). Let $c = 1$. Then for any $a \in \mathcal{A}$, $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$, $F(c, a, x, u)$ is a recording of the same shout, but starting at time $t = 1$ and ending at time $t = 2$. Therefore, $\boldsymbol{shift}_c\left(F(c, a, x, u)\right) = x$ (more generally, for any $T \in \mathbb{T}$ $\boldsymbol{shift}_T \circ \boldsymbol{delay}_T$ is the identity on a stream space). But $\boldsymbol{shift}_c(x)$, on the other hand, is simply the zero stream, and no matter how much we delay it, we can never get the "Echo!" part back.

So for any $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$,

$$F(c, a, \boldsymbol{shift}_c(x), \boldsymbol{shift}_c(u)) = \boldsymbol{delay}_c\left(\boldsymbol{shift}_c(x)\right) = \mathbf{0} \neq x = \boldsymbol{shift}_c\left(F(c, a, x, u)\right)$$

(where $\mathbf{0}$ is the zero stream). Thus, F does not satisfy ***Invar***, and moreover, we've done everything possible to make it satisfy ***Invar*** without changing its behaviour.

[14] Or to be less colourful, take x to be any nonzero stream with support $[0, 1)$.

Chapter 2

Research Project #1: Solving Network Equations Using Vanishing Delays

2.1 Overview

Tucker and Zucker's theory centres around their construction of the fixed point along with a set of complementary properties (***Caus***, ***Contr***, ***Invar***, as well as several others without special names). All four of their main theorems use that construction as theoretical scaffolding to draw conclusions about the fixed point of a stream operator. I thought of an altogether different sort of construction for the fixed point and attempted to emulate their work using that. My construction involves introducing a delay in the stream transformer—making an operator with a guaranteed fixed point that is much easier to find—and then letting that delay approach zero, sort of like a homotopy in operator space. The main challenge I set for myself was to find sufficient (and ideally necessary) criteria to guarantee that the fixed point of the delayed transformer converges to a fixed point of the original stream transformer. Overcoming this challenge would give me an analogue of Theorem TZ1 (existence and uniqueness of the fixed point) from [TZ11].

Unfortunately, I never did overcome that challenge—at least not to my satisfaction. I was able to show that if the fixed point of the delayed transformer converges (to a stream), then indeed it converges to a fixed point of the original transformer. I was also able to show that my vanishing delay construction does work under the same conditions ($F \in$ ***Caus*** $\cap$ ***Contr***) that Tucker and Zucker's construction works. But what I really wanted to find was my own set of properties—specifically tailored for my construction—that would serve the same function as ***Caus*** and ***Contr*** (i.e. to test whether the construction will work at all for a given stream transformer). The closest I came was to devise a set of properties which I *think* might work, and to sketch out the beginning of a proof, but I got stuck (for several months) trying to finish that proof and realized I was probably going to drown in it if I didn't abandon ship.

2.2 Imposing a Delay on the Network Model

There is a (potentially profound) simplification built into Tucker and Zucker's network model, along with most similar models: the omission of propagation delay. Streams are carried from module to module over channels instantaneously, and this will obviously not be true in any physical implementation of a network. The delay usually makes qualitatively little difference in a purely feed-forward network with modules consisting of total functions. The output is always well-defined and perhaps only slightly phase-shifted, but when feedback is involved, the situation changes.

Example 2.2.1. Consider the following network in which $f : \mathcal{C}[\mathbb{T},\mathbb{R}] \to \mathcal{C}[\mathbb{T},\mathbb{R}]$ is some linear (and total) function that satisfies ***WCaus*** (see Definition 1.3.13):

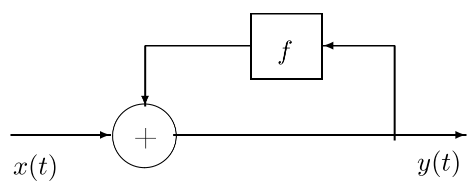

Figure 2.2.1: A Simple Feedback Network

The network output, if indeed it is well-defined, becomes an infinite regress if we attempt to solve it directly:

$$\begin{aligned} y(t) &= x(t) + f(y)(t) \\ &= x(t) + f(x)(t) + f^2(y)(t) \\ &= x(t) + f(x)(t) + f^2(x)(t) + f^3(x)(t) + f^4(y)(t) \\ &= \vdots \end{aligned}$$

Expressed in the notation of our network models, the network behaviour would be given by the (hopefully unique) solution of the following equation—if such a solution exists:

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix}(t) &= F \begin{bmatrix} x \\ y \end{bmatrix}(t) \\ &= \begin{bmatrix} x \\ x + f(y) \end{bmatrix}(t) \end{aligned}$$

But what if the solution doesn't exist? What does that mean about the system for which it serves as a model? Aside, perhaps, from pathological examples like black holes and misfortunate cats imprisoned opaquely with poisonous, nuclear-triggered deathtraps in paradoxical gedankenexperiments, there are no "undefined" values in nature (which is, in fact, the very raison d'être of the aforementioned felines). The system *will* exhibit some sort of behaviour, whether or not the equation model has a solution, and this disparity indicates a deficiency in the model.

Now, suppose we introduce a delay of $\gamma \in \mathbb{R}^+$ on every channel. We would have (for $t \geq \gamma$)

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix}(t) &= F \begin{bmatrix} x \\ y \end{bmatrix}(t-\gamma) \\ &= \begin{bmatrix} x \\ x + f(y) \end{bmatrix}(t-\gamma) \end{aligned} \tag{2.2.1}$$

and for $t \leq \gamma$, there would be constants $x_0, y_0 \in \mathbb{R}$ such that

$$\begin{bmatrix} x \\ y \end{bmatrix}(t) = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \tag{2.2.2}$$

This system leads to only finite regress. We can solve it directly for any value of $t \in \mathbb{T}$. If $t \leq \gamma$, the solution is given directly by (2.2.2). If $n\gamma \leq t \leq (n+1)\gamma$ for some $n \in \mathbb{Z}^+$, then we can use the constant solution on $[0, \gamma]$ together with (2.2.1) to find the solution on $[\gamma, 2\gamma]$, which we can then use to find the solution on $[2\gamma, 3\gamma]$, and so on, until we reach our target interval: $[n\gamma, (n+1)\gamma]$. As long as F is total, the network behaviour is *always* well-defined with the delay imposed (see the Delayed Operator Theorem on page 45).

Of course, to be even more physically accurate, we should equip each channel with its own delay, $\gamma_x, \gamma_y > 0$, and solve the system on the sequence of (possibly irregular) intervals with the endpoints given by multiples of γ_x and γ_y, but that level of generality is beyond the scope of our discussion.

2.3 The Problem with Imposing Delays and the Concept of Vanishing Delays

Imposing a mandatory delay (even an arbitrary delay of $\gamma > 0$) on every channel would make our mathematical model somewhat more accurate if the system being modelled directly resembles the network. For example, if we build an electronic circuit that looks exactly like a network diagram, it will indeed exhibit some latency as the signal travels from module to module. The latency would, of course, be different on each channel, but it still would be nonzero, so a uniformly delayed model would be at least somewhat closer to the real thing.

Analog computation involves building *analogies* of a real system, however, not building scale models. Consider, for example, the model of a mass-spring-damper system covered Chapter 3 (or in [TZ11]). We cover three different models, and each of them uses a different number of channels. Even if we imposed the same delay on each of them, the three systems would exhibit different solutions. But they're all supposed to be a model of the same system! Moreover, the system for which they serve as models would exhibit no such delays because information carried on separate channels of the models are, in fact, different physical properties (position, velocity, and acceleration). The idea of "propagation delay" between any two of those properties in the physical system is simply nonsensical. It is only when that network represents an actual circuit (and one that resembles the network exactly) that the delay makes any sense. Thus, while an explicit "delay module" could certainly be a useful addition to our modelling toolbox (along with adders, integrators, multipliers, etc.), forcing a delay into the very calculus of models would be a mistake.

The fact remains that a network with a delay on every channel always has a solution (and one that can be found directly) but a network without delays may not. So what if we introduce the delay temporarily, find a solution to the delayed system, and then see what happens to that solution as the delay approaches zero? This is the question I explored in Project #1.

2.4 Case Study: Linear Homogeneous Systems

Before engaging in the development of a theory based on vanishing delays, it seemed prudent to test the idea on a simple type of system with a known solution—just to serve as a proof-of-concept. Linear homogeneous systems fit the bill, and indeed, everything fell into place as I had hoped (as I'll now demonstrate).

In this section, we'll take $\mathcal{A} = \mathbb{R}$ (although $\mathcal{A} = \mathbb{C}$ would work just as well), choose some $m \in \mathbb{N}$. Let $A \in \mathbb{R}^{m \times m}$ and $\mathbf{c} \in \mathbb{R}^m$. Take $F : \mathcal{C}[\mathbb{T}, \mathbb{R}]^m \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^m$ to be,

$$F(\mathbf{u})(t) = \int_0^t A\mathbf{u}(s)ds + \mathbf{c} \tag{2.4.1}$$

It is well known from the theory of ordinary differential equations (see [BD01], for example) that F has a unique fixed point, $\mathbf{u}_0 \in \mathcal{C}[\mathbb{T}, \mathbb{R}]^m$, given by, $\mathbf{u}_0(t) =$

$e^{At}\mathbf{c}$, where

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

What we need to do is introduce a delay of $\gamma > 0$ to F, find the fixed point (if there is one[1]) of this delayed F, and then check whether that fixed point approaches $e^{At}\mathbf{c}$ as $\gamma \to 0^+$.

Given any $\gamma > 0$, any $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathbb{R}]^m$, and any $t \in \mathbb{T}$,

$$\boldsymbol{delay}_{\gamma} F(\mathbf{u})(t) = \begin{cases} \mathbf{c} & \text{if } t \leq \gamma \\ \int_0^{t-\gamma} A\mathbf{u}(s)ds + \mathbf{c} & \text{if } t \geq \gamma \end{cases} \tag{2.4.2}$$

Lemma 2.4.1. *$\boldsymbol{delay}_{\gamma} F$ (as defined by Equation (2.4.2)) has a unique fixed point:*

$$\mathbf{u}_{\gamma}(t) = \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t - k\gamma)^k}{k!} A^k \mathbf{c} \tag{2.4.3}$$

Proof. We must first verify that $\mathbf{u}_{\gamma}$ is actually a stream; in particular, it must be continuous. It is obviously continuous on any interval of the form, $[n\gamma, (n+1)\gamma)$ (where $n \in \mathbb{N}$) because it is defined as a sum of a fixed number of continuous terms there ($\lfloor t/\gamma \rfloor + 1$ of them). Once t crosses into the next such interval, however, a new term is added. So we need to check only that for $n \in \mathbb{N}$, $\mathbf{u}_{\gamma}(t) \to \mathbf{u}_{\gamma}((n+1)\gamma)$ as $t \to (n+1)\gamma^-$ (i.e. from the left). This is readily apparent since that new term contains the factor, $t - (n+1)\gamma$, which is zero at the left endpoint of the next interval (when $t = (n+1)\gamma$). With that formality out of the way, we can demonstrate that $\mathbf{u}_{\gamma}$ is actually a fixed point for $\boldsymbol{delay}_{\gamma} F$.

Let $t \in \mathbb{T}$. If $t < \gamma$ then,

$$\boldsymbol{delay}_{\gamma} F(\mathbf{u}_{\gamma})(t) = \mathbf{c} = \sum_{k=0}^{0} \frac{(t - k\gamma)^k}{k!} A^k \mathbf{c} = \mathbf{u}_{\gamma}(t) \tag{2.4.4}$$

So suppose $t \geq \gamma$ and let $N = \lfloor t/\gamma \rfloor - 1$ (making $(N+1)\gamma \leq t < (N+2)\gamma$).

[1] Admittedly, this seems somewhat coy in light of Sections 2.6 and 2.7. By inspection, it is clear that F satisfies $\boldsymbol{Caus}$ and has a unique 0-approximate fixed point—namely, $\mathbf{c}$. Therefore, by The Delayed Operator Theorem on page 45, the delayed operator, $\boldsymbol{delay}_{\gamma} F$, has a unique fixed point $\mathbf{u}_{\gamma}$ (given by the construction in Proof 2 of that theorem) for every $\gamma > 0$. The work in this section, however, was a feasibility study which necessarily preceded all that.

Then,

$$\begin{aligned}
\boldsymbol{delay}_\gamma F(\mathbf{u}_\gamma)(t) &= \int_0^{t-\gamma} A\mathbf{u}_\gamma(s)\,ds + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \int_{n\gamma}^{(n+1)\gamma} A\mathbf{u}_\gamma(s)\,ds + \int_{N\gamma}^{t-\gamma} A\mathbf{u}_\gamma(s)\,ds + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \int_{n\gamma}^{(n+1)\gamma} A\left(\sum_{k=0}^{\lfloor s/\gamma \rfloor} \frac{(s-k\gamma)^k}{k!} A^k\mathbf{c}\right) ds \\
&\quad + \int_{N\gamma}^{t-\gamma} A\left(\sum_{k=0}^{\lfloor s/\gamma \rfloor} \frac{(s-k\gamma)^k}{k!} A^k\mathbf{c}\right) ds + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \int_{n\gamma}^{(n+1)\gamma} \sum_{k=0}^{n} \frac{(s-k\gamma)^k}{k!} A^{k+1}\mathbf{c}\,ds \\
&\quad + \int_{N\gamma}^{t-\gamma} \sum_{k=0}^{N} \frac{(s-k\gamma)^k}{k!} A^{k+1}\mathbf{c}\,ds + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \sum_{k=0}^{n} \int_{n\gamma}^{(n+1)\gamma} \frac{(s-k\gamma)^k}{k!}\,ds A^{k+1}\mathbf{c} \\
&\quad + \sum_{k=0}^{N} \int_{N\gamma}^{t-\gamma} \frac{(s-k\gamma)^k}{k!}\,ds A^{k+1}\mathbf{c} + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \sum_{k=0}^{n} \left.\frac{(s-k\gamma)^{k+1}}{(k+1)!}\right|_{n\gamma}^{(n+1)\gamma} A^{k+1}\mathbf{c} \\
&\quad + \sum_{k=0}^{N} \left.\frac{(s-k\gamma)^{k+1}}{(k+1)!}\right|_{N\gamma}^{t-\gamma} A^{k+1}\mathbf{c} + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \sum_{k=0}^{n} \frac{((n+1)\gamma-k\gamma)^{k+1} - (n\gamma-k\gamma)^{k+1}}{(k+1)!} A^{k+1}\mathbf{c} \\
&\quad + \sum_{k=0}^{N} \frac{(t-\gamma-k\gamma)^{k+1} - (N\gamma-k\gamma)^{k+1}}{(k+1)!} A^{k+1}\mathbf{c} + \mathbf{c} \\
&= \sum_{n=0}^{N-1} \sum_{k=0}^{n} \frac{(n-k+1)^{k+1} - (n-k)^{k+1}}{(k+1)!} \gamma^{k+1} A^{k+1}\mathbf{c} \\
&\quad + \sum_{k=0}^{N} \frac{(t/\gamma-(k+1))^{k+1} - (N-k)^{k+1}}{(k+1)!} \gamma^{k+1} A^{k+1}\mathbf{c} + \mathbf{c}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{N-1} \left(\sum_{n=k}^{N-1} (n-k+1)^{k+1} - (n-k)^{k+1} \right) \frac{\gamma^{k+1} A^{k+1}}{(k+1)!}\mathbf{c} \\
&\qquad + \sum_{k=0}^{N} \left(\left(\frac{t}{\gamma} - k - 1 \right)^{k+1} - (N-k)^{k+1} \right) \frac{\gamma^{k+1} A^{k+1}}{(k+1)!}\mathbf{c} + \mathbf{c} \\
&= \sum_{k=0}^{N-1} \left(\left(\sum_{n=k}^{N-1} (n-k+1)^{k+1} - (n-k)^{k+1} \right) - (N-k)^{k+1} \right) \frac{\gamma^{k+1} A^{k+1}}{(k+1)!}\mathbf{c} \\
&\qquad + \sum_{k=0}^{N} \left(\left(\frac{t}{\gamma} - (k+1) \right)^{k+1} \right) \frac{\gamma^{k+1} A^{k+1}}{(k+1)!}\mathbf{c} + \mathbf{c}
\end{aligned}
$$

The nested sum above (the one indexed by n) telescopes:

$$
\begin{aligned}
\sum_{n=k}^{N-1} (n-k+1)^{k+1} - (n-k)^{k+1} &= 1^{k+1} - 0^{k+1} + 2^{k+1} - 1^{k+1} + \cdots \\
&\qquad \cdots + (N-k)^{k+1} - (N-1-k)^{k+1} \\
&= (N-k)^{k+1} - 0^{k+1}
\end{aligned}
$$

Thus, the summand in the first summation is zero for every value of k from 0 to $N-1$, which leaves only,

$$
\begin{aligned}
\boldsymbol{delay}_\gamma F(\mathbf{u}_\gamma)(t) &= \sum_{k=0}^{N} \left(\left(\frac{t}{\gamma} - (k+1) \right)^{k+1} \right) \frac{\gamma^{k+1} A^{k+1}}{(k+1)!}\mathbf{c} + \mathbf{c} \\
&= \sum_{k=1}^{N+1} \left(\left(\frac{t}{\gamma} - k \right)^{k} \right) \frac{\gamma^{k} A^{k}}{k!}\mathbf{c} + \mathbf{c} \\
&= \sum_{k=1}^{N+1} \frac{(t-k\gamma)^k}{k!} A^k \mathbf{c} + \mathbf{c} \\
&= \sum_{k=0}^{N+1} \frac{(t-k\gamma)^k}{k!} A^k \mathbf{c} \\
&= \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t-k\gamma)^k}{k!} A^k \mathbf{c} \\
&= \mathbf{u}_\gamma(t)
\end{aligned}
$$

$\mathbf{u}_\gamma$ can easily be shown to be unique by induction on N. In Equation (2.4.4), it can be seen that the value of $\boldsymbol{delay}_\gamma F(\mathbf{u}_\gamma)$ on $[0, \gamma)$ is independent of $\mathbf{u}_\gamma$— so we know at least that portion of the fixed point is unique. The rest of the calculations show that given any $N \in \mathbb{N}$, the value of $\boldsymbol{delay}_\gamma F(\mathbf{u}_\gamma)$ on $[(N+1)\gamma, (N+2)\gamma)$ depends on $\mathbf{u}_\gamma(t)$ only for $t < (N+1)\gamma$. Therefore, $\mathbf{u}_\gamma$ is the unique fixed point of $\boldsymbol{delay}_\gamma F$. $\square$

What must be shown next is that this solution approaches the stream, $\mathbf{u}_0(t) = e^{At}\mathbf{c}$ (which is the fixed point of F) as $\gamma \to 0^+$. First, a quick lemma. Recall that for a real number x, the power series for e^x is given by

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$$

What the following lemma shows is that we can "pollute" the terms of this expansion (in a particular way that suits our purposes) and affect only the rate of convergence, but not the end result.

Lemma 2.4.2. *For all* $x \in \mathbb{R}^{\geq 0}$

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \left(1 - \frac{k}{n}\right)^k = e^x$$

Proof. Let $x \geq 0$ and $\varepsilon > 0$. By definition,

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = e^x$$

Hence, there is an $N > 0$ such that $\forall n \geq N$ $\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| < \varepsilon/2$. Clearly for any fixed $k \geq 0$,

$$\lim_{n\to\infty} \left(1 - \frac{k}{n}\right)^k = 1$$

So, $\forall k > 0\ \forall \varepsilon' > 0\ \exists M_k > 0\ \forall n \geq M_k$

$$\left| \left(\frac{x^k}{k!}\right) \left(1 - \left(1 - \frac{k}{n}\right)^k\right) \right| < \varepsilon'$$

Let $M_0 = 1$ and for $k = 1, 2, \ldots, N$, let M_k be such that $\forall n \geq M_k$

$$\left| \left(\frac{x^k}{k!}\right) \left(1 - \left(1 - \frac{k}{n}\right)^k\right) \right| < \frac{\varepsilon}{2N+2}$$

Let $M = \max(\{M_k\}_{k=0}^{N} \cup \{N\})$. Then $\forall n \geq M$,

$$\begin{aligned}
\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!}\left(1-\frac{k}{n}\right)^k \right| &\leq \left| e^x - \sum_{k=0}^{N} \frac{x^k}{k!}\left(1-\frac{k}{n}\right)^k \right| \\
&\quad \left(\text{since } 0 \leq \frac{x^k}{k!}\left(1-\frac{k}{n}\right)^k \leq \frac{x^k}{k!} \text{ for } k = N, N+1, \ldots, n\right) \\
&= \left| e^x - \sum_{k=0}^{N} \frac{x^k}{k!} + \sum_{k=0}^{N} \frac{x^k}{k!} - \sum_{k=0}^{N} \frac{x^k}{k!}\left(1-\frac{k}{n}\right)^k \right| \\
&\leq \left| e^x - \sum_{k=0}^{N} \frac{x^k}{k!} \right| + \left| \sum_{k=0}^{N} \frac{x^k}{k!}\left(1-\left(1-\frac{k}{n}\right)^k\right) \right| \\
&< \frac{\varepsilon}{2} + (N+1)\frac{\varepsilon}{2N+2} = \varepsilon
\end{aligned}$$

□

Now we can finish the job by showing that $\mathbf{u}_\gamma \to \mathbf{u}_0$ as $\gamma \to 0^+$. The following theorem shows that my technique of introducing vanishing delays is capable of solving linear homogeneous systems (and with that, we bring the case study to a close). This is reassuring, certainly, and a modest victory, but it is not terribly exciting since there are much better ways to solve these systems already.

Theorem 2.4.3 (Vanishing Delay Theorem for Linear Homogeneous Systems)**.** *If A is an $m \times m$ matrix (real or complex) then $\forall t \in \mathbb{T}$*

$$\lim_{\gamma \to 0^+} \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t-k\gamma)^k}{k!} A^k = e^{At}$$

Proof. It is convenient to use a matrix norm to show this. The particular one chosen is unimportant so long as it is submultiplicative ($\|A_1 A_2\| \leq \|A_1\|\|A_2\|$). The operator norm, which is defined as follows for any matrix B, is such a norm:

$$\|B\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|B\mathbf{x}\|}{\|\mathbf{x}\|}$$

The symbol $\|\cdot\|$ is overloaded here, representing the vector norm on the right-hand side and the operator norm on the left. In addition to being submulti-

plicative, the operator norm satisfies all the usual norm axioms[2]
Let $t \in \mathbb{T}$, and $\varepsilon > 0$. We must find a $\Gamma > 0$ such that $\forall \gamma < \Gamma$,

$$\left\| e^{At} - \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t - k\gamma)^k}{k!} A^k \right\| < \varepsilon$$

Since the operator norm is subadditive,

$$\begin{aligned}
&\left\| e^{At} - \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t - k\gamma)^k}{k!} A^k \right\| \\
&\qquad = \left\| \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k - \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{(t - k\gamma)^k}{k!} A^k \right\| \\
&\qquad = \left\| \sum_{k=\lfloor t/\gamma \rfloor + 1}^{\infty} \frac{t^k}{k!} A^k + \sum_{k=0}^{\lfloor t/\gamma \rfloor} \left(t^k - (t - k\gamma)^k \right) \frac{A^k}{k!} \right\| \\
&\qquad \leq \left\| \sum_{k=\lfloor t/\gamma \rfloor + 1}^{\infty} \frac{t^k}{k!} A^k \right\| + \left\| \sum_{k=0}^{\lfloor t/\gamma \rfloor} \left(t^k - (t - k\gamma)^k \right) \frac{A^k}{k!} \right\| \qquad (2.4.5)
\end{aligned}$$

Since $\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = e^{At}$, it follows that, $\exists N_1 > 0 \; \forall n \geq N_1$

$$\left\| \sum_{k=n}^{\infty} \frac{t^k}{k!} A^k \right\| < \frac{\varepsilon}{2} \qquad (2.4.6)$$

Let N_1 be as such, and let $\Gamma_1 < t/N_1$. Then $\forall \gamma \in (0, \Gamma_1)$, the first term of (2.4.5) is less than $\varepsilon/2$. We now turn to the second term.

[2] For matrices B, B_1, B_2, and scalars α,

$$\begin{aligned}
&i. && B = \mathbf{0} \iff \|B\| = 0 \\
&ii. && \|\alpha B\| = |\alpha| \, \|B\| \\
&iii. && \|B_1 + B_2\| \leq \|B_1\| + \|B_2\|
\end{aligned}$$

$$\left\| \sum_{k=0}^{\lfloor t/\gamma \rfloor} \left(t^k - (t-k\gamma)^k \right) \frac{A^k}{k!} \right\| \leq \sum_{k=0}^{\lfloor t/\gamma \rfloor} \left\| \left(t^k - (t-k\gamma)^k \right) \frac{A^k}{k!} \right\| \tag{2.4.7}$$

$$= \sum_{k=0}^{\lfloor t/\gamma \rfloor} \left| \left(t^k - (t-k\gamma)^k \right) \frac{1}{k!} \right| \left\| A^k \right\| \tag{2.4.8}$$

$$\leq \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{\|A\|^k}{k!} \left| t^k - (t-k\gamma)^k \right| \tag{2.4.9}$$

Inequality (2.4.7) follows since the matrix norm is subadditive. Equation (2.4.8) follows since $\|\alpha B\| = |\alpha| \, \|B\|$ (for all scalars α and all matrices B). Inequality (2.4.9) follows since the matrix norm is submultiplicative. For convenience, let $Q(\gamma)$ represent the summation on line (2.4.9).

If $t < \gamma$, then $Q(\gamma) = 0$ which would allow us to ignore the whole term, but we're looking for an *upper* bound for γ that ensures $Q(\gamma) < {}^{\varepsilon}\!/_{2}$ when γ is sufficiently small. So we must assume $\gamma \leq t$ (in fact, we should assume $\gamma \ll t$). Let $q(\gamma) = \lfloor t/\gamma \rfloor$. Then (considering the $k\gamma$ near the end of line (2.4.9)),

$$k\gamma \leq k \frac{t}{q(\gamma)}$$

and hence $\forall \gamma \leq t$,

$$\begin{aligned} Q(\gamma) &\leq \sum_{k=0}^{q(\gamma)} \frac{\|A\|^k}{k!} \left| t^k - \left(t - k\frac{t}{q(\gamma)} \right)^k \right| \\ &= \sum_{k=0}^{q(\gamma)} \frac{(\|A\|t)^k}{k!} \left(1 - \left(1 - \frac{k}{q(\gamma)} \right)^k \right) \end{aligned}$$

From Lemma 2.4.2 on page 37 and the fact that e^x can also be written as $\sum_{k=0}^{\infty} \frac{x^k}{k!}$, it follows that $\exists N_2 > 0 \; \forall n \geq N_2$

$$\begin{aligned} \sum_{k=0}^{n} \frac{(\|A\|t)^k}{k!} - \sum_{k=0}^{n} \frac{(\|A\|t)^k}{k!} \left(1 - \frac{k}{n} \right)^k &= \sum_{k=0}^{n} \frac{(\|A\|t)^k}{k!} \left(1 - \left(1 - \frac{k}{n} \right)^k \right) \\ &< \frac{\varepsilon}{2} \end{aligned}$$

So choose $\Gamma_2 = t/N_2$. Then $\forall \gamma \leq \Gamma_2 \; q(\gamma) \geq N_2$, and hence $Q(\gamma) < \varepsilon/2$, as desired.

Therefore, if we select $\Gamma = \min\{\Gamma_1, \Gamma_2\}$ it follows that,

$$0 < \gamma < \Gamma \;\Rightarrow\; \left\| e^{At} - \sum_{k=0}^{\lfloor t/\gamma \rfloor} \frac{t - k\gamma}{k!} A^k \right\| < \varepsilon$$

$\square$

2.5 Properties of the Delay Operator

Having established that the vanishing delay technique works in some situations, we can proceed to studying it theoretically. We begin with some elementary properties of ***delay***, which will be essential for the investigation.

2.5.1 $\boldsymbol{delay}_\gamma$ is nonexpansive and preserves $\boldsymbol{Lip}$

Lemma 2.5.1. $\forall u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]\ \forall \gamma, T \in \mathbb{T}$

$$d_{T+\gamma}(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) = d_T(u, v)$$

Proof. Obvious (see Figure 2.5.1). Here is a proper proof, though:

$$\begin{aligned}
d_T(u, v) &= \max_{0 \le t \le T} d_{\mathcal{A}}(u(t), v(t)) \\
&= \max_{0 \le t \le T} d_{\mathcal{A}}(u(t - \gamma + \gamma), v(t - \gamma + \gamma)) \\
&= \max_{\gamma \le t \le T+\gamma} d_{\mathcal{A}}(u(t - \gamma), v(t - \gamma)) \\
&= \max_{\gamma \le t \le T+\gamma} d_{\mathcal{A}}\left(\left(\boldsymbol{delay}_\gamma u\right)(t), \left(\boldsymbol{delay}_\gamma v\right)(t)\right) \\
&= \max_{0 \le t \le T+\gamma} d_{\mathcal{A}}\left(\left(\boldsymbol{delay}_\gamma u\right)(t), \left(\boldsymbol{delay}_\gamma v\right)(t)\right) \\
&= d_{T+\gamma}\left(\left(\boldsymbol{delay}_\gamma u\right), \left(\boldsymbol{delay}_\gamma v\right)\right)
\end{aligned}$$

The second last equation holds since $\boldsymbol{delay}_\gamma u$ and $\boldsymbol{delay}_\gamma v$ are constant on $[0, \gamma]$. $\square$

Lemma 2.5.2. *For any $\gamma \ge 0$, $\boldsymbol{delay}_\gamma$ is nonexpansive (i.e. it is Lipschitz—in the traditional sense—with a Lipschitz constant of $\alpha = 1$). That is, $\forall u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]\ \forall \gamma \ge 0\ d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) \le d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(u, v)$.*

Proof. Let $\gamma \ge 0$. For any $T \le \gamma$,

$$\begin{aligned}
d_T(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) &= d_{\mathcal{A}}(u(0), v(0)) \\
&\le d_T(u, v)
\end{aligned}$$

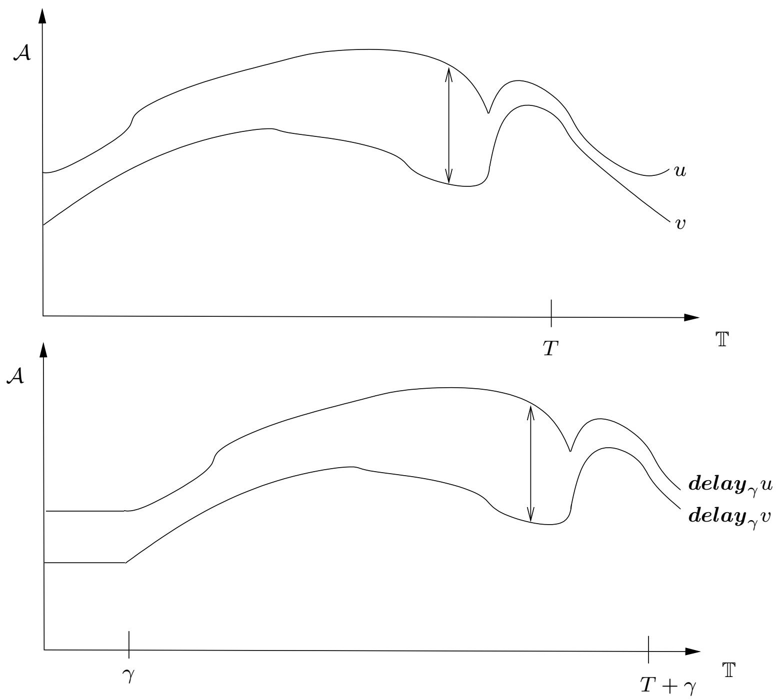

Figure 2.5.1: An Illustration of why $d_{T+\gamma}(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) = d_T(u, v)$

By Lemma 2.5.1, for any $T > \gamma$,

$$\begin{aligned} d_T(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) &= d_{T-\gamma}(u, v) \\ &\leq d_T(u, v) \end{aligned}$$

The result then follows from the definition of $d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}$ (each term is individually nonexpansive, so the summation is as well). □

Corollary 2.5.3. *Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$, let $\lambda, \tau \in \mathbb{R}^+$, and suppose $F \in \boldsymbol{Lip}(\lambda, \tau)$. Then $\boldsymbol{delay}_\gamma F \in \boldsymbol{Lip}(\lambda, \tau + \gamma)$.*

Proof. Trivial, using Lemma 2.5.1:

$$d_{T+\tau+\gamma}(\boldsymbol{delay}_\gamma F(u), \boldsymbol{delay}_\gamma F(v)) = d_{T+\tau}(F(u), F(v))$$

□

2.5.2 *delay* preserves *Caus*, *WCaus*

Lemma 2.5.4. *Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ and $\gamma > 0$, then $F \in \boldsymbol{WCaus} \Rightarrow \boldsymbol{delay}_\gamma F \in \boldsymbol{WCaus}$ and $F \in \boldsymbol{Caus} \Rightarrow \boldsymbol{delay}_\gamma F \in \boldsymbol{Caus}$.*

Proof. We'll skip ahead here and use the Building Block Lemma (Lemma 3.2.1 on page 55) together with Lemma 2.5.2 on page 41. By the latter, for all $T, \gamma \in \mathbb{T}$, $u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$,

$$d_T(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v) \leq d_T(u, v)$$

So if $d_T(u, v) = 0$, so does $d_T(\boldsymbol{delay}_\gamma u, \boldsymbol{delay}_\gamma v)$. Thus, $\boldsymbol{delay}_\gamma$ satisfies ***WCaus***. The result follows from Part 3 of the Building Block Lemma. ☐

Remark 2.5.5. In fact, we might say that $\boldsymbol{delay}_\gamma F$ is "supercausal" when F is causal since the value of $\boldsymbol{delay}_\gamma F(u)$ at any point in time cannot even depend upon values of u that are too recent, let alone upon present or future values of u. For $T \leq \gamma$,

$$u(0) = v(0) \Rightarrow \boldsymbol{delay}_\gamma F(u)(T) = \boldsymbol{delay}_\gamma F(v)(T)$$

and for $T \geq \gamma$,

$$d_{T-\gamma}(u, v) = 0 \Rightarrow \boldsymbol{delay}_\gamma F(u)(T) = \boldsymbol{delay}_\gamma F(v)(T)$$

2.5.3 *delay* is continuous

The following lemma deals with uniform continuity rather than (nonuniform) continuity and while the latter generalizes to topological spaces—and thus can be easily defined using pseudometrics—the most general setting for the former is uniform spaces, which is a topic that requires a fair bit of development. As a result, I'll use the metric defined in Definition 1.3.5 on page 11.

Lemma 2.5.6 (***delay*** is uniformly continuous on cross-sections)**.** *For any given $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$, the stream delay operator is uniformly continuous on $\mathbb{T} \times \{u\}$. That is, $\forall u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall \gamma_1, \gamma_2 \in \mathbb{T}$*

$$|\gamma_1 - \gamma_2| < \delta \ \Rightarrow\ d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(\boldsymbol{delay}_{\gamma_1} u, \boldsymbol{delay}_{\gamma_2} u) < \varepsilon$$

Proof. Let $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ and $\varepsilon > 0$. Let N be the smallest integer such that $2^{-N} < \varepsilon/2$. Define $\widehat{u} : \mathbb{R} \to \mathcal{A}^m$

$$\widehat{u}(t) = \begin{cases} u(0) & \text{if } t \leq 0 \\ u(t) & \text{if } t \geq 0 \end{cases}$$

Then $\forall \gamma, t \in \mathbb{T}\ \ \boldsymbol{delay}_\gamma u(t) = \widehat{u}(t - \gamma)$. Since $\widehat{\mathbf{u}}$ is continuous on $\mathbb{R}$, it is uniformly continuous on any closed interval. Moreover, this coupled with the fact that $\widehat{u}$ is constant on $(-\infty, 0]$, ensures that it is also uniformly continuous on any half-open interval of the form, $(-\infty, x]$. In particular, $\exists \delta > 0\ \forall t_1, t_2 \in (-\infty, N]$

$$|t_1 - t_2| < \delta \ \Rightarrow\ d_{\mathcal{A}}(\widehat{u}(t_1), \widehat{u}(t_2)) < \frac{\varepsilon}{2N}$$

Let $\gamma_1, \gamma_2 \in \mathbb{T}$ such that $|\gamma_1 - \gamma_2| < \delta$. Then,

$$\begin{aligned}
d_{\mathcal{C}[\mathbb{T},\mathcal{A}]}(\boldsymbol{delay}_{\gamma_1} u, \boldsymbol{delay}_{\gamma_2} u) &= \sum_{k=1}^{\infty} \min\left(2^{-k}, \max_{t\in[0,k]} d_{\mathcal{A}}\left(\boldsymbol{delay}_{\gamma_1} u(t), \boldsymbol{delay}_{\gamma_2} u(t)\right)\right) \\
&= \sum_{k=1}^{\infty} \min\left(2^{-k}, \max_{t\in[0,k]} d_{\mathcal{A}}\left(\widehat{u}(t-\gamma_1), \widehat{u}(t-\gamma_2)\right)\right) \\
&\le \sum_{k=1}^{N} \max_{t\in[0,k]} d_{\mathcal{A}}\left(\widehat{u}(t-\gamma_1), \widehat{u}(t-\gamma_2)\right) + \sum_{k=N+1}^{\infty} 2^{-k} \\
&\le N \max_{t\in[0,N]} d_{\mathcal{A}}\left(\widehat{u}(t-\gamma_1), \widehat{u}(t-\gamma_2)\right) + 2^{-N} \\
&< N\frac{\varepsilon}{2N} + \frac{\varepsilon}{2} \\
&\quad (\uparrow \text{ since } (t-\gamma_1), (t-\gamma_2) \in (-\infty, N] \\
&\qquad \text{and } |(t-\gamma_1) - (t-\gamma_2)| < \delta) \\
&= \varepsilon
\end{aligned}$$

□

In 1821, Cauchy infamously stated that a function continuous in each of its variables separately is continuous [Cau21]. While this is strictly false, we *can* prove something similar.

Lemma 2.5.7. *If X, Y, Z are metric spaces, $f : X \times Y \to Z$ is continuous in each of its variables separately, and f is equicontinuous*[3] *in one of them (i.e. taking f as an X-indexed family of functions from Y into Z, or as a Y-indexed family of functions from X into Z), then f itself is continuous (with respect to the product topology on $X \times Y$).*

Proof. Suppose, without loss of generality, that the family, $\{f(\cdot, y)\}_{y\in Y}$ is equicontinuous. That is, suppose there exists a function $\delta_X : X \times \mathbb{R}^+ \to \mathbb{R}^+$ such that $\forall x_0, x \in X\ \forall \varepsilon > 0\ \forall y \in Y$,

$$d_X(x_0, x) < \delta_X(x_0, \varepsilon) \ \Rightarrow\ d_Z(f(x_0, y), f(x, y)) < \varepsilon$$

Since f is continuous in Y separately, there is a function $\delta_Y : X \times Y \times \mathbb{R}^+$ such that $\forall \varepsilon > 0\ \forall x_0 \in X\ \forall y_0, y \in Y$,

$$d_Y(y_0, y) < \delta_Y(x_0, y_0, \varepsilon) \ \Rightarrow\ d_Z(f(x_0, y_0), f(x_0, y)) < \varepsilon$$

We choose the most convenient metric for our purposes that induces the product topology on $X \times Y$, namely the "maximum" metric:

$$d_{X\times Y}((x, y), (x', y')) = \max\{d_X(x, x'), d_Y(y, y')\}$$

[3] A family of functions is equicontinuous if they all share the same modulus of continuity.

Let $(x_0, y_0), (x, y) \in X \times Y$ and let $\varepsilon > 0$. Suppose

$$d_{X\times Y}\left((x_0, y_0), (x, y)\right) < \min\left\{\delta_X\left(x_0, \frac{\varepsilon}{2}\right), \delta_Y\left(x_0, y_0, \frac{\varepsilon}{2}\right)\right\}$$

Then it follows that,

$$\begin{aligned} d_Z(f(x_0, y_0), f(x, y)) &\leq d_Z(f(x_0, y_0), f(x_0, y)) + d_Z(f(x_0, y), f(x, y)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \end{aligned}$$

Therefore, f is continuous at (x_0, y_0). $\square$

Corollary 2.5.8. *The stream delay operator* $\boldsymbol{delay} : \mathbb{T} \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ *is continuous.*

Proof. By Lemma 2.5.2 on page 41, $\boldsymbol{delay}$ is equicontinuous in its second variable separately (in fact, more than that, it's globally Lipschitz with Lipschitz constant $\alpha = 1$ for all values of γ). That is, the family $\left\{\boldsymbol{delay}_\gamma\right\}_{\gamma\in\mathbb{T}}$ is equicontinuous. By Lemma 2.5.6 on page 43, $\boldsymbol{delay}$ is continuous (uniformly so) in its first variable. That is, all the functions in the family $\{\boldsymbol{delay}(\cdot, u)\}_{u\in\mathcal{C}[\mathbb{T},\mathcal{A}]}$ are uniformly continuous. Thus, $\boldsymbol{delay}$ is continuous in each of its variables separately, and is equicontinuous in one of them. By Lemma 2.5.7 on the preceding page, $\boldsymbol{delay}$ is continuous. $\square$

Corollary 2.5.9. *If* $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ *is continuous then* $\boldsymbol{delay} \circ \tilde{F} : \mathbb{T} \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ *(which is given by* $\boldsymbol{delay} \circ \tilde{F}(\gamma, u) = \boldsymbol{delay}_\gamma F(u)$*) is also continuous.*

Proof. $\tilde{F}$ is continuous by Corollary 1.4.14 on page 27, $\boldsymbol{delay}$ is continuous by Corollary 2.5.8, and a composition of continuous functions is continuous.

$\square$

2.6 Delayed Operators Always Have Unique Fixed Points

Theorem 2.6.1 (Delayed Operator Theorem)**.** *Let* $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ *satisfy* ***Caus*** *with* $F(u)(0) = c \in \mathcal{A}$ *for all* $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$. *Then* $\forall\gamma > 0\ \exists! u_\gamma \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ *which satisfies the system,*

$$\mathbf{u}_\gamma = \boldsymbol{delay}_\gamma F(\mathbf{u}_\gamma) \tag{2.6.1}$$

$$\mathbf{u}_\gamma(0) = c \tag{2.6.2}$$

Proof. [Short Version] Let $u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ $T \in \mathbb{T}$, and $\gamma \in \mathbb{R}^+$. Suppose $d_T(u, v) = 0$. Then, by Lemma 2.5.1 on page 41

$$d_{T+\gamma}(\boldsymbol{delay}_\gamma F(u), \boldsymbol{delay}_\gamma F(v)) = d_T(F(u), F(v))$$

And since $F \in \boldsymbol{Caus}$,

$$\begin{aligned} d_T(F(u), F(v)) &= d_T(u, v) \\ &= 0 \end{aligned}$$

Therefore, $\boldsymbol{delay}_\gamma F$ satisfies $\boldsymbol{Lip}(0, \gamma)$ (and thus $\boldsymbol{Contr}$). By Lemma 2.5.4 on page 42, $\boldsymbol{delay}_\gamma F$ satisfies $\boldsymbol{Caus}$. Hence, by Theorem TZ1 on page 19, $\boldsymbol{delay}_\gamma F$ has a unique fixed point. □

Remark 2.6.2. It is interesting to note that the continuity of F is not required to establish the existence of a fixed point. In fact, if F is continuous, a direct proof of the Delayed Operator Theorem (that does not invoke Theorem TZ1) becomes fairly trivial.

2.7 The Delay Vanishes[4]

2.7.1 Why the Vanishing Delay Construction Produces the Fixed Point of F

The Delayed Operator Theorem on the preceding page tells us that every stream operator that satisfies $\boldsymbol{Caus}$ has an associated family of streams, $\{\mathbf{u}_\gamma\}_{\gamma \in \mathbb{R}^+}$, each of which satisfy Equations 2.6.1 on the previous page and (2.6.2). It is convenient, then, to define a corresponding structure:

Definition 2.7.1. Let $F : \mathcal{C}[\mathbb{T}, \mathcal{A}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ satisfy $\boldsymbol{Caus}$, and suppose $F(\mathbf{u})(0) = \mathbf{c}$ (for all $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$). Define $\mathfrak{U} : \mathbb{R}^+ \to \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ as, $\mathfrak{U}(\gamma) = \mathbf{u}_\gamma$ (as defined in the Delayed Operator Theorem). Then the pair, $(F, \mathfrak{U})$, is a *delay system*. This provides a context for the symbol, $\mathbf{u}_\gamma$ (and similar variations), which will often be used in place of $\mathfrak{U}(\gamma)$ without explicitly acknowledging it.

Remark 2.7.2. Given a delay system $(F, \mathfrak{U})$, we're hoping to find a fixed point for F by finding the limit of $\mathfrak{U}(\gamma)$ as $\gamma \to 0^+$. So our big question is "when does that limit exist?" But before we get to that question, how do we know this limit will even work? That is, even if $\lim_{\gamma \to 0^+} \mathfrak{U}(\gamma)$ exists (in $\mathcal{C}[\mathbb{T}, \mathcal{A}]$), how do we know it's a fixed point of F? This is addressed by our next theorem, and here we do use continuity (although I have a hunch it's not necessary).

[4]No copyright infringement here since the syllables of "lady" have been reversed.

Theorem 2.7.3 (The Vanishing Delay Theorem)**.** *Suppose* $F : \mathcal{C}[\mathbb{T},\mathcal{A}]^m \to \mathcal{C}[\mathbb{T},\mathcal{A}]^m$ *is continuous and satisfies* ***Caus****. Let* $\{\gamma_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^+$ *be a sequence such that* $\gamma_k \to 0$ *as* $k \to \infty$*, and suppose that* $\lim_{k\to\infty} \mathfrak{U}(\gamma_k) = \mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{A}]^m$*. Then* $\mathbf{u} = F(\mathbf{u})$*.*

Proof.

$$\begin{aligned}
\mathbf{u} &= \lim_{k\to\infty} \mathbf{u}_{\gamma_k} && \text{(by hypothesis)}\\
&= \lim_{k\to\infty} (\boldsymbol{delay} \circ \tilde{F})(\mathbf{u}_{\gamma_k}, \gamma_k)\\
&= (\boldsymbol{delay} \circ \tilde{F})(\lim_{k\to\infty} (\mathbf{u}_{\gamma_k}, \gamma_k)) && \text{(since } \boldsymbol{delay} \circ \tilde{F} \text{ is continuous)}\\
&= (\boldsymbol{delay} \circ \tilde{F})(\mathbf{u}, 0)\\
&= F(\mathbf{u}) && \text{(since } \boldsymbol{delay}(\cdot, 0) \text{ is the identity on } \mathcal{C}[\mathbb{T},\mathcal{A}]^m)
\end{aligned}$$

The operative step is the third one, in which $\boldsymbol{delay} \circ \tilde{F}$ is moved outside the limit. Some explanation is warranted here. Since F is continuous, so is $\tilde{F}$ (by Corollary 1.4.14 on page 27). Thus, since ***delay*** is continuous (by Lemma 2.5.6 on page 43), $\boldsymbol{delay}\circ\tilde{F}$ is too. According to a well-known theorem[5] in Topology, if $f : X \to Y$ is continuous, X is metrizable, and $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ is convergent, then $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$. In our case, $X = Y = \mathcal{C}[\mathbb{T},\mathcal{A}]^m$, which is metrizable, and $\forall n \in \mathbb{N}\ x_n = (\mathbf{u}_{\gamma_n}, \gamma_n)$. □

2.7.2 When Does the Limit Exist?

This is the first major question about the vanishing delay construction, and the biggest obstacle I faced during this project (indeed, it was big enough that I never quite overcame it). While my efforts failed to provide a satisfactory answer, they did lead indirectly to the Generalized Theorem TZJ2 on page 102, which I consider to be among the most significant results of this thesis.

Since my vanishing delay construction is meant to be an alternative to the construction presented in [TZ11], an obvious question is whether it is *at least* as widely applicable. That is, if Tucker and Zucker's Theorem TZ1 guarantees the existence of a unique fixed point for an operator $F : \mathcal{C}[\mathbb{T},\mathcal{A}] \to \mathcal{C}[\mathbb{T},\mathcal{A}]$, will the vanishing delay construction necessarily converge to it?

I wasn't able to answer even this question completely, but I came close (falling short by having to assume continuity in addition to the antecedents of Theorem TZ1). Furthermore, it was in the process of answering this question that I developed Theorem TZJ2 on page 27—which led to the Generalized Theorem TZJ2 on page 102.

[5] See Theorem 10.3 of [Mun75], for example.

Theorem 2.7.4 (Theorem TZJ1 for Vanishing Delays)**.** *Let $(F, \mathfrak{U})$ be a continuous delay system (where $F : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$). Suppose F satisfies **Caus** and **Contr**. Then the unique fixed point of F (guaranteed to exist by Theorem TZ1) is given by $u = \lim_{\gamma \to 0^+} \mathfrak{U}(\gamma)$.*

Proof. The idea is to use Theorem TZJ2 on page 27, taking our parameter to be γ in the operator $\boldsymbol{delay} \circ \tilde{F}$, and then $\mathfrak{U}$ essentially becomes the Φ in Theorem TZJ2. Let $P = \mathbb{R}^{\geq 0}$. Define $G : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ as follows:

$$G(\gamma, u) = \boldsymbol{delay}_\gamma F(u) = \boldsymbol{delay} \circ \tilde{F}(\gamma, u)$$

By Lemma 2.5.4 on page 42, G satisfies ***Caus***. By Corollary 2.5.9 on page 45, G is everywhere continuous, and in particular it is continuous at $(0, u)$ $\forall u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$. Finally, by Corollary 2.5.3 on page 42, G satisfies ***Contr***.

Thus G satisfies all the conditions of the operator in Theorem TZJ2 for $p = 0$, and hence the Φ function for G is continuous at 0. The relevance of this observation is the fact that the Φ function for G is simply $\mathfrak{U}$, continuously extend from $\mathbb{R}^+$ to $\mathbb{R}^{\geq 0}$. Therefore, $\lim_{\gamma \to 0^+} \mathfrak{U}(\gamma)$ exists and by the Vanishing Delay Theorem on the previous page, this limit is the fixed point of F. $\square$

2.7.3 Addendum: What happens if $\boldsymbol{delay}_\gamma$ commutes with F?

Many of the results would be rendered fantastically simpler if only $\boldsymbol{delay}_\gamma$ would be so kind as to commute with F. In the last proof, we constructed a function, $\mathbf{u}_\gamma$, such that for any $t \in \mathbb{T}$, as long as we take a sufficiently large $n \in \mathbb{N}$, $\mathbf{u}_\gamma(t) = (\boldsymbol{delay}_\gamma F)^n(v_1)(t)$. If we could interchange $\boldsymbol{delay}_\gamma$ and F, we would have,

$$\mathbf{u}_\gamma(t) = (\boldsymbol{delay}_\gamma^n F^n)(v_1)(t) = (\boldsymbol{delay}_{n\gamma} F^n)(v_1)(t)$$

As it happens, this is not only a surprisingly unrealistic expectation, but it also causes big trouble.

Proposition 2.7.5. *Let $F, \gamma, \mathbf{u}_\gamma$ be as in the Delayed Operator Theorem on page 45, and suppose that*[6]

$$\mathbf{u}_\gamma = \boldsymbol{delay}_\gamma F \mathbf{u}_\gamma = F \boldsymbol{delay}_\gamma \mathbf{u}_\gamma$$

Then $\mathbf{u}_\gamma$ is constant ($\mathbf{u}_\gamma \equiv \mathbf{c}$, where $\mathbf{c}$ is the 0-afp of F, to be specific).

[6] Alternatively, we could assume that $\forall \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$ $\mathbf{u} = \boldsymbol{delay}_\gamma F \mathbf{u} = F \boldsymbol{delay}_\gamma \mathbf{u}$, but we needn't go so far for this proposition.

Proof. Let $t \in \mathbb{T}$ and let $n \in \mathbb{Z}^+$ such that $n\gamma \geq t$. Then,

$$\begin{aligned}
\mathbf{u}_\gamma(t) &= (\boldsymbol{delay}_\gamma F)^n(\mathbf{u}_\gamma)(t) && \text{(since } \mathbf{u}_\gamma \text{ is a fixed point of } \boldsymbol{delay}_\gamma F\text{)}\\
&= \boldsymbol{delay}_\gamma^n F^n(\mathbf{u}_\gamma)(t) && \text{(since } \boldsymbol{delay}_\gamma \text{ commutes with } F \text{ by assumption)}\\
&= \boldsymbol{delay}_{n\gamma} F^n(\mathbf{u}_\gamma)(t) && \text{(since } \boldsymbol{delay}_{\gamma_1}\boldsymbol{delay}_{\gamma_2} = \boldsymbol{delay}_{\gamma_1+\gamma_2}\text{)}\\
&= F^n(\mathbf{u}_\gamma)(0) && \text{(since } t \leq n\gamma\text{)}\\
&= \mathbf{c} && \text{(since } \mathbf{c} \text{ is a 0-afp of } F\text{)}
\end{aligned}$$

□

Remark 2.7.6. Proposition 2.7.5 shows that imposing such a condition on F drastically undermines the potential power of the theory, and yet it seems like a such a natural and benign property that may well apply to several common stream operators. It is, however, a much more restrictive condition than it might appear.

Example 2.7.7. Take $F(u)(t) = \int_0^t au(s)ds + c$ (for some $a, c \in \mathbb{R}$), and let u be the constant function, c. Then,

$$\boldsymbol{delay}_\gamma F(u)(t) = \boldsymbol{delay}_\gamma c \cdot (at+1) = \begin{cases} c & \text{if } t \leq \gamma \\ ca \cdot (t-\gamma) + c & \text{if } t \geq \gamma \end{cases}$$

Conversely,

$$F \circ \boldsymbol{delay}_\gamma(u)(t) = F(u)(t) = c \cdot (at+1) \neq \boldsymbol{delay}_\gamma F(u)(t) \ \text{(unless } \gamma \cdot t \cdot a \cdot c = 0\text{)}$$

Hence, even the members of this simple, general class of operators don't commute with $\boldsymbol{delay}_\gamma$.

Chapter 3

Research Project #2: Exploring the Special Case in Which $\mathcal{A}$ is a Banach Space

In this chapter, I will present the main portions of the paper [JZ12] which has been published in *The Computer Journal*. Both the paper and this thesis were written to be relatively self-contained, so if I were to paste the paper here, wholesale, much of the preliminary content in Chapter 1 would needlessly come with it. So in addition to omitting the redundant sections, I've made some minor edits to smooth the exposition from paper to thesis chapter. Furthermore, there are a few proofs that were omitted from the paper for the sake of brevity, and in place of those proofs, I refer the reader to this thesis. Hence, those have been included here.

3.1 Introduction

In [TZ11], Tucker and Zucker show that an operator which satisfies ***Caus*** and ***Contr*** has a unique fixed point, but which operators satisfy those properties? The authors offer two mass-spring-damper systems as examples, which is certainly very helpful, but it still leaves us with little intuition about which operators would have those properties and which ones wouldn't. Without imposing some restrictions on $\mathcal{A}$, there likely isn't much to be done about this. There just isn't enough to work with if we want to be more specific. If we restrict our attention to the case in which $\mathcal{A} = \mathbb{R}$, however, then we have a rich algebraic structure upon which to build operators that satisfy the properties. That's going a little further than necessary, though. It turns out that there is quite a lot we can say about the properties if we go only as far as making $\mathcal{A}$ a Banach space.

There are two main parts to this chapter. In the first part I conduct a thorough inventory of the pointwise stream operations induced by the algebraic operations of the Banach space, and examine the way each of them affects the stream properties covered in Chapter 1. While the pointwise operations yield a wide assortment of operators that satisfy the Lipschitz condition, the real engine behind the results is integration. An operator which satisfies the Lipschitz condition is all well and good, but in order to work with the two fixed point theorems, the operator must be *contracting,* and that is what integration provides. The integral (with respect to time) of a Lipschitz operator satisfies ***Contr***. All of these results are consolidated into a pair of lemmas (the Building Block and Continuity lemmas) and a single main theorem (the General Form Theorem on page 65).

In the second part, I move on to discuss two applications from mechanical physics. The first is the mass-spring-damper system described in [TZ07, TZ11], which the general form is more than powerful enough to handle on its own. The second—which is only a simple pendulum—neatly highlights the limitations of that form, as it is apparently not general enough to apply to that system. If we introduce a predefined operator (the sin function, in this case), however, we can still apply the two main lemmas separately to do the work the theorem cannot.

3.1.1 Algebra of Streams over a Banach Space

The operators with which we are concerned in this chapter operate on streams from $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$, where $\mathcal{B}$ is a Banach space over a field of scalars $\mathcal{S}$. The norm on $\mathcal{B}$ will be denoted using double bars, $\|\cdot\|$, and it induces a metric $d_{\mathcal{B}}(x,y)=\|x-y\|$. The same m-tuple convention used for the stream metric will be used for the norms on both $\mathcal{B}$ and $\mathcal{S}$: $\|(u_1,\ldots,u_m)\|=\max_{1\le k\le m}\|u_k\|$ and $|(a_1,\ldots,a_k)|=\max_{1\le k\le m}|a_k|$. Furthermore, corresponding to each pseudometric d_T (for $T\in\mathbb{T}$) is a seminorm (or a “pseudonorm,” using the vernacular in [Roy63]) $\|\mathbf{u}\|_T=d_T(\mathbf{u},\mathbf{0})$.

$\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ inherits several properties directly from $\mathcal{B}$—almost enough to make it a Banach space itself. The addition operation on $\mathcal{B}$ naturally induces (pointwise) addition on $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ (the continuity of the sum of two streams is assured by the subadditivity of the norm on $\mathcal{B}$). Scalars from $\mathcal{S}$ operate on $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ as they do on $\mathcal{B}$ (e.g. $a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$, $(ab)\mathbf{u}=a(b\mathbf{u})$, etc.). It is shown in [TZ11] that if $\mathcal{B}$ is separable and complete (which it is, being a Banach space), then so too is $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$. Similarly (although not addressed in [TZ11]), the local convexity of $\mathcal{B}$ assures the local convexity of $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$.

This collection of properties ensures that $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ is at least a Fréchet space[1]

[1] A Fréchet space is like a Banach space, except it lacks a norm. In its place, however, a

over $\mathcal{S}$, but since the origin of $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ does not necessarily contain an open bounded neighbourhood[2], it follows from Theorem 1.39 in [Rud91] that $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ is not normable. Hence it is not, itself, a Banach space.

For our purposes, however, a more useful observation is that $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ could *almost* serve as the set of scalars for the Fréchet space $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$. Addition and multiplication on $\mathcal{S}$ induce corresponding pointwise operations under which $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ is closed, and which commute, associate, and distribute according to the field axioms. Pointwise multiplication of a stream from $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ with a stream from $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ produces a stream from $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$. Being rife with zero divisors, however, $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ is not a field (it is only a commutative ring), and thus it cannot serve as a proper field of scalars in a topological vector space.

Despite this shortcoming, pairing $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ with $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ produces a useful algebra of pointwise operations—one which lays the foundation for matrix multiplication of streams in $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ by matrices in $\mathcal{C}[\mathbb{T},\mathcal{S}]^{m\times m}$. In fact, membership in a commutative ring is all that is required of the entries of a matrix in order to define a determinant (see [HK71]). That fact, in and of itself, is not immediately relevant to our research here, but it does suggest promising avenues of exploration in future research.

Most of the observations noted above follow readily, but we will take care to prove that pointwise multiplication between $\mathcal{C}[\mathbb{T},\mathcal{B}]^m$ and $\mathcal{C}[\mathbb{T},\mathcal{S}]^m$ works as we have claimed because that statement, in particular, is not completely trivial.

Lemma 3.1.1. *If $\boldsymbol{a} \in \mathcal{C}[\mathbb{T},\mathcal{S}]^m$ and $\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$, then $\boldsymbol{a}\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$, where $\boldsymbol{a}\mathbf{u}$ is the pointwise multiplication of $\boldsymbol{a}$ and $\mathbf{u}$:*

$$(\boldsymbol{a}\mathbf{u})\,(t) = (a_1(t)u_1(t), \ldots, a_m(t)u_m(t))$$

Proof. What must be shown is that $\boldsymbol{a}\mathbf{u}$ is continuous. Let $t_0 \in \mathbb{T}$ and $\varepsilon > 0$. Let,

$$\varepsilon' = \frac{1}{2}\left(\sqrt{(|\boldsymbol{a}(t_0)| + \|\mathbf{u}(t_0)\|)^2 + 4\varepsilon} - |\boldsymbol{a}(t_0)| - \|\mathbf{u}(t_0)\|\right)$$

Since $\boldsymbol{a}$ and $\mathbf{u}$ are continuous, $\exists \delta_{\boldsymbol{a}}, \delta_{\mathbf{u}} > 0$ such that $\forall t \in \mathbb{T}$,

$$\begin{aligned} |t - t_0| < \delta_{\boldsymbol{a}} \;&\Rightarrow\; |\boldsymbol{a}(t) - \boldsymbol{a}(t_0)| < \varepsilon' \text{ and} \\ |t - t_0| < \delta_{\mathbf{u}} \;&\Rightarrow\; \|\mathbf{u}(t) - \mathbf{u}(t_0)\| < \varepsilon' \end{aligned}$$

Fréchet space has a countable collection of seminorms that induce its topology. See [Rud91] for details.

[2] In this context a subset $X \subseteq \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ is *bounded* if for every neighbourhood B of $\mathbf{0} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ there is an $R > 0$ such that for all $r \in \mathcal{S}$ with $|r| > R$, $X \subseteq rB$. This means that unless $\mathcal{B}$ is the trivial space $\mathcal{B} = \{\mathbf{0}\}$ (or perhaps a rather esoteric and pathological space) we have that for every $T \in \mathbb{T}$ and $\varepsilon > 0$, $B_{T,\varepsilon}(\mathbf{0}) = \{\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m : d_T(\mathbf{u},\mathbf{0}) < \varepsilon\}$ is unbounded. This is because for any $r \in \mathcal{S}$ there is (for all the common Banach spaces, at least) a stream $\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ such that $\|\mathbf{u}(T+1)\| > |r|$, and hence $B_{T,\varepsilon}(\mathbf{0}) \not\subseteq r\,B_{T+1,\varepsilon}(\mathbf{0})$.

Let $\delta = \min\{\delta_{\boldsymbol{a}}, \delta_{\mathbf{u}}\}$. Then for t such that $|t - t_0| < \delta$,

$$\begin{aligned}
\|(\boldsymbol{a}\mathbf{u})(t) - (\boldsymbol{a}\mathbf{u})(t_0)\| &= \|(\boldsymbol{a}(t) - \boldsymbol{a}(t_0))(\mathbf{u}(t) - \mathbf{u}(t_0)) \\
&\quad + (\boldsymbol{a}(t) - \boldsymbol{a}(t_0))\mathbf{u}(t_0) + \boldsymbol{a}(t_0)(\mathbf{u}(t) - \mathbf{u}(t_0))\| \\
&\leq \|(\boldsymbol{a}(t) - \boldsymbol{a}(t_0))(\mathbf{u}(t) - \mathbf{u}(t_0))\| \\
&\quad + \|(\boldsymbol{a}(t) - \boldsymbol{a}(t_0))\mathbf{u}(t_0)\| + \|\boldsymbol{a}(t_0)(\mathbf{u}(t) - \mathbf{u}(t_0))\| \\
&= |\boldsymbol{a}(t) - \boldsymbol{a}(t_0)|\,\|\mathbf{u}(t) - \mathbf{u}(t_0)\| \\
&\quad + |\boldsymbol{a}(t) - \boldsymbol{a}(t_0)|\,\|\mathbf{u}(t_0)\| + |\boldsymbol{a}(t_0)|\,\|(\mathbf{u}(t) - \mathbf{u}(t_0))\| \\
&< (\varepsilon')^2 + \varepsilon'(\|\mathbf{u}(t_0)\| + |\boldsymbol{a}(t_0)|) \\
&= \varepsilon \text{ (after simplification)}
\end{aligned}$$

□

Corollary 3.1.2. *If $A \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m\times m}$ and $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, then $A\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*

While the algebraic operations on $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ facilitate the construction of many interesting stream operators, they would be of rather limited utility to the theory without integration (or something like it).

Lemma 3.1.3. *The Riemann integral[3] is well-defined on $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ and $\forall \mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ $\forall a, b \in \mathbb{T}$,*

$$\left\| \int_a^b \mathbf{u}(s)\, ds \right\| \leq \int_a^b \|\mathbf{u}(s)\|\, ds$$

Proof. See Theorems 2.1 and 5.1 in [Fea99] for the definition and the inequality, respectively. □

Remark 3.1.4. Iterated integrals are of particular importance to our theory, but standard integral notation becomes a little cumbersome for representing them. So we'll be using the following notational conventions. Given $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, $a, t \in \mathbb{T}$ (with $a \leq t$), and $n \in \mathbb{N}$,

$$\begin{aligned}
\int_a^{(0)} \mathbf{u}(t) &= \mathbf{u}(t) \\
\int_a^{(n+1)} \mathbf{u}(t) &= \int_a^t \left(\int_a^{(n)} \mathbf{u}(s) \right) ds
\end{aligned}$$

Equivalently,

$$\int_a^{(n)} \mathbf{u}(t) = \int_a^t \int_a^{s_1} \int_a^{s_2} \cdots \int_a^{s_{n-1}} \mathbf{u}(s_n)\, ds_n\, ds_{n-1} \ldots ds_1$$

[3] More accurately, the *generalized* Riemann integral, as defined by Feauveau [Fea99]. For an exposition of generalized integrals, see [Bar01].

Lemma 3.1.5. *Let* $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*,* $n \in \mathbb{Z}^+$*, and* $a, b \in \mathbb{T}$ *with* $a < b$*. Then,*

$$\left\| \int_a^{(n)} \mathbf{u}(b) \right\| \leq \frac{(b-a)^n}{n!} \max_{a \leq t \leq b} \|\mathbf{u}(t)\|$$

Proof. The base case, for $n = 1$, follows from Lemma 3.1.3, along with the fact that for a real function, $f : \mathbb{R} \to \mathbb{R}$ (like the $\|\mathbf{u}(s)\|$ from the right-hand side of the inequality in Lemma 3.1.3), $\int_a^b f(s)\, ds \leq (b-a) \max_{a \leq s \leq b} |f(s)|$. Now if we suppose that the inequality holds for all $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, $a, b \in \mathbb{T}$ ($a < b$), and for some $n > 0$ then

$$\begin{aligned}
\left\| \int_a^{(n+1)} \mathbf{u}(b) \right\| &= \left\| \int_a^b \int_a^{(n)} \mathbf{u}(s)\, ds \right\| \\
&\leq \int_a^b \left\| \int_a^{(n)} \mathbf{u}(s) \right\| ds \quad \text{(by Lemma 3.1.3)} \\
&\leq \int_a^b \frac{(s-a)^n}{n!} \max_{a \leq t \leq s} \|\mathbf{u}(t)\|\, ds \quad \text{(by the inductive hypothesis)} \\
&\leq \max_{a \leq t \leq b} \|\mathbf{u}(t)\| \int_a^b \frac{(s-a)^n}{n!}\, ds \\
&= \frac{(b-a)^{n+1}}{(n+1)!} \max_{a \leq t \leq b} \|\mathbf{u}(t)\|
\end{aligned}$$

□

3.2 Operators Which Satisfy the Fixed Point Theorems

Having established in Section 3.1.1 some of the basic operations we can use to create stream operators, we can now proceed to examine the way the properties discussed in Section 1.3.3 are affected by these operations. In the Building Block and Continuity Lemmas (Lemmas 3.2.1 on the following page and 3.2.2 on page 59 below), we will simply audit the effects of the algebraic operations so that when building operators from them or deconstructing operators in terms of them, we can directly calculate their properties. In the General Form Theorem on page 65, all these results are consolidated into the most general class of operators definable using these algebraic operations exclusively. The Building Block Lemma and the Continuity Lemma can also be used à la carte, however, with predefined operators that cannot be expressed using only the algebraic operations from Section 3.1.1 (see Section 3.3.2 on page 74 for an example).

Lemma 3.2.1 (The Building Block Lemma). *Given stream operators $F, G : \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, scalar stream $\boldsymbol{a} = (a_1, a_2, \ldots, a_m) \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^m$, and matrix stream $A \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$, the properties of **Caus**, **WCaus**, and **Lip**(λ, τ) are preserved by the basic stream operations as follows:*

1. **Primitive Operators**

   (a) Given $\mathbf{w} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, the constant operator $F_{\mathbf{w}}(\mathbf{v}) = \mathbf{w}$ satisfies $\boldsymbol{Caus}$ and $\boldsymbol{Lip}(0, \tau)$ for any $\tau \geq 0$.

   (b) The identity on $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ satisfies $\boldsymbol{WCaus}$ and $\boldsymbol{Lip}(1, \tau)$ for all $\tau \geq 0$.

2. **Addition of Operators**

   (a) $F, G \in \boldsymbol{WCaus} \Rightarrow (F + G) \in \boldsymbol{WCaus}$

   (b) $F, G \in \boldsymbol{Caus} \Rightarrow (F + G) \in \boldsymbol{Caus}$

   (c) $F \in \boldsymbol{Lip}(\lambda_F, \tau_F), G \in \boldsymbol{Lip}(\lambda_G, \tau_G) \Rightarrow (F + G) \in \boldsymbol{Lip}\left(\lambda_F + \lambda_G, \min\left\{\tau_F, \tau_G\right\}\right)$

3. **Composition of Operators**

   (a) $F, G \in \boldsymbol{WCaus} \Rightarrow (F \circ G) \in \boldsymbol{WCaus}$

   (b) $F \in \boldsymbol{Caus}$ and $G \in \boldsymbol{WCaus} \Rightarrow (F \circ G), (G \circ F) \in \boldsymbol{Caus}$

   (c) If $F \in \boldsymbol{Lip}(\lambda_F, \tau_F)$, $G \in \boldsymbol{Lip}(\lambda_G, \tau_G)$, and $F, G \in \boldsymbol{WCaus}$ then $(F \circ G) \in \boldsymbol{Lip}\left(\lambda_F \lambda_G, \min\left\{\tau_F, \tau_G\right\}\right)$

4. **Pointwise Multiplication by a Scalar Stream**

   (a) $F \in \boldsymbol{WCaus} \Rightarrow \boldsymbol{a}F \in \boldsymbol{WCaus}$

   (b) $F \in \boldsymbol{Caus} \Rightarrow \boldsymbol{a}F \in \boldsymbol{Caus}$

   (c) Let $\alpha \geq 0$. If $F \in \boldsymbol{Lip}(\lambda, \tau)$ and $\forall t \in \mathbb{T} \ \max\limits_{1 \leq i \leq m} |a_i(t)| \leq \alpha$ then $\boldsymbol{a}F \in \boldsymbol{Lip}(\alpha\lambda, \tau)$

5. **Pointwise Multiplication by a Scalar Matrix**

   (a) $F \in \boldsymbol{WCaus} \Rightarrow AF \in \boldsymbol{WCaus}$

   (b) $F \in \boldsymbol{Caus} \Rightarrow AF \in \boldsymbol{Caus}$

   (c) $F \in \boldsymbol{WCaus}$ and $A(0) = \mathbf{0} \Rightarrow AF \in \boldsymbol{Caus}$

   (d) Let $\alpha \geq 0$. If $F \in \boldsymbol{Lip}(\lambda, \tau)$ and $\forall t \in \mathbb{T} \ \|A(t)\| \leq \alpha$ then $AF \in \boldsymbol{Lip}(\alpha\lambda, \tau)$

6. **Integration**
Define $F_{\int} : \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ *as follows:*

$$F_{\int}(\mathbf{u})(t) = \int_0^t \mathbf{u}(s)\,ds = \left(\int_0^t u_1(s)\,ds, \int_0^t u_2(s)\,ds, \ldots, \int_0^t u_m(s)\,ds\right)$$

where $\mathbf{u} = (u_1, u_2, \ldots, u_m) \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*. Then,*

(a) $F_{\int} \in \boldsymbol{Caus}$
(b) $F_{\int} \in \boldsymbol{Lip}(\lambda, \lambda)\ \forall \lambda \in \mathbb{R}^+$

Proof. □

(1a) $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \left[F_{\mathbf{w}}(\mathbf{u}) = F_{\mathbf{w}}(\mathbf{v})\right]$, so both results follow trivially.

(1b) $\forall T \in \mathbb{T}\ \forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \Big[\big(\mathbf{u}\restriction_{[0,T]} = \mathbf{v}\restriction_{[0,T]} \Rightarrow\ \mathbf{id}(\mathbf{u})(T) = \mathbf{id}(\mathbf{v})(T)\big)$ and $d_{T+\tau}(\mathbf{u}, \mathbf{v}) \leq 1 \cdot d_{T+\tau}(\mathbf{id}(\mathbf{u}), \mathbf{id}(\mathbf{v}))\Big]$.

(2a) $d_T(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow (F + G)(\mathbf{u})(T) = F(\mathbf{u})(T) + G(\mathbf{u})(T) = F(\mathbf{v})(T) + G(\mathbf{v})(T) = (F + G)(\mathbf{v})$

(2b) By Remark 1.3.15, all that remains to be shown (given Part (2a)) is that $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \left[(F + G)(\mathbf{u})(0) = (F + G)(\mathbf{v})(0)\right]$. This follows directly from the fact that $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \big[F(\mathbf{u})(0) = F(\mathbf{v})(0)$ and $G(\mathbf{u})(0) = G(\mathbf{v})(0)\big]$.

(2c) Let $\tau = \min\{\tau_F, \tau_G\}$. By Lemma 1.4.6, $F \in \boldsymbol{Lip}(\lambda_F, \tau)$ and $G \in \boldsymbol{Lip}(\lambda_G, \tau)$. The result follows readily by taking $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ and expanding

$$d_{T+\tau}((F + G)(\mathbf{u}), (F + G)(\mathbf{v}))$$

into

$$\max_{0 \leq t \leq T+\tau} \|F(\mathbf{u})(t) + G(\mathbf{u})(t) - F(\mathbf{v})(t) - G(\mathbf{v})(t)\|$$

Then finally rearranging the terms and using the subadditivity of $\|\cdot\|$ to obtain the result.

(3a) $F, G \in \boldsymbol{WCaus} \Rightarrow \forall T \in \mathbb{T}\,\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \big[d_T(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow d_T(G(\mathbf{u}), G(\mathbf{v})) = 0 \Rightarrow F(G(\mathbf{u}))(T) = F(G(\mathbf{v}))(T)\big]$.

(3b) Given any $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, it may be the case that $G(\mathbf{u})(0) \neq G(\mathbf{v})(0)$, but under F the image of all streams (including those two) at time $t = 0$ is the same. Thus $F \circ G \in \boldsymbol{Caus}$. As for $G \circ F$, we do know that $F(\mathbf{u})(0) = F(\mathbf{v})(0)$, and since $G \in \boldsymbol{WCaus}$, that equality "up to 0" is preserved: $G(F(\mathbf{u}))(0) = G(F(\mathbf{v}))(0)$.

(3c) By Lemma 1.4.6, $F \in \boldsymbol{Lip}(\lambda_F, \tau)$ and $G \in \boldsymbol{Lip}(\lambda_G, \tau)$, where $\tau = \min\{\tau_F, \tau_G\}$. So, given $T \in \mathbb{T}$ and $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $d_T(\mathbf{u}, \mathbf{v}) = 0$, it follows from (3a) that $d_T\left(F(G(\mathbf{u})), F(G(\mathbf{v}))\right) = 0$ also. Hence, $d_{T+\tau}\left(F(G(\mathbf{u})), F(G(\mathbf{v}))\right) \leq \lambda_1 d_{T+\tau}\left(G(\mathbf{u}), G(\mathbf{v})\right) \leq \lambda_F \lambda_G d_{T+\tau}(\mathbf{u}, \mathbf{v})$.

(4a) $d_T(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow F(\mathbf{u})(T) = F(\mathbf{v})(T) \Rightarrow \boldsymbol{a}F(\mathbf{u})(T) = \boldsymbol{a}F(\mathbf{v})(T)$

(4b) $F(\mathbf{u})(0) = F(\mathbf{v})(0) \Rightarrow \boldsymbol{a}F(\mathbf{u})(0) = \boldsymbol{a}F(\mathbf{v})(0)$

(4c) Consider F as an m-tuple of functions: for $\mathbf{w} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ $F(\mathbf{w}) = (F_1(\mathbf{w}), \ldots, F_m(\mathbf{w}))$. Then,

$$
\begin{aligned}
d_{T+\tau}(\boldsymbol{a}F(\mathbf{u}), \boldsymbol{a}F(\mathbf{v})) &= \max_{\substack{0 \leq t \leq T+\tau \\ 1 \leq k \leq m}} \|a_k(t)F_k(\mathbf{u})(t) - a_k(t)F_k(\mathbf{v})(t)\| \\
&= \max_{\substack{0 \leq t \leq T+\tau \\ 1 \leq k \leq m}} |a_k(t)| \, \|F_k(\mathbf{u})(t) - F_k(\mathbf{v})(t)\| \\
&\leq \max_{\substack{0 \leq t \leq T+\tau \\ 1 \leq k \leq m}} |a_k(t)| \max_{\substack{0 \leq t \leq T+\tau \\ 1 \leq k \leq m}} \|F_k(\mathbf{u})(t) - F_k(\mathbf{v})(t)\| \\
&\leq \alpha d_{T+\tau}(F(\mathbf{u}), F(\mathbf{v})) \\
&\leq \alpha \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v})
\end{aligned}
$$

(5a, 5b) Same as (4a, 4b).

(5c) $A(0) = \mathbf{0} \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m} \Rightarrow (AF)(\mathbf{u})(0) = (AF)(\mathbf{v})(0) = \mathbf{0} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$. The rest is given by (5a).

(5d) Similar to (4c), but using the matrix norm $\|A(t)\|$ in place of $|a_k(t)|$.

(6a) Let $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $\forall t < T$ $\mathbf{u}(t) = \mathbf{v}(t)$. Then, using the norm on $\mathcal{B}^m$,

$$
\begin{aligned}
\left\|F_\int(\mathbf{u})(T) - F_\int(\mathbf{v})(T)\right\| &= \left\| \int_0^T \mathbf{u}(s)\, ds - \int_0^T \mathbf{v}(s)\, ds \right\| \\
&= \left\| \int_0^T (\mathbf{u}(s) - \mathbf{v}(s))\, ds \right\| \\
&\leq \int_0^T \|(\mathbf{u}(s) - \mathbf{v}(s))\|\, ds \\
&= 0
\end{aligned}
$$

The linearity of the integral in the second line comes from Theorem 3.1 in [Fea99], and the inequality arises from 3.1.3. Since $\|\cdot\|$ is a norm (rather than a mere seminorm), it follows that $F_\int(\mathbf{u})(T) = F_\int(\mathbf{v})(T)$.

(6b) Let $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $d_T(\mathbf{u}, \mathbf{v}) = 0$. Let $\lambda \geq 0$. Then,

$$
\begin{aligned}
d_{T+\lambda}\left(F_{\int}(\mathbf{u}), F_{\int}(\mathbf{v})\right) &= \max_{0 \leq t \leq T+\lambda} \left\| \int_0^t \mathbf{u}(s)\, ds - \int_0^t \mathbf{v}(s)\, ds \right\| \\
&= \max_{0 \leq t \leq T+\lambda} \left\| \int_0^t \left(\mathbf{u}(s) - \mathbf{v}(s)\right)\, ds \right\| && (3.2.1) \\
&\leq \max_{0 \leq t \leq T+\lambda} \int_0^t \|\mathbf{u}(s) - \mathbf{v}(s)\|\, ds && (3.2.2) \\
&= \int_0^{T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\|\, ds && (3.2.3) \\
&= \int_T^{T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\|\, ds && (3.2.4) \\
&\leq \lambda \max_{T \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\| && (3.2.5) \\
&= \lambda \max_{0 \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\| && (3.2.6) \\
&= \lambda d_{T+\lambda}(\mathbf{u}, \mathbf{v})
\end{aligned}
$$

Step Justifications (unnumbered steps require no further elucidation):

(3.2.1) By Theorem 3.1 in [Fea99].

(3.2.2) By Lemma 3.1.3. Note that this converts the Banach-valued integral to an ordinary integral over $\mathbb{R}$.

(3.2.3) Since the integrand is nonnegative, the maximum will be at $t = T + \lambda$.

(3.2.4) $d_T(\mathbf{u}, \mathbf{v}) = 0 \Rightarrow \int_0^T \|\mathbf{u}(s) - \mathbf{v}(s)\|\, ds = 0$.

(3.2.5) By Lemma 3.1.5.

(3.2.6) Since

$$
\begin{aligned}
\max_{0 \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\| &= \max \left\{ \max_{0 \leq s \leq T} \|\mathbf{u}(s) - \mathbf{v}(s)\|, \right. \\
&\qquad \left. \max_{T \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\| \right\} \\
&= \max \left\{ 0, \max_{T \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\| \right\} \\
&= \max_{T \leq s \leq T+\lambda} \|\mathbf{u}(s) - \mathbf{v}(s)\|
\end{aligned}
$$

Lemma 3.2.2 (The Continuity Lemma)**.** *Let (P, d_P) be a metric space (which will serve as a parameter space) and let $p \in P$. Let $F, G : P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ and suppose that for all $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ F and G are continuous at $(p, \mathbf{u})$. Let $A : P \to \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$ be continuous at p. Then the functions $H : P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ defined below for $(r, \mathbf{u}) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, $t \in \mathbb{T}$, are all continuous at every point in $\{p\} \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \subseteq P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$:*

1. Addition: $H(r, \mathbf{u})(t) = (F + G)(r, \mathbf{u})(t) = F(r, \mathbf{u})(t) + G(r, \mathbf{u})(t)$
2. Composition: $H(r, \mathbf{u})(t) = F(r, G(r, \mathbf{u}))(\mathrm{t})$
3. Matrix Multiplication: $H(r, \mathbf{u})(t) = (AF)(r, \mathbf{u})(t) = A(r)(t)\, F(r, \mathbf{u})(t)$
4. Integration: $H(r, \mathbf{u})(t) = \int_0^t F(r, \mathbf{u})(s)\, ds$

Proof.

(1) Follows from the subadditivity of the seminorms on $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$: $\|\mathbf{u}\|_T = d_T(\mathbf{u}, \mathbf{0})$.

(2) Let $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$. Let $\varepsilon > 0$. Since F is continuous on $\{p\} \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, there is a $\delta_F > 0$ such that $\forall \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ $\forall r \in P$,

$$d_{P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m}\left((p, G(p, \mathbf{u})), (r, \mathbf{v})\right) < \delta_F \Rightarrow d_C\left(F\left(p, G(p, \mathbf{u})\right), F(r, \mathbf{v})\right) < \varepsilon$$

Since G is continuous on $\{p\} \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, there is a $\delta_G > 0$ such that $\forall (r, \mathbf{w}) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$,

$$d_{P \times \mathcal{C}[\mathbb{T}, \mathcal{A}]^m}\left((p, \mathbf{u}), (r, \mathbf{w})\right) < \delta_G \Rightarrow d_C\left(G(p, \mathbf{u}), G(r, \mathbf{w})\right) < \delta_F$$

(3) Let $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$. Let $\varepsilon > 0$, $T \in \mathbb{T}$. For the sake of tidiness, we'll overload the symbol $\|\cdot\|_T$ using it as a seminorm on both $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ and $\mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$. In the latter case, $\|A(p)\|_T = \max_{0 \le t \le T} \|A(p)(t)\|$, where $\|\cdot\|$ is the matrix norm on $\mathcal{B}^{m \times m}$. Let

$$\varepsilon' = \frac{1}{2}\left(\sqrt{\left(\|A(p)\|_T + \|F(p, \mathbf{u})\|_T\right)^2 + 4\varepsilon} - \|A(p)\|_T - \|F(p, \mathbf{u})\|_T\right)$$

Then $\exists \delta_F, \delta_A > 0$ $\exists T_F, T_A \in \mathbb{T}$ such that $\forall (r, \mathbf{v}) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$,

$$\begin{aligned} d_{T_F}\left((p, \mathbf{u}), (r.\mathbf{v})\right) < \delta_F &\quad\Rightarrow\quad d_T\left(F(p, \mathbf{u}), F(r.\mathbf{v})\right) < \varepsilon' \\ d_{T_A}\left((p, \mathbf{u}), (r.\mathbf{v})\right) < \delta_A &\quad\Rightarrow\quad d_T\left(A(p, \mathbf{u}), A(r.\mathbf{v})\right) < \varepsilon' \end{aligned}$$

Let $T' = \max\{T_A, T_F\}$ and take $(r, \mathbf{v}) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $d_{T'}((p, \mathbf{u}), (r.\mathbf{v})) < \min\{\delta_F, \delta_A\}$. Then,

$$\begin{aligned} d_T(A(p)F(p,\mathbf{u}), A(r)F(r,\mathbf{v})) &= \|A(p)F(p,\mathbf{u}) - A(r)F(r,\mathbf{v})\|_T \\ &\leq \|A(p) - A(r)\|_T \|F(p,\mathbf{u}) - F(r.\mathbf{v})\|_T \\ &\quad + \|A(p) - A(r)\|_T \|F(p,\mathbf{u})\|_T \\ &\quad + \|A(p)\|_T \|F(p,\mathbf{u}) - F(r.\mathbf{v})\|_T \\ &< \varepsilon \end{aligned}$$

(4) Choose $\varepsilon > 0$, $T \in \mathbb{T}$, and $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$. Since F is continuous on $\{p\} \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, there is an open neighbourhood of $(p, \mathbf{u})$ in $P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $T \cdot d_T(F(p, \mathbf{u}), F(r, \mathbf{v})) < \varepsilon$ holds for all points, $(r, \mathbf{v})$, in the neighbourhood. Reusing several steps from the proof of the Building Block Lemma (6b), it is easy to show that $d_T(F(p, \mathbf{u}), F(r, \mathbf{v})) \leq T \cdot d_T((p, \mathbf{u}), (r, \mathbf{v}))$.

□

Remark 3.2.3. The Building Block Lemma and the Continuity Lemma naturally complement Theorems TZ1 and TZJ2 (on page 19 and page 27), respectively. The former suggests ways to construct operators that satisfy Theorem TZ1 and the latter merely assures us that there will be no unpleasant surprises when we hope for them to satisfy Theorem TZJ2. The key observation here is that most operators we might build from these theorems—starting with the identity operator as our foundation—will satisfy only ***WCaus*** and ***Lip***(λ, τ) for some $\lambda \geq 1$. There are only two operations in the list that can be applied to modify such an operator into one which will satisfy ***Caus*** and ***Lip***(λ, τ) for a $\lambda < 1$:

- integration, and
- multiplication by a matrix stream $A(t)$ that begins at $\mathbf{0}$ (at $t = 0$) and whose norm remains bounded by some $\lambda < 1$.

Remark 3.2.4. This suggests the following class of operators, at least as a starting point.

Corollary 3.2.5. *Let (P, d_P) be a metric space (of parameter values), let $p \in P$, and let $V \subseteq P$ be a neighbourhood of p. Let $\boldsymbol{y} : P \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ be continuous at p. Let $A, B : P \to \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$ be functions such that*

- *A and B are continuous at p*
- $\forall r \in V \;\; B(r)(0) = \mathbf{0} \in \mathcal{B}^{m \times m}$

- $\exists M_A, M_B \in \mathbb{R}^+ \; \forall t \in \mathbb{T} \; \forall r \in V,$
  $\|A(r)(t)\| \leq M_A$ *and* $\|B(r)(t)\| \leq M_B < 1$

Define $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*, as follows for* $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*,* $r \in P$*:*

$$F(r, \mathbf{u})(t) = \boldsymbol{y}(r)(t) + B(r)(t)\mathbf{u}(t) + \int_0^t A(r)(s)\mathbf{u}(s)\, ds$$

Then for every $r \in V$*, the function* $F(r, \cdot) : \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ *has a fixed point* $\mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*, and its fixed point function* $\Phi : V \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ *(as described in (1.3.3) on page 15) is continuous at* p*.*

Remark 3.2.6. Since $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ and $\mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$ are closed under their various algebraic operations and since integration is linear, a single $\boldsymbol{y}$, B, and A are clearly sufficient here (e.g. the sum of two constant streams $\boldsymbol{y}_1(r)(t) + \boldsymbol{y}_2(r)(t)$ could obviously be expressed using a single constant stream $\boldsymbol{y}(r)(t)$ and likewise for the other terms). Nested integrals, however, cannot be simplified into a single integral. So Corollary 3.2.5 can be generalized further in the following way.

Corollary 3.2.7. *Let* (P, d_P)*,* p*,* V*,* $\boldsymbol{y}$*, and* B *be as defined in Corollary 3.2.5. Let* $n \in \mathbb{Z}^+$ *and let* $A_1, A_2, \ldots, A_n$ *all be as* A *is defined (all continuous at* p *and all having bounded norms throughout* V *and* $\mathbb{T}$*). Then the same results can be obtained by defining* F *as follows (using the notation introduced in Remark 3.1.4):*

$$F(r, \mathbf{u})(t) \;\; = \;\; \boldsymbol{y}(r)(t) + B(r)(t)\mathbf{u}(t) + \sum_{k=1}^{n} \int_0^{(k)} \left(A_k(r)\mathbf{u}\right)(t)$$

Remark 3.2.8. Corollary 3.2.7 is the most general result that can be obtained directly (and exclusively) from the Building Block Lemma and the Continuity Lemma, but with a bit of extra work, we can go further to tackle infinite series of nested integrals instead of merely finite sums of them.

Lemma 3.2.9. *Let*[4] $M \in \mathbb{R}^+$ *and let* $A_1, A_2, A_3, \ldots \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$ *be a sequence of matrix streams such that* $\forall t \in \mathbb{T} \; \forall k \in \mathbb{Z}^+ \; \|A_k(t)\| \leq M$*. Then the following operator is well-defined on* $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$*:*

$$F(\mathbf{u})(t) \;\; = \;\; \sum_{k=1}^{\infty} \int_0^{(k)} \left(A_k\mathbf{u}\right)(t) \qquad (3.2.7)$$

[4]In fact, the lemma holds if M is any function of the form $M : \mathbb{T} \to \mathbb{R}^+$, but we can't make use of this generality here and it becomes merely inconvenient for our purposes.

Proof. For $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, $n \in \mathbb{Z}^+$, and $t \in \mathbb{T}$, define the partial sum F_n as follows:

$$F_n(\mathbf{u})(t) \;=\; \sum_{k=1}^{n} \int_0^{(k)} (A_k\mathbf{u})(t)$$

Then by Lemma 3.1.5, for any $T \in \mathbb{T}$, and any $N > 0$ and $n > N$,

$$\begin{aligned}
d_T(F_n(\mathbf{u}), F_N(\mathbf{u})) &= \|F_n(\mathbf{u}) - F_N(\mathbf{u})\|_T \\
&= \left\| \sum_{k=N+1}^{n} \int_0^{(k)} (A_k\mathbf{u}) \right\|_T \\
&\leq \sum_{k=N+1}^{n} \left\| \int_0^{(k)} (A_k\mathbf{u}) \right\|_T \\
&= \sum_{k=N+1}^{n} \max_{0\leq t\leq T} \left\| \int_0^{(k)} (A_k\mathbf{u})(t) \right\| \\
&\leq M\,\|\mathbf{u}\|_T \sum_{k=N+1}^{n} \max_{0\leq t\leq T} \frac{t^k}{k!} \\
&= M\,\|\mathbf{u}\|_T \sum_{k=N+1}^{n} \frac{T^k}{k!}
\end{aligned}$$

Given $\mathbf{u}$ and T, this distance can be made arbitrarily small by making N sufficiently large (and keeping $n > N$). Thus, for every $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, $\{F_n(\mathbf{u})\}_{n=1}^{\infty}$ is a locally uniform Cauchy sequence and since $\mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ is complete, $\lim_{n\to\infty} F_n(\mathbf{u})$ exists (and hence defines $F(\mathbf{u})$). ☐

Lemma 3.2.10. *For all $\lambda > 0$ the operator $F : \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ defined in (3.2.7) on the preceding page satisfies* ***Caus*** *and* ***Lip***(λ, τ) *with $\tau = \frac{\lambda}{M+\lambda}$ (where M is the upper bound for $\|A_k(t)\|$ indicated in Lemma 3.2.9).*

Proof. By the Building Block Lemma, Parts (1b), (2b), (3b), and (6a), $F_n \in$ ***Caus*** for every $n \in \mathbb{Z}^+$. Locally uniform convergence implies pointwise convergence, so if for some $T \in \mathbb{T}$, $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ $\forall n \in \mathbb{Z}^+$ $F_n(\mathbf{u})(T) = F_n(\mathbf{v})(T)$, then the same is true of the limits of each side of the equation as well. This is true whether $T = 0$ or $T > 0$. Thus, it follows that the limit F also satisfies ***Caus***.

Let $T \in \mathbb{T}$ and $\mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ such that $d_T(\mathbf{u}, \mathbf{v}) = 0$. Let $\lambda > 0$ and $\tau = \frac{\lambda}{M+\lambda}$. Then by the Building Block Lemma (1b), (3c), (5d), and (6b), $\forall n \in \mathbb{Z}^+$ the operator $\mathbf{u} \mapsto \int_0^{(n)} (A_n\mathbf{u})$ satisfies ***Lip***$(\tau^n M, \tau)$. Thus, using (2c),

$$F_n \in \boldsymbol{Lip}\left(M \sum_{k=1}^{n} \tau^k, \tau\right)$$

Hence, $\forall \mathbf{u}, \mathbf{v} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \ \forall T \in \mathbb{T}$, if $d_T(\mathbf{u}, \mathbf{v}) = 0$ then

$$d_{T+\tau}\left(F_n(\mathbf{u}), F_n(\mathbf{v})\right) \leq M \sum_{k=1}^{n} \tau^k d_{T+\tau}(\mathbf{u}, \mathbf{v})$$

Now, $d_{T+\tau}\left(F(\mathbf{u}), F(\mathbf{v})\right) = d_{T+\tau}\left(\lim_{n\to\infty} F_n(\mathbf{u}), \lim_{n\to\infty} F_n(\mathbf{v})\right)$, and since $d_{T+\tau}$ is continuous,

$$\begin{aligned} d_{T+\tau}\left(\lim_{n\to\infty} F_n(\mathbf{u}), \lim_{n\to\infty} F_n(\mathbf{v})\right) &= \lim_{n\to\infty} d_{T+\tau}\left(F_n(\mathbf{u}), F_n(\mathbf{v})\right) \\ &\leq M \sum_{k=1}^{\infty} \tau^k d_{T+\tau}(\mathbf{u}, \mathbf{v}) \\ &= \frac{M\tau}{1-\tau} d_{T+\tau}(\mathbf{u}, \mathbf{v}) \\ &= \lambda d_{T+\tau}(\mathbf{u}, \mathbf{v}) \end{aligned}$$

□

Remark 3.2.11. Lemmas 3.2.9 and 3.2.10 offer conditions sufficient to guarantee a fixed point for integral series operator (3.2.7) using Theorem TZ1. We now wish to augment the domain of this operator with a parameter space and determine a (ideally modest) set of conditions to be imposed on the matrix streams $\{A_n\}_1^\infty$ to ensure such an operator is continuous at a given point in its parameter space (the main requirement demanded by Theorem TZJ2).

Lemma 3.2.12 (The Equicontinuity Lemma). *Let $(X, \mathfrak{T}_X)$ be a topological space (where $\mathfrak{T}_X$ is the topology on X) and (Y, d_Y) be a metric space. Let $\{f_n\}_{n=1}^\infty$ be a sequence of functions $f_n : X \to Y$ that converges pointwise to a function $f : X \to Y$. If $\{f_n\}_{n=1}^\infty$ is equicontinuous at a point $x \in X$, then f is continuous at x.*

Proof. If $\{f_n\}_{n=1}^\infty$ is equicontinuous at x, then $\exists \delta_x : \mathbb{R}^+ \to \mathfrak{T}_X$ such that $\forall \varepsilon > 0$ $x \in \delta_x(\varepsilon)$ and $\forall n \in \mathbb{Z}^+ \ \forall y \in X \ y \in \delta_x(\varepsilon) \Rightarrow d_Y\left(f_n(x), f_n(y)\right) < \varepsilon$. Since f_n converges pointwise to f, $\exists N : X \times \mathbb{R}^+ \to \mathbb{N}$ such that $\forall y \in X \ \forall \varepsilon > 0$ $\forall k > N(y, \varepsilon) \ d(f_k(y), f(y)) < \varepsilon$. Let $\varepsilon > 0$. Let $y \in \delta(\varepsilon/3)$. Choose any $n > \max\{N(x, \varepsilon/3), N(y, \varepsilon/3)\}$. Then,

$$\begin{aligned} d_Y\left(f(x), f(y)\right) &\leq d_Y\left(f(x), f_n(x)\right) + d_Y\left(f_n(x), f_n(y)\right) + d_Y\left(f_n(y), f(y)\right) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

□

Lemma 3.2.13. *Let (P, d_P) be a metric space and let $\{A_n\}_{n=1}^{\infty}$ be a sequence of functions $A_n : P \to \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m\times m}$. For each $n \in \mathbb{Z}^+$ define $H_n : P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ as $H_n(r, \mathbf{v})(t) = (A_n(r)(t))\,(\mathbf{v}(t))$ (pointwise matrix multiplication). If $\{A_n\}_{n=1}^{\infty}$ are equicontinuous at a point $p \in P$ and $\exists M : \mathbb{T} \to \mathbb{R}^+$ such that $\forall T \in \mathbb{T}\ \forall n \in \mathbb{Z}^+\ \|A_n\|_T \le M(T)$, then $\{H_n\}_{n=1}^{\infty}$ are equicontinuous at $(p, \mathbf{u})$ for every $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$.*

Proof. Let $\delta_A : \mathbb{R}^+ \times \mathbb{T} \to \mathbb{R}^+$ be the modulus of continuity for $\{A_n\}_{n=1}^{\infty}$ at p. That is, $\forall \varepsilon > 0\ \forall T \in \mathbb{T}\ \forall n \in \mathbb{Z}^+ \forall r \in P$,

$$d_P(r, p) < \delta_A(\varepsilon, T) \Rightarrow \|A_n(p) - A_n(r)\|_T < \varepsilon$$

We can then derive a modulus of continuity for $\{H_n\}_{n=1}^{\infty}$ using only δ_A, $M(T)$, and a stream $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ (and, in particular, *not* using n) by following the proof of the Continuity Lemma part (3), taking F to be the projection function $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, defined for $(r, \mathbf{v}) \in P \times \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ as $F(r, \mathbf{v}) = \mathbf{v}$. Specifically, given $\varepsilon > 0$, $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$, and $T \in \mathbb{T}$, we take

$$\varepsilon' = \frac{1}{2}\left(\sqrt{\left(M(T) + \|\mathbf{u}\|_T\right)^2 + 4\varepsilon} - M(T) - \|\mathbf{u}\|_T\right)$$

(cf. proof of the Continuity Lemma (3.2.2), part (3)). Then define $\delta(\varepsilon, T, \mathbf{u}) = \min\{\varepsilon', \delta_A(\varepsilon', T)\}$. $\square$

Lemma 3.2.14. *Let (P, d_P) be a metric space and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions $f_n : P \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ which is equicontinuous at every point in some set $Q \subseteq P$. Define $F_n : P \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ as follows for $r \in P$, $n \in \mathbb{Z}^+$, and $t \in \mathbb{T}$:*

$$F_n(r) \;=\; \sum_{k=1}^{n} \int_0^{(k)} (f_k(r))\,(t)$$

Then $\{F_n\}_{n=1}^{\infty}$ is equicontinuous on Q.

Proof. Since $\{f_n\}_{n=1}^{\infty}$ is equicontinuous on Q, there is a modulus of continuity function $\delta_f : \mathbb{R}^+ \times \mathbb{T} \times Q \to \mathbb{R}^+$ such that $\forall \varepsilon > 0\ \forall n \in \mathbb{Z}^+\ \forall T \in \mathbb{T}\ \forall q \in Q\ \forall p \in P\ d_P(p, q) < \delta_f(\varepsilon, T, q) \Rightarrow d_T(f_n(p), f_n(q)) < \varepsilon$. Define $\delta(\varepsilon, T, p) = \delta_f(e^{-T}\varepsilon, T, p)$. Then for any $T \in \mathbb{T}$, $n \in \mathbb{Z}^+$, and $q \in Q$ such that $d_P(p, q) <$

$\delta(\varepsilon, T, p)$,

$$
\begin{aligned}
d_T\left(F_n(p), F_n(q)\right) &= \left\| \sum_{k=1}^{n} \int_0^{(k)} \left(f_k(p)\right) - \sum_{k=1}^{n} \int_0^{(k)} \left(f_k(q)\right) \right\|_T \\
&= \left\| \sum_{k=1}^{n} \int_0^{(k)} \left(f_k(p) - f_k(q)\right) \right\|_T \\
&\leq \sum_{k=1}^{n} \left\| \int_0^{(k)} \left(f_k(p) - f_k(q)\right) \right\|_T \\
&= \sum_{k=1}^{n} \max_{0 \leq t \leq T} \left\| \int_0^{(k)} \left(f_k(p) - f_k(q)\right)(t) \right\| \\
&\leq \sum_{k=1}^{n} \max_{0 \leq t \leq T} \frac{t^k}{k!} \max_{0 \leq s \leq t} \left\| \left(f_k(p) - f_k(q)\right)(s) \right\| \qquad (3.2.8) \\
&= \sum_{k=1}^{n} \max_{0 \leq t \leq T} \frac{t^k}{k!} d_t\left(f_k(p), f_k(q)\right) \\
&\leq \sum_{k=1}^{n} \frac{T^k}{k!} d_T\left(f_k(p), f_k(q)\right) \\
&< e^{-T} \varepsilon \sum_{k=1}^{n} \frac{T^k}{k!} \\
&< e^{-T} \varepsilon \sum_{k=1}^{\infty} \frac{T^k}{k!} \\
&= e^{-T} \varepsilon e^{T} = \varepsilon
\end{aligned}
$$

The inequality in (3.2.8) is from Lemma 3.1.5 on page 54. ☐

Theorem 3.2.15 (The General Form Theorem)**.** *Let (P, d_p) be a metric space (of parameters). Let $V \subseteq P$ be a neighbourhood of a point $p \in P$. Let $\boldsymbol{y} : P \to \mathcal{C}[\mathbb{T}, \mathcal{B}]^m$ be continuous at p. Let $B, A_1, A_2, \ldots : P \to \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$ be functions such that*

- *B is continuous at p, and $\{A_n\}_{n=1}^{\infty}$ are equicontinuous at p,*
- *$\forall r \in V\ \ B(r)(0) = \mathbf{0} \in \mathcal{C}[\mathbb{T}, \mathcal{S}]^{m \times m}$, and*
- *$\exists M_A, M_B \in \mathbb{R}^+\ \forall r \in V\ \forall t \in \mathbb{T}\ \forall n \in \mathbb{Z}^+\ \|A_n(r)(t)\| \leq M_A$ and $\|B(r)\| \leq M_B < 1$*

Define $F : P \times \mathcal{C}[\mathbb{T},\mathcal{B}]^m \to \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ as follows for $r \in P$, $\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$, and $t \in \mathbb{T}$:

$$F(r,\mathbf{u})(t) \quad = \quad \boldsymbol{y}(r)(t) + B(r)(t)\mathbf{u}(t) + \sum_{k=1}^{\infty} \int_0^{(k)} \left(A_k(r)\mathbf{u}\right)(t) \qquad (3.2.9)$$

For each $r \in P$ define $F_r : \mathcal{C}[\mathbb{T},\mathcal{B}]^m \to \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ as $F_r(\mathbf{u}) = F(r,\mathbf{u})$.
Then for each $r \in P$, F_r has a unique fixed point $\Phi(r)$, and the fixed point function $\Phi : V \to \mathcal{C}[\mathbb{T},\mathcal{B}]^m$ for F is continuous at p.

Proof. First we'll show that $\forall r \in P$ F_r satisfies ***Caus*** and ***Contr***. Theorem TZ1 on page 19 informs us that these conditions are sufficient to guarantee that F_r has a unique fixed point for all $r \in P$. Finally we show that for every $\mathbf{u} \in \mathcal{C}[\mathbb{T},\mathcal{B}]^m$, F is continuous at $(p,\mathbf{u})$, and thus, Theorem TZJ2 on page 27 provides the conclusion.

[$F_r \in$ ***Caus***] Lemma 3.2.9 establishes the fact that $F_r(\mathbf{u})$ converges to a stream for all $(r,\mathbf{u}) \in P \times \mathcal{C}[\mathbb{T},\mathcal{B}]^m$. Using the Building Block Lemma, Parts (1a), (1b), (3b), (5a), (5c), and (3.2.10), we find that each of the main three terms satisfies ***Caus*** for any fixed $r \in P$. Part (2b) assembles them to show that F_r, itself, satisfies ***Caus***.

[$F_r \in$ ***Contr***] Let $\lambda_\Sigma = \frac{1-M_B}{2}$ and $\tau = \frac{\lambda_\Sigma}{M_A+\lambda_\Sigma}$. From (1a) the first term of F_r satisfies ***Lip***$(0,\tau)$. From (1b) and (5c) the second term satisfies ***Lip***(M_B,τ). From Lemma 3.2.10, the third term (the summation) satisfies ***Lip***(λ_Σ,τ). Putting the three results together, we conclude from (2c) that for all $r \in P$, F_r satisfies ***Lip***(λ,τ) with $\lambda = 0 + M_B + \lambda_\Sigma = \frac{1+M_B}{2} < 1$.

[F continuous at $(p,\mathbf{u})$] By Lemma 3.2.13, the set of integrands is equicontinuous at every point in the set $Q = \{p\} \times \mathcal{C}[\mathbb{T},\mathcal{B}]^m$. Thus, by Lemma 3.2.14, the set of partial sums $\left\{\sum_{k=1}^{n} \int_0^{(k)} \left(A_k(r)\mathbf{u}\right)(t)\right\}_{n=1}^{\infty}$ is equicontinuous at every point in Q. Since the series converges pointwise, the Equicontinuity Lemma then asserts that its limit is continuous at every point of Q. It is then trivial to use the Continuity Lemma to show that F is continuous at every point in Q, and hence by Theorem TZJ2, Φ is continuous at p. □

3.3 Applications

3.3.1 The Mass-Spring-Damper System Revisited

3.3.1.1 Case Study 1

The simple mass-spring-damper system (see Figure 3.3.1) was introduced in [TZ07] as an analog network case study. The system is typically expressed as a second-order, homogeneous ODE with constant coefficients:

$$M\ddot{x}(t) + D\dot{x}(t) + Kx(t) \quad = \quad f(t)$$

where M is the mass, D is the damping coefficient, K is the spring constant, f is the forcing function, and x is the displacement. The initial conditions are given as

$$\begin{aligned} x(0) &= x_0 \in \mathbb{R} \ \text{(initial displacement)} \\ \dot{x}(0) &= v_0 \in \mathbb{R} \ \text{(initial velocity)} \end{aligned}$$

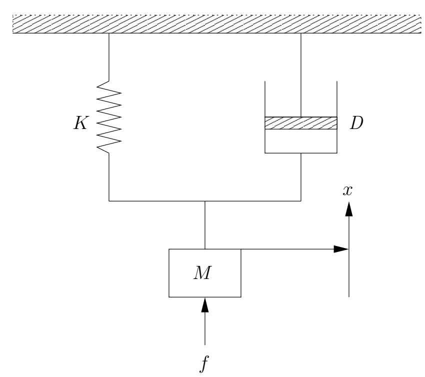

Figure 3.3.1: Mass-Spring-Damper System

It is typical to reduce the second-order equation to a first-order system using the substitutions $v(t) = \dot{x}(t)$ and $a(t) = \dot{v}(t)$. Integrating this system with respect to t and solving for the initial conditions gives us a system of integral equations equivalent to the original initial value problem:

$$\begin{aligned} a(t) &= \frac{f(t) - Dv(t) - Kx(t)}{M} \\ v(t) &= \int_0^t a(s)\, ds + v_0 \\ x(t) &= \int_0^t v(s)\, ds + x_0 \end{aligned}$$

This system is the mass-spring-damper system as it is represented in [TZ07] and [TZ11] as their first case study. For each parameter choice $p = (M, K, D, v_0, x_0, f)$, it induces the operator $F_p : \mathcal{C}[\mathbb{T}, \mathbb{R}]^3 \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^3$ defined for $\mathbf{u}(t) = (a, v, x)^\intercal (t)$ as

$$F_p \begin{bmatrix} a \\ v \\ x \end{bmatrix} (t) = \begin{bmatrix} \frac{1}{M}\left(f(t) - Dv(t) - Kx(t)\right) \\ \int_0^t a(s)\, ds + v_0 \\ \int_0^t v(s)\, ds + x_0 \end{bmatrix} \tag{3.3.1}$$

A fixed point of this operator represents both a solution the original initial value problem (with the given parameters), and the semantics for the analog network shown in Figure 3.3.2 (which is a slightly less formal version of the one used by Tucker and Zucker):

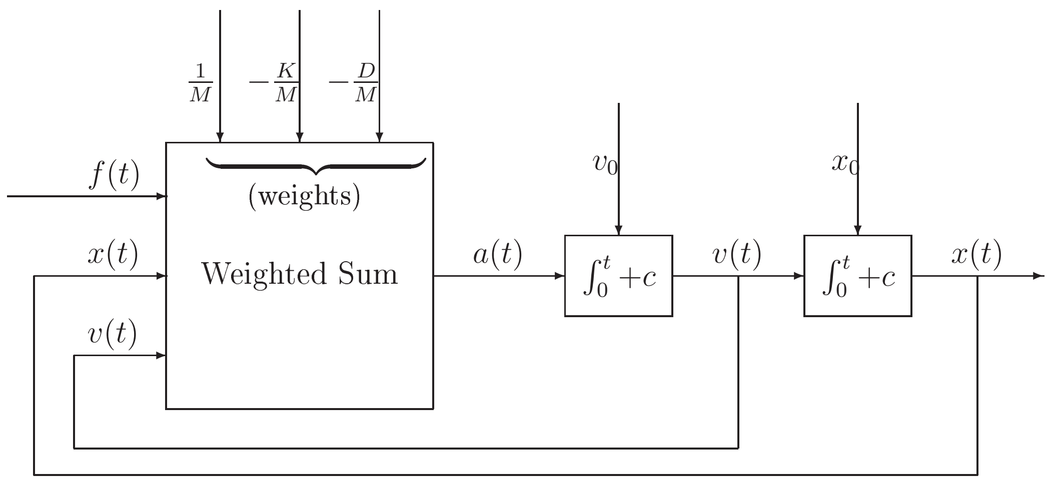

Figure 3.3.2: Analog Network for Simple Mass-Spring-Damper System

Tucker and Zucker prove that this operator F_p satisfies the ***Contr*** condition if $M > \max\{K, 2D\}$, and hence their theory guarantees the existence of a fixed point under that condition. It is unclear whether this is a necessary condition, however, but while it may be weakened to some degree, it cannot be disposed altogether, as the following example demonstrates.

Example 3.3.1. Let $T \geq 0$. Take the constants $M = D = K = 1$, $v_0 = x_0 = 0$, and let $\mathbf{u}_1 = (a_1, v_1, x_1)^\intercal, \mathbf{u}_2 = (a_2, v_2, x_2)^\intercal$ be stream tuples such that $x_1 = x_2 = a_1 = a_2$ (for all time), and $v_1\restriction_{[0,T]} = v_2\restriction_{[0,T]}$ but $\exists t \in (T, T+\tau]$ such that $v_1(t) \neq v_2(t)$. For convenience, write $(a_i', v_i', x_i')^\intercal = F_p(\mathbf{u}_i)$ for $i = 1, 2$.

Then for any $\tau > 0$, $\lambda < 1$, and any input stream f,

$$\begin{aligned}
d_{T+\tau}(F_p\mathbf{u}_1, F_p\mathbf{u}_2) &= \max\{d_{T+\tau}(a_1', a_2'),\ d_{T+\tau}(v_1', v_2'),\ d_{T+\tau}(x_1', x_2')\} \\
&\geq d_{T+\tau}(a_1', a_2') \\
&= d_{T+\tau}\left(\frac{f - Dv_1 - Kx_1}{M}, \frac{f - Dv_2 - Kx_2}{M}\right) \\
&= d_{T+\tau}\left((f - v_1 - x_1), (f - v_2 - x_2)\right) \\
&= \max_{0 \leq t \leq T+\tau} |(f(t) - v_1(t) - x_1(t)) - (f(t) - v_2(t) - x_2(t))| \\
&= \max_{T \leq t \leq T+\tau} |(f(t) - v_1(t) - x_1(t)) - (f(t) - v_2(t) - x_2(t))| \\
&= \max_{T \leq t \leq T+\tau} |v_2 - v_1| \\
&= d_{T+\tau}(v_1, v_2) \\
&= \max\{d_{T+\tau}(v_1, v_2),\ 0,\ 0\} \\
&= \max\{d_{T+\tau}(v_1, v_2),\ d_{T+\tau}(x_1, x_2),\ d_{T+\tau}(a_1, a_2)\} \\
&= d_{T+\tau}(\mathbf{u}_1, \mathbf{u}_2) \\
&> \lambda d_{T+\tau}(\mathbf{u}_1, \mathbf{u}_2)
\end{aligned}$$

Thus, F_p with $p = (1, 1, 1, 0, 0, f)$ does not satisfy the ***Contr*** condition.

3.3.1.2 A More Robust Formulation

Example 3.3.1 shows that there *are* parameter values which cause Tucker and Zucker's model of the mass-spring-damper system to fail to satisfy ***Contr***, and hence also to fail to satisfy their special condition, $M > \max\{K, 2D\}$. In other words, the special condition is not simply an artifact of calculation (or an "idle threat," as it were); it does identify systems which do not satisfy ***Contr***. While somewhat disappointing, it is not completely unexpected that such systems would exist. In particular, it is conceivable to think we might see the ***Contr*** condition fail in regions of the parameter space in which the system behaves erratically or in which the system is most sensitive to parameter variation. Oddly enough, that does not appear to be the case.

Recall that there are three types of behaviour a mass-spring-damper system can exhibit (see [BD01], for example, or almost any elementary text on ordinary differential equations): overdamped, critically damped, and underdamped. An overdamped system behaves as if submerged in molasses—if the mass is displaced (and no other forcing function acts on it), it gradually and monotonically returns to the equilibrium position. A critically damped system monotonically returns to its equilibrium position as well, but as quickly as possible (like an optimized overdamped system). An underdamped system will oscillate with exponentially decreasing amplitude.

The value of the damping ratio $\zeta = D/\sqrt{4MK}$ determines which behaviour a system will exhibit. If $\zeta > 1$ the system is overdamped, if $\zeta < 1$ the system is underdamped, and if $\zeta = 1$ it is critically damped. Since the motion of an underdamped system is the least constrained, we might expect that if Tucker and Zucker's condition ($M > \max\{K, D\}$) is to fail, an underdamped system is where it would happen; and likewise, if it ever holds, surely it would hold for an overdamped system. In fact, for each type of behaviour there is a system which satisfies the special condition and a system which doesn't.

Example 3.3.2. Let $D \in \mathbb{R}^+$ and set $M = 3D$. Then

$$\zeta = \frac{D}{\sqrt{12DK}} = \sqrt{\frac{D}{12K}}$$

The system is overdamped if $K < {}^{D}/_{12}$, critically damped if $K = {}^{D}/_{12}$, and underdamped if $K > {}^{D}/_{12}$. As long as $K < 3D$ (which leaves plenty of wiggle room), we have $M > \max\{K, 2D\}$. So there are systems of every type which satisfy the condition.

Now let $K \in \mathbb{R}^+$, and let $M = K$. Then

$$\zeta = \frac{D}{2\sqrt{MK}} = \frac{D}{2K}$$

The system is overdamped if $D > 2K$, critically damped if $D = 2K$, and underdamped if $D < 2K$. Regardless of the value of D, $M \leq \max\{K, 2D\}$. So there are also systems of every type which do not satisfy the condition.

Fortunately, by simply making the acceleration stream implicit, we can rearrange the system into an equivalent one that satisfies the ***Contr*** condition for any choice of $M, K, D > 0$ (so while the special condition was not merely an artifact of calculation, it was only an idiosyncracy of that particular model of the system).

Define the operator $G : P \times \mathcal{C}[\mathbb{T}, \mathbb{R}]^2 \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$ as follows for $p = (M, K, D, v_0, x_0, f) \in (\mathbb{R}^+)^3 \times \mathbb{R}^2 \times \mathcal{C}[\mathbb{T}, \mathbb{R}] = P$ and $(v, x)^\intercal \in \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$ (cf. (3.3.1)):

$$G\left(p, \begin{bmatrix} v \\ x \end{bmatrix}\right)(t) = \begin{bmatrix} \frac{1}{M}\int_0^t (f(s) - Dv(s) - Kx(s))\, ds + v_0 \\ \int_0^t v(s)\, ds + x_0 \end{bmatrix} \tag{3.3.2}$$

For convenience, we'll use the notation $G_p(\mathbf{u}) = G(p, \mathbf{u})$ for $\mathbf{u} \in \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$ and $p \in P$.

The corresponding network is shown in Figure 3.3.3. We will now show that G satisfies the conditions demanded of F from the General Form Theorem. Define

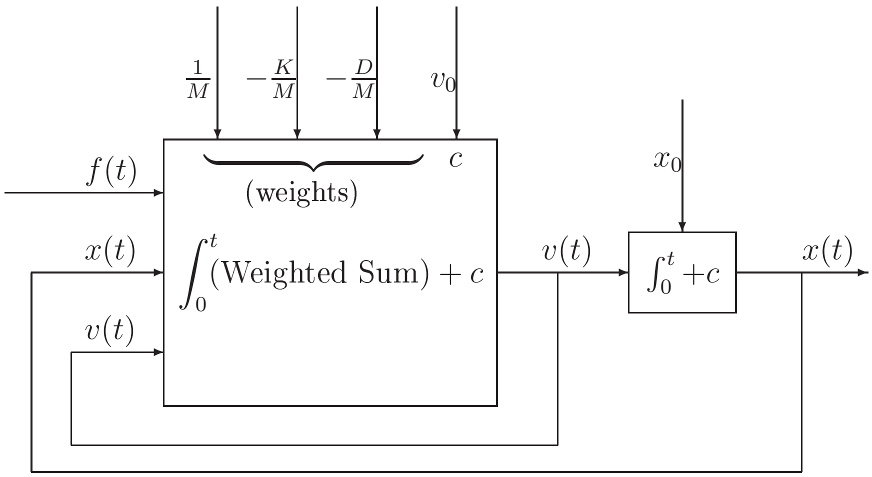

Figure 3.3.3: Revised Mass-Spring-Damper Network

$$A_1(p)(t) = \begin{bmatrix} -\frac{D}{M} & -\frac{K}{M} \\ 1 & 0 \end{bmatrix} \text{ and } \boldsymbol{y}(p)(t) = \begin{bmatrix} \frac{1}{M}\int_0^t f(s)\,ds + v_0 \\ x_0 \end{bmatrix}$$

Let all the other matrices from the General Form Theorem (B and A_k for $k = 2, 3, \ldots$) be zero. Rewrite Equation (3.3.2) as follows to put it in the form of (3.2.9):

$$G\left(p, \begin{bmatrix} v \\ x \end{bmatrix}\right)(t) \quad = \quad \boldsymbol{y}(p)(t) + \int_0^{(1)} \left(A_1(p)\begin{bmatrix} v \\ x \end{bmatrix}\right)(t)$$

It is relatively straightforward to show that $\boldsymbol{y}$ and A_1 are continuous on P—and hence, on any neighbourhood $V \subseteq P$ of p. So take V to be the open ball of radius $\frac{M}{2}$, centred at $p = (M, K, D, v_0, x_0, f)$. More precisely,

$$\begin{aligned} V \quad = \quad & P \cap \left(\frac{M}{2}, \frac{3M}{2}\right) \times \left(K - \frac{M}{2}, K + \frac{M}{2}\right) \times \left(D - \frac{M}{2}, D + \frac{M}{2}\right) \\ & \times \left(v_0 - \frac{M}{2}, v_0 + \frac{M}{2}\right) \times \left(x_0 - \frac{M}{2}, x_0 + \frac{M}{2}\right) \\ & \times \left\{ g \in \mathcal{C}[\mathbb{T}, \mathbb{R}] \,:\, d_C(f, g) < \frac{M}{2} \right\} \end{aligned}$$

Let $M_A = 1 + \frac{2D+2K}{M}$. Then, as required by the General Form Theorem,

$\forall p' = (M', K', D', v_0', x_0', f') \in V \ \forall t \in \mathbb{T}$,

$$\begin{aligned} \|A_1(r)(t)\| &= \sup\left\{ \left\| \begin{bmatrix} -\frac{D'}{M'} & -\frac{K'}{M'} \\ 1 & 0 \end{bmatrix} \mathbf{u} \right\| : \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{R}^2 \text{ and } \|\mathbf{u}\| \leq 1 \right\} \\ &= \max\left\{ \frac{D' + K'}{M'}, 1 \right\} \\ &\leq \max\left\{ \frac{D + \frac{M}{2} + K + \frac{M}{2}}{M - \frac{M}{2}}, 1 \right\} = M_A \end{aligned}$$

There is no matrix stream B, so using the proof of the General Form Theorem, a straightforward calculation reveals that $G \in \boldsymbol{Lip}(\lambda, \tau)$ for

$$\lambda = \frac{1}{2} \quad \text{and} \quad \tau = \frac{M}{4D + 4K + 3M}$$

The remaining antecedents of the General Form Theorem follow trivially for G. Hence, for every $p = (M, K, D, v_0, x_0, f)$, G_p has a unique fixed point $\Phi(p)$, and the corresponding fixed-point function

$$\Phi : \left(\mathbb{R}^+\right)^3 \times \mathbb{R}^2 \times \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$$

for G is also continuous.

The characterization of G as a "formulation" of F is justified by the fact that any fixed point of G_p uniquely specifies a fixed point for F_p and vice versa. In particular,

$$\begin{bmatrix} v \\ x \end{bmatrix} \text{ is a fixed point for } G_p \ \Leftrightarrow \ \begin{bmatrix} \frac{1}{M}(f - Dv - Kx) \\ v \\ x \end{bmatrix} \text{ is a fixed point for } F_p$$

Hence, as intuition would suggest, Tucker and Zucker's theory can indeed be applied to mass-spring-damper systems with any positive values for K, D, and M. Admittedly G is not strictly equivalent to F (being two-dimensional), but if an explicit acceleration stream is desired, it can introduced to the system like this:

$$\begin{aligned}
A'(p)(t) &= \begin{bmatrix} 0 & \frac{D^2}{M^2} - \frac{K}{M} & \frac{DK}{M^2} \\ 0 & -\frac{D}{M} & -\frac{K}{M} \\ 0 & 1 & 0 \end{bmatrix} \\
\boldsymbol{y}'(p)(t) &= \begin{bmatrix} \frac{1}{M}\left(f(t) - \frac{D}{M}\int_0^t f(s)\,ds - v_0 - x_0\right) \\ \frac{1}{M}\int_0^t f(s)\,ds + v_0 \\ x_0 \end{bmatrix} \\
F'_p\begin{bmatrix} a \\ v \\ x \end{bmatrix}(t) &= \int_0^t A'(p)(s)\begin{bmatrix} a \\ v \\ x \end{bmatrix}(s)\,ds + \boldsymbol{y}'(p)(t)
\end{aligned}$$

Alternatively, we could skip the order-reduction step and simply integrate the original ODE twice with respect to t, solving for the constants of integration using the initial conditions to yield

$$\begin{aligned}
x(t) &= \frac{1}{M}\int_0^t \left(\int_0^s (f(r) - Kx(r))\,dr + Dx(s) + v_0\right) ds + x_0 \\
&= \frac{1}{M}\int_0^t \left(-K\int_0^s x(r)\,dr + Dx(s)\right) ds + \frac{1}{M}\left(\int_0^t\int_0^s f(r)\,dr\,ds + tv_0\right) + x_0
\end{aligned}$$

In this case we use 1×1 “matrices,” setting

$$\begin{aligned}
A_1(p) &= \frac{D}{M} \\
A_2(p) &= -\frac{K}{M} \\
B(p) &= A_3 = A_4 = \cdots = 0 \\
\boldsymbol{y}(p)(t) &= \frac{1}{M}\left(\int_0^{(2)} f(t) + tv_0\right) + x_0
\end{aligned}$$

Finally, returning to the issue of molasses-submerged systems and similarly whimsical contrivances (along with more practical ones), observe that the matrices employed in this application have made no use of the dimension of time, which is built into the model. Thus, K, D, and M can be made to vary smoothly over time if, for example, one wishes to model such systems as a mass-spring-damper in a medium of varying viscosity and/or temperature.

3.3.1.3 Case Study 2

The second case study in [TZ07] involves a coupled mass-spring-damper system: two MSD systems with one connected to the mass of the other. The authors derive a similar system of integral equations (with two of everything involved in Case Study 1) and determine that the system satisfies ***Contr*** as long as $M_1 > max(2K_1, 2D_1)$ and $M_2 > max(2K_1 + 2K_2, 2D_2)$. Fortunately, this can be modified in the same way as Case Study 1 to yield an equivalent system that satisfies ***Contr*** for any parameter values. Just as in the simpler version, the corresponding parametrized operator is continuous, and hence, Theorem TZJ2 can be applied to it to obtain a continuous fixed-point function $\Phi : (\mathbb{R}^+)^6 \times \mathbb{R}^4 \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^k$ (where k can be chosen to be 2, 4, or 6, depending on whether acceleration and velocity are to be explicitly represented by streams).

3.3.2 Simple Pendulum

The simple, frictionless pendulum with a single, rigid arm, constrained to move within a vertical plane is another staple of elementary mechanics. It is represented using the following second-order ODE (see [Ach97]):

$$\ddot{\theta}(t) \quad = \quad -\frac{g}{\ell} \sin(\theta(t)) \tag{3.3.3}$$

where $\theta(t)$ is the angle formed by the bob and its equilibrium position at time t, g is the gravitational constant, and ℓ is the length of the arm. Using the order-reduction trick from the last example, let $\phi = \dot{\theta}$. Then (3.3.3) can be represented by the following equivalent system:

$$\begin{aligned} \phi(t) &= -\int_0^t \frac{g}{\ell} \sin(\theta(s))\, ds + \phi_0 \\ \theta(t) &= \int_0^t \phi(s)\, ds + \theta_0 \end{aligned}$$

Our parameter space is $P = \mathbb{R}^+ \times \mathbb{R}^2$ (condense g/ℓ into a single, positive parameter, leaving ϕ_0 and θ_0 as real numbers). In this case, the General Form Theorem is of no help at all since the sin function is nonlinear. We can, however, still use the Building Block Lemma directly and treat the sin function as a sort of magically-bestowed, primitive operator like the identity and the constant functions from (1a) and (1b) of the Building Block Lemma. Define[5] $G : P \times \mathcal{C}[\mathbb{T}, \mathbb{R}]^2 \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$ and $\boldsymbol{y} : P \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^2$ as follows:

$$G\left(p, \begin{bmatrix} \phi \\ \theta \end{bmatrix}\right)(t) = \begin{bmatrix} -\frac{g}{\ell}\sin(\theta(t)) \\ \phi(t) \end{bmatrix} \text{ and } \boldsymbol{y}(p)(t) = \begin{bmatrix} \phi_0 \\ \theta_0 \end{bmatrix}$$

[5] Note that G could instead be defined using the simpler form $G : P \times \mathbb{R}^2 \to \mathbb{R}^2$, but such a definition—while certainly more elegant here—introduces awkwardness in the next step.

We can then define $F : P \times \mathcal{C}[\mathbb{T},\mathbb{R}]^2 \to \mathcal{C}[\mathbb{T},\mathbb{R}]^2$ like this:

$$F\left(p, \begin{bmatrix} \phi \\ \theta \end{bmatrix}\right)(t) = \int_0^{(1)} G\left(p, \begin{bmatrix} \phi \\ \theta \end{bmatrix}\right)(t) + \boldsymbol{y}(p)(t)$$

While G is defined above as a stream operator, it actually uses only the current value of the input stream (see Footnote 5). Hence, it clearly satisfies ***WCaus***. It takes a bit of work to develop the formal details, but differentiating G_p with respect to ϕ and θ reveals that G_p satisfies $\boldsymbol{Lip}(\lambda_G, \tau)$ for any $\tau \in \mathbb{R}^+$ and $\lambda_G = \max\{1, {}^g\!/\!_\ell\}$. This is because the magnitude of the slope of the first component of G (with respect to θ rather than t) never exceeds ${}^g\!/\!_\ell$, and the slope of the second (with respect to ϕ) is always 1.

Thus, we can apply the Building Block Lemma to deduce that F_p satisfies ***Caus*** and $\boldsymbol{Lip}({}^1\!/\!_2, \tau)$ for $\tau = \frac{1}{2}\min\{1, {}^\ell\!/\!_g\}$. It is clear by inspection that G is continuous, and hence, by the Continuity Lemma, so is F. Hence, by Theorem TZJ2, so is the fixed point function for F.

The continuity of the fixed point may be somewhat surprising in this case since, for any ${}^g\!/\!_\ell$, there is a certain critical angular velocity (or position/velocity pair) which will be precisely the right amount to turn the bob upright and leave it there forever in its unstable equilibrium position. Even the slightest amount less and the bob falls back down on the side from which it approached the vertical. The slightest amount more, and it goes over the top, swinging back down on the other side. This would seem to represent a discontinuity at that point of critical velocity, but in fact, it doesn't. Theorem TZJ2 assures us of this, but it offers little in the way of insight.

What drives our perception of a discontinuity is the abrupt change in the *asymptotic* behaviour of the system in response to arbitrarily small changes in the initial conditions. Qualitatively speaking, there is a profound difference between a pendulum that falls back down and one that remains upright. What this observation fails to consider is the length of time the bob spends in a near-upright position. As the initial velocity approaches that critical value which leaves the pendulum upright forever, the bob spends more and more time in that very slow-moving limbo state in which it would appear to have an uncertain future.

Now consider this fact in light of the topology on $\mathcal{C}[\mathbb{T},\mathbb{R}]$. Increasingly large values of T must be used to encounter any significant difference (with respect to the pseudometrics d_T) between the trajectory of the perpetually upright bob, and those with sufficiently similar initial velocity. This phenomenon is plotted in Figure 3.3.4. Each curve corresponds to the trajectory of the bob, starting at $\theta(0) = 0$ (hanging straight down initially), with a certain initial velocity. The trajectory marked "$\approx$" is the one corresponding to the perfect amount of initial velocity to push the bob upright and leave it there forever. The

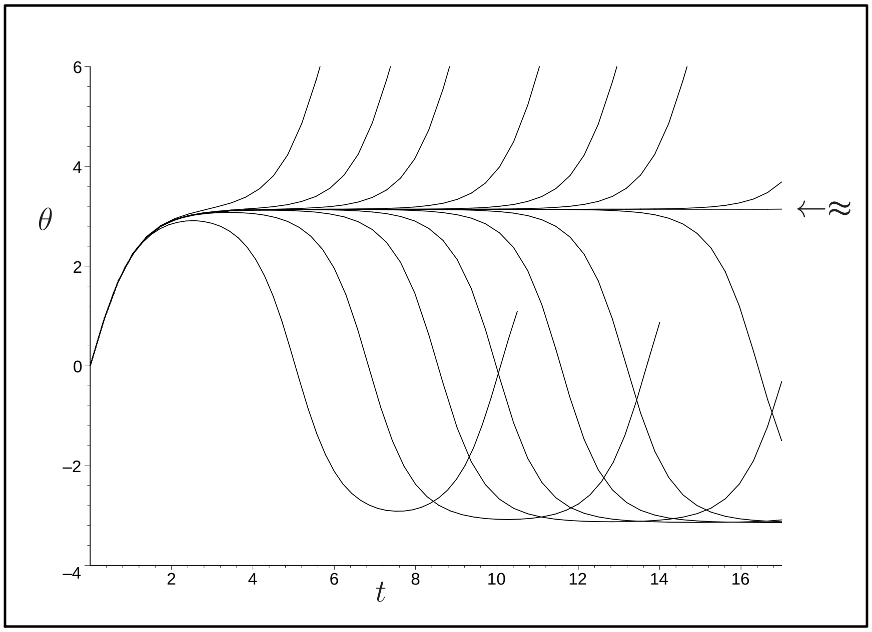

Figure 3.3.4: Pendulum trajectories approaching perfect equilibrium

trajectories that slope downward are produced by less initial velocity (some are truncated in the plot for the sake of clarity), and those which slope upward are produced by an excessive initial velocity, which pushes the bob over the top. The important thing to note is that—regardless of whether too little or too much initial velocity is involved—the time at which the non-upright trajectories distinguish themselves becomes later and later, the closer their initial velocity is to the critical value. From this, we may conclude that while instability likely always results in a (locally) smaller modulus of continuity, it does not necessarily imply actual discontinuity—i.e. the modulus of continuity will get very small around an unstable equilibrium point, but it may still remain strictly positive.

3.4 Future Work

3.4.1 Develop Building Blocks to Handle Cases Like the Pendulum

It is, of course, disappointing that for all the power of the General Form Theorem, it is still insufficient to handle an application as basic as the most simple pendulum from elementary mechanics. This is one price we pay for keeping our Banach space $\mathcal{B}$ distinct from its set of scalars $\mathcal{S}$. By letting $\mathcal{B} = \mathcal{S}$ (which holds for many common vector spaces anyway), we can introduce exponentiation, which in turn, allows for power series and hence trigonometric and exponential operators. This is not, by any means, a straightforward addition to the theory, however. Consider, for example, the prospect of including the rather modest, pointwise squaring operator to $\mathcal{C}[\mathbb{T},\mathbb{R}]$:

Example 3.4.1. Let $\mathbf{id}^2 : \mathcal{C}[\mathbb{T},\mathbb{R}] \to \mathcal{C}[\mathbb{T},\mathbb{R}]$ be defined as follows for $u \in \mathcal{C}[\mathbb{T},\mathbb{R}]$:

$$\mathbf{id}^2(u)(t) = (u(t))^2$$

$\mathbf{id}^2$ certainly satisfies $\boldsymbol{WCaus}$, but what about $\boldsymbol{Lip}(\lambda,\tau)$? Let $\lambda,\tau \in \mathbb{R}^+$ and consider the following two streams:

$$\begin{aligned} u(t) &= \frac{\lambda+1}{\tau}t \\ v(t) &= 0 \end{aligned}$$

Then $d_0(u,v) = 0$, but $d_{0+\tau}(\mathbf{id}^2(u),\mathbf{id}^2(v)) = (\lambda+1)^2 > \lambda(\lambda+1) = \lambda d_{0+\tau}(u,v)$. Thus, $\forall \lambda,\tau \in \mathbb{R}^+$ $\mathbf{id}^2 \notin \boldsymbol{Lip}(\lambda,\tau)$.

The problem here is ultimately due to the fact that the derivative of $f(x) = x^2$ is unbounded on $\mathbb{R}$. No matter how leniently we choose λ (and τ), we can always find a steep enough stream to deny $\mathbf{id}^2$ its coveted membership in the class of $\boldsymbol{Lip}(\lambda,\tau)$ operators. One way we might be able to circumvent this problem is by developing a nested exhaustion $\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \cdots \subseteq \mathcal{B}$ of the codomain of the streams. After all, given any $R,\tau \in \mathbb{R}^+$ and any $u,v \in \mathcal{C}[\mathbb{T},[-R,R]] \subseteq \mathcal{C}[\mathbb{T},\mathbb{R}]$, we see that

$$\begin{aligned} d_{T+\tau}(\mathbf{id}^2(u),\mathbf{id}^2(v)) &= \max_{0\le t\le T+\tau} \left|u^2(t) - v^2(t)\right| \\ &= \max_{0\le t\le T+\tau} |u(t)+v(t)|\,|u(t)-v(t)| \\ &\le 2R \max_{0\le t\le T+\tau} |u(t)-v(t)| \\ &= 2R d_{T+\tau}(u,v) \end{aligned}$$

Therefore, $\mathbf{id}^2$ can be said to satisfy the $\boldsymbol{Lip}(2R,\tau)$ condition on $\mathcal{C}[\mathbb{T},[-R,R]]$. Returning to the example of the pendulum, recall the Maclaurin series for $\sin(t)$:

$$\sin(t) = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}$$

Let s_n be the derivative of the n^{th} partial sum:

$$s_n(t) = \sum_{k=0}^{n} \frac{(-1)^k t^{2k}}{(2k)!}$$

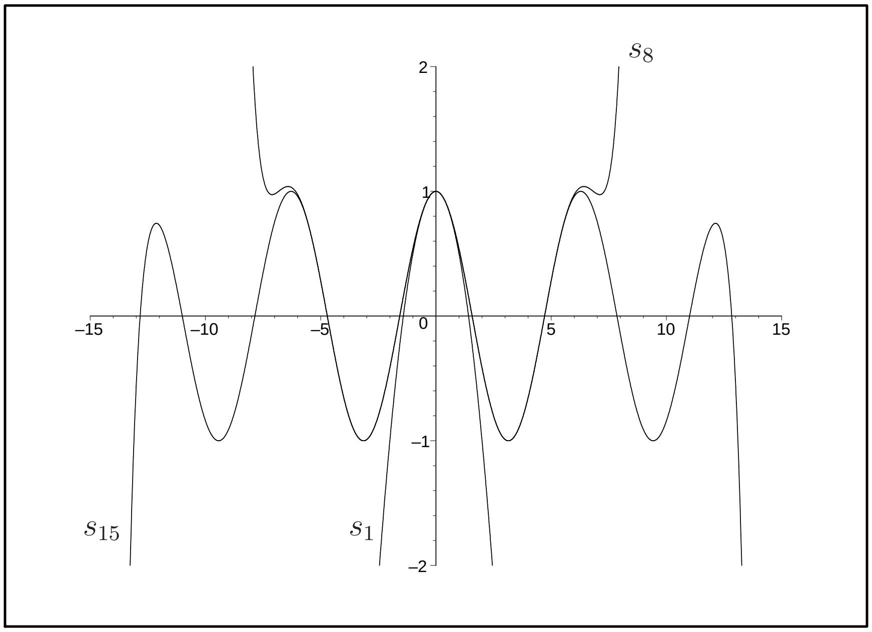

Figure 3.4.1: Partial sums in the Maclaurin expansion for $\cos(t)$

The preimage of any bounded interval centred at 0 (say, $[-1, 1]$) continues to expand as we examine successively larger partial sums. Turning our attention to Figure 3.4.1, we see that the preimage of $[-1, 1]$ under s_1 is $[-2, 2]$. Under s_8, it's roughly $[-6.1, 6.1]$, and by s_{15}, the preimage has expanded to approximately $[-13.1, 13.1]$. As $n \to \infty$, this preimage of $[-1, 1]$ under s_n approaches $\mathbb{R}$.

So while none of these partial sums satisfy the ***Lip*** condition on $\mathcal{C}[\mathbb{T}, \mathbb{R}]$, the operator to which they converge does. This observation offers some hope that with a bit of care, power series might be incorporated into the theory. The main value of doing so is in their tremendous versatility. We could, of course, simply throw specific analytic functions like sin into the theory individually, but it would be far more powerful (and elegant) to catch them all in a single net. Furthermore, it allows for greater generality. The sin function traditionally assumes only real or complex values, but its power series expansion could be used to define versions of it (along with several other functions) on more exotic spaces.

3.4.2 Other Lines of Inquiry

- A network model (or rather a mathematical model that can be interpreted as a network model) along with a system of fixed-point semantics is presented in [TZ11]. The authors describe a set of conditions sufficient to guarantee that the model operates properly. The abstractness of the model necessarily imposes a corresponding level of abstractness on these conditions. This chapter is meant to be a companion work in which some of that abstractness is sacrificed in an attempt to get closer to a more concrete, GPAC-like result—a result in which a tangible class of functions is identified that satisfies Tucker and Zucker's conditions. [TZ11] is, however, only the first of a two-part series, the second of which examines their model from the framework of computable analysis (specifically, the computable analysis covered in [TZ04]). Hence, a natural second step will be to follow [JZ12] with a corresponding entry that applies computable analysis to the Building Block and Continuity Lemmas, and to the General Form Theorem.

- Even with the pendulum included, this theory, as it stands currently, cannot be applied to most of the dynamical systems from elementary physics (instances of the wave equation, heat diffusion, and even merely the double pendulum). Its reliance on explicit formulas is perhaps the biggest limiting factor. Most of the common systems of partial differential equations and differential-algebraic equations cannot be represented explicitly the way the pendulum and the mass-spring-damper system can be written: with isolated (stream) variables exclusively on the left-hand side and potentially more complicated expressions on the right. While it is certainly more powerful than a direct application of Banach's Contraction Mapping Principle, the dependence upon this form is quite frustrating. It would enhance the theory tremendously if it could be adapted somehow to be applicable to some of the implicit forms. Note that, unlike the GPAC, there is no obvious reason the model presented here could not be applied to functions of more than one variable. While we insist on having at least one nonnegative real variable, others could easily be included in the Banach space and the parameter space (e.g. our "streams" could be continuous functions of the form $u : \mathbb{T} \to \mathcal{L}^2(\mathbb{R})$).

- As mentioned in the introductory remarks of Section 3.1.1 on page 51, it is a somewhat intriguing coincidence that our theory involves the use of square matrices whose elements are taken from what turns out be a commutative ring with identity ($\mathcal{C}[\mathbb{T}, \mathcal{S}]$), and that this just happens to be the minimal algebraic structure necessary to define determinants [HK71]. Whether any of the myriad uses for determinants is applicable to the theory is unknown to us, but it would seem to warrant at least a cursory investigation.

Chapter 4

Research Project #3: Exploring the More General Case in Which $\mathbb{T}$ is Replaced by an Arbitrary σ-Compact Space

While most of [TZ11, TZ12] concerns "streams," in $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ (where $\mathbb{T}$ is a representation of time, taken to be either $\mathbb{N}$ or $\mathbb{R}^{\geq 0}$, and $\mathcal{A}$ is a topological algebra), the first paper begins more generally—looking at $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ where $\mathbb{X}$ is merely an arbitrary σ-compact[1] space with some extra conditions imposed on it.

On page 3380 of [TZ11], shortly after mentioning causality (see Definition 1.3.13 on page 16) for the first time, Tucker and Zucker write,

> *"It is not clear how to define (or even make sense of) the concept of causality in the general case for $\mathbb{X}$ (taking, for example, $\mathbb{X} = \mathbb{Z}^2$ or $\mathbb{X} = \mathbb{R}^3$)."*

Hence, in Section 3 of their paper, where causality becomes essential to assume, they restrict $\mathbb{X}$ to the special case in which it serves as a model of *time*, renaming it to $\mathbb{T}$ and taking $\mathbb{T} = \mathbb{R}^{\geq 0}$ (as well as often addressing the case in which $\mathbb{T} = \mathbb{N}$).

Their footnote is incontrovertible. Causality is an inherently temporal phenomenon, and we, as human beings, appear to be hardwired to perceive time as a strictly one-dimensional entity. Nevertheless, there does seem to be a way to generalize their properties ***Caus*** and ***WCaus*** for spaces like $\mathbb{Z}^2$ or $\mathbb{R}^3$ while preserving their operational role within the theory—even if the new definitions could no longer be described as having anything to do with our intuitive notion

[1] A set is *σ-compact* if it is a countable union of compact sets.

of causality. Essentially they are purely abstract properties which do the same job causality does.

Loosely speaking, what is special about $\mathbb{T}$ within their theory—the reason they appear to have chosen it—is the fact that it is totally ordered and has a first element (zero). But this is true of their compact exhaustion of $\mathbb{X}$ as well. Granted, $\mathbb{X}$ itself doesn't necessarily have those two properties, but it turns out that this is not actually required. We can trade $\mathbb{T}$ for a compact exhaustion similar to the way we might use a compact exhaustion as a model of $\mathbb{N}$. There's one property (shift invariance) that seems somewhat tricky to transpose in this way, but in fact, we can circumvent it completely via the Generalized Theorem TZJ2 on page 102—which is quite a stroke of luck, given that I hadn't conceived of the project in this chapter back when I proved its precursor (Theorem TZJ2 on page 27).

While struggling with what I thought was a problem in their computability results (but turned out to be an embarrassing misunderstanding on my part), I thought of a few variations of their main "workhorse" property $\boldsymbol{Contr}(\lambda, \tau)$ that would produce similar results and would work with the more general (smoothie-based) operators as well. I defer those for the Future Work section

4.1 Smoothie Space

$\mathcal{C}[\mathbb{X}, \mathcal{A}]$ represents the set of continuous functions from $\mathbb{X}$ into $\mathcal{A}$, equipped with the compact-open topology. The elements of $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ are appropriately named "streams" when $\mathbb{X}$ is a model of time, but the metaphor falls apart when it isn't. For what is, apparently, an acute lack of imagination, I've adopted the word *smoothie* to describe the elements of $\mathcal{C}[\mathbb{X}, \mathcal{A}]$. They're continuous functions, but they lack the total ordering we naturally associate with the word "stream." So I refer to $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ as *smoothie space.*

Throughout this chapter, as in previous chapters, we assume $\mathcal{A}$ is a complete metric space with metric $d_{\mathcal{A}}$. We assume $\mathbb{X}$ is a σ-compact topological space (i.e. it is a countable union of compact sets) with a compact exhaustion $\mathfrak{X}$.

Definition 4.1.1 (Compact Exhaustion)**.** Let $\mathfrak{X} = \{X_n\}_{n \in \mathbb{N}}$ be a family of compact subsets of $\mathbb{X}$. Then we say $\mathfrak{X}$ is *compact exhaustion* if

$$X_0 \subseteq X_1 \subseteq \cdots \subseteq \mathbb{X}$$
$$\textstyle\bigcup_{n=0}^{\infty} X_n = \mathbb{X}$$

and for every compact set $K \subseteq \mathbb{X}$ there is an $n \in \mathbb{N}$ such that $K \subseteq X_n$ (see Remark 4.1.4 on page 84 if this last condition seems unusual).

4.1.1 Pseudometrics

As before, we define the sequence of pseudometrics $\{d_n\}_{n\in\mathbb{N}}$ as follows:

$$d_n(u,v) = \sup_{x\in X_n} d_{\mathcal{A}}(u(x),v(x))$$

There are some situations in which it is convenient to apply these pseudometrics not to $\mathcal{C}[\mathbb{X},\mathcal{A}]$, but to $\mathcal{C}[X_n,\mathcal{A}]$ (for some $n\in\mathbb{N}$). So here we take the domain of d_n to be $\mathcal{C}[\mathbb{X},\mathcal{A}]^2 \cup \bigcup_{k=n}^{\infty} \mathcal{C}[X_k,\mathcal{A}]^2$.
And again, since u, v, and d_n are continuous (the latter via Lemma 1.4.1 on page 21) and X_n is compact, it follows that $\exists y \in X_n$ such that $d_{\mathcal{A}}(u(y),v(y)) = d_n(u,v)$, and hence we can write,

$$d_n(u,v) = \max_{x\in X_n} d_{\mathcal{A}}(u(x),v(x))$$

The same sort of metric we constructed for $\mathcal{C}[\mathbb{T},\mathcal{A}]$ works here as well, although we have no need for it (the pseudometrics are more convenient):

$$d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u,v) = \sum_{n=0}^{\infty} \min\left\{2^{-n}, d_n(u,v)\right\}$$

It is worthwhile, at this point, to review the use of pseudometrics in a metric space.

Lemma 4.1.2. *A sequence $\{u_n\}_{n\in\mathbb{N}}$ in $\mathcal{C}[\mathbb{X},\mathcal{A}]$ is Cauchy iff $\exists N : \mathbb{N}\times\mathbb{R}^+ \to \mathbb{N}$ such that $\forall M \in \mathbb{N}\ \forall\varepsilon > 0$,*

$$n,m \geq N(M,\varepsilon) \ \Rightarrow\ d_M(u_m,u_n) < \varepsilon$$

Proof. ($\Rightarrow$) Suppose $\{u_n\}_{n\in\mathbb{N}}$ is Cauchy. Then $\exists N' : \mathbb{R}^+ \to \mathbb{N}$ such that $\forall\varepsilon > 0$,

$$n,m \geq N'(\varepsilon) \ \Rightarrow\ d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u_m,u_n) < \varepsilon$$

Define $N : \mathbb{N}\times\mathbb{R}^+ \to \mathbb{N}$ as follows:

$$N(M,\varepsilon) = \max\left\{N'(\varepsilon), N'(2^{-M-1})\right\}$$

Let $M \in \mathbb{N}$, $\varepsilon > 0$, and $m,n \geq N(M,\varepsilon)$. Being just a member of the summation of nonnegative terms,

$$\begin{aligned} \min\left\{2^{-M}, d_M(u_m,u_n)\right\} &\leq d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u_m,u_n) \\ &\leq 2^{-M-1} \\ &< 2^{-M} \end{aligned}$$

Thus,

$$\begin{aligned} d_M(u_m,u_n) &= \min\left\{2^{-M}, d_M(u_m,u_n)\right\} \\ &\leq d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u_m,u_n) \\ &< \varepsilon \text{ (since } m,n \geq N'(\varepsilon)) \end{aligned}$$

($\Leftarrow$) Now suppose $\exists N : \mathbb{N}\times\mathbb{R}^{\geq 0} \to \mathbb{N}$ as described in the statement of the lemma. Define $N' : \mathbb{R}^+ \to \mathbb{N}$ as follows:

$$N'(\varepsilon) = N\left(\lceil 1-\log_2(\varepsilon)\rceil, \frac{\varepsilon}{2\lceil 1-\log_2(\varepsilon)\rceil}\right)$$

Let $\varepsilon > 0$. Let $m,n \geq N'(\varepsilon)$ and for convenience, let $M = \lceil 1-\log_2(\varepsilon)\rceil$. Then,

$$\begin{aligned} d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u_m,u_n) &\leq \sum_{i=0}^{M}\min\left\{2^{-i}, d_i(u_m,u_n)\right\} + \sum_{i=M+1}^{\infty} 2^{-i} \\ &\leq M d_M(u_m,u_n) + \frac{2^{-(M+1)}}{1-1/2} \\ &= M d_M(u_m,u_n) + 2^{-M} \end{aligned}$$

By definition of N and M,

$$\begin{aligned} M d_M(u_m,u_n) &< M\cdot\frac{\varepsilon}{2\lceil 1-\log_2(\varepsilon)\rceil} \\ &= \frac{\varepsilon}{2} \end{aligned}$$

and

$$\begin{aligned} 2^{-M} &= 2^{-\lceil 1-\log_2(\varepsilon)\rceil} \\ &= 2^{-\lceil \log_2(2/\varepsilon)\rceil} \\ &\leq 2^{\log_2(\varepsilon/2)} \\ &= \frac{\varepsilon}{2} \end{aligned}$$

Therefore, $d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(u_m,u_n) < \varepsilon$. □

4.1.2 Completeness of $\mathcal{C}[\mathbb{X},\mathcal{A}]$

Lemma 4.1.3. *If $\mathcal{A}$ is complete and $\mathbb{X}$ is a σ-compact space with a compact exhaustion, then $\mathcal{C}[\mathbb{X},\mathcal{A}]$ is complete.*

Proof. See Lemma 2.3.11, Part (c) from [TZ11]. □

Remark 4.1.4. The third property of a compact exhaustion (that every compact set be contained within some member of the exhaustion) is somewhat nonstandard. Some authors omit it, others omit it but insist that for each $n \in \mathbb{N}$, X_n is a subset, not just of X_{n+1}, but of the interior of X_{n+1}. It—or at least something like it—is necessary for Lemma 4.1.3 on the preceding page. It also ensures that the topology generated by $\{d_n\}_{n\in\mathbb{N}}$ is the compact-open (and the local uniform and the inverse limit) topology, although that, in and of itself, is not mandatory.

Example 4.1.5. It is possible that some other condition would suffice to ensure completeness relative to $\mathcal{A}$ (perhaps that every member of the exhaustion be simply connected?), but it's clear that without any third condition, we cannot be certain $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ will be complete. For example, going back to our old standby, $\mathbb{X} = \mathbb{R}^{\geq 0}$ with its usual topology, let $\mathcal{A} = \mathbb{R}$ and for each $n \in \mathbb{N}$ define

$$X_n = \left[0, \frac{n}{n+1}\right] \cup [1, n+1]$$

Then each X_n is compact, $X_0 \subseteq X_1 \subseteq \cdots \subseteq \mathbb{X}$ and $\bigcup_{n\in\mathbb{N}} X_n = \mathbb{X}$, but of course there are several compact sets that aren't contained in any X_n (e.g. $[0,1]$). For each $n \in \mathbb{N}$, define $u_n \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ as follows:

$$u_n(x) \quad = \quad \begin{cases} 1 & \text{if } 0 \leq x < \frac{n}{n+1} \\ (n+1)(1-x) & \text{if } \frac{n}{n+1} \leq x < 1 \\ 0 & \text{if } x \geq 1 \end{cases}$$

Define $N : \mathbb{N} \times \mathbb{R}^{\geq 0} \to \mathbb{N}$ as follows:

$$N(M, \varepsilon) = M$$

Let $\varepsilon > 0$ and let $M \in \mathbb{N}$. Let $m, n \geq N(M, \varepsilon)$. Then, for $x < \frac{M}{M+1}$, $u_m(x) = u_n(x) = 1$, and for $x \geq 1$, $u_m(x) = u_n(x) = 0$. Thus, u_m and u_n agree on X_M and hence,

$$d_M(u_m, u_n) = 0 < \varepsilon$$

Therefore, by Lemma 4.1.2 on page 82, $\{u_n\}_{n\in\mathbb{N}}$ is Cauchy. It is clear, however, that $\{u_n\}_n$ converges (pointwise) to the following function:

$$u(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 0 & \text{if x}\geq\text{1} \end{cases}$$

But $u \notin \mathcal{C}[\mathbb{X}, \mathcal{A}]$.

4.2 $\boldsymbol{Caus}$, $\boldsymbol{Lip}$, and $\boldsymbol{Contr}$ for $\mathcal{C}[\mathbb{X}, \mathcal{A}]$

Definition 4.2.1 ($\boldsymbol{Caus}(\mathfrak{X})$)**.** Let $\mathbb{X}$ be a σ-compact space with compact exhaustion $\mathfrak{X} = \{X_k\}_{k\in\mathbb{N}}$. Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$. Then F satisfies $\boldsymbol{Caus}(\mathfrak{X})$ if the following conditions[2] hold:

1. $\forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]\ \forall x \in X_0\ F(u)(x) = F(v)(x)$
2. $\forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]\ \forall n \in \mathbb{N}\ d_n(u, v) = 0 \Rightarrow d_n(F(u), F(v)) = 0$ (or, in other words, $u \restriction_{X_k} = v \restriction_{X_k} \Rightarrow F(u) \restriction_{X_k} = F(v) \restriction_{X_k}$)

Remark 4.2.2. It is, perhaps, helpful at this point to recall Fact 1.3.15 on page 16, which provides a definition of $\boldsymbol{Caus}$ equivalent to Definition 1.3.13 on page 16 and upon which the definition of $\boldsymbol{Caus}(\mathfrak{X})$ is based. Fact 1.3.15 states that $\boldsymbol{Caus}$ can be expressed as two separate conditions: (i) the condition that the image of every stream under F is the same at time $t = 0$, and (ii) that $F \in \boldsymbol{WCaus}$. The first condition in the definition of $\boldsymbol{Caus}(\mathfrak{X})$ is analogous to Condition (i), where the "X_0" of Condition (i) is simply the singleton set $\{0\}$. The second condition in the definition of $\boldsymbol{Caus}(\mathfrak{X})$ is clearly analogous to Condition (ii) ($\boldsymbol{WCaus}$).

Example 4.2.3. Even when $\mathbb{X} = \mathbb{R}^{\geq 0}$, $\boldsymbol{Caus}(\mathfrak{X})$ is more general than the property $\boldsymbol{Caus}$ from Definition 1.3.13 (but anything that satisfies $\boldsymbol{Caus}$ does satisfy $\boldsymbol{Caus}(\mathfrak{X})$), as the following example shows. Let $\mathbb{T}$ be the nonnegative real numbers with the standard compact exhaustion $\mathfrak{X} = \{[0, k]\}_{k\in\mathbb{N}}$. Define $F : \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]$ as follows for $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$, $t \in \mathbb{T}$:

$$F(u)(t) = u(\lfloor t \rfloor) + (t - \lfloor t \rfloor)\ (u(\lceil t \rceil) - u(\lfloor t \rfloor))$$

Whenever t is not a natural number, the value of $F(u)(t)$ depends on the value of u at the *next* natural number above t (of course, it also depends on the value at the previous natural number, but that's perfectly consistent with causality)—i.e. at a "future" point of u (see Figure 4.2.1).

The fact that this F really isn't "causal," by any usual definition of the word is a not a problem. The essential feature of causality within this theory is to ensure we can get convergence of a fixed point in "bite-sized pieces," and for this, $\boldsymbol{Caus}(\mathfrak{X})$ works just as well as $\boldsymbol{Caus}$. In particular, it allows for the following lemma.

Lemma 4.2.4. *Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfy $\boldsymbol{Caus}(\mathfrak{X})$. Then F induces a sequence of unique functions $\{F_n : \mathcal{C}[X_n, \mathcal{A}] \to \mathcal{C}[X_n, \mathcal{A}]\}_{n\in\mathbb{N}}$ defined as follows:*

$$F_n(u\restriction_{X_n}) = F(u)\restriction_{X_n}$$

[2] Frankly, I would greatly prefer to define $\boldsymbol{Caus}(\mathfrak{X})$ using only the second condition, making the first one a distinct property (like *initially constant*, or something similar), but I feel it's probably better to be consistent with [TZ11] here to avoid confusion.

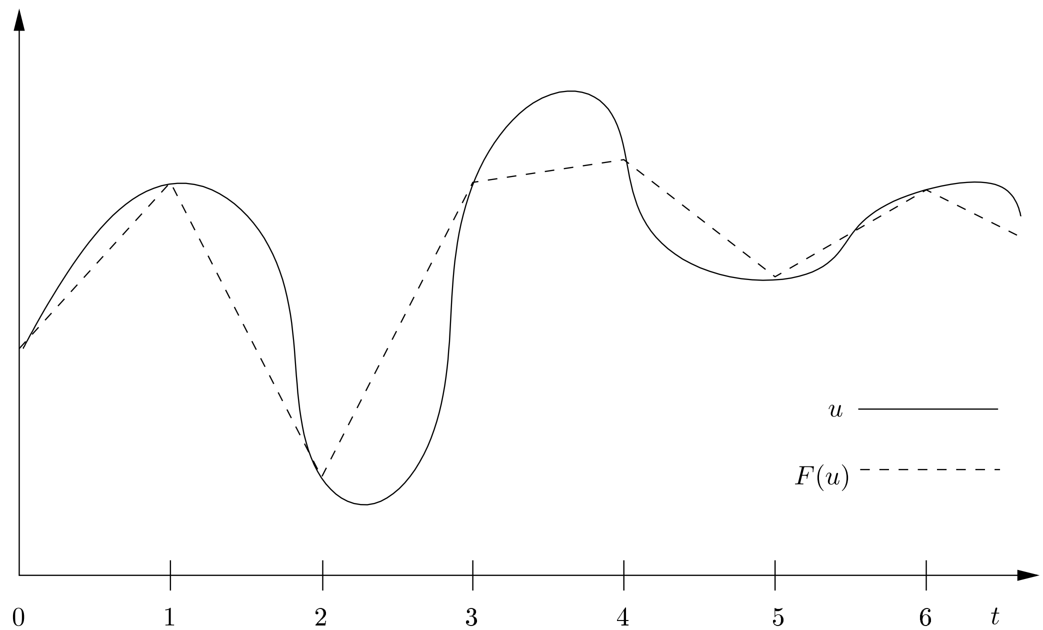

Figure 4.2.1: A stream operator that satisfies $\boldsymbol{Caus}(\mathfrak{X})$ but not $\boldsymbol{Caus}$

Proof. Follows directly from the second property of $\boldsymbol{Caus}(\mathfrak{X})$. ☐

Remark 4.2.5. While $\boldsymbol{Caus}(\mathfrak{X})$ serves the same role in the theory as $\boldsymbol{Caus}$, it seems to have little to do with causality (at least as we perceive it). So I've been tempted to name it *progressive* (and use the notation $\boldsymbol{Prog}(\mathfrak{X})$ instead of $\boldsymbol{Caus}(\mathfrak{X})$) since an operator $F \in \boldsymbol{Caus}(\mathfrak{X})$ operates "progressively" on the compact exhaustion. The value of $F(u)$ on X_n depends on the value of u only on X_n. I've opted to stick with "backwards-compatibility" (i.e. $\boldsymbol{Caus}(\mathfrak{X})$) for now to avoid a surfeit of new, made-up notation and vernacular.

Definition 4.2.6 **($\boldsymbol{Lip}(\lambda, \mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$).** Let $\mathbb{X}$ be a σ-compact space with compact exhaustion $\mathfrak{X} = \{K_k\}_{k\in\mathbb{N}}$. Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$. Let $\lambda \in \mathbb{R}^+$. If $\forall k \in \mathbb{N}\ \forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$,

$$d_k(u, v) = 0 \ \Rightarrow \ d_{k+1}(F(u), F(v)) \leq \lambda d_{k+1}(u, v)$$

Then we say $F \in \boldsymbol{Lip}(\lambda, \mathfrak{X})$. If $\lambda < 1$, then we may say simply, $F \in \boldsymbol{Contr}(\mathfrak{X})$.

4.3 Generalizing the *hold* Operator for Smoothies

Now that we have versions of $\boldsymbol{Caus}$ and $\boldsymbol{Lip}$ that work on $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, we have almost everything we need to generalize Theorem TZ1 on page 19. The only thing missing is the $\boldsymbol{hold}$ operator from Definition 1.3.8 on page 12 (or the **ext** operator from [TZ11]). Recall that $\boldsymbol{hold}_T(u)$ agrees with u on $[0, T]$, and is constant on $[T, \infty)$, with the value $u(T)$. It is essential for the construction of the fixed point (see Construction 1.3.24 on page 20). How can we generalize this idea for $\mathbb{X}$?

Obviously we can change $T \in \mathbb{T}$ to $k \in \mathbb{N}$ (giving us $\boldsymbol{hold}_k : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ for every $k \in \mathbb{N}$ instead of $\boldsymbol{hold}_T : \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ for every $T \in \mathbb{T}$), and then define $\boldsymbol{hold}_k(u)$ so that it agrees with u on X_k. But what happens outside X_k? If we made it constant outside X_k, there would be no way to guarantee $\boldsymbol{hold}_k(u)$ would be continuous—and how would we choose the constant anyway? Clearly another approach is needed.

The reason $\boldsymbol{hold}_T(u)$ is set to the constant $u(T)$ outside the interval $[0, T]$ is mainly to ensure that for any $t \geq 0$ and any $u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$, we have $d_{T+t}(\boldsymbol{hold}_T(u), \boldsymbol{hold}_T(v)) = d_T(u, v)$. The existence of $\lim_{k\to\infty} \Psi(n, k)$ (from Construction 1.3.24 on page 20) depends on this equation. So how can we emulate that behaviour on a wild space like $\mathbb{X}$ instead of the nice, orderly space $\mathbb{T}$?

The key (or *a* key, at least) is to look at $\boldsymbol{hold}_T(u)$, not as a piecewise function (equal to $u(t)$ for $t \leq T$, and equal to $u(T)$ otherwise), but as a composition of functions: $\boldsymbol{hold}_T(u) = (u \circ \rho_T)$, where $\rho_T : \mathbb{T} \to \mathbb{T}$ is defined as $\rho_T(t) = \min\{t, T\}$. Since it's a composition of continuous functions, the result is continuous, and it should be immediately apparent to any student of topology what sort of function ρ_T is. Recall the following definition (see [Mun75], for example):

Definition 4.3.1 (Retract)**.** If X is a subspace of Y, then we say X is a *retract* of Y if there exists a continuous function, $\rho : Y \to X$ such that $\forall x \in X$ $\rho(x) = x$. The function ρ is called a *retraction* of Y onto X.

We are now in a position to generalize the $\boldsymbol{hold}$ operator in a way that will facilitate the construction.

Definition 4.3.2 (Retractable Exhaustion)**.** Let $\mathbb{X}$ be a σ-compact space with compact exhaustion $\mathfrak{X} = \{X_k\}_{k\in\mathbb{N}}$. Then $\mathfrak{X}$ is *retractable* if $\forall k \in \mathbb{N}$ X_k is a retract of $\mathbb{X}$. In other words, there exists a sequence of retractions $\{\rho_k : \mathbb{X} \to X_k\}_{k\in\mathbb{N}}$ of $\mathbb{X}$ onto X_k.

Definition 4.3.3. Let $\mathfrak{X}$ be a retractable compact exhaustion of a space, $\mathbb{X}$, with retractions $\{\rho_k : \mathbb{X} \to X_k\}_{k\in\mathbb{N}}$. For $k \in \mathbb{N}$, define $\boldsymbol{hold}_k : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ as follows:

$$\boldsymbol{hold}_k(u) = u \circ \rho_k$$

Remark 4.3.4. In fact, we have constructed $\boldsymbol{hold}$ here only to demonstrate the application of the retractions and their relation to the original construction. It turns out (as I learned the hard way) that the proofs are a little nicer if we stick with using the retractions (ρ_k) directly and drop the ***hold*** notation altogether. This has the added benefit of eliminating the rather unintuitive notation, "$\boldsymbol{hold}_n$." It makes sense in the context of $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ as we allow $\boldsymbol{hold}_T(u(t))$ to vary up until $t = T$, and then "hold" it, fixed at that value forever after. The function we have defined above, however, isn't "holding" anything. It would be more accurate to say it's propagating or smearing values of u taken from X_k throughout $\mathbb{X}$, but we don't actually have to say anything if we just use ρ_k.

Still, while we'll generally avoid the ***hold*** notation, we need to establish its continuity.

Lemma 4.3.5. *$\forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ $\forall n, k \in \mathbb{N}$ $d_k(u \circ \rho_n, v \circ \rho_n) = d_j(u, v)$ where $j = \min\{n, k\}$.*

Proof. If $k \leq n$, then

$$\begin{aligned} d_k(u \circ \rho_n, v \circ \rho_n) &= \max_{x \in X_k} \{d_{\mathcal{A}}(u(\rho_n(x)), v(\rho_n(x)))\} \\ &= \max_{x \in X_k} \{d_{\mathcal{A}}(u(x), v(x))\} \\ &= d_k(u, v) \end{aligned}$$

Otherwise,

$$\begin{aligned} d_k(u \circ \rho_n, v \circ \rho_n) &= \max_{x \in X_k} \{d_{\mathcal{A}}(u(\rho_n(x)), v(\rho_n(x))\} \\ &= \max_{y \in X_n} \{d_{\mathcal{A}}(u(y), v(y))\} \\ &= d_n(u, v) \end{aligned}$$

□

Lemma 4.3.6. *Let $n \in \mathbb{N}$ and define $\boldsymbol{hold}_n : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ as in Definition 4.3.3. Then $\boldsymbol{hold}_n$ is (uniformly) continuous.*

Proof. Let $\varepsilon > 0$, $k \in \mathbb{N}$. Let $\delta = \varepsilon$ and $j = \min\{n, k\}$. Let $u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ (or in $\mathcal{C}[X_i, \mathcal{A}]$ for some $i \geq \max\{n, k\}$) such that $d_j(u, v) < \delta$. Then by Lemma 4.3.5

$$\begin{aligned} d_k(\boldsymbol{hold}_n u, \boldsymbol{hold}_n v) &= d_k(u \circ \rho_n, v \circ \rho_n) \\ &= d_j(u, v) \\ &\leq d_k(u, v) \\ &< \varepsilon \end{aligned}$$

□

Both equations in the following lemma can be proven trivially (given that $\rho_n(\mathbb{X}) = X_n$), but when they appear in the middle of a proof with complicated expressions in the place of u and v, the steps don't seem quite so obvious. So the lemma is stated here without proof.

Lemma 4.3.7. *If $u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $n \in \mathbb{N}$, then*

$$\begin{aligned} \sup_{x \in \mathbb{X}} d_{\mathcal{A}} \left((u \circ \rho_n)(x), (v \circ \rho_n)(x) \right) &= d_n(u \circ \rho_n, v \circ \rho_n) \\ &= d_n(u, v) \end{aligned}$$

Lemma 4.3.8. *If $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $n \in \mathbb{N}$, then $\boldsymbol{hold}_n(u)$ is bounded (that is, the range of $\boldsymbol{hold}_n(u)$ is a bounded subset of $\mathcal{A}$).*

Proof. Let $x_0 \in X_n$. Define $f : \mathbb{X} \to \mathbb{R}$ as $f(x) = d_{\mathcal{A}}(u(x), u(x_0))$. Since u and $d_{\mathcal{A}}$ are continuous (the former by definition of $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, and the latter by Lemma 1.4.1 on page 21), f is continuous on $\mathbb{X}$. Thus, f is continuous on X_n, which is compact. Therefore, $f(X_n)$ is compact and since $f(X_n) \subseteq \mathbb{R}$, it follows that $f(X_n)$ is bounded (and closed). Hence, $\exists M > 0$ such that $\forall x \in X_n\ f(x) \leq M$. □

4.3.1 Do We Have the "Right" Retractable Exhaustion?

Remark 4.3.9 (Alternative Definition of *Retractable Exhaustion*). In some situations, it would be convenient to have a different sequence of retractions: $\{\rho'_k : X_{k+1} \to X_k\}_{k \in \mathbb{N}}$. It is, of course, easy to define ρ'_k as simply, $\rho_k \restriction_{X_{k+1}}$ for any $k \in \mathbb{N}$ if we already have the sequence from Definition 4.3.2 on page 87. It seems as though we should be able to go the other way too, however—that is, to start with the sequence $\{\rho'_k : X_{k+1} \to X_k\}_{k \in \mathbb{N}}$, and from it, define the sequence $\{\rho_k : \mathbb{X} \to X_k\}_{k \in \mathbb{N}}$. It is clear how this would be done:

$$\rho_k(x) = \begin{cases} \left(\rho'_k \circ \rho'_{k+1} \circ \cdots \circ \rho'_{q(x)-1} \right)(x) & \text{if } q(x) > k \\ x & \text{if } q(x) \leq k \end{cases}$$

where $q(x) = \min \{ i \in \mathbb{N} : x \in X_i \}$ for any $x \in \mathbb{X}$.

While this ρ_k is clearly well-defined, unfortunately I'm not sure whether it is necessarily continuous. Superficially it appears as though it is not defined the same way at every point (it is defined above as a piecewise function with a countably infinite number of pieces), but recall that retractions behave as the identity on their ranges. If $x \in X_i$ then $\forall j \geq i\ \rho'_j(x) = x$. Hence, retractions of higher index can be composed indefinitely without moving the image of x. Hence, it would appear that ρ_k is actually defined consistently across its domain as a composition of continuous functions. The problem is that it's not necessarily a *finite* composition over an open set. One thing I do know is that if it *is* continuous, the proof will depend on that third condition of a compact exhaustion (that every compact set be contained within a member of the exhaustion), as the following counterexample[3] illustrates.

Example 4.3.10. Let $\mathbb{X} = [0, 1]$ with the usual subspace topology of $\mathbb{R}$. For $k \in \mathbb{N}$, let

$$X_k = \{0\} \cup \left[\frac{1}{k+1}, 1\right]$$

Clearly $\{X_k\}_{k\in\mathbb{N}}$ would be a compact exhaustion of $\mathbb{X}$ if only it satisfied the extra requirement that every compact set be contained within a member of the exhaustion (and $[0, 1/2]$ isn't, for example). A definition for $\rho'_k : X_{k+1} \to X_k$ is immediately apparent:

$$\rho'_k(x) = \begin{cases} \frac{1}{k+1} & \text{if } x \in \left[\frac{1}{k+2}, \frac{1}{k+1}\right] \\ x & \text{otherwise} \end{cases}$$

According to the proposed definition of $\rho_k : \mathbb{X} \to X_k$ as a composition, $\rho_k = \rho'_k \circ \rho'_{k+1} \circ \cdots$, we obtain (after putting them all together),

$$\rho_k(x) = \begin{cases} \frac{1}{k+1} & \text{if } x \in \left(0, \frac{1}{k+1}\right] \\ x & \text{otherwise} \end{cases}$$

This ρ_k is discontinuous, however (for any $k \in \mathbb{N}$), and thus it is not a retraction. Thus, if the construction of $\{\rho_k\}_{k\in\mathbb{N}}$ from $\{\rho'_k\}_{k\in\mathbb{N}}$ described in Remark 4.3.9 is guaranteed to produce continuous functions, this guarantee depends upon the extra condition in our definition of a compact exhaustion. The whole question, however, is admittedly somewhat moot when we can simply use Definition 4.3.2 on page 87 and have everything we need.

4.4 Existence of a Unique Fixed Point for Operators Which Satisfy $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$

We now have all the tools necessary to generalize Theorem TZ1 and show that an operator $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ that satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$ has a unique fixed point, regardless of whether $\mathbb{X}$ is a model of time or not. There is a quick and easy way to do this proof that simply invokes Banach's Fixed Point Theorem—allowing it to do all the heavy lifting—and a longer, more

[3]Thanks to Prof. Jacques Carette for providing the compact exhaustion which inspired the counterexample!

involved proof that directly reproduces many of the steps in the proof of that theorem. Each has a slightly different fixed point construction associated with it. I present the second one only because it uses a construction that is more consistent with the one used by Theorem TZJ2 on page 27—lest the reader think I am pulling some sleight-of-hand.

In each case, the explication is a little clearer if the construction is shown first (after some notation is defined), and the proof that it works is presented afterwards. We start with the simpler one. To be consistent with Construction 1.3.24 on page 20, we would define $\Psi : \mathbb{N}^2 \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$, but in the simpler proof we use the form $\Psi : \mathbb{N}^2 \to \bigcup_{n\in\mathbb{N}} \mathcal{C}[X_n, \mathcal{A}]$. Then, since this no longer produces elements of $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, we define our main convergent sequence, $\{\psi_n\}_{n\in\mathbb{N}} \subseteq \mathcal{C}[\mathbb{X}, \mathcal{A}]$, from Ψ.

Definition 4.4.1. Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$, $n \in \mathbb{N}$, and $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$. If $d_n(u, F(u)) = 0$ (i.e. $u{\restriction_{X_n}} = F(u){\restriction_{X_n}}$) then u is said to be an X_n-*approximate fixed point* of F.

Lemma 4.4.2 (F_n: Truncations of F)**.** *Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfy $\boldsymbol{Caus}(\mathfrak{X})$. Then F induces a sequence of unique functions*

$$\{F_n : \mathcal{C}[X_n, \mathcal{A}] \to \mathcal{C}[X_n, \mathcal{A}]\}_{n\in\mathbb{N}}$$

defined as follows:

$$F_n(w) = F(w \circ \rho_n){\restriction_{X_n}}$$

Moreover, the value of $F_n(w)$ does not depend on the definition of ρ_n (it depends only on ρ_n being a retraction from $\mathbb{X}$ to X_n).

Proof. Since F, w, and ρ_n are continuous, so is $F(w \circ \rho_n)$. The restriction of this function to X_n is obviously continuous (or see Theorem 7.2, page 107 of [Mun75] if it doesn't seem obvious). So given any $n \in \mathbb{N}$ and $w \in \mathcal{C}[X_n, \mathcal{A}]$, $F(w \circ \rho_n){\restriction_{X_n}} \in \mathcal{C}[X_n, \mathcal{A}]$. Now suppose $\rho_n' : \mathbb{X} \to X_n$ is another retraction. Since both ρ_n and ρ_n' behave as the identity on X_n, $\forall w \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$

$$w \circ \rho_n{\restriction_{X_n}} = w \circ \rho_n'{\restriction_{X_n}}$$

So $d_n(w\circ\rho_n, w\circ\rho_n') = 0$. Thus, since $F \in \boldsymbol{Caus}(\mathfrak{X})$, $d_n(F(w\circ\rho_n), F(w\circ\rho_n')) = 0$. Hence, F_n does not depend on the retraction used. □

Construction 4.4.3. (Construction for Simpler Proof that F has a Unique Fixed Point)

Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfy $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$. Then it has a unique fixed point, which can be constructed as follows (see the Generalized Theorem TZ1 on page 94 for proof):

1. Let $u_0 \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and set $\psi_0 = F(u_0) \circ \rho_0$.
   Since $F \in \boldsymbol{Caus}(\mathfrak{X})$, ψ_0 is an X_0-approximate fixed point of F.

2. Let $n \in \mathbb{N}$ and suppose ψ_n is an X_n-approximate fixed point of F. Define $\Psi(n, 0) = \psi_n \restriction_{X_{n+1}}$, making $\Psi(n, 0)$ an X_n-approximate fixed point of F_{n+1}.

3. For all $k \in \mathbb{Z}^+$ define $\Psi(n, k) = F^k_{n+1}(\Psi(n, 0))$.
   We will show this sequence converges to a unique element of $\mathcal{C}[X_{n+1}, \mathcal{A}]$.

4. Given an $n \in \mathbb{N}$ for which the sequence $\{\Psi(n, k)\}_{k\in\mathbb{N}}$ exists and converges, define $\psi_{n+1} = (\lim_{k\to\infty} \Psi(n, k)) \circ \rho_{n+1}$.

5. Define $v = \lim_{n\to\infty} \psi_n$
   This will be the unique fixed point of F.

Remark 4.4.4. There are really only three things to prove: that the fixed point is unique, and that the two limits (in Steps 3 and 5) exist. We'll cover uniqueness first (in Lemma 4.4.5). To prove the other statements, we'll need to go over a few lemmas and Banach's celebrated Fixed Point Theorem first.

Lemma 4.4.5 (If F has a fixed point, it's unique)**.** *Suppose $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$, and has a fixed point, $v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$. Then v is unique.*

Proof. Suppose u and v are fixed points of F. By definition of $\boldsymbol{Caus}(\mathfrak{X})$ (Part 1, in particular), it follows that u and v agree on X_0 since they're both in the range of F. Hence, $d_0(u, v) = 0$. Since $F \in \boldsymbol{Contr}(\mathfrak{X})$, there is a $\lambda < 1$ such that $F \in \boldsymbol{Lip}(\lambda, \mathfrak{X})$. Let $n \in \mathbb{N}$ and suppose $d_n(u, v) = 0$. Then, since $F \in \boldsymbol{Lip}(\lambda, \mathfrak{X})$,

$$\begin{aligned} d_{n+1}(u, v) &= d_{n+1}(F(u), F(v)) \\ &\leq \lambda \cdot d_{n+1}(u, v) \end{aligned}$$

But $0 < \lambda < 1$, so the only way this is possible is for $d_{n+1}(u, v)$ to be zero. Hence, $\forall k \in \mathbb{N}\ d_k(u, v) = 0$, and therefore $u = v$. □

Theorem 4.4.6 (Banach Fixed Point Theorem)**.** *Let X be a complete metric space and $f : X \to X$. Suppose that $\exists \lambda \in \mathbb{R}^+$ such that $\lambda < 1$ and $\forall x, y \in X\ d(f(x), f(y)) \leq \lambda d(x, y)$. Then f has a unique fixed point given by $\lim_{n\to\infty} f^n(x)$, where x is any element of X.*

Proof. See any introductory text on real analysis (e.g. Theorem 9.23 in [Rud76]). Also known as the "contraction mapping principle." □

Lemma 4.4.7. *Let $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfy $\boldsymbol{Caus}(\mathfrak{X})$. Let $n \in \mathbb{N}$. Then,*

1. *If $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is an X_n-approximate fixed point of F then $u{\restriction_{X_n}}$ is a fixed point of F_n.*

2. *If $u \in \mathcal{C}[X_n, \mathcal{A}]$ is a fixed point of F_n then $u \circ \rho_n$ is an X_n-approximate fixed point of F.*

Proof. (1) If $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is an X_n-approximate fixed point of F, then $d_n(u, F(u)) = d_n(u{\restriction_{X_n}}, F(u){\restriction_{X_n}}) = 0$. Since d_n is a metric on $\mathcal{C}[X_n, \mathcal{A}]$, this implies that $u{\restriction_{X_n}} = F(u){\restriction_{X_n}}$. Since $F \in \boldsymbol{Caus}(\mathfrak{X})$, $F(u){\restriction_{X_n}} = F(u{\restriction_{X_n}})$.
(2) If $u \in \mathcal{C}[X_n, \mathcal{A}]$ is a fixed point of F_n then $u = F_n(u)$. By definition (in Lemma 4.4.2, $F_n(u) = F(u \circ \rho_n){\restriction_{X_n}}$. Since ρ_n behaves as the identity on X_n, $u = u \circ \rho_n{\restriction_{X_n}}$. So,

$$\begin{aligned} d_n(u \circ \rho_n, F(u \circ \rho_n)) &= d_n(u \circ \rho_n{\restriction_{X_n}}, F(u \circ \rho_n){\restriction_{X_n}}) \\ &= d_n(u, F_n(u)) \\ &= 0 \end{aligned}$$

□

Remark 4.4.8. In [TZ11], Lemma 2.1.2 states that if K is compact and $\mathcal{A}$ is complete, then $\mathcal{C}[K, \mathcal{A}]$ is complete. This is used to prove that the result holds even if K is not compact, but is σ-compact (and $\mathcal{A}$ is complete). Later in the paper, however, it is also used in the proof of Theorem TZ1 to establish the convergence of a particular Cauchy sequence (loosely speaking, it's the sequence I've called $\Psi(n, 0)$, $\Psi(n, 1)$, $\ldots$). Using the following lemma in place of Lemma 2.1.2 is what allows us to invoke Banach's Fixed Point Theorem directly, instead of producing a similar proof from the ground up:

Lemma 4.4.9. *Let X be a compact (or σ-compact) metric space, and let K be a compact subset of X. Let Y be a complete metric space. Let $f : K \to Y$ be continuous. Let $C_f[X, Y] = \{g \in \mathcal{C}[X, Y] : g{\restriction_K} = f\}$. That is, let $C_f[X, Y]$ be the set of continuous functions from X into Y which agree with f on K. Then $C_f[X, Y]$ (endowed with the subspace topology) is complete.*

Proof. Let $\{g_k\}_{k \in \mathbb{N}}$ be a Cauchy sequence in $C_f[X, Y]$. Since $C_f[X, Y] \subseteq \mathcal{C}[X, Y]$ (which is complete, by Lemma B.0.3 on page 133), there exists a unique $g \in \mathcal{C}[X, Y]$ such that $g_k \to g$ as $k \to \infty$. If g_k converges to g, then it certainly converges pointwise to g, and since $\forall k \in \mathbb{N}\ g_k{\restriction_K} = f$, it follows that $g{\restriction_K} = f$ also. □

Lemma 4.4.10 (Convergence in Step 3)**.** *Suppose $F \in \boldsymbol{Caus}(\mathfrak{X}) \cap \boldsymbol{Contr}(\mathfrak{X})$. Let $n \in \mathbb{N}$ and suppose $w \in \mathcal{C}[X_{n+1}, \mathcal{A}]$ is an X_n-approximate fixed point of F_{n+1}. Then F_{n+1} has a unique fixed point, which is given by the limit of the sequence $\left\{F_{n+1}^k(w)\right\}_{k \in \mathbb{N}}$.*

Proof. First note that $\boldsymbol{Caus}(\mathfrak{X})$ is used implicitly in the statement of the lemma to establish the existence of F_{n+1} (via Lemma 4.4.2). Let

$$C_w[X_{n+1}, \mathcal{A}] = \{v \in \mathcal{C}[X_{n+1}, \mathcal{A}] : d_n(v, w) = 0\}$$

By Lemma 4.4.9, $C_w[X_{n+1}, \mathcal{A}]$ is complete. Since w is an X_n-approximate fixed point of F_{n+1}, it follows that $C_w[X_{n+1}, \mathcal{A}]$ is closed under F_{n+1}. Thus, restricting the domain of F_{n+1} to $C_w[X_{n+1}, \mathcal{A}]$ yields a function of the form

$$\hat{F}_{n+1} : C_w[X_{n+1}, \mathcal{A}] \to C_w[X_{n+1}, \mathcal{A}]$$

Now from the definition of d_{n+1} and from the fact that $F \in \boldsymbol{Caus}(\mathfrak{X})$, for all $u, v \in C_w[X_{n+1}, \mathcal{A}]$,

$$d_{n+1}(\hat{F}_{n+1}(u), \hat{F}_{n+1}(v)) \;=\; d_{n+1}(F(u \circ \rho_{n+1}), F(v \circ \rho_{n+1}))$$

Since $F \in \boldsymbol{Contr}(\mathfrak{X})$, $\exists \lambda < 1$ such that $\forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$,

$$d_n(u, v) = 0 \;\Rightarrow\; d_{n+1}(F(u), F(v)) \le \lambda d_{n+1}(u, v)$$

But for *all* $u, v \in \mathcal{C}_w[X_{n+1}, \mathcal{A}]$ $d_n(u, v) = 0$. Since ρ_{n+1} behaves as the identity on X_n (and on X_{n+1}, for that matter, but that's not relevant at the moment), it follows that $d_n(u \circ \rho_{n+1}, v \circ \rho_{n+1}) = 0$. Thus, $\forall u, v \in \mathcal{C}_w[X_{n+1}, \mathcal{A}]$,

$$\begin{aligned} d_{n+1}(\hat{F}_{n+1}(u), \hat{F}_{n+1}(v)) &= d_{n+1}(F(u \circ \rho_{n+1}), F(v \circ \rho_{n+1})) \\ &\le \lambda d_{n+1}(u \circ \rho_{n+1}, v \circ \rho_{n+1}) \\ &= \lambda d_{n+1}(u, v) \end{aligned}$$

Therefore, $\hat{F}_{n+1}$ is a *contraction* (in the usual Banach sense[4]) on a complete metric space, $\mathcal{C}_w[X_{n+1}, \mathcal{A}]$. By Banach's Fixed Point Theorem, it has a unique fixed point $v = \lim_{k\to\infty} \hat{F}^k_{n+1}(u) = \lim_{k\to\infty} F^k_{n+1}(u)$ (where $u \in \mathcal{C}_w[X_{n+1}, \mathcal{A}]$ is an arbitrary initial point). ☐

Theorem 4.4.11 (Generalized Theorem TZ1)**.** *Construction 4.4.3 on page 91 works as advertized. That is, if $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Contr}(\mathfrak{X})$, then it has a unique fixed point given by the limit in Step 5 of Construction 4.4.3.*

Proof. As suggested by the construction, we show by induction on n, that there is a sequence $\psi_0, \psi_1, \ldots \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ such that for any $n \in \mathbb{N}$, ψ_n is an X_n-approximate fixed point of F. $F \in \boldsymbol{Caus}(\mathfrak{X})$ yields our basis step: an

[4]Definition: If (X, d) is a metric space and $f : X \to X$, then f is a *contraction* if it satisfies the antecedent of Banach's Fixed Point Theorem. That is, $\exists \lambda < 1$ $(\lambda \in \mathbb{R}^+)$ such that $\forall x, y \in X$ $d(f(x), f(y)) \le \lambda d(x, y)$.

X_0-approximate fixed point, ψ_0. For the inductive step, let $n \in \mathbb{N}$ and assume $\psi_n \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is an X_n-approximate fixed point of F. Then $\Psi(n, 0) = \psi_n\restriction_{X_{n+1}}$ is an X_n-approximate fixed point of F_{n+1}. By Lemma 4.4.10 on page 93, F_{n+1} has a unique fixed point given by the limit of sequence $\Psi(n, 0)$, $\Psi(n, 1)$, ... described in Step 3 of the construction. Extending the domain of this fixed point of F_{n+1} from X_{n+1} to $\mathbb{X}$ by composing it with ρ_{n+1} (as suggested by Lemma 4.4.7 on page 92) yields an X_{n+1}-approximate fixed point of F, which we call ψ_{n+1}. This concludes the induction used to show the existence of $\{\psi_n\}_{n\in\mathbb{N}}$.

What remains to be shown is that $\{\psi_n\}_{n\in\mathbb{N}}$ is convergent and that it converges to a fixed point of F. Define $N : \mathbb{R}^+ \times \mathbb{N} \to \mathbb{N}$ as follows:

$$N(\varepsilon, M) = M$$

Let $\varepsilon > 0$ and $M \in \mathbb{N}$. Then $\forall n, m \geq M$, both ψ_n and ψ_m are X_M-approximate fixed points of F (as our induction in the beginning of the proof showed). Hence,

$$d_M(\psi_m, \psi_n) = 0 < \varepsilon$$

Thus, by Lemma 4.1.2 on page 82, $\{\psi_n\}_{n\in\mathbb{N}}$ is Cauchy. Since $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ is complete (by Lemma 4.1.3 on page 83), $\{\psi_n\}_{n\in\mathbb{N}}$ is convergent. Now, given any $x \in \mathbb{X}$, there is a $k \in \mathbb{N}$ such that $x \in X_k$. For any $j \geq k$, ψ_j is an X_k-approximate fixed point of F. Therefore, ψ is also an X_k-approximate fixed point of F (where $\psi = \lim_{n\to\infty} \psi_n$). Thus, $\psi(x) = F(\psi)(x)$. Since this holds for every point $x \in \mathbb{X}$, it follows that ψ is a fixed point of F. □

4.4.1 An Alternative Construction

I developed Construction 4.4.3 on page 91 long after I proved Theorem TZJ2 on page 27 for stream operators, and I obviously want to generalize that theorem to work with smoothie operators. I'm confident that I could write a continuity proof based on Construction 4.4.3 if I had more time, but it would take me too long to adapt it now. What I can do instead is present my original proof which uses a construction that is a strict generalization of Construction 1.3.24. From there, it is easy to generalize Theorem TZJ2.

While our sequence of retractions does the job of the ***hold*** function from Definition 1.3.8 on page 12 well enough, it doesn't work quite the same way that original ***hold*** function does. In particular, for any $T, t \in \mathbb{T}$, $\boldsymbol{hold}_{T+t} \circ \boldsymbol{hold}_T = \boldsymbol{hold}_T$, but this is not a feature shared by the retracts of every retractable exhaustion. That is, it is not necessarily the case that for $n, j \in \mathbb{N}$, $\rho_n \circ \rho_{n+j} = \rho_n$ (although that is the case for the standard retractable exhaustion of $\mathbb{T}$). This property of the original ***hold*** is necessary for showing that Construction 1.3.24 works (it is used when comparing $\Psi(n, 0)$ with $\Psi(n, 1)$).

Recall that we had defined $\Psi(0,0) = \boldsymbol{hold}_0(F(u))$ (where u is any stream), but for all $k \in \mathbb{N}$, we had $\Psi(0,k+1) = \boldsymbol{hold}_\tau F(\Psi(0,k))$. Likewise, we had defined $\Psi(n,0)$ as the limit of functions which are constant outside $[0,n\tau]$, but $\Psi(n,k+1) = \boldsymbol{hold}_{n+1} F(\Psi(n,k))$ (which is constant only outside $[0,(n+1)\tau]$). To make the construction (or at least the proof for it) work for an arbitrary sequence of retractions, we'll have to apply ρ_{n+1} explicitly to every $\Psi(n,0)$.

Construction 4.4.12.

1. Let $u_0 \in \mathcal{C}[\mathbb{X},\mathcal{A}]$ and set $\Psi(0,0) = F(u_0) \circ \rho_1$.
   Since $F \in \boldsymbol{Caus}(\mathfrak{X})$, every smoothie in the range of F agrees with $\Psi(0,0)$ on X_0 (thus, $\Psi(0,0)$ is an X_0-approximate fixed point of F).

2. Let $n \in \mathbb{N}$ and suppose $\Psi(n,0)$ is an X_n-approximate fixed point of F.
   For all $k \in \mathbb{Z}^+$ define $\Psi(n,k) = F^k(\Psi(n,0)) \circ \rho_{n+1}$.
   We will show this sequence converges to an X_{n+1}-approximate fixed point of F.

3. Given $n \in \mathbb{N}$, define $\Psi(n+1,0) = (\lim_{k\to\infty} \Psi(n,k)) \circ \rho_{n+2}$.
   This will be the first X_{n+1}-approximate fixed point of F encountered in the construction.

4. Define $v = \lim_{n\to\infty} \Psi(n,0)$
   As before, this will be the fixed point of F.

As before, if F is of the form $F : P \times \mathcal{C}[\mathbb{X},\mathcal{A}] \to \mathcal{C}[\mathbb{X},\mathcal{A}]$ (where P is some parameter space), then we define $\Psi : P \times \mathbb{N}^2 \to \mathcal{C}[\mathbb{X},\mathcal{A}]$ as above for each $p \in P$, along with $\Phi : P \to \mathcal{C}[\mathbb{X},\mathcal{A}]$ to be the function such that $\forall p \in P$,

$$\Phi(p) = \lim_{n\to\infty} \Psi(p,n,0) = F(p,\Phi(p)) \tag{4.4.1}$$

Lemma 4.4.13. *Let $n, k_1, k_2 \in \mathbb{N}$ and suppose $\Psi(n,k_1)$ and $\Psi(n,k_2)$ are defined as above. Then*

$$\sup_{x\in\mathbb{X}} \{d_{\mathcal{A}}(\Psi(n,k_1)(x), \Psi(n,k_2)(x))\} = d_{n+1}(\Psi(n,k_1), \Psi(n,k_2))$$

Proof. If $\Psi(n,k_1)$ and $\Psi(n,k_2)$ are defined as above, then there are smoothies u_1 and u_2 such that $\Psi(n,k_i) = u_i \circ \rho_{n+1}$ (for $i = 1,2$). Since ρ_{n+1} is idempotent, $\Psi(n,k_i) = \Psi(n,k_i) \circ \rho_{n+1}$. Hence, the ordered pairs being compared are the following:

$$\begin{aligned}
&\{(\Psi(n,k_1)(x), \Psi(n,k_2)(x)) \,:\, x \in \mathbb{X}\} \\
&\quad = \{(\Psi(n,k_1)(\rho_{n+1}(x)), \Psi(n,k_2)(\rho_{n+1}(x))) \,:\, x \in \mathbb{X}\} \\
&\quad = \{(\Psi(n,k_1)(y), \Psi(n,k_2)(y)) \,:\, y \in X_{n+1}\}
\end{aligned}$$

$\square$

Lemma 4.4.14. *If* $F \in \boldsymbol{Caus}(\mathfrak{X})$ *and* $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$, *then* $\forall n \in \mathbb{N}$ $F(u) \circ \rho_n = F(u \circ \rho_n) \circ \rho_n$

Proof. Let $n \in \mathbb{N}$. Since ρ_n behaves as the identity on X_n, $d_n(u, u \circ \rho_n) = 0$. Thus, since $F \in \boldsymbol{Caus}(\mathfrak{X})$ $d_n(F(u), F(u \circ \rho_n)) = 0$. So $\forall x \in X_n$, $F(u)(x) = F(u \circ \rho_n)(x)$. Since the range of ρ_n is X_n, $\forall x \in \mathbb{X}$ $(F(u) \circ \rho_n)(x) = (F(u \circ \rho_n) \circ \rho_n)(x)$. □

Theorem 4.4.15 (Generalized TZ1 for the Alternate Construction)**.** *Let* $\mathbb{X}$ *be a* σ*-compact space with a retractable exhaustion,* $\mathfrak{X} = \{X_k, \rho_k\}_{k \in \mathbb{N}}$ *and let* $(\mathcal{A}, d_{\mathcal{A}})$ *be a metric space. Let* $\{d_k\}_{k \in \mathbb{N}}$ *be the sequence of pseudometrics corresponding to* $\mathfrak{X}$ *and* $d_{\mathcal{A}}$ *(i.e.* $d_k(u, v) = \max_{x \in K_k} \{d_{\mathcal{A}}(u(x), v(x))\}$. *Let* $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ *satisfy* $\boldsymbol{Caus}(\mathfrak{X})$ *and* $\boldsymbol{Lip}(\lambda, \mathfrak{X})$ *for some positive* $\lambda < 1$. *Then* F *has a unique fixed point.*

Proof. The proof is, of course, modelled after Theorem TZ1, but at least superficially it appears very different. Uniqueness has already been covered in a construction-independent way by Lemma 4.4.5 on page 92, so we need only show the fixed point exists.

We must show that the limits in Construction 4.4.3 on page 91 exist and that $v = F(v)$. First we'll show that, assuming $\Psi(n, 0)$ exists, $\lim_{k \to \infty} \Psi(n, k)$ exists (in $\mathcal{C}[\mathbb{X}, \mathcal{A}]$). We do this by demonstrating that the sequence $\{\Psi(n, k)\}_{k \in \mathbb{N}}$ is uniformly Cauchy (see Definition B.0.4 on page 133). In other words, for any $n, j \in \mathbb{N}$, $\sup_{x \in \mathbb{X}} \{d_{\mathcal{A}}(\Psi(n, k), \Psi(n, k + j))\}$ can be made arbitrarily small by making k sufficiently large.

Let $n \in \mathbb{N}$ and assume[5] $\Psi(n, 0) \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$. For all $k > 0$, let $\Psi(n, k)$ be defined as indicated above.

Define the quantity

$$D_1 \quad = \quad \max \{1, d_{n+1} (\Psi(n, 0), \Psi(n, 1))\}$$

Define $N : \mathbb{R}^+ \to \mathbb{R}^+$ as follows:

$$N(\varepsilon) = \log_\lambda \left(\frac{\varepsilon(1 - \lambda)}{D_1} \right)$$

Let $\varepsilon > 0$ and let $k, j \in \mathbb{N}$ with $k \geq N(\varepsilon)$. Then, by Lemma 4.4.13 on the previous page,

$$\sup_{x \in \mathbb{X}} d_{\mathcal{A}} (\Psi(n, k)(x), \Psi(n, k + j)(x)) = d_{n+1} (\Psi(n, k), \Psi(n, k + j))$$

[5] At this point in the proof, we have established the existence of only $\Psi(0, k)$ (for all $k \in \mathbb{N}$), so for $n > 0$, it is necessary to assume $\Psi(n, 0)$ exists. It's almost a Catch 22 induction.

Using the triangle inequality,

$$
\begin{aligned}
d_{n+1}(\Psi(n,k),\Psi(n,k+j)) &\leq \sum_{i=k}^{k+j-1} d_{n+1}(\Psi(n,i),\Psi(n,i+1)) \\
&= \sum_{i=k}^{k+j-1} d_{n+1}\left(F^i \circ \Psi(n,0) \circ \rho_{n+1}, F^i \circ \Psi(n,1) \circ \rho_{n+1}\right) \\
&= \sum_{i=k}^{k+j-1} d_{n+1}\left(F^i \circ \Psi(n,0), F^i \circ \Psi(n,1)\right) \\
&\leq D_1 \sum_{i=0}^{j-1} \lambda^{i+k} \\
&= D_1 \lambda^k \frac{1-\lambda^j}{1-\lambda} \\
&< D_1 \frac{\lambda^k}{1-\lambda}
\end{aligned}
$$

Since $k \geq N(\varepsilon)$ and $\lambda < 1$, it follows that

$$
\begin{aligned}
D_1 \frac{\lambda^k}{1-\lambda} &\leq D_1 \frac{\lambda^{N(\varepsilon)}}{1-\lambda} \\
&= D_1 \frac{\lambda^{\log_\lambda\left(\frac{\varepsilon(1-\lambda)}{D_1}\right)}}{1-\lambda} \\
&= D_1 \frac{\left(\frac{\varepsilon(1-\lambda)}{D_1}\right)}{1-\lambda} \\
&= \varepsilon
\end{aligned}
$$

Therefore, $\{\Psi(n,k) : k \in \mathbb{N}\}$ is uniformly Cauchy and hence by Corollary B.0.6 on page 134, $\lim_{k\to\infty} \Psi(n,k)$ converges to some $\psi_n \in \mathcal{C}[\mathbb{X},\mathcal{A}]$. Since $\forall i,j \in \mathbb{N}$ $d_n(\Psi(n,i),\Psi(n,j)) = 0$, the limit ψ_n agrees with every member of the sequence on X_n as well. That is, $\forall k \in \mathbb{N}$ $d_n(\psi_n, \Psi(n,k)) = 0$. Since $F \in \boldsymbol{Lip}(\lambda, \mathfrak{X})$, $\forall k \in \mathbb{N}$

$$
d_n(\psi_n,\Psi(n,k)) = 0 \;\Rightarrow\; d_{n+1}\left(F(\psi_n), F(\Psi(n,k))\right) \leq \lambda d_{n+1}(\psi_n, \Psi(n,k))
$$

Thus, since $\Psi(n,k) \to \psi_n$ as $n \to \infty$, we can use the same modulus of convergence to show that $F(\Psi(n,k)) \circ \rho_{n+1} \to F(\psi_n) \circ \rho_{n+1}$.
By Lemma 4.4.14 on the previous page,

$$
\begin{aligned}
F(\Psi(n,k)) \circ \rho_{n+1} &= F\left(F^k(\Psi(n,0)) \circ \rho_{n+1}\right) \circ \rho_{n+1} \\
&= F^{k+1}(\Psi(n,0)) \circ \rho_{n+1} \\
&= \Psi(n,k+1)
\end{aligned}
$$

Therefore,

$$\begin{aligned} F(\psi_n) \circ \rho_{n+1} &= \lim_{k\to\infty} F(\Psi(n,k)) \circ \rho_{n+1} \\ &= \lim_{k\to\infty} \Psi(n,k) \\ &= \psi_n \end{aligned}$$

Hence, $d_{n+1}(\psi_n, F(\psi_n)) = 0$. In other words, $\forall n \in \mathbb{N}$, ψ_n is an X_{n+1}-approximate fixed point of F. The remainder of the proof is identical to the proof of the Generalized Theorem TZ1 on page 94 $\square$

4.5 The *shift* Operator for Smoothies

Given that I was unable to show (see Remark 1.4.16 on page 28 and Example 1.4.17 on page 29) that my Theorem TZJ2 is a strict generalization of Tucker and Zucker's Theorem TZ2 (from [TZ11]), there may still be a place for shift invariance in this theory. This is one of the most challenging concepts to generalize to arbitrary σ-compact spaces, but I see one way it might be done that I believe would still allow it to perform its intended role in (a generalized version of) the proof of Theorem TZ2.

Definition 4.5.1 (Shiftable)**.** Let $\mathbb{X}$ be a σ-compact space with compact exhaustion $\mathfrak{X} = \{X_k\}_{k\in\mathbb{N}}$. Then $\mathfrak{X}$ is *shiftable* if there is a continuous function $\zeta : \mathbb{X} \to \mathbb{X}$ such that $\forall k \in \mathbb{N}$,

$$\zeta(X_{k+1} \setminus X_k) = X_{k+2} \setminus X_{k+1}$$

In other words, if for all $k \geq 1$ we define ζ_k as the restriction of ζ to $X_k \setminus X_{k-1}$, then $\zeta_k : X_k \setminus X_{k-1} \to X_{k+1} \setminus X_k$ is surjective.

Definition 4.5.2 (***shift***)**.** Let $\mathbb{X}$ be a σ-compact space with a shiftable compact exhaustion $\mathfrak{X} = \{X_k\}_{k\in\mathbb{N}}$. Define $\boldsymbol{shift} : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ as

$$\boldsymbol{shift}(u) = u \circ \zeta$$

Remark 4.5.3. There may be a more general way to define ***shift*** on $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, and the condition that ζ be continuous could possibly be relaxed, but its continuity *does* ensure the range of ***shift*** is $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ since the composition of continuous functions is again continuous.

Example 4.5.4. We can see that this definition of ***shift*** is a generalization of the previous one from Definition 1.3.8 on page 12. Let $\mathbb{X} = \mathbb{T}$, fix $T \in \mathbb{R}^+$, and choose $\mathfrak{X} = \{[0, kT]\}_{k\in\mathbb{N}}$. Take $\zeta : \mathbb{T} \to \mathbb{T}$ as the function $\zeta(t) = t + T$. Then it's clear that the two definitions of $\boldsymbol{shift}_T$ agree.

4.6 Continuity Theorem for Smoothie Operators

Lemma 4.6.1. *Let X and Y be topological spaces and suppose X is metrizable. Let $f : X \to Y$. Then f is continuous if and only if for every convergent sequence $x_n \to x$, the sequence $f(x_n)$ converges to $f(x)$.*

Proof. See Theorem 10.3, page 128 of [Mun75] ☐

Lemma 4.6.2. *Let $\{u_k\}_{k\in\mathbb{N}} \subseteq \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and suppose $\exists u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ $\exists n \in \mathbb{N}$ such that $\lim_{k\to\infty} (u_k \circ \rho_n) = u$. Then $u \circ \rho_n = u$.*

Proof. Recall that $\boldsymbol{hold}_n : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is defined as $\boldsymbol{hold}_n(v) = v \circ \rho_n$. Since ρ_n is idempotent,

$$\begin{aligned} u &= \lim_{k\to\infty} (u_k \circ \rho_n) \\ &= \lim_{k\to\infty} (u_k \circ \rho_n \circ \rho_n) \\ &= \lim_{k\to\infty} (\boldsymbol{hold}_n (u_k \circ \rho_n)) \end{aligned}$$

Since $\boldsymbol{hold}_n$ is continuous (see Lemma 4.3.6 on page 88) and $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ is metrizable, Lemma 4.6.1 asserts that $\boldsymbol{hold}_n$ commutes with $\lim_{k\to\infty}$. Thus,

$$\begin{aligned} \lim_{k\to\infty} (\boldsymbol{hold}_n (u_k \circ \rho_n)) &= \boldsymbol{hold}_n \left(\lim_{k\to\infty} (u_k \circ \rho_n) \right) \\ &= \boldsymbol{hold}_n(u) \\ &= u \circ \rho_n \end{aligned}$$

☐

Lemma 4.6.3. *Let $\Psi : P \times \mathbb{N}^2 \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $\Phi : P \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ be given as in Construction 4.4.12 on page 96 (the parametrized versions), and let $n \in \mathbb{N}$, $p, p' \in P$. Then $d_n(\Phi(p), \Phi(p')) = d_{n+1}(\Psi(p, n, 0), \Psi(p', n, 0))$.*

Proof. The $n = 0$ is almost identical to the $n > 0$ case, but it's simpler. So we'll skip directly to the latter case. Let $p, p' \in P$ and suppose $n > 0$. Then,

$$\begin{aligned}
d_n(\Phi(p), \Phi(p')) &= d_n(\Psi(p,n,0), \Psi(p',n,0)) && (4.6.1)\\
&= d_n\left(\left(\lim_{k\to\infty} \Psi(p,n-1,k)\right) \circ \rho_{n+1}, \left(\lim_{k\to\infty} \Psi(p',n-1,k)\right) \circ \rho_{n+1}\right) \\
&= d_n\left(\lim_{k\to\infty} \Psi(p,n-1,k), \lim_{k\to\infty} \Psi(p',n-1,k)\right) && (4.6.2)\\
&= d_n\left(\left(\lim_{k\to\infty} F^k(\Psi(p,n-1,0)) \circ \rho_n\right),\right.\\
&\qquad \left.\left(\lim_{k\to\infty} F^k(\Psi(p',n-1,0)) \circ \rho_n\right)\right)\\
&= d_n\left(\left(\lim_{k\to\infty} F^k(\Psi(p,n-1,0)) \circ \rho_n\right) \circ \rho_n,\right.\\
&\qquad \left.\left(\lim_{k\to\infty} F^k(\Psi(p',n-1,0)) \circ \rho_n\right) \circ \rho_n\right) && (4.6.3)\\
&= d_{n+1}\left(\left(\lim_{k\to\infty} F^k(\Psi(p,n-1,0)) \circ \rho_n\right) \circ \rho_n,\right.\\
&\qquad \left.\left(\lim_{k\to\infty} F^k(\Psi(p',n-1,0)) \circ \rho_n\right) \circ \rho_n\right) && (4.6.4)\\
&= d_{n+1}\left(\left(\lim_{k\to\infty} F^k(\Psi(p,n-1,0)) \circ \rho_n\right),\right.\\
&\qquad \left.\left(\lim_{k\to\infty} F^k(\Psi(p',n-1,0)) \circ \rho_n\right)\right) && (4.6.5)\\
&= d_{n+1}(\Psi(p,n,0), \Psi(p',n,0))
\end{aligned}$$

Step Justifications:

(4.6.1) It was shown in the proof of Theorem 4.4.15 on page 97 that for all $j \geq n$ (and for any $r \in P$), $\Psi(r,j,0)$ is an X_n-approximate fixed point of $F(r,\cdot)$. Since $\Psi(r,j,0) \to \Phi(r)$, $\Phi(r)\restriction_{X_n} = \Psi(r,n,0)\restriction_{X_n}$.

(4.6.2) By definition of d_n and ρ_n, $\forall u,v \in \mathcal{C}[\mathbb{X},\mathcal{A}]\ \forall j \geq n\ d_n(u \circ \rho_j, v \circ \rho_j) = d_n(u,v)$.

(4.6.3) By Lemma 4.6.2 on the previous page, composing each of the arguments of d_n

(4.6.4) Similar justification to line (4.6.2), but this time we're using the fact that $\forall u,v \in \mathcal{C}[\mathbb{X},\mathcal{A}]\ \forall j \geq n\ d_j(u \circ \rho_n, v \circ \rho_n) = d_n(u \circ \rho_n, v \circ \rho_n)$.

(4.6.5) Again, using Lemma 4.6.2 on the preceding page.

The other lines follow by definition of Ψ. □

The proofs of the following two lemmas are routine, and thus, have been omitted.

Lemma 4.6.4. *Let X and Y be topological spaces, each with two topologies: $\mathfrak{T}_X$ and $\mathfrak{T}'_x$ for X, and $\mathfrak{T}_Y$ and $\mathfrak{T}'_Y$ for Y. Suppose $\mathfrak{T}_X$ is coarser than $\mathfrak{T}'_X$ (i.e. $\mathfrak{T}_X \subseteq \mathfrak{T}'_X$), and that $\mathfrak{T}_Y$ is coarser than $\mathfrak{T}'_Y$. Let $f : X \to Y$ and $x \in X$. Then we have the following local continuity properties for f, relative to the topologies on its domain and co-domain:*

1. *If f is continuous at x when X is equipped with $\mathfrak{T}_X$ then it is also continuous at x when X is equipped with $\mathfrak{T}'_X$.*

2. *If f is continuous at x when Y is equipped with $\mathfrak{T}'_Y$ then it is also continuous at x when Y is equipped with $\mathfrak{T}_Y$.*

Lemma 4.6.5. *Let X be a topological space, let $x \in X$, and let $f : X \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$. If f is continuous at x with respect to every pseudometric in $\{d_n\}_{n\in\mathbb{N}}$, then f is continuous at x.*

Theorem 4.6.6 (Generalized Theorem TZJ2)**.** Let (P, d_P) be a metric space and let $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$. Let $p \in P$ and let $V \subseteq P$ be a neighbourhood of p. Let $\lambda \in \mathbb{R}^+$ with $\lambda < 1$. Using the notation $F_r(u) = F(r, u)$, suppose $\forall r \in V$ F_r satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Lip}(\lambda, \mathfrak{X})$, and that $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ F is continuous at (p, u). Then $\Phi : V \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ (as described in (4.4.1) on page 96, whose existence is assured by the Generalized Theorem TZ1 on page 94) is continuous at p.

Proof. We begin by showing that Φ is continuous with respect to the topology induced by the d_0 pseudometric. We then proceed by induction, showing that for any $k \in \mathbb{N}$, if Φ is continuous with respect to the topology induced by d_n then it is continuous with respect to the topology induced by d_{n+1}. Pairing this with Lemma 4.6.5 completes the proof.

Basis Step

Since F_r satisfies $\boldsymbol{Caus}(\mathfrak{X})$ for any $r \in V$, it follows that $\forall u, v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$,

$$F(r, F(r, u))\restriction_{X_0} = F(r, u)\restriction_{X_0} = F(r, v)\restriction_{X_0}$$

In other words, $\forall r \in V$, $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$,

$$\Phi(r)\restriction_{X_0} = F(r, u)\restriction_{X_0} \tag{4.6.6}$$

By Lemma 4.6.4, since $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ F is continuous at (p, u) with respect to the local uniform topology on $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, it is also continuous at (p, u) with respect to the topology induced by the d_0 pseudometric (which is coarser than the local uniform topology). Thus, there exists a function, $\delta : \mathbb{R}^+ \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathbb{R}^+$ such that $\forall \varepsilon > 0$ $\forall p' \in V$ $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$,

$$d_P(p,p') < \delta(\varepsilon,u) \quad \Rightarrow \quad d_0(F(p,u),F(p',u)) < \varepsilon$$

In fact, since the choice of u is irrelevant when we're using the d_0 pseudometric, δ is constant on its second parameter. So choose an arbitrary $u \in \mathcal{C}[\mathbb{X},\mathcal{A}]$ and define $\delta_0 : \mathbb{R}^+ \to \mathbb{R}^+$ as simply $\delta_0(\varepsilon) = \delta(\varepsilon,u)$.
By Equation (4.6.6),

$$d_0(\Phi(p),\Phi(p')) \;=\; d_0(F(p,u_0),F(p',u_0))$$

Thus, $\forall\varepsilon > 0\ \forall p' \in V$,

$$d_P(p,p') < \delta_0(\varepsilon) \;\Rightarrow\; d_0(\Phi(p),\Phi(p')) < \varepsilon$$

Therefore, Φ is continuous at p with respect to the d_0 pseudometric.

Inductive Step

For the inductive hypothesis, assume that for some $n \in \mathbb{N}$, there is a function $\delta_n : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\forall\varepsilon > 0\ \forall p' \in V$,

$$d_P(p,p') < \delta_n(\varepsilon) \;\Rightarrow\; d_n(\Phi(p),\Phi(p')) < \varepsilon \tag{4.6.7}$$

We must show that there is a function $\delta_{n+1} : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\forall\varepsilon > 0$ $\forall p' \in V$,

$$d_P(p,p') < \delta_{n+1}(\varepsilon) \;\Rightarrow\; d_{n+1}(\Phi(p),\Phi(p')) < \varepsilon$$

To do this, we will analyze $\Phi(p)$ using Construction 4.4.12 on page 96 (because it's closer to the construction I used to prove the original, special case of this theorem). The trick is to observe that a d_{n+1}-modulus of continuity for $\Psi(\cdot,n+1,0)$ at p will serve as the desired δ_{n+1} modulus for Φ, since $\Phi(r)\restriction_{X_{n+1}} = \Psi(r,n+1,0)\restriction_{X_{n+1}}$ for all $r \in V$. We can get that modulus of continuity by beating the construction utterly senseless with a countably infinite application of the triangle inequality. Essentially, we're building a ladder of moduli of continuity between the two sequences: $\Psi(p,n,0)$, $\Psi(p,n,1)$,$\Psi(p,n,2)$,... and $\Psi(p',n,0)$, $\Psi(p',n,1)$,$\Psi(p',n,2)$,.... We build only a finite portion of the ladder—up to the N^{th} rung—where N is a carefully chosen number which depends upon λ, upon the fixed distance, $d_{n+1}(\Psi(p,n,0),\Psi(p,n,1))$, and upon the $\varepsilon > 0$ desired. Finally, using a pair of geometric series together with that N^{th} rung, we can construct the final rung (between $\Psi(p,n+1,0)$ and $\Psi(p',n+1,0)$), using the triangle inequality with $\Psi(p,n,N)$ and $\Psi(p',n,N)$ as intermediate points. See Figure 4.6.1 on page 108 for an overview (the "ladder" is bent only to avoid the suggestion that F is some dull, orderly isometry that just moves everything in one direction and never does anything interesting).

To begin, we need the bottom rectangle of the ladder. We can simply record the distance (with respect to the d_{n+1} pseudometric) between $\Psi(p,n,0)$ and $\Psi(p,n,1)$—although to avoid a potential problem with inequalities, we'll record a strictly positive number (1 works as well any) if that distance happens to be zero; all we really need is a positive upper bound for it, and any one will do. The modulus of continuity δ_n given in the inductive hypothesis provides the lowest rung (between $\Psi(p,n,0)$ and $\Psi(p',n,0)$). A single application of F gives us the next rung. Finally, all three can be put together with the triangle inequality to secure an upper bound on the distance between $\Psi(p',n,0)$ and $\Psi(p',n,1)$, thus completing the bottom rectangle (or more accurately, the quadrilateral).

In accordance with our first task, let,

$$D_p \quad = \quad \max\left\{1, d_{n+1}\left(\Psi(p,n,0), \Psi(p,n,1)\right)\right\} \tag{4.6.8}$$

Next, observe that for any $r \in V$,

$$\Psi(r,n,0)\restriction_{X_n} = \Phi(r)\restriction_{X_n}$$

By Lemma 4.6.3, $\forall p' \in V$,

$$d_n(\Phi(p), \Phi(p')) \quad = \quad d_{n+1}(\Psi(p,n,0), \Psi(p',n,0))$$

Thus, we can rewrite (4.6.7) as follows: $\forall \varepsilon > 0\ \forall p' \in V$,

$$d_P(p,p') < \delta_n(\varepsilon) \;\Rightarrow\; d_{n+1}(\Psi(p,n,0), \Psi(p',n,0)) < \varepsilon \tag{4.6.9}$$

Before we proceed, there is a bit notation that will greatly assist with the exposition. We define a family of functions, $\{H_k : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]\}_{k\in\mathbb{N}}$ as follows:

$$\begin{aligned} H_0(r,u) &= u \quad \text{(i.e. } H_0 \text{is the projection function } \pi_2) \\ H_1(r,u) &= F(r,u) \circ \rho_{n+1} \\ H_k(r,u) &= H_1 \circ \left(\widetilde{H}_1^{k-1}\right) \;\text{ for } k > 1 \end{aligned}$$

The purpose of defining H_k is that $\forall k, m \in \mathbb{N}\ \forall r \in V$,

$$\Psi(r,n,k+m) = H_k(r, \Psi(r,n,m))$$

F is continuous at (p,u) $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ by hypothesis and $\boldsymbol{hold}_{n+1}$ (for lack of a better name) is continuous everywhere by Lemma 4.3.6. Therefore, H_1—being a composition of these functions—is continuous at (p, u) $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$. From Corollary 1.4.14, for any $k > 0$, H_k is also continuous at (p, u) $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$. Thus, there exists[6] a function $\delta_n^{(1)} : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\forall (p', u) \in V \times \mathcal{C}[\mathbb{X}, \mathcal{A}]$ $\forall \varepsilon > 0$,

$$\begin{aligned} d_{n+1}\left(H_1(p, \Psi(p, n, 0)), H_1(p', u)\right) &= d_{n+1}\left(\Psi(p, n, 1), H_1(p', u)\right) \\ &< \varepsilon \end{aligned} \tag{4.6.10}$$

whenever,

$$\max\left\{d_P(p, p'),\, d_{n+1}\left(\Psi(p, n, 0), u\right)\right\} < \delta_n^{(1)}(\varepsilon) \tag{4.6.11}$$

We now have the first and second rungs (δ_n and $\delta_n^{(1)}$, respectively), along with the strut that joins them on the p side. All we need now is the strut that joins them on the p' side: a radius around p which will ensure a fixed upper bound (of, say, $2D_p$) on the distance between $\Psi(p', n, 0)$ and $\Psi(p', n, 1)$, which we'll call $D_{p'}$. More precisely, we need a number, $R \in \mathbb{R}^+$such that,

$$d_P(p, p') < R \;\Rightarrow\; D_{p'} = d_{n+1}(\Psi(p', n, 0), \Psi(p', n, 1)) < 2D_p$$

This is easy to obtain by going around the back, using $\delta_n^{(1)}$, δ_n, and the triangle inequality. Choose,

$$R = \min\left\{\delta_n\left(\frac{D_p}{2}\right), \delta_n^{(1)}\left(\frac{D_p}{2}\right), \delta_n^{(1)}\left(\delta_n\left(\frac{D_p}{2}\right)\right)\right\} \tag{4.6.12}$$

Then, given any $p' \in V$ such that $d_p(p, p') < R$, we get the following two inequalities:

$$d_{n+1}\left(\Psi(p, n, 0), \Psi(p', n, 0)\right) < \frac{D_p}{2} \tag{4.6.13}$$

$$d_{n+1}\left(\Psi(p, n, 1), \Psi(p', n, 1)\right) < \frac{D_p}{2} \tag{4.6.14}$$

Equation 4.6.13 comes directly from 4.6.9 and 4.6.12. Equation 4.6.14 is somewhat more tricky. Recall from line 4.6.11 that we need both $d_P(p, p')$ and $d_{n+1}\left(\Psi(p, n, 0), \Psi(p', n, 0)\right)$ to be less than $\delta_n^{(1)}\left(\frac{D_p}{2}\right)$ in order to ensure that $d_{n+1}\left(H_1\left(p, \Psi(p, n, 0)\right), H_1\left(p', \Psi(p', n, 0)\right)\right)$ (which is simply $d_{n+1}\left(\Psi(p, n, 1), \Psi(p', n, 1)\right)$)

[6] This step would obviously not be possible if F were continuous only w.r.t. P—as Tucker and Zucker require in their theorem. We need it to be continuous on a portion of its entire domain; continuity on only a projection of the domain is insufficient. This is what we sacrifice in order to liberate F from the necessity of being shift invariant.

is less than $\frac{D_p}{2}$. The second term in the definition of R ensures that $d_P(p,p') < \delta_n^{(1)}\left(\frac{D_p}{2}\right)$, and the third term ensures that $d_{n+1}\left(\Psi(p,n,0),\Psi(p',n,0)\right) < \delta_n^{(1)}\left(\frac{D_p}{2}\right)$. Therefore, $\forall p' \in V$, if $d_P(p,p') < R$ then,

$$\begin{aligned}
D_{p'} &= d_{n+1}\left(\Psi(p',n,0),\Psi(p',n,1)\right) \\
&\leq d_{n+1}\left(\Psi(p',n,0),\Psi(p,n,0)\right) + d_{n+1}\left(\Psi(p,n,0),\Psi(p,n,1)\right) \\
&\quad + d_{n+1}\left(\Psi(p,n,1),\Psi(p',n,1)\right) \\
&< \frac{D_p}{2} + D_p + \frac{D_p}{2} \\
&= 2D_p
\end{aligned}$$

Since D_p and λ are fixed and $\lambda < 1$, there exists a function $N : \mathbb{R}^+ \to \mathbb{Z}^+$ such that, given any $\varepsilon > 0$,

$$\frac{\lambda^{N(\varepsilon)}}{1-\lambda} 2D_p < \varepsilon \tag{4.6.15}$$

The reason for the expression above will become clear soon enough (if the reader hasn't guessed it already). We will now begin to apply the real star of the show: the contraction property! Since $F_r \in \boldsymbol{Lip}(\lambda,\mathfrak{X})$ for all $r \in V$, it follows that $\forall u,v \in \mathcal{C}[\mathbb{X},\mathcal{A}]$,

$$d_n(u,v) = 0 \;\Rightarrow\; d_{n+1}\left(H_k(r,u),H_k(r,v)\right) \leq \lambda^k d_{n+1}(u,v)$$

Thus, since $\lim_{j\to\infty}\Psi(p,n,j) = \Psi(p,n+1,0)$, it follows that $\forall\varepsilon > 0$,

$$\begin{aligned}
&d_{n+1}\left(\Psi(p,n,N(\varepsilon)),\Psi(p,n+1,0)\right) \\
&\quad\leq \sum_{j=N(\varepsilon)}^{\infty} d_{n+1}\left(\Psi(p,n,j),\Psi(p,n,j+1)\right) \\
&\quad= \sum_{j=N(\varepsilon)}^{\infty} d_{n+1}\left(H_j\left(p,\Psi(p,n,0)\right),H_j\left(p,\Psi(p,n,1)\right)\right) \\
&\quad\leq \sum_{j=0}^{\infty} d_{n+1}\left(H_{j+N(\varepsilon)}\left(p,\Psi(p,n,0)\right),H_{j+N(\varepsilon)}\left(p,\Psi(p,n,1)\right)\right) \\
&\quad\leq \sum_{j=0}^{\infty} \lambda^{j+N(\varepsilon)} d_{n+1}\left(\Psi(p,n,0),\Psi(p,n,1)\right) \\
&\quad\leq \lambda^{N(\varepsilon)} d_{n+1}\left(\Psi(p,n,0),\Psi(p,n,1)\right) \sum_{j=0}^{\infty} \lambda^j \\
&\quad\leq \frac{\lambda^{N(\varepsilon)}}{1-\lambda} D_p
\end{aligned}$$

Similarly, $\forall p' \in V$, if $d_P(p, p') < R$ (thus ensuring $D_{p'} < 2D_p$), then $\forall \varepsilon > 0$,

$$d_{n+1}(\Psi(p', n, N(\varepsilon)), \Psi(p', n+1, 0)) \leq \frac{\lambda^{N(\varepsilon)}}{1-\lambda} D_{p'} < \frac{\lambda^{N(\varepsilon)}}{1-\lambda} 2D_p \tag{4.6.16}$$

Since every H_k is continuous at (p, u) $\forall u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$, there is a $\delta_n^H : \mathbb{R}^+ \times \mathbb{Z}^+ \to \mathbb{R}^+$ such that, $\forall \varepsilon > 0$ $\forall (p', u) \in V \times \mathcal{C}[\mathbb{X}, \mathcal{A}]$ $\forall k \in \mathbb{Z}^+$,

$$\begin{aligned} d_{n+1}\left(H_k(p, \Psi(p, n, 0)), H_k(p', u)\right) &= d_{n+1}\left(\Psi(p, n, k), H_k(p', u)\right) \\ &< \varepsilon \end{aligned} \tag{4.6.17}$$

whenever,

$$\max\{d_P(p, p'),\ d_{n+1}(\Psi(p, n, 0), u)\} < \delta_n^H(\varepsilon, k) \tag{4.6.18}$$

We now use δ_n^H to obtain the function $\delta_n^* : \mathbb{R}^+ \to \mathbb{R}^+$, which will allow us to make the N^{th} rung (loosely speaking, since N is a function of ε), partway up the ladder, arbitrarily short.

$$\delta_n^*(\varepsilon) = \min\left\{\delta_n^H(\varepsilon, N(\varepsilon)), \delta_n\left(\delta_n^H(\varepsilon, N(\varepsilon))\right)\right\}$$

Now, $\forall p' \in V$ $\forall \varepsilon > 0$, if $d_P(p, p') < \delta_n^*(\varepsilon)$ then,

$$d_{n+1}(\Psi(p, n, 0), \Psi(p', n, 0)) < \delta_n^H(\varepsilon, N(\varepsilon))$$

(thanks to the second term in $\delta_n^*(\varepsilon)$), and therefore,

$$\max\{d_P(p, p'),\ d_{n+1}(\Psi(p, n, 0), \Psi(p', n, 0))\} < \delta_n^H(\varepsilon, N(\varepsilon))$$

So by (4.6.18) and (4.6.17), it follows that whenever $d_P(p, p') < \delta_n^*(\varepsilon)$, we get,

$$\begin{aligned} & d_{n+1}\left(\Psi(p, n, N(\varepsilon)), H_{N(\varepsilon)}(p', \Psi(p', n, 0))\right) \\ & = d_{n+1}\left(\Psi(p, n, N(\varepsilon)), \Psi(p', n, N(\varepsilon))\right) \\ & < \varepsilon \end{aligned} \tag{4.6.19}$$

Finally, define $\delta_{n+1} : \mathbb{R}^+ \to \mathbb{R}^+$ as follows:

$$\delta_{n+1}(\varepsilon) = \min\left\{R, \delta_n^*\left(\frac{\varepsilon}{3}\right)\right\}$$

Let $\varepsilon > 0$ and let $M = N(\varepsilon/3)$. Then $\forall p' \in V$ such that $d_P(p', p) < \delta_{n+1}(\varepsilon)$, we obtain the following three inequalities:

$$\begin{aligned} d_{n+1}(\Psi(p, n+1, 0), \Psi(p, n, M)) &< \frac{\lambda^M}{1-\lambda} D_p < \frac{\varepsilon}{3} \\ d_{n+1}(\Psi(p, n, M), \Psi(p', n, M)) &< \frac{\varepsilon}{3} \\ d_{(n+1)\tau}(\Psi(p', n, M), \Psi(p', n+1, 0)) &< \frac{\lambda^M}{1-\lambda} 2D_p < \frac{\varepsilon}{3} \end{aligned}$$

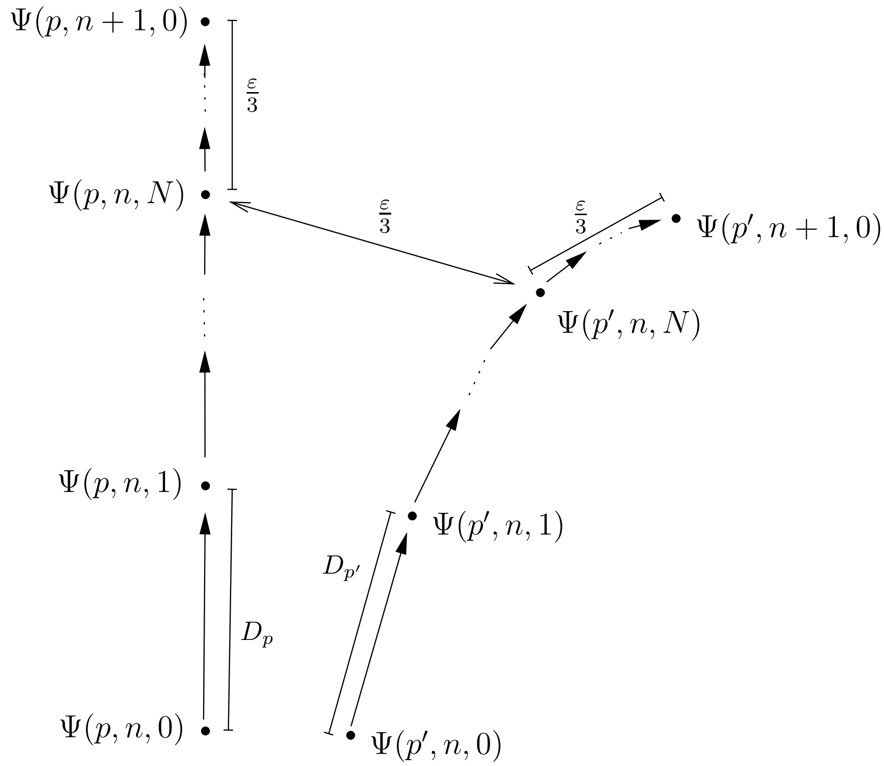

Figure 4.6.1: Inequalities used in the Generalized Theorem TZJ2

The first and third come from (4.6.15) and the second comes from (4.6.19). Merging the left-hand sides using two applications of the triangle inequality yields the final result: if $p' \in V$ and $d_P(p, p') < \delta_{n+1}(\varepsilon)$ then,

$$d_{n+1}(\Psi(p, n+1, 0), \Psi(p', n+1, 0)) < \varepsilon$$

□

We can now present the proof of Theorem TZJ2 on page 27, which follows as a corollary to Generalized Theorem TZJ2.

Theorem 1.4.15 *(Theorem TZJ2) Let (P, d_P) be a metric space and let $F : P \times \mathcal{C}[\mathbb{T}, \mathcal{A}] \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$. Let $p \in P$ and let $V \subseteq P$ be a neighbourhood of p. Let $\tau, \lambda \in \mathbb{R}^+$ with $\lambda < 1$. Using the notation $F_r(u) = F(r, u)$, suppose that for all $r \in V$ F_r satisfies* ***Caus*** *and* ***Lip***(λ, τ)*, and that for all $u \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ F is continuous at (p, u). Then $\Phi : V \to \mathcal{C}[\mathbb{T}, \mathcal{A}]$ (as described in (1.3.3) on page 15, whose existence is assured by Theorem TZ1 on page 19) is continuous at p.*

Proof. To prove that this is merely a special case of the Generalized Theorem TZJ2, we must find a compact exhaustion $\mathfrak{X}$ for $\mathbb{T}$, and show that F satisfies ***Caus***$(\mathfrak{X})$ and ***Lip***$(\lambda, \mathfrak{X})$.

Let $\mathbb{X} = \mathbb{T}$ $(= \mathbb{R}^{\geq 0})$ and for all $n \in \mathbb{N}$, let $X_n = [0, n\tau]$. Let $\mathfrak{X} = \{X_n\}_{n\in\mathbb{N}}$. Obviously, $X_0 \subseteq X_1 \subseteq \cdots \subseteq \mathbb{X}$ and $\bigcup_{n\in\mathbb{N}} X_n = \mathbb{X}$. The third property follows from the Heine-Borel Theorem[7]: every compact subset of $\mathbb{X}$ is bounded, and every bounded set is contained within a member of the exhaustion (i.e. if K is bounded, then $\exists n \in \mathbb{N}$ such that $K \subseteq X_n$). Thus $\mathfrak{X}$ satisfies Definition 4.1.1 on page 81.

Let $r \in V$, $n \in \mathbb{N}$, and let $u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$ such that $d_{n\tau}(u, v) = 0$. Then $\forall t \leq n\tau$ $d_t(u, v) = 0$. Since $F_r \in \boldsymbol{Caus}$, it follows that $\forall t \leq n\tau$ $F_r(u)(t) = F_r(v)(t)$. Therefore, $d_{n\tau}(F_r(u), F_r(v)) = 0$, and hence $F_r \in \boldsymbol{Caus}(\mathfrak{X})$. Furthermore, since $F_r \in \boldsymbol{Lip}(\lambda, \tau)$ it follows that $d_{(n+1)\tau}(F_r(u), F_r(v)) \leq \lambda d_{(n+1)\tau}(u, v)$. Thus, $F_r \in \boldsymbol{Lip}(\lambda, \mathfrak{X})$. By the Generalized Theorem TZJ2 on page 102, Φ is continuous at p. □

4.7 Concrete Computability of Φ

The theory we have developed is part of a general framework for studying analog computation. The prevailing notion in analog computation research is that the Church-Turing Thesis extends to all manner of computation (see [BCGH06, TZ04, Wei00], for example). Part of the job of testing this variant of the Church-Turing thesis is to verify that anything "computable" within our framework is computable in others as well. In our models of analog computation, a function is (implicitly defined as being) "computable" if it is the fixed point of a smoothie operator.

Given that $\mathbb{X}$ and $\mathcal{A}$ are fairly abstract spaces, how can we relate the objects in this model to classical computability theory? One way is to determine whether an operator and its fixed point can be codified somehow, using only natural numbers and computable functions on natural numbers. This is essentially a form of meta-computation. The idea behind concrete computability is somewhat similar to the central idea in many areas of mathematics (category theory, in particular). It is possible to develop an abstract mathematical structure and discover that there are morphisms which "translate" this structure to another one which appears to be completely unrelated, developed from within an entirely different context, as though it is a distorted mirror image of the original structure. Our original structure is $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ (or a multisorted algebra which includes $\mathcal{C}[\mathbb{X}, \mathcal{A}]$) along with a parameter space P, and the "mirror image" is classical computability theory on $\mathbb{N}$.

This analysis has already been done for $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ (where $\mathbb{T} = \mathbb{R}^{\geq 0}$ or $\mathbb{N}$) in [TZ12]. Tucker and Zucker identify a set of conditions, sufficient to ensure

[7]For $K \subseteq \mathbb{R}^m$, K is compact if and only if K is closed and bounded. See [Rud76], or any elementary text on real analysis for details.

that the fixed point function Φ of F is concretely computable, and it appears that their arguments can be extended naturally to the more general case in which $\mathbb{T}$ is replaced by a σ-compact space $\mathbb{X}$ with a retractable exhaustion. To show this in full detail, however, I would need to border on plagiarism since little of the theory from [TZ12] needs to be changed. So instead, I will present a summary of the core ideas and put particular emphasis on the few details that need to be changed to accommodate smoothies. One of the main aspects of their theory I will be glossing over is the allowance of *partial* stream operators. I do not address that here simply because of a lack of time. There is nothing $\mathbb{T}$-specific about it, however. Similarly, they insist on the effective local uniform continuity[8] of their streams and I impose no such requirement here. It is not strictly necessary for establishing the concrete computability of Φ, but it is a useful condition to have when defining interesting operators. I omit this treatment for both the lack of time and for the sake of brevity and simplicity.

Before we continue, we must encumber our spaces with a few additional properties. Up until now, P has been an arbitrary metric space, $\mathcal{A}$ has been an arbitrary complete metric space, and $\mathbb{X}$ has been a retractable σ-compact topological space. There was no need to assume anything more about them in this chapter. For the following material, however, we require each of these spaces to be complete, separable metric spaces. Recall that a space is complete if every Cauchy sequence converges, and a space is separable if it contains a countable dense[9] subset. The reason for this will be made clear as we go along, but it is helpful to know this in advance.

4.7.1 The Codes: α-computability

To model computation on $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ using computation on $\mathbb{N}$, we must encode the spaces and operators in our theory using natural numbers. The problem, of course, is that (in all but trivial cases) there simply aren't enough natural numbers to go around. If $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, $\mathbb{X}$, $\mathcal{A}$, and/or P are uncountable (as we typically imagine them to be), then most of the elements in these spaces and the functions on them won't be lucky enough to get their own code numbers and hence cannot be represented exactly using our $\mathbb{N}$-based model of computation for smoothies. Hence, we must settle for encoding only countable subsets

[8] Loosely speaking, the streams in $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ are *effectively locally uniformly continuous* (with respect to an exhaustion) if there is a single computable parametrized modulus of local continuity that works for all streams and all members of the exhaustion. That is, there is a computable function that accepts a (code for a) stream, a member of the exhaustion, and an $\varepsilon > 0$ (used in its traditional sense with respect to continuity), and it returns a corresponding δ sufficient for the specified stream, restricted to the specified member of the exhaustion.

[9] A subset X is "dense" in a topological space Y if the closure of X is Y.

$\mathcal{Z}_S \subseteq S$, where $S = \mathbb{X}, \mathcal{A}, P, \mathcal{C}[\mathbb{X}, \mathcal{A}]$, and we do this using surjective functions of the form $\alpha_S : \mathbb{N} \to \mathcal{Z}_S$.

There are all sorts of countable subsets we could choose for this purpose, but since we're trying to encode as much of these spaces as we can, we need subsets that are dense. That way, even if an element is not in $\mathcal{Z}_S$, we can approximate it arbitrarily closely using elements that are in $\mathcal{Z}_S$. Thus, we need $\mathbb{X}$, $\mathcal{A}$, P, and $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ to be separable (i.e. have a countable dense subset).

In fact, it can be shown (see [TZ12]) that if both $\mathbb{X}$ and $\mathcal{A}$ are separable, then $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ is too. So this one comes "for free," but it should be noted that in [TZ12] the authors require a special form of separability which must be assumed of $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ even if both $\mathbb{T}$ and $\mathcal{A}$ are separable. This assumption, however, can be made for $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ as just easily as it can be for $\mathcal{C}[\mathbb{T}, \mathcal{A}]$, if necessary for a particular purpose (in [TZ12], the authors need it to ensure that there are particularly nice Cauchy sequences in $\mathcal{C}[\mathbb{T}, \mathcal{A}]$ that facilitate some useful operations like integration to be proven computable).

We develop these α_S functions to analyze the computability of functions among the four spaces above rather than the spaces themselves. Some of the functions of interest have domains and/or codomains which are products of these spaces. In particular, we obviously need to look at functions of the form $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $\Phi : P \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$. The details are somewhat involved, so I will indulge in a bit of hand-waving and simply state that we can assume there is a single, universal encoding function $\alpha : \mathbb{N} \to \mathcal{Z}$, where $\mathcal{Z}$ is the union of all finite products of $\mathcal{Z}_{\mathbb{X}}$, $\mathcal{Z}_{\mathcal{A}}$, $\mathcal{Z}_P$, and $\mathcal{Z}_{\mathcal{C}[\mathbb{X},\mathcal{A}]}$. This is possible since a countable union of countable sets is countable.

With α-computable elements in hand (those in $\mathcal{Z}$), we proceed to define α-computable sequences. The following definition is adapted from [PER89].

Definition 4.7.1 (α-computable sequence)**.** A sequence $\{x_n\} \subseteq \mathcal{Z}$ is *α-computable* if there is a recursive function $e : \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$,

$$x_n = \alpha\left(e(n)\right)$$

Using the limited encoding we have developed so far, we can already introduce a primitive notion of operator computability:

Definition 4.7.2 (*α-computable function*)**.** Let S_1 and S_2 be finite products of $\mathbb{X}$, $\mathcal{A}$, P, and $\mathcal{C}[\mathbb{X}, \mathcal{A}]$, and suppose $f : S_1 \to S_2$. Then f is *α-computable* if there is a computable[10] function (called a *tracking function*) $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\forall k \in \alpha^{-1}\left(S_1 \cap \mathcal{Z}\right)$

$$f\left(\alpha(k)\right) = \alpha\left(\varphi(k)\right)$$

[10] "Computable" in the usual sense, i.e. recursive.

The problem with this form of computability is that $\mathcal{Z}$ is a relatively "sparse" subset (in most common choices, it will exclude infinitely many limit points which could be easily encoded), and this permits only a relatively limited encoding. We can do better than α.

4.7.2 The Computable Closure of $\mathcal{Z}$ and $\overline{\alpha}$-computability

Let S be a finite product of $\mathbb{X}$, $\mathcal{A}$, P, and $\mathcal{C}[\mathbb{X}, \mathcal{A}]$. Since each of these spaces is complete and separable, so is S (although I have omitted the proof). Furthermore, it can be shown that a product of countable dense subsets is dense in the corresponding product of spaces (proof also omitted). Therefore, $S \cap \mathcal{Z}$ is a countable dense subset of S. For convenience, let $\mathcal{Z}_S = S \cap \mathcal{Z}$. Since S is complete and $\mathcal{Z}_S \subseteq S$, every Cauchy sequence in $\mathcal{Z}_S$ converges to an element of S. Since $\mathcal{Z}_S$ is dense in S, every element in S has such a Cauchy sequence. So we can refer to any element of S using a Cauchy sequence in $\mathcal{Z}_S$ (i.e. there exists a surjection from the set of Cauchy sequences in $\mathcal{Z}_S$ onto S).

Now, since we have an encoding $\alpha : \mathbb{N} \to \mathcal{Z}$ of $\mathcal{Z}$, Cauchy sequences (and any other sequences, for that matter) in $\mathcal{Z}_S \subseteq \mathcal{Z}$ can be represented by total functions of the form $e : \mathbb{N} \to \mathbb{N}$. For any sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}_S$ there is a function $e : \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$, $u_n = \alpha(e(n))$. Here is where classical computability theory enters the picture. Some of these functions on $\mathbb{N}$ will be (classically) computable and some of them won't be. It is plainly the former class with which we are concerned, and it is these functions, together with α, that determine the *computable closure* of $\mathcal{Z}_S$ (which we write as $\mathcal{C}_\alpha(\mathcal{Z}_S)$). Once we have $\mathcal{C}_\alpha(\mathcal{Z}_S)$ for every S, we can define a new (and better) encoding $\overline{\alpha} : \mathbb{N} \to \mathcal{C}_\alpha(\mathcal{Z})$, and then define $\overline{\alpha}$-computability exactly as we did for α-computability in Definition 4.7.2 above. There is also one further subtlety to be addressed: it is not enough for the functions $e : \mathbb{N} \to \mathbb{N}$ representing Cauchy sequences to be computable; the modulus of convergence of the Cauchy sequence each e represents must also be computable (see Remark 4.7.4).

Definition 4.7.3 (α-effective Cauchy Sequence)**.** Let $\{u_n\}_{n\in\mathbb{N}} \subseteq \mathcal{Z}$ be a Cauchy sequence. Then $\{u_n\}_{n\in\mathbb{N}}$ is an *α-effective Cauchy sequence* if the following two conditions hold:

1. The sequence itself is α-computable. That is, there is a total computable function $e : \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$
$$u_n = \alpha(e(n))$$
2. The convergence of the sequence is *effective* (it has a computable modulus of convergence). That is, there is a total computable function $M : \mathbb{N} \to \mathbb{N}$ such that for all $j, k, n \in \mathbb{N}$,
$$j, k \geq M(n) \;\Rightarrow\; d\left(u_j, u_k\right) < 2^{-n}$$

Remark 4.7.4. The first condition is obvious. Each of the elements in the sequence must be computable (w.r.t. α), and so too must be the sequence itself. Otherwise it could hardly be regarded as an α-effective sequence of any kind. The second condition is less obvious. As long as we can compute the sequence, and we know that it converges, why must we be able to compute in advance how far out in the sequence we must go to get within a certain radius of the limit? It may seem to be a superfluous condition, but without it, we cannot legitimately claim that the limit is computable.

It's not a matter of being able to compute in advance how far along we must go in the sequence; it's a matter of being able to determine—at any point—whether we're even remotely close to the limit. If we claim that an element is computable, we mean that we have a mechanism for generating a point arbitrarily close to it. A Cauchy sequence will eventually contain such points, but it is under no obligation to begin marching steadily toward its limit right from the start. The first million points of the sequence may appear to be steadily converging within a tiny portion of the space, and then suddenly, in the next point, it might spontaneously veer quite far away and appear to begin converging in a region very distant from the previous one. This may happen any number of times before it begins to converge in earnest. Without being able to compute its modulus of convergence, how can we be at all justified in claiming the sequence is effective? We have a way of generating our sequence and we know that it will eventually generate a satisfactory point (one close enough), but unless its modulus is computable, we have no way of selecting such a point. Hence, we may as well be generating a completely random sequence of points.

That is the reason for insisting on a computable modulus of convergence.

Lemma 4.7.5 ("Fast" Cauchy Sequences)**.** *Without loss of generality, we can assume that the modulus of convergence of an α-effective Cauchy sequence (in some metric space with metric d) is simply the identity function.*

That is, suppose $\{x_n\}_{n\in\mathbb{N}}$ is an α-effective Cauchy sequence with associated recursive functions $e, M : \mathbb{N} \to \mathbb{N}$ such that $\forall n \in \mathbb{N}\ x_n = e(n)$ and $\forall j, k, \ell \in \mathbb{N}$, $j, k \geq M(\ell) \Rightarrow d(x_j, x_k) < 2^{-\ell}$. Then there exists another α-effective Cauchy sequence $\{x'_n\}_{n\in\mathbb{N}}$ such that $\forall j, k, \ell \in \mathbb{N}$, $j, k \geq \ell \Rightarrow d(x'_j, x'_k) < 2^{-\ell}$ (i.e. M', if it were to be defined, would be merely the identity).

Proof. First, without loss of generality, we can assume that M is monotonic (increasing). This is a fairly standard assumption for moduli of convergence and continuity in any context, and it is easy to show that computability is not threatened by it.

Any finite composition of recursive functions is recursive, so simply define $e' : \mathbb{N} \to \mathbb{N}$ as follows:

$$e'(n) = e(M(n))$$

and, of course, set $x'_n = e'(n)$ for all n. Then e' is recursive and $\forall j,k,\ell \in \mathbb{N}$, if $j,k \geq \ell$, it follows that $M(j), M(k) \geq M(\ell)$. Thus,

$$\begin{aligned} d(x'_j, x'_k) &= d(e'(j), e'(k)) \\ &= d\left(e(M(j)), e(M(k))\right) \\ &< 2^{-\ell} \end{aligned}$$

□

Notation 4.7.6. Let $\mathcal{C}_\alpha(\mathcal{Z})$ be the set of all limits of α-effective Cauchy sequences in $\mathcal{Z}$ (and likewise for $\mathcal{C}_\alpha(\mathcal{Z}_S)$, given any product S of spaces).

Definition 4.7.7 ($\Omega_{\overline{\alpha}}$ and $\overline{\alpha}$)**.** By definition, for every element of $\mathcal{C}_\alpha(\mathcal{Z})$, there is an α-effective Cauchy sequence with two associated computable functions (e and M). Every computable function can be uniquely represented as a Gödel number in $\mathbb{N}$, and every pair of natural numbers can be encoded as a single natural number (using, for example, a second Gödel numbering). Therefore, there is a set $\Omega_{\overline{\alpha}} \subseteq \mathbb{N}$ with a surjective function $\overline{\alpha} : \Omega_{\overline{\alpha}} \to \mathcal{C}_\alpha(\mathcal{Z})$ that encodes $\mathcal{C}_\alpha(\mathcal{Z})$.

Definition 4.7.8 ($\overline{\alpha}$-computable function)**.** As in Definition 4.7.2, let S_1 and S_2 be finite products of $\mathbb{X}$, $\mathcal{A}$, P, and $\mathcal{C}[\mathbb{X},\mathcal{A}]$, and suppose $f : S_1 \to S_2$. Then f is *$\overline{\alpha}$-computable* if there is a computable (tracking) function $\varphi : \mathbb{N} \to \mathbb{N}$ such that $\forall k \in \overline{\alpha}^{-1}\left(\mathcal{C}_\alpha(\mathcal{Z}_{S_1})\right)$

$$f\left(\overline{\alpha}\left(k\right)\right) = \overline{\alpha}\left(\varphi\left(k\right)\right)$$

Remark 4.7.9. It is natural, at this point, to wonder whether we need to be concerned with $\overline{\alpha}$-effective Cauchy sequences (and consequently, $\overline{\overline{\alpha}}$-computable operators). Fortunately, the answer is no. $\mathcal{C}_\alpha(\mathcal{Z})$ is "$\overline{\alpha}$-computably closed."

Lemma 4.7.10. *Let S be any finite product of our four spaces (as in Definition 4.7.2), and let $\{s_n\}_{n\in\mathbb{N}} \subseteq \mathcal{C}_\alpha(\mathcal{Z}_S)$ be an $\overline{\alpha}$-effective Cauchy sequence (i.e. the sequence satisfies Definition 4.7.3 when α is replaced by $\overline{\alpha}$) which converges to an element $s \in S$. Then there is an α-effective Cauchy sequence which also converges to s.*

Proof. If $\{s_n\}_{n\in\mathbb{N}}$ is an $\overline{\alpha}$-computable sequence, then there is a recursive function $e : \mathbb{N} \to \mathbb{N}$ such that for each $n \in \mathbb{N}$, $s_n = \overline{\alpha}(e(n))$. Now, any such $e(n)$ is actually the Gödel number for a pair of other recursive functions: e_n and M_n. e_n is the function which defines the α-effective Cauchy sequence $\{s_{nk}\}_{k\in\mathbb{N}}$ and M_n is its modulus of convergence (however, as we observed in Lemma 4.7.5, we can assume without loss of generality that each M_n is simply the identity and thus ignore it). Decoding a Gödel number for a recursive function and evaluating it at a given point is, itself, recursive (e.g. consider the Universal

Turing Machine). Therefore, there is a recursive function $\overline{e} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that $\forall n, k \in \mathbb{N}$, $\overline{e}(n,k) = s_{nk}$ and $\forall n, j, k, \ell \in \mathbb{N}$,

$$j, k \geq \ell \;\Rightarrow\; d_S(s_{nj}, s_{nk}) < 2^{-\ell}$$

Thus, $\{s_{nk}\}_{n,k\in\mathbb{N}}$ is an α-computable double sequence, each row of which converges at a brisk minimum rate (at least as fast as $2^{-n} \to 0$ as $n \to \infty$).

While Lemma 4.7.5 speaks only of α-effective Cauchy sequences, it is equally applicable to $\overline{\alpha}$-computable Cauchy sequences with $\overline{\alpha}$-effective moduli of convergence. Thus, we can assume that $\forall j, k, \ell \in \mathbb{N}$,

$$j, k \geq \ell \;\Rightarrow\; d_S(s_j, s_k) < 2^{-\ell}$$

Since $\{s_{nk}\}_{n,k\in\mathbb{N}}$ is α-computable, so is the sequence $\{r_n = s_{nn}\}_{n\in\mathbb{N}}$. For this sequence, we can use the (obviously recursive) modulus of convergence $M(\ell) = \ell + 2$, as we now demonstrate. For any $n \in \mathbb{N}$,

$$\begin{aligned} d_S(r_n, s_n) &= d_S(r_n, \lim_{k\to\infty} s_{nk}) \\ &= \lim_{k\to\infty} d_S(r_n, s_{nk}) \text{ (since } d_S \text{ is continuous by Lemma 1.4.1)} \\ &\leq 2^{-n} \end{aligned}$$

Therefore, given any $\ell \in \mathbb{N}$, for all $j, k \geq \ell + 2$,

$$\begin{aligned} d_S(r_j, r_k) &\leq d_S(r_j, s_j) + d_S(s_j, s_k) + d_S(s_k, r_k) \\ &< 2^{-\ell-2} + 2^{-\ell-2} + 2^{-\ell-2} \\ &= 3 \cdot 2^{-\ell-2} \\ &< 2^{-\ell} \end{aligned}$$

□

4.7.3 The $\overline{\alpha}$-computability of Φ

The objective in this line of inquiry is to establish a set of conditions on an operator $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$, along with the spaces comprising its domain and codomain sufficient to ensure that if F has a fixed point function Φ defined on P, then this Φ is concretely computable.

Theorem 4.7.11 (Concrete Computability Theorem)**.** *Suppose the antecedents of the Generalized TZ1 Theorem for the Alternate Construction on page 97 are satisfied by some operator $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ at every point $p \in P$. That is,* *(a)* (P, d_P) is a metric space.

(b) $\forall p \in \mathcal{C}_\alpha(\mathcal{Z}_P)$ λ_p is a real number with $0 < \lambda_p < 1$.

(c) $\mathbb{X}$ *is a* σ*-compact space with a retractable compact exhaustion* $\mathfrak{X} = \{X_n\}_{n\in\mathbb{N}}$ *and retractions* $\{\rho_n\}_{n\in\mathbb{N}}$.

(d) $\forall p \in P$, $F(p, \cdot) : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ satisfies $\boldsymbol{Caus}(\mathfrak{X})$ and $\boldsymbol{Lip}(\lambda_p, \mathfrak{X})$ for all $u \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$.

And further, suppose

(e) P*,* $\mathbb{X}$*, and* $\mathcal{A}$ *are complete separable metric spaces.*

(f) *For each* $n \in \mathbb{N}$*,* $\mathcal{C}_\alpha(\mathcal{Z}_{\mathcal{C}[\mathbb{X},\mathcal{A}]})$ *is closed under* $\boldsymbol{hold}_n$*.*

(g) $\boldsymbol{hold} : \mathbb{N} \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ *is* $\overline{\alpha}$*-computable*[11] *(this implicitly requires (e), of course).*

(h) F *is* $\overline{\alpha}$*-computable.*

(i) *The parametrized pseudometric* $d : \mathbb{N}\times\mathcal{C}[\mathbb{X}, \mathcal{A}]^2 \to \mathbb{R}^{\geq 0}$ *(where* $d(n, u, v) = d_n(u, v)$*) is* $\overline{\alpha}$*-computable.*

(j) *There is an* $\overline{\alpha}$*-computable function* $\Lambda : P \to \mathbb{R}^+$ *such that* $\forall p \in P$ $\Lambda(p) = \lambda_p$

Then the fixed-point function $\Phi : P \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ *for* F *is* $\overline{\alpha}$*-computable.*

Proof. It is easiest to use Construction 4.4.12 to prove this since it involves fewer spaces (it never uses any of the $\mathcal{C}[X_n, \mathcal{A}]$ spaces) and no induced operators (truncations of F), both of which would require extra care to be taken at each step.

To show that Φ is $\overline{\alpha}$-computable, we must first show that Ψ is $\overline{\alpha}$-computable. In Construction 4.4.12, we chose an arbitrary initial point u_0 and set $\Psi(p, 0, 0) = F(p, u_0) \circ \rho_1 = \boldsymbol{hold}_1(F(p, u_0))$ for all $p \in P$. If we wish for $\Psi(\cdot, 0, 0) : P \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ to be $\overline{\alpha}$-computable, however, u_0 must obviously be chosen from $\mathcal{C}_\alpha(\mathcal{Z})$. Since $F : P \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $\boldsymbol{hold}_1$ are $\overline{\alpha}$-computable (by hypothesis) and the projection function $\pi_1 : \mathbb{N}^2 \to \mathbb{N}$ (which maps $(i, j) \to i$ for all $(i, j) \in \mathbb{N}^2$) is recursive, it follows that both $F(\cdot, u_0) : P \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and $\Psi(\cdot, 0, 0)$ are $\overline{\alpha}$-computable (since a composition of finitely many recursive

[11] I am uncertain whether the $\overline{\alpha}$-computability of $\boldsymbol{hold} : \mathbb{N} \times \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ would necessarily follow from the $\overline{\alpha}$-computability of $\rho : \mathbb{N} \times \mathbb{X} \to \mathbb{X}$ (if we were to insist on it instead), but the former computability is the one required for this theorem, so I have included it in the antecedent directly.

functions is recursive). From the former it follows that, for any $k \in \mathbb{N}$, $F^k(\cdot, u_0)$ is $\overline{\alpha}$-computable, and thus so is $\Psi(\cdot, 0, k+1) = \boldsymbol{hold}_1\left(F^k(\cdot, \Psi(\cdot, 0, k))\right)$.

We must be cautious with the notation and the vernacular here since there is a difference between showing that every point in the range of a function is $\overline{\alpha}$-computable, and showing that the function itself is $\overline{\alpha}$-computable. We have shown both above (for $\Psi(\cdot, 0, \cdot) : P \times \mathbb{N} \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$), although the latter only loosely. To see it more clearly, note that if $\varphi_F : \mathbb{N} \to \mathbb{N}$ is an $\overline{\alpha}$-tracking function for $F(\cdot, u_0)$, then the $\overline{\alpha}$-tracking function φ_G for $G(p,k) = F^k(p, u_0)$ is actually primitive recursive (not just recursive):

$$\begin{aligned} \varphi_G(j, 0) &= \varphi_F(j) \\ \varphi_G(j, k+1) &= \varphi_F(\varphi_G(j,k)) \end{aligned}$$

Alternatively, we could invoke the Church-Turing Thesis and express the tracking function for $\Psi(p, 0, k)$ using a programming language together with φ_F and the tracking function for $\boldsymbol{hold}_1$.

All of the above is clearly applicable to $\Psi(p, n, k)$ for any values of $n, k > 0$— provided that $\Psi(p, n, k)$ is $\overline{\alpha}$-computable for $k = 0$. We've shown above that $\Psi(p, n, k)$ is $\overline{\alpha}$-computable for $k = 0$ when $n = 0$, but getting $\Psi(p, n, 0)$ for $n > 0$ is more challenging. Recall that for $n > 0$, we have defined

$$\Psi(p, n+1, 0) = \boldsymbol{hold}_{n+2}\left(\lim_{k \to \infty} \Psi(p, n, k)\right)$$

From the argument above, $\{\Psi(p, n, k)\}_{k \in \mathbb{N}}$ is certainly an $\overline{\alpha}$-computable sequence, but we must show that it is also an α-effective Cauchy sequence to ensure its limit is $\overline{\alpha}$-computable. That is, we must show there is an $\overline{\alpha}$-computable $M : P \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ such that $\forall p \in P$, $\forall k_1, k_2, n, \ell \in \mathbb{N}$,

$$k_1, k_2 \geq M(p, n, \ell) \;\Rightarrow\; d_{\mathcal{C}[\mathbb{X}, \mathcal{A}]}(\Psi(p, n, k_1), \Psi(p, n, k_2)) < 2^{-\ell}$$

Almost exactly this was done already in the development of $N(\varepsilon)$ in the proof of the Generalized TZ1 Theorem for the Alternate Construction on page 97, but our requirements here are a little more stringent. In particular, we must ensure that

- M has the form $P \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, instead of $\mathbb{R}^+ \to \mathbb{R}^+$ (as N has in that proof).
- M is an $\overline{\alpha}$-computable function (which is obviously a requirement N didn't have to satisfy).
- M is developed with respect to the metric $d_{\mathcal{C}[\mathbb{X}, \mathcal{A}]}$ rather than the pseudometric d_{n+1}. This is, however, required merely by the exposition. For

> the sake of simplicity, I've neglected to develop a "local" version of the $\overline{\alpha}$ theory in this chapter (that would allow for a pseudometric modulus of convergence), but Tucker and Zucker develop this for streams in [TZ12] and their work appears to carry over to smoothies naturally (although I would need more time to confirm that there are no snags along the way).

It is possible to show (although I won't do it here), that the following definition for $M : P \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is $\overline{\alpha}$-computable:

$$M(p, n, \ell) = \left\lfloor \log_{\Lambda(p)} \left(\frac{\ell - n + 1}{D(n)} (1 - \Lambda(p))\, 2^{-\ell-1} \right) \right\rfloor + 1$$

where $\forall n \in \mathbb{N}$,

$$D(n) = \max\{1, d(n+1, \Psi(p, n, 0), \Psi(p, n, 1))\}$$

We know M is $\overline{\alpha}$-computable because d and Λ are $\overline{\alpha}$-computable by hypothesis (for precisely this purpose, in fact), and the rest is composed of elementary real functions which can be shown to be $\overline{\alpha}$-computable.

We now prove that M is a modulus of convergence for $\{\Psi(p, n, k)\}_{k \in \mathbb{N}}$. Let $\ell, n \in \mathbb{N}$, and assume without loss of generality that $\ell \geq n$. Let $k_1, k_2 \geq M(p, n, \ell)$, and assume (again without loss of generality) that $k_1 \leq k_2$. Then,

$$\begin{aligned}
&d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(\Psi(p, n, k_1), \Psi(p, n, k_2)) \\
&\quad = \sum_{i=0}^{\infty} \min\left(2^{-i}, d_i(\Psi(p, n, k_1), \Psi(p, n, k_2))\right) \\
&\quad = \sum_{i=n+1}^{\infty} \min\left(2^{-i}, d_i(\Psi(p, n, k_1), \Psi(p, n, k_2))\right) && (4.7.1) \\
&\quad \leq \sum_{i=n+1}^{\infty} \min\left(2^{-i}, d_{n+1}(\Psi(p, n, k_1), \Psi(p, n, k_2))\right) && (4.7.2)
\end{aligned}$$

(4.7.1) follows from the fact that $\Psi(p, n, k_1)$ and $\Psi(p, n, k_2)$ agree on X_n and (4.7.2) follows from Lemma 4.4.13.

The remaining steps are familiar from several earlier proofs in the thesis. Continuing from (4.7.2),

$$
\begin{aligned}
&\sum_{i=n+1}^{\infty} \min\left(2^{-i}, d_{n+1}(\Psi(p,n,k_1), \Psi(p,n,k_2))\right) \\
&\quad \leq \sum_{i=n+1}^{\ell+1} d_{n+1}(\Psi(p,n,k_1), \Psi(p,n,k_2)) + \sum_{i=\ell+2}^{\infty} 2^{-i} \\
&\quad = (\ell - n + 1)\, d_{n+1}(\Psi(p,n,k_1), \Psi(p,n,k_2)) + 2^{-\ell-1} \\
&\quad \leq (\ell - n + 1) \sum_{i=k_1}^{k_2-1} d_{n+1}(\Psi(p,n,i), \Psi(p,n,i+1)) + 2^{-\ell-1} \\
&\quad \leq (\ell - n + 1)\, D(n) \sum_{i=k_1}^{k_2-1} \lambda_p^i + 2^{-\ell-1} \\
&\quad = (\ell - n + 1)\, D(n)\, \lambda_p^{k_1} \frac{1 - \lambda_p^{k_2-k_1}}{1-\lambda_p} + 2^{-\ell-1} \\
&\quad \leq (\ell - n + 1)\, D(n)\, \lambda_p^{k_1} \frac{1}{1-\lambda_p} + 2^{-\ell-1} \\
&\quad < 2^{-\ell-1} + 2^{-\ell-1} \\
&\quad = 2^{-\ell}
\end{aligned}
$$

Therefore, Ψ is $\overline{\alpha}$-computable.

What remains to be shown is that there is also an $\overline{\alpha}$-computable modulus of convergence

$$M' : P \times \mathbb{N} \to \mathbb{N}$$

for $\lim_{n\to\infty} \Psi(p,n,0)$. Mercifully, this is *much* more straightforward: $\{\Psi(p,n,0)\}_{n\in\mathbb{N}}$ is already (almost) a "fast" Cauchy sequence! Its modulus of convergence is given by $M'(p,\ell) = \ell{+}1$, which we will now show as the final step.

As we have established previously, given any $n \in \mathbb{N}$, $\forall m \geq n$, $\Psi(p,m,0)$ is an X_n-approximate fixed point of F. Thus, if $\ell \in \mathbb{N}$ and $m, n \geq \ell + 1$,

$$
\begin{aligned}
d_{\mathcal{C}[\mathbb{X},\mathcal{A}]}(\Psi(p,n,0), \Psi(p,m,0)) &= \sum_{i=0}^{\infty} \min\left(2^{-i}, d_i\left(\Psi(p,n,0), \Psi(p,m,0)\right)\right) \\
&= \sum_{i=\ell+2}^{\infty} \min\left(2^{-i}, d_i\left(\Psi(p,n,0), \Psi(p,m,0)\right)\right) \\
&\leq \sum_{i=\ell+2}^{\infty} 2^{-i} \\
&= 2^{-\ell-1} \\
&< 2^{-\ell}
\end{aligned}
$$

□

Remark 4.7.12. There is a somewhat major weakness in the Concrete Computability Theorem that prevents it from being a generalization of Theorem 1 from [TZ12] (which was my original goal): it relies on a single retractable exhaustion, $\mathfrak{X}$. In Tucker and Zucker's paper, there is a family of contraction moduli $\lambda_{c,a,x}$ and a family of increments $\tau_{c,a,x}$ such that F locally satisfies $\boldsymbol{Contr}(\lambda_{c,a,x}, \tau_{c,a,x})$. In the theorem above, we do have a parametrized family of contraction moduli λ_p, but essentially we have only the one "increment" (exhaustion). I do believe it would be relatively easy to expand the theorem, allowing for a family of compact exhaustions $\{\mathfrak{X}_p\}_{p\in P}$ such that for each $p \in P$, F satisfies $\boldsymbol{Caus}(\mathfrak{X}_p)$ and $\boldsymbol{Contr}(\lambda_p, \mathfrak{X}_p)$, but I haven't taken the time to attempt the theorem this way. Hence, I will relegate that project for future work, along with the following additional ideas.

4.8 Future Work

4.8.1 Study the Abstract Computability of Φ

In [TZ12], the computability of the model presented in [TZ11] is analyzed from two different perspectives: concrete computability and abstract computability. I believe I have done the bulk of the work in generalizing concrete computability to $\mathcal{C}[\mathbb{X}, \mathcal{A}]$ (although, clearly much remains to be done before that work can be considered complete), and it seems to hold up very well. It would be interesting to see whether the same is true of abstract computability. In abstract computability, a more algebraic approach is taken (verses the analytic approach of concrete computability) and the stream/smoothie operators are approximated using a simple imperative language that is independent of the data representation and is augmented by the operations defined on the data types being used. The language used by Tucker and Zucker is called *WhileCC**, and it includes "while" loops, a nondeterministic countable choice function (the "CC" part of the name), and arrays of arbitrary length (the "*" part of the name).

4.8.2 Generalize $\mathcal{A}$ from metric spaces to uniform Hausdorff spaces

There is a way to generalize Banach's Fixed Point Theorem so that it doesn't require a metric. I thought of a way to do this, myself, but E. Tarafdar appears to have beat me by a few decades [Tar74] (although, admittedly, with a much more thoroughly-developed idea than I had). Rather than working within a metric space, we work within a *uniform space*–which is a type of topological space strictly more general than a metric space. In a uniform space, we don't (necessarily) have anything like a metric; instead we have a family of "entourages." An entourage of a uniform space X is a collection of subsets of X^2 that satisfy certain properties devised to impart a notion of proximity without necessitating actual "distance."

If (X, d) is a metric space, the uniformity induced by the metric consists of one entourage for every $r \in \mathbb{R}^+$. The entourage associated with r is the set of all pairs of points no further than r of each other. That is,

$$E_r = \left\{(x, y) \in X^2 \,:\, d(x, y) \leq r\right\}$$

With a system of entourages, it is possible to define contractions and nonexpansions in a few different ways, each of which permits a variation of Banach's Fixed Point Theorem. Some approaches are outlined in [Tar74], and I believe they might be applicable here. Generalizing the concrete computability of Φ to uniform spaces would require the use of Cauchy filters in place of Cauchy sequences, so this could be a major undertaking, but it seems quite feasible.

4.8.3 An Alternative to *Contr*

Another generalization of Banach's Fixed Point Theorem occurred to me as well: the theorem would still hold for an operator f that isn't contracting, as long as there is some $n \in \mathbb{N}$ such that f^n is contracting. Again, this was too obvious not to have been studied already. The obvious name for a such a property would be "eventually contracting," and quick search reveals the following definition from [HK03]:

Definition 4.8.1. Let X be a metric space, $C \in \mathbb{R}^+$, $\lambda \in (0, 1)$, and $f : X \to X$. Then f is *eventually contracting* if $\forall n \in \mathbb{N}\ \forall x, y \in X$,

$$d\left(f^n(x), f^n(y)\right) \leq C\lambda^n f(x, y)$$

This is my definition (which I suspect is roughly equivalent):

Definition 4.8.2. $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is *progressively contracting (or $F \in$* $\boldsymbol{PContr}(\lambda, \eta, \mathfrak{X})$) if there is a function $\eta : \mathbb{N} \to \mathbb{N}$ and a constant λ (with $0 < \lambda < 1$) such that $\forall N \in \mathbb{N}\ \forall u, v \in \mathcal{C}[\mathbb{T}, \mathcal{A}]$,

$$d_N\left(F^{\eta(N)}u, F^{\eta(N)}v\right) \leq \lambda d_N(u, v)$$

Furthermore, we say F is *effectively progressively contracting* if η is recursive.

Example 4.8.3. $F : \mathcal{C}[\mathbb{T}, \mathbb{R}] \to \mathcal{C}[\mathbb{T}, \mathbb{R}]$ where $F(u)(t) = \int_0^t u(s)\, ds + f(t)$ (and $f \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$) is effectively progressively contracting.

Proof. First we demonstrate that $\forall u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]\ \forall N, k \in \mathbb{N}$,

$$d_N(F^k u, F^k v) \leq \frac{N^k}{k!} \cdot d_N(u, v)$$

Since F^0 is simply the identity on $\mathcal{C}[\mathbb{T}, \mathbb{R}]$ and $\frac{N^0}{0!} = 1$, the statement holds for $k = 0$.

Now let $k \in \mathbb{N}$ and suppose that $\forall N \in \mathbb{N}\ \forall u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]\ d_N\left(F^k(u), F^k(v)\right) \leq \frac{N^k}{k!} d_N(u, v)$. Then, $\forall N \in \mathbb{N}\ \forall u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$,

$$\begin{aligned}
d_N\left(F^{k+1}(u), F^{k+1}(v)\right) &= \max_{0 \leq t \leq N} \left| \left(\int_0^t F^k(u)(s)\, ds + f(t) \right) \right. \\
&\qquad \left. - \left(\int_0^t F^k(v)(s)\, ds + f(t) \right) \right| \\
&= \max_{0 \leq t \leq N} \left| \int_0^t \left(F^k(u)(s) - F^k(v)(s) \right) ds \right| \\
&\leq \max_{0 \leq t \leq N} \int_0^t \left| F^k(u)(s) - F^k(v)(s) \right| ds \\
&= \int_0^N \left| F^k(u)(s) - F^k(v)(s) \right| ds \\
&\leq \int_0^N \max_{0 \leq r \leq s} \left| F^k(u)(r) - F^k(v)(r) \right| ds \\
&= \int_0^N d_s\left(F^k(u), F^k(v)\right) ds \\
&\leq \int_0^N \frac{s^k}{k!} d_s(u, v)\, ds \\
&\leq d_N(u, v) \int_0^N \frac{s^k}{k!}\, ds \\
&= \frac{N^{k+1}}{(k+1)!} d_N(u, v)
\end{aligned}$$

Define $\eta(N) = \max\{3N, 1\}$. Then $\forall N \in \mathbb{N}^+$, with $k = \eta(N)$ (for convenience), we observe that

$$\frac{N^k}{k!} = \left(\frac{k}{3}\right)^k \frac{1}{k!} < \left(\frac{k}{e}\right)^k \frac{1}{k!} < \sqrt{k} \left(\frac{k}{e}\right)^k \frac{1}{k!} = \frac{1}{\sqrt{2\pi}} \left(\sqrt{2\pi k} \left(\frac{k}{e}\right)^k \frac{1}{k!} \right)$$

Stirling's Formula provides the following inequality for any $k \in \mathbb{N}$,

$$\sqrt{2\pi k} \left(\frac{k}{e}\right)^k < k!$$

Therefore, letting $\lambda = \frac{1}{\sqrt{2\pi}}$, we compute,

$$\frac{N^k}{k!} < \lambda \left(\sqrt{2\pi k} \left(\frac{k}{e} \right)^k \frac{1}{k!} \right) < \lambda < 1$$

Hence, $\forall u, v \in \mathcal{C}[\mathbb{T}, \mathbb{R}]$ $\forall N \in \mathbb{N}^+$,

$$d_N(F^{\eta(N)}u, F^{\eta(N)}v) \le \frac{N^{\eta(N)}}{\eta(N)!} \cdot d_N(u, v) \le \lambda d_N(u, v)$$

For $N = 0$, $d_N(F^{\eta(N)}u, F^{\eta(N)}v) = d_0(Fu, Fv) = 0 \le \lambda d_0(u, v)$.
Thus, F is progressively contracting, and since η is clearly recursive, this contraction is effective. $\square$

Remark 4.8.4. I'm sure this argument can be adapted to work for any $F : \mathcal{C}[\mathbb{T}, \mathbb{R}]^m \to \mathcal{C}[\mathbb{T}, \mathbb{R}]^m$ of the form,

$$F(\mathbf{u})(t) = \int_0^t A\mathbf{u}(s)\, ds + \mathbf{f}(t)$$

where $A \in \mathbb{R}^{m \times m}$ and $\mathbf{f} \in \mathcal{C}[\mathbb{T}, \mathbb{R}]^m$. I just wanted to check that the simpler version works first.
Hence, this version of the theory—while possibly not quite as broad as the versions which use $\mathbf{Contr}(\lambda, \tau)$ and $\boldsymbol{Contr}(\lambda, \mathfrak{X})$—should still work with the two mass-spring-damper case studies in [TZ11] and it offers a diagonal construction which will obviously converge to the same stream as the "ω^2" process from that paper and the other constructions in this thesis do.

Theorem 4.8.5 (Progressive Contraction Theorem)**.** *If $F : \mathcal{C}[\mathbb{X}, \mathcal{A}] \to \mathcal{C}[\mathbb{X}, \mathcal{A}]$ is progressively contracting, then it has a unique fixed point.*

Proof. Let $v_0 \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ and define the sequence, $\left\{v_k = F^k\left(v_0\right)\right\}_{k \in \mathbb{N}}$, which we will show is locally uniformly Cauchy. To do so, we must show that $\forall N \in \mathbb{N}$ $\forall \varepsilon > 0$ $\exists M \in \mathbb{N}$ such that $\forall n, m \ge M$,

$$d_N(v_m, v_n) < \varepsilon$$

Without loss of generality, assume $\eta(N) \ge 2$ and let[12],

$$r = \max_{0 \le m, n \le \eta(N)} d_N\left(v_m, v_n\right)$$

[12]It may appear at first glance (if you see where this is going) that r should be defined as $\max_{0 \le m, n \le \eta(N)-1} d_N\left(v_k, v_j\right)$ since $v_{\eta(N)} = F^{\eta(N)} v_0$. The inclusion of $v_{\eta(N)}$ itself seems superfluous, but this inclusion is actually deliberate and essential.

Since F is progressively contracting, $\forall m, n \in \{0, 1, \ldots, \eta(N) - 1\}$ $\forall k \geq 0$,
$$d_N\left(F^{k\cdot\eta(N)}(v_m), F^{k\cdot\eta(N)}(v_n)\right) \leq \lambda^k d_N(v_m, v_n) \leq \lambda^k r$$
Let $M \in \mathbb{N}$ be a number such that,
$$\lambda^M < \frac{\varepsilon(1-\lambda)}{r}$$
For example, we could take,
$$M = \left\lceil \log_\lambda \frac{\varepsilon(1-\lambda)}{r} \right\rceil$$
Let $m, n \in \mathbb{N}$. Then $\exists m_1, n_1 \in \mathbb{N}$ $\exists m_2, n_2 \in \{0, 1, \ldots, \eta(N) - 1\}$ such that $m = m_1 \cdot \eta(N) + m_2$, $n = n_1 \cdot \eta(N) + n_2$. Without loss of generality, assume $m_1 \leq n_1$ and let $q = n_1 - m_1$. Then,
$$\begin{aligned}
d_N(v_m, v_n) &= d_N\left(F^{m_1\cdot\eta(N)}(v_{m_2}), F^{n_1\cdot\eta(N)}(v_{n_2})\right) \\
&= d_N\left(F^{m_1\cdot\eta(N)}(v_{m_2}), F^{m_1\cdot\eta(N)}\left(F^{q\cdot\eta(N)}(v_{n_2})\right)\right) \\
&\leq \lambda^{m_1} d_N\left(v_{m_2}, F^{q\cdot\eta(N)}(v_{n_2})\right) \\
&\leq \lambda^{m_1}\Bigg(d_N\left(v_{m_2}, v_{\eta(N)}\right) + \sum_{i=1}^{q-1} d_N\left(v_{i\cdot\eta(N)}, v_{(i+1)\eta(N)}\right) \\
&\qquad + d_N\left(v_{q\cdot\eta(N)}, v_{q\cdot\eta(N)+n_2}\right)\Bigg) \\
&\leq \lambda^{m_1}\Bigg(r + \sum_{i=1}^{q-1} d_N\left(F^{i\cdot\eta(N)}(v_0), F^{i\cdot\eta(N)}\left(v_{\eta(N)}\right)\right) \\
&\qquad + d_N\left(F^{q\cdot\eta(N)}(v_0), F^{q\cdot\eta(N)}(v_{n_2})\right)\Bigg) \\
&\leq \lambda^{m_1}\left(r + \sum_{i=1}^{q-1} \lambda^i d_N\left(v_0, v_{\eta(N)}\right) + \lambda^q d_N(v_0, v_{n_2})\right) \\
&\leq \lambda^{m_1}\left(r + \sum_{i=1}^{q-1} \lambda^i r + \lambda^q r\right) \\
&\leq \lambda^{m_1}\left(\sum_{i=0}^{q} \lambda^i r\right) \\
&= \lambda^{m_1}\frac{1-\lambda^{q+1}}{1-\lambda} r \\
&< \lambda^{m_1}\frac{1}{1-\lambda} r \\
&\leq \lambda^{M}\frac{1}{1-\lambda} r \\
&\leq \varepsilon
\end{aligned}$$
By Corollary B.0.6 on page 134, $\exists v \in \mathcal{C}[\mathbb{X}, \mathcal{A}]$ such that $v_n \to v$ as $n \to \infty$. $\square$

Chapter 5

Conclusion and Discussion

The highlights and successes of my Ph.D. research were covered adequately in the Chapter Summary (Section 1.2), so in this section, I will take the opportunity to examine some of the shortcomings of the work and look ahead to see how it might be improved.

The three research projects covered in the thesis extend the work in [TZ11] in different directions. In the first research project, I thought of a different way to construct a fixed point and tried to replicate the approach in [TZ11] using the new construction in place of Tucker and Zucker's. I consider this to be the most original work in the thesis (to the best of my knowledge), but also the least successful of the three projects. The underlying idea seems obvious (to me), so its apparent absence from the literature arouses my suspicion. My guess is that a few people have toyed with it in the past and dismissed it as impractical. In most nontrivial cases, the limit of the delayed fixed-point function probably becomes too unwieldy to be of any use. I do, however, think it's likely that there is a much better fixed point theorem for it (than Theorem TZJ1 for Vanishing Delays on page 48). I believe this would be the most useful next step for the project if anyone were to pursue it in the future: finding a set of conditions (ideally which do not include ***Contr***) on a stream operator, sufficient to ensure the operator has a unique fixed point.

The second project began as a reformulation of the mass-spring-damper case study (as seen in Section 3.3.1.2), motivated by the unusual condition on the parameters M, K, and D required by [TZ07, TZ11]. After the reforumulation, it seemed natural to ask what other sorts of operators would satisfy ***Contr***, and the choice of a Banach space offered an ideal venue to begin answering that question. While the answer I was able to provide was somewhat disappointing (*very* few dynamical systems can be expressed in the required form), I do believe it was at least somewhat illuminating and it attained a level of generality beyond what I had initially aspired to reach. It is obviously much too restrictive, however, essentially allowing for only one kind of module in an analog network (the kind shown in the General Form Theorem). Even the feedforward network from Example 1.3.11 on page 14 is excluded by the General Form Theorem. There should be a way to relax that theorem somewhat to allow a greater diversity of modules.

The third project was, in my opinion, the most successful, but the research was done in a relatively short period of time. As a result, it's a bit messy having two equivalent constructions. I have little doubt that there is a way to prove the Generalized Theorem TZJ2 on page 102 and the Concrete Computability Theorem on page 115 using Construction 4.4.3 on page 91, rather than Construction 4.4.12 on page 96. If so, the latter construction would be rendered entirely superfluous (as it should be). I simply didn't have time to attempt these proofs. That, as well as fully generalizing the computability theory from [TZ12], I feel is the (relatively) easy part. The hard part is finding a suitable case study like the mass-spring-damper system to which the smoothie theory can be applied, but to which the stream theory cannot. I spent a rather significant amount of time trying to find one. I looked at Nash Equilibrium and physical models involving partial differential equations—paying particular attention to the rather simple model of heat diffusion along a fixed-length rod whose ends are held constant at 0°C. The sequence of retractions I developed for that system were fairly elaborate (at least for what was meant to be a simple case study), but ultimately I failed to represent the physical model with a contracting operator[1].

[1] I have Prof. Jacques Carette to thank for rescuing me from the potentially endless pursuit down that blind alley. I may have still been trying (with red-rimmed eyes and grinding teeth) to make it work today if not for him).

Appendix A

Thoughts on Hadamard's Principle

A.1 Continuity isn't doing quite what we want

Continuity pervades every nook and cranny of both [TZ11] and this thesis. Much of this is due to the mathematical convenience afforded by continuity: continuous functions have very nice properties which make them easy to work with. If one is presented with both a continuous model of a phenomenon and a discontinuous one—the former is nearly always preferable. Furthermore, since the field of computable analysis typically defines computable functions as being continuous, it's much easier to compare analog computation with digital computation if the analog models are continuous as well.

In [TZ11], however, the authors offer a different reason for the importance they place on continuity: Hadamard's Principle. On page 3380 of [TZ11] they introduce this principle:

> *The significance of Theorems 1 and 2 is that continuity implies the stability of the fixed point solution Φ to the specification given by F with respect to the system parameters, initial values and input streams. This means that small changes in tuples of system parameters $\mathbf{c} \in \mathcal{A}^r$, initial values $\mathbf{a} \in A^s$ and input streams $\mathbf{x} \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^p$ will result in small changes in the behaviour of the systems as defined by $\Phi(c, a, x) \in \mathcal{C}[\mathbb{T}, \mathcal{A}]^m$. Here "small" is measured by any topology chosen for the task in hand. The significance of continuity is expressed in Hadamard's principle which, in the present context, can be (re-)formulated in the form: for a model of a physical system to be acceptable, the behaviour of the model must depend continuously on the data. This principle formalises the fact that if the system's behaviour depends significantly on small perturbations in its data, then it cannot behave in a stable fashion and its physical observation cannot be reliable. This is because, for example, repeating an experiment or computation will involve small variations of physical data, and for the system to be observable the corresponding variation in behaviour must also be small.*

On page 3402, they continue,

> *An important aspect of Hadamard's principle is that it can be viewed as making classical experimental physics possible. Suppose, for example, that one wants to verify any of the well-known relations of classical physics — Hooke's Law or Charles's Law, for example—by taking measurements and drawing a graph of the relationship between the "independent" and "dependent variables"—force vs displacement of a spring in the first example, and temperature vs volume of a gas (at constant pressure) in the second. ... The experimental results, and consequent graph, only make sense on the assumption that the function that one is attempting to plot is continuous, so that small discrepancies or inaccuracies in the inputs produce only small variations in the outputs. Moreover, this is needed to guarantee repeatability of experiments.*

I agree with the spirit and the motivation behind Hadamard's Principle, but not with its prescription of continuity. Certainly small variations in the input data must yield small changes in the behaviour of the system, but this is significantly different from insisting that *arbitrarily* small changes in the behaviour of the system always be attainable via sufficiently small variations in the input.

In the context of experimental science, "small" will depend on our measuring instruments and the object under study. A light-year is "small" when measuring the diameter of a galaxy, while even a nanometre is not when measuring an atom. Suppose our instruments are capable of taking measurements to within $\varepsilon > 0$ of the "true value" of the quantity (if, indeed, such a value even exists).

Now suppose we have a mathematical model $f : X \to Y$ of some physical system. That is, if we take a measurement $x \in X$ from the system, the value $f(x)$ can be calculated and yields a prediction about the system's behaviour which can be compared with a measurement. And remember: we have a margin of error of ε in both the input and the output measurements. Consider the following examples of models we might have.

Example A.1.1.

$$f : \mathbb{R} \to \mathbb{R}$$

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{x}{\varepsilon} & \text{if } 0 < x \leq \varepsilon \\ 1 & \text{if } x > \varepsilon \end{cases}$$

This system is continuous but it is experimentally indistinguishable from the (discontinuous) step function. The discontinuity in the step function would be unmeasurable if it were present in the physical system, and if it weren't present, that too would be an unmeasurable aspect of its behaviour. If we were to verify the accuracy of this model experimentally, it would be indistinguishable from the step function. Both functions would be either confirmed or falsified together by any conceivable experiment. They are effectively both members of the same experimental equivalence class.

Yet according to Hadamard's Principle, the step function would be an "unacceptable" model of a physical system (or as Courant and Hilbert would say, the problem which produced it was "ill-posed"). I can appreciate that we might have reasons to prefer one model over the other, depending on the situation, but to reject the step function reflexively as part of philosophical moratorium on all discontinuity for its own sake seems absurd to me.

Now consider a more extreme example:

Example A.1.2.

$$f : \mathbb{R} \to \mathbb{R}$$

$$f(x) = \begin{cases} \sin(x) + \frac{\varepsilon}{2} & \text{if } x \text{ is rational} \\ \sin(x) & \text{if } x \text{ is irrational} \end{cases}$$

This system is nowhere continuous, but that seems to have almost no adverse effects on its predictive capabilities. The only impact these discontinuities have on the viability of the model is that they slightly enlarge the margin of error. For example, a measurement of $5\varepsilon/4$ at $x = 0$ (which would be attainable with our hypothetical measuring instrument and read as being different from a measurement of 0) would be consistent with this model, but not consistent with the model $\sin(x)$.

Obviously we'd *prefer* to work with $\sin(x)$ over $f(x)$ because it's much simpler and far more well-behaved. All else being equal, there would certainly be no reason to favour the discontinuous model. It goes out of its way to be unwieldy and it does so for no apparent reason, offering nothing but slightly fuzzier predictions. That is hardly grounds for dismissing such a model as having no experimental value, however.

Example A.1.3.

$$f : \mathbb{R} \to [-1, 1]$$
$$f(x) = \sin\left(\frac{2\pi x}{\varepsilon}\right)$$

This system is everywhere continuous and even infinitely differentiable (on $(0, 1)$), but it assumes *every* possible value in its range within the input margin of error. Thus, it has absolutely no predictive capabilities whatsoever. It couldn't be more continuous, yet even unmeasurably small changes in the input result in arbitrarily large changes in the output.

Even worse, perhaps, consider the (in)famous example of the "Topologist's Sine Curve" (but restricted to $\mathbb{R}^+$):

$$f(x) = \sin\left(\frac{1}{x}\right)$$

No matter how precise your measuring instruments are, if you need a measurement near 0, you're out of luck.

Remark A.1.4. Examples A.1.1 on the preceding page and A.1.2 on the previous page show that continuity is not necessary to ensure that a model is experimentally viable, and Example A.1.3 shows that nor is continuity sufficient. Thus, it appears it has no role to play as a criterion of experimental applicability.

Richard Courant and David Hilbert write (in [CH53])

> *The [requirement of continuity] is necessary if the mathematical formulation is to describe observable natural phenomena. Data in nature cannot possibly be conceived as rigidly fixed; the mere process of measuring them involves small errors. For example, prescribed values for space or time coordinates are always given within certain margins of precision. Therefore, a mathematical problem cannot be considered as realistically corresponding to physical phenomena unless a variation of the given data in a sufficiently small range leads to an arbitrarily small change in the solution.*

This paragraph appears to me to contradict itself. If the mere process of measuring data necessarily involves small errors, then why must the solution to a mathematical problem corresponding to physical phenomena be required to exhibit *arbitrarily* small changes? It is impossible (and we can only assume it will always be impossible) to measure arbitrarily small changes, so this is much too extreme a limitation to impose on mathematical models of physical phenomena.

In a later section entitled "Remarks about 'Improperly Posed' Problems," Courant and Hilbert write,

> *Nonlinear phenomena, quantum theory, and the advent of powerful numerical methods have shown that "properly posed" problems are by far not the only ones which appropriately reflect real phenomena.*

With this, I agree, and I struggle to see how it is consistent with their earlier statement.

A.2 If not continuity, then what?

I believe the aim of Hadamard's Principle is to ensure that in any scientific model, an unmeasurable difference between two input values should not result in a measurable[1] difference between their images. We might codify this mathematically as follows[2]:

[1]Just to dispel any possible confusion, I use the term "measurable" here in the ordinary sense rather than the mathematical sense. There are no σ-algebras or measures involved.

[2]This definition has surely been proposed before, but by whom and what it has been named, I have no idea.

Definition A.2.1. Let $(X, d_X), (Y, d_Y)$ be metric spaces, let $\varepsilon_X, \varepsilon_Y > 0$, and let $f : X \to Y$. Then f is $(\varepsilon_X, \varepsilon_Y)$-*stable* if $\forall x, y \in X$

$$d_X(x, y) < \varepsilon_X \;\Rightarrow\; d_Y(f(x), f(y)) < \varepsilon_Y$$

The values ε_X and ε_Y might represent the precision of our measuring instruments. That is, ε_X is so small that if $d_X(x, y) < \varepsilon_X$, we lack the technology to determine it (and likewise for ε_Y). This is obviously much messier and far less satisfying than continuity since the precision of our measuring instruments is always improving, but I feel this pair of precisions is absolutely essential to the mandate we are attempting to draft. I don't believe continuity is fulfilling the role that Hadamard intended for it, so its elegance is moot.

One way we might liberate this admittedly awkward condition from the precision of actual measuring devices is to assume theoretical limits such as the Planck length on the quantities involved rather than technological ones. Another possibility (which I think is insufficient, but at least a step in the right direction) would be to replace continuity with bounded variation.

Remark A.2.2. Even using this alternative criterion, I feel the principle is too strict, as it suggests we completely dismiss any model or solution that does not conform to it—that we would be utterly wasting our time with any such models as they have no scientific value. On this point I defer to Karl Popper's philosophy of science and maintain that a scientific statement need only be falsifiable to have scientific value. As long as it makes *some* prediction about a system that can be proven incorrect in the face of the right observation, it should not be rejected as being experimentally worthless. It may obviously be replaced by a superior model that makes stronger predictions or has nicer properties, but that's rather different from rejecting a model altogether.

In the case of Example A.1.3 on page 130, that model makes *no* falsifiable predictions. There is no measurement of the system that would be inconsistent with the model. Therefore, I agree that it should be rejected from the realm of experimental science. The case of the Topologist's Sine Curve is rather different since it does make falsifiable predictions when we move sufficiently far away from zero. It may still be consistent with several different possible measurements, but as long as there is at least *one* measurement that the model rules out as being impossible, it is an experimentally viable model.

At the opposite extreme, when a model rules out all *but* one measurement as being impossible (i.e. it makes predictions with the same or greater precision than our measuring devices or our assumptions of theoretical limits), that is an *ideal* model in the sense of experimental viability. I believe there is a spectrum of models in between the two extremes, and models that make stronger (more easily falsifiable) predictions should typically be favoured over those which make weaker predictions, but the latter should not be dismissed altogether the way Hadamard's Principle suggests.

Appendix B

Supplementary Propositions

Lemma B.0.3. *If K is a compact metric space and $\mathcal{A}$ is a complete metric space, then $\mathcal{C}[K, \mathcal{A}]$ (with the compact-open topology) is complete.*

Proof. See [TZ11, Mun75]. $\square$

Definition B.0.4 (Uniformly Cauchy Sequence)**.** Let X be a set and Y be a metric space (with the metric d_Y). Let $\mathcal{F} = \{f_n : X \to Y\}_{n\in\mathbb{N}}$ be a sequence of functions. We say $\mathcal{F}$ is *uniformly Cauchy* if there exists a function $N : \mathbb{R}^+ \to \mathbb{N}$ such that $\forall\varepsilon > 0\ \forall m, n \in \mathbb{N}$,

$$m, n \geq N(\varepsilon) \ \Rightarrow\ \sup_{x\in X}\{d_Y\,(f_n(x), f_m(x))\} < \varepsilon$$

I *know* the following lemma must be in some textbook, somewhere, and I'd rather just refer to it, but I couldn't find a solid reference that stated it at this level of generality (without the domain or codomain being $\mathbb{R}^n$). At this point, I'm thinking I'll waste less time by just re-inventing the wheel here.

Lemma B.0.5. *Let X be a set and Y be a complete metric space (with the metric d_Y). Let $\{f_n : X \to Y\}_{n\in\mathbb{N}}$ be a sequence of uniformly Cauchy functions. Then there exists a unique function $f : X \to Y$ such that f_n converges uniformly to f.*

Proof. It's clear from Definition B.0.4 that for any $x \in X$, the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in Y, and since Y is complete, that sequence must converge. Hence, the sequence $\{f_n\}_{n\in\mathbb{N}}$ converges pointwise to a unique function $f : X \to Y$. What is perhaps not entirely obvious (albeit, thoroughly unsurprising) is that the convergence is uniform.

Let $N : \mathbb{R}^+ \to \mathbb{N}$ be the (uniformly Cauchy) modulus function from Definition B.0.4 for the sequence $\{f_n\}_{n\in\mathbb{N}}$ and define $N_f : \mathbb{R}^+ \to \mathbb{N}$ as $N_f(\varepsilon) = N(\varepsilon/2)$. We will show that N_f is modulus of convergence for $\{f_n\}_{n\in\mathbb{N}}$.

Since $f_n \to f$ (pointwise) as $n \to \infty$, there is another (pointwise convergence) modulus function for f_n. Call it $N'_f : X \times \mathbb{R}^+ \to \mathbb{N}$. Then $\forall x \in X \ \forall \varepsilon > 0 \ \forall n \in \mathbb{N}$,

$$n \geq N'_f(x, \varepsilon) \ \Rightarrow \ d_Y\left(f_n(x), f(x)\right) < \varepsilon$$

Let $\varepsilon > 0$ and let $n > N_f(\varepsilon)$. Let $x \in X$ and let $m \geq \max\left\{n, N'_f(x, \varepsilon/2)\right\}$. Then,

$$\begin{aligned} d_Y\left(f_n(x), f(x)\right) &\leq d_Y\left(f_n(x), f_m(x)\right) + d_Y\left(f_m(x), f(x)\right) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Thus, $f_n \to f$ uniformly as $n \to \infty$ with (uniform) modulus of continuity N_f. □

Corollary B.0.6. *Let X be a topological space, and Y be a complete metric space. Let $\{f_n : X \to Y\}_{n\in\mathbb{N}}$ be a uniformly Cauchy sequence of continuous functions. Then there exists a unique, continuous function $f : X \to Y$ such that $f_n \to f$ as $n \to \infty$.*

Proof. By Lemma B.0.5, there exists a unique $f : X \to Y$ such that $f_n \to f$ uniformly as $n \to \infty$. Since all the f_n functions are continuous, the Uniform Limit Theorem (see [Mun75], for example) states that f will be continuous as well. □

Bibliography

[Ach97] David Acheson, *From calculus to chaos: An introduction to dynamics*, Oxford University Press, 1997.

[Bar01] Robert Gardner Bartle, *A modern theory of integration (graduate studies in mathematics)*, vol. 22, American Mathematical Society, 2001.

[BCGH06] O. Bournez, M. L. Campagnolo, D. S. Graça, and E. Hainry, *The general purpose analog computer and computable analysis are two equivalent paradigms of analog computation*, Theory and Applications of Models of Computation: Third International Conference, TAMC 2006, Beijing, China, May 15-20, 2006. Proceedings (S.B. & Li A. Cai, J.-Y.; Cooper, ed.), vol. 3959, Springer-Verlag, 2006, pp. 631–643.

[BCS97] Peter Buergisser, Michael Clausen, and Amin Shokrollahi, *Algebraic Complexity Theory*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, 1997.

[BD01] William E. Boyce and Richard C. DiPrima, *Elementary differential equations and boundary value problems*, 7 ed., John Wiley & Sons, Inc., 2001.

[Cau21] Augustin-Louis Cauchy, *Cours d'analyse de l'Ecole Royale Polytechnique*, Académie des Sciences, 1821.

[CH53] R. Courant and D. Hilbert, *Methods of mathematical physics, vol. ii*, Interscience, 1953.

[Cor36] Gaspard-Gustave Coriolis, *Note sur un moyen de tracer des courbes données par des équations différentielles*, Journal de Mathématiques Pures et appliquées **1** (1836), 5–9.

[F$^+$06] T. Freeth et al., *Decoding the ancient Greek astronomical calculator known as the Antikythera Mechanism*, Nature **444** (2006), 587–591.

[Fea99] Jean-Christophe Feauveau, *A generalized Riemann integral for Banach-valued functions*, Real Analysis Exchange **25** (1999), no. 2, 919–930.

[GC03] D.S. Graça and J.F. Costa, *Analog computers and recursive functions over the reals*, Journal of Complexity **19** (2003), 644–664.

[Grz55] A. Grzegorczyk, *Computable functionals*, Fundamenta Mathematicae **42** (1955), 168–202.

[Grz57] ______, *On the defintions of computable real continuous functions*, Fundamenta Mathematicae **44** (1957), 61–71.

[Had52] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*, Dover, 1952.

[HK71] Kenneth Hoffman and Ray Kunze, *Linear algebra*, 2 ed., Prentice Hall, 1971.

[HK03] Boris Hasselblatt and A. B. Katok, *A first course in dynamics: With a panorama of recent developments*, Cambridge University Press, 2003.

[JZ12] Nick D. James and Jeff Zucker, *A class of contracting stream operators*, The Computer Journal (2012), Oxford University Press, DOI: 10.1093/comjnl/bxs054.

[Kem10] Achim Kempf, *Spacetime could be simultaneously continuous and discrete, in the same way that information can be*, New Journal of Physics **12** (2010).

[Ko91] K.-I. Ko, *Complexity theory of real functions*, Birkhäuser, 1991.

[Lac55] D. Lacombe, *Extension de la notion de fonction recursive aux fonctions d'une ou plusieurs variables reelles, i, ii, iii*, C.R. Acad. Sci. Paris **240, 241** (1955), 2478–2480, 13–14, 151–153.

[LCR10] Mathieu Lihoreau, Lars Chittka, and Nigel E. Raine, *Travel optimization by foraging bumblebees through readjustments of traplines after discovery of new feeding locations*, The American Naturalist **176** (2010), no. 6, 744–757.

[MPH+06] Jonathan W. Mills, Matt Parker, Bryce Himebaugh, Craig Shue, Brian Kopecky, and Chris Weilemann, *Empty space computes: The evolution of an unconventional supercomputer*, ACM International Conference on Computing Frontiers, 2006.

[Mun75] James R. Munkres, *Topology: a first course*, Prentice Hall, 1975.

[PE74] M.B. Pour-El, *Abstract computability and its relation to the general-purpose analog computer*, Journal of the American Mathematical Society **199** (1974), 1–28.

[PER89] M.B. Pour-El and J.I. Richards, *Computability in analysis and physics*, Springer-Verlag, 1989.

[Pol80] Anthony Pollen, *The great gunnery scandal: The mystery of jutland*, Wm Collins & Sons & Co, 1980.

[Rab03] A. Rabinovich, *Automata over continuous time*, Theoretical Computer Science **300** (2003), 331–363.

[Roy63] H.L. Royden, *Real analysis*, Macmillan, 1963.

[Rub93] Lee A. Rubel, *The extended analog computer*, Advances in Applied Mathematics **14** (1993), 39–50.

[Rud76] W. Rudin, *Principles of mathematical analysis ,(3rd ed.)*, McGraw-Hill, 1976.

[Rud91] Walter Rudin, *Functional analysis*, 2 ed., McGraw-Hill, 1991.

[Sha41] C Shannon, *Mathematical theory of the differential analyser*, Journal of Mathematics and Physics **20** (1941), 337–354.

[SHT99] V. Stoltenberg-Hansen and J. V. Tucker, *Concrete models of computation for topological algebras*, Theoretical Computer Science **219** (1999), 347–378.

[Son90] E.D. Sontag, *Mathematical control theory: Deterministic finite dimensional systems*, Springer-Verlag, 1990.

[Tar74] E. Tarafdar, *An approach to fixed point theorems on uniform spaces*, Transactions of the American Mathematical Society **191** (1974), 209–225.

[Tho76] James Thomson, *An integrating machine having a new kinematic principle*, Proceedings of the Royal Society, vol. 24, 1876, pp. 262–265.

[Tra99] Boris Trakhtenbrot, *Automata and their interaction: Definitional suggestions*, Fundamentals of Computation Theory (Gabriel Ciobanu and Gheorghe Paun, eds.), Lecture Notes in Computer Science, vol. 1684, Springer Berlin / Heidelberg, 1999, pp. 82–82.

[TZ04] J. V. Tucker and J. I. Zucker, *Abstract versus concrete computation on metric partial algebras*, ACM Transactions on Computational Logic **5** (2004), 611–668.

[TZ07] ______, *Computability of analog networks*, Theoretical Computer Science **371** (2007), 115–146.

[TZ11] ______, *Continuity of operators on continuous and discrete time streams*, Theoretical Computer Science **412** (2011), no. 28, 3378–3403.

[TZ12] ______, *Computability of operators on continuous and discrete time streams*, submitted to Computability (2012).

[Wei00] K. Weihrauch, *Computable analysis: An introduction*, Springer-Verlag, 2000.

[YMTK95] H. Yokoi, T. Mizuno, M. Takita, and Y. Kakazu, *Euclidean TSP using characteristics of slime mold*, IEEE Conference on Evolutionary Computation, vol. 2, 1995, pp. 689–694.