Project Thesis/Report

1 Introduction

2 Discrete-time Dynamical Systems

3 The Learning Problem

3.1 Taken Embedding Theorem

4 Driven Dynamical Systems to Forecast Problems

In this chapter, we discuss results pertaining to the mapping of temporal data obtained from a discrete-dynamical system onto a different space through the notion of a driven dynamical system. We also consider the conditions a driven system should have to avoid adding distortion to its state space representation, as well as then ensuring that the single-delay dynamics (SDD) are conjugate (or at least semi-conjugate to the system). The SDD can then be used to forecast and reconstruct the underlying attractor. **Is this attractor assumed to exist?**

4.1 Nonautonomous and Driven Dynamical Systems

A nonautonomous dynamical system is simply a dynamical system (as defined before in **link to above**) where the input u from the input space U, a topological space, is time-dependent. **This means it can no longer be generated by a flow?**

We immediately recall the figure first defined in **??** and remind ourselves of our ultimate goal to forecast the evolution of an unknown nonautonomous dynamical system (U, T) by constructing a conjugate system such as (V, S) below.

U U

V V.

To do this, the need arises to consider a driven dynamical system where input is taken from both an input space U, and an underlying state space X. This is done to account, in a manner more similar to the true scenario, for the influence of both the input and the actual state of the system at a specific timestep.

Definition 1 (**Driven and Compactly Driven Dynamical System).** *A driven dynamical system comprises two topological metric spaces U, X and a continuous function $g : U \times X \to X$ where $g(u_n, x_n) = x_{n+1}$. The dynamics on X are generated by the update equation $x_{n+1} = g(u_n, x_n)$ where $n \in \mathbb{Z}$, input u_n from U and state x_n belonging to X, where X is compact. If the input space U is compact, we refer to the system as compactly driven. Abbreviated, we shall refer only to the* driven system *g, with all other entities being understood implicitly.*

In particular, a nonautonomous dynamical system may be generated from U; any input $\overline{u}$, a bi-infinite sequence from U, gives rise to the sequence of self-maps $\{g(u_n, \cdot)\}_{n \in \mathbb{Z}}$ contained in X. Physically, one may think of this bi-infinite sequence as referring to a system that has been running for an incredibly long period at the time of the first measurement taken from the system (alternatively, the first time a probe is inserted into the system to take an observation).

Definition 2 (**Entire Solution**). *A sequence* $\{x_n\}_{n\in\mathbb{Z}} \subseteq U$ *is called an entire solution(or simply a solution) to the driven system* g *with input* $\overline{u}$ *when it satisfies*

$$g(u_{n-1}, x_{n-1}) = x_n$$

for all $n \in \mathbb{Z}$

It is important to emphasise that a sequence satisfying the update equation above can only be a solution if $x_n \in U$ for all $n \in \mathbb{Z}$. Consider the example below.

Example 1. *The only solution* $\{x_n\}_{n\in\mathbb{Z}}$ *to the driven system* $g(u,x) = \frac{ux}{2}$*, where* $X = [0,1]$*,* $U = [0,1]$*, is the zero solution* $x_n \equiv 0$*. To see this, consider any* $x_n = a \in [0,1]$ *where* $a \neq 0$*. Let* $\overline{u} \in U$ *be an non-zero constant sequence , say* $u_n = 0.5$*. The driven system may be rewritten as* $x_{n-1} = \frac{2x_n}{u_{n-1}} = 4x_n$ *and the iterates of* x_n *in backward time will increase by a factor of 4 at each timestep. Thus for some* $m \leq n$*,* $1 < x_m$ *i.e.* $x_m \notin X$*. So* $\{x_n\}_{n\in\mathbb{Z}}$ *is not a solution and it follows that the only possible solution is the zero solution.*

A system may also have multiple solutions as is evidenced in the example below.

Example 2. *Consider the driven system* $g(u_n, x_n) = x^n$ *for* $X = [0,1]$*,* $U = \mathbb{R}$*. The system has an uncountable number of solutions, as there exists a solution for every* $x \in X$ *which also passes through the point* x *and* $lim_{n\to\infty} x_n = 0$*,* $lim_{n\to-\infty} x_n = 1$*. The proof is deferred to immediately after the next paragraph.*

As the solutions to a driven system are often considered, we next identify a subspace X_U of X that contains all possible solutions. To realize such a subspace of a driven system g, the concept of a reachable set is defined.

Definition 3 (**Reachable Set**). *The reachable set of a driven system* g *is exactly the union of all the elements of all the solutions, i.e.,*

$$X_U := \Big\{ x \in X : x = x_k \text{ where } \{x_n\} \text{ is a solution for some } \bar{u} \Big\}.$$

The set of all reachable states at a specific time n *for input* $\overline{u}$ *is denoted by* $X_n(\overline{u})$

We note that $x \in X_n(\overline{u})$ if and only g has a solution $\{x_k\}$ for $x_n = x$ and input $\overline{u}$. **Cite**. This will lead to a result established later, but which is worth taking note of now: *g being a topological contraction is equivalent to the existence of a unique entire solution.* To this end we define the notion of a topological contraction.

Definition 4 (**Topological Contraction**). *A function* $g : U \times X \to X$ *is a topological contraction if for all* $n \in \mathbb{Z}$ *and all* $\overline{u} \subseteq U$*,* $X_n(\overline{u})$ *is a singleton subset of X.*

Proof of Example2. To see this, we show that for every input $\overline{u} = \{u_n\}_{n\in\mathbb{Z}}$, $g(u_n, \cdot)$ is a contraction map on X for every $k \in \mathbb{Z}$. Indeed $g(u_n, x_n) = x^n$ depends only on the state at time n and so for each input $\overline{u}$, $X_n(\overline{u}) = \{x^n\}$ is a singleton subset. We then easily conclude that $\{x_n\} \equiv 0$ is the only solution. □

Thus far it has been demonstrated that a system may have one or more solutions; one may ask if a driven system always has a solution and, if so, whether it satisfies certain properties such as uniqueness. Should the driven system be compact, existence follows immediately as shown in the following result.

Theorem 1. *If X is compact then for each input $\overline{u}$, there exists at least one solution to the driven system $g(.\,,x)$*

Proof. Consider an input $\overline{u} = \{u_n\}_{n\in\mathbb{Z}} \subseteq U$ and driven system $g : U \times X \to X$ generating a sequence $\{x_n\}_{n\in\mathbb{Z}}$ in the compact space X. Since X is a metric space, it follows immediately from a well-known result in Analysis (**Cite**) that $\{x_n\}$ has a convergent subsequence which then is a solution. □

We may easily construct many systems with trivial solution-sets, such as $g(u, x) = x$ which has only the constant solution x and so for $U = [-1, 1]$, the system would have no solution if $|x| > 1$. To refine the scenario, we consider only systems with unique solutions.

4.2 Unique Solution Property

Definition 5 (**Unique Solution Property**)**.** *A driven system g is said to have the Unique Solution Property (USP) if for each input $\overline{u}$ there exists exactly one solution. Alternatively we may formulate the USP as follows: g has the Unique Solution Property if there exists a well-defined map $\Psi : U \to X$ with $\Psi(\overline{u})$ denoting the unique solution.*

(More discussion on USP?)

Having already defined the reachable set X_U, we pause for a moment to fix additional notation. Letting $\overleftarrow{u}^n := (\ldots, u_{n-2}, u_{n-1})$ be the left-infinite subsequence of an input, $\overleftarrow{U}$ then denotes all these left-infinite sequences in U. Moreover, $\overleftarrow{u}^n v := (\ldots, u_{n-2}, u_{n-1}, v)$ is to symbolise the input up to time n with $v \in U$ being the specific input value at time n. The introduction of a new input at time n can be described by a mapping $\sigma_v : \overleftarrow{u}^n \mapsto \overleftarrow{u}^n v$.

The question now becomes whether we may establish a conjugacy as presented below for g driven

$$\begin{array}{cc} \overleftarrow{U} & \overleftarrow{U} \\ \\ X_U & X_U. \end{array} \tag{1}$$

Restricting our attention more and more, we now consider a specific subclass of conjugacies.

Definition 6 **(Universal Semi-Conjugacy).** *Given a driven system g, we call a continuous and surjective map $h : \overleftarrow{U} \to X_U$ a universal semi-conjugacy if diagram 1 commutes for all $v \in U$.*

If the universal semi-conjugacy h exists (i.e. the diagram in 1 commutes) then the solution Ψ will be at least a coarse-grained representation of the input u.

Does such a function h for the driven system defined above exist? If g has the USP and $\Psi(u) = \{x_n\}_{n\in\mathbb{Z}}$ then h, defined by $h(\overleftarrow{u}_n) := g(u_n, x_{n-1}) = x_n$, will satisfy the semi-conjugacy in the graph above 1. Regrettably, the mapping h is not guaranteed to exist in general.

Re-sketching the graph1 above by fixing the input v in g and replacing X_U by its left-infinite sequence space $\overleftarrow{U}$, we obtain the graph below. In this case, the function $H : \overleftarrow{U} \to \overleftarrow{X}_U$, a map that is both continuous and surjective, is called a *causal mapping.*

Definition 7 **(Causal Mapping).** *A continuous, surjective map $H : \overleftarrow{U} \to \overleftarrow{X}_U$ such that*

$$H \circ \tilde{g}_v = \sigma \circ H$$

where $\tilde{g}_v$ maps $(\dots, u_{-2}, u_{-1})$ to $(\dots, u_{-2}, u_{-1}, g(v, u_{-1}))$.

$$\begin{array}{cc} \overleftarrow{U} & \overleftarrow{U} \\ \\ \overleftarrow{X}_U & \overleftarrow{X}_U. \end{array} \tag{2}$$

Theorem 2. *For a compactly driven system, a causal mapping H exists if and only if g has the USP.*

Proof. *See proof from p.15 of [**?**]* **Expand.** □

Note that even when h does exist, we are not guaranteed its injectivity. Considering again example 1, we see that even if h were to exist, it could not be injective as $X_U = \{0\}$.

An embedding of the space $\overleftarrow{U}$ would allow one to establish topological conjugacy, which in turn provides stronger results than merely obtaining a coarse-grained representation via a causal mapping.

Before formalising this, we make mention of the concept of the inverse-limit system of a dynamical system (U, T). In broad sweeps, the inverse-limit system of a dynamical system is a self-map on a subset of an infinite-dimensional space **(Expand)** where each point in the inverse-limit space corresponds to a backward orbit of the map T. In literature, the inverse-limit space is denoted by $\hat{U}_T \subseteq \overleftarrow{U}$ [**?**]. The map T, then, induces a self-map $\hat{T} : \widehat{T} : (\ldots, u_{-2}, u_{-1}) \mapsto (\ldots, u_{-2}, u_{-1}, T(u_{-1}))$ on $\widehat{U}_T$

Definition 8 (**Causal Embedding**). *A driven system g is said to causally embed the dynamical system (U, T) if*

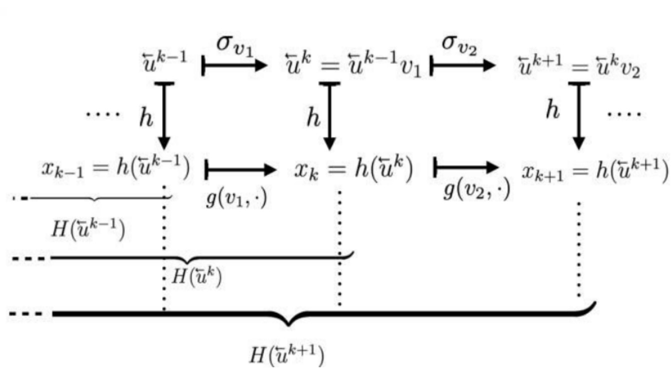
i. The diagram in 2 commutes (i.e. g is a universal semi-conjugacy),
ii. $H_2(\overline{u}) := (h(r\overline{u}, h(\overline{u})))$ embeds the inverse-limit space $(\hat{U}, \hat{T})$ of (U, T) in $X \times X$.

The driven system g can induce an embedding of $\overleftarrow{U}$ in $\overleftarrow{X}$ as follows: If $H : \overleftarrow{U} \to \overleftarrow{X}_U$ is also injective (in addition to being surjective, and hence bijective), it becomes the embedding of $\overleftarrow{U}$ in $\overleftarrow{X}$ induced by the driven system g and we refer to H as a *causal embedding.*

Theorem 3. *If $g(.,x)$ is invertible and has the USP then H is a causal embedding.*

Proof. *See proof from [**?**]* **Expand.** □

When g has the USP, the diagram below illustrates the operation of the the mappings h and H. The mapping $h : \overleftarrow{U} \to X_U$ is an observable as discussed in an above section (**??**).



The next result is of relative import but to fully comprehend its influence, some discussion still remains. Before stating it, we define a concept which will directly be shown to be equivalent to the USP.

Definition 9 (**Uniform Attraction Property**). *A driven system g has the Uniform Attraction Property (UAP) if, regardless of starting position, all trajectories converge to a single trajectory as time flows forward. This trajectory is also the unique solution sequence x to the input sequence u as mentioned above* **??**.

Theorem 4. *The following statements are equivalent:*

i. g has the USP
ii. g is a topological contraction
iii. g has the Unique Attraction Property (UAP).

Proof. (we'll show this) [11] eqn.6 □

We have not as of yet treated the UAP properly, and establish the necessary result here (without proof) so as not to break up the completeness of theorem.

The proof of the result for the UAP is left to literature. Why? Cite.

4.3 Choosing the driven system g

One must be careful to avoid a choice of g which would add complexity to the obtained solution.

When a causal embedding H exists for the driven system g, one can map an arbitrary input u onto the solution space X without additional distortion or information-loss. **(Cite.)** When an embedding is established, the question of possible additional complexity in the solution is removed by guaranteeing that, since the systems are conjugate (semi-conjugate, (refer)), g does not add any (some) complexity to the system. It is, however, a balancing act as it also undesirable to choose a function g that quenches the temporal structure in u by contracting to such a degree that the ability to recover information from the original system is lost completely. To obtain a suitably complex function g, it is desired that the the reachable set of a driven system be large enough to relate to the input. **(Expand here. It stops a bit abruptly.)**

In the example above (1), the input's temporal variation cannot be related to the reachable set as X_U consists of a single element and so little, if not no, information is encoded. Rather, the reachable set of a driven system must therefore be such that the inverse-limit space of U_T can be embedded in some finite self-product of the reachable set of T **(refer or cite)**. To this end, consider the notion of State-Input (SI) Invertibility.

Definition 10 (**SI-Invertibility**)**.** *A map g is said to be SI-Invertible if g is invertible for all $x \in X$. Alternatively it may be said that if, given x_n and x_{n-1}, u_n can be uniquely determined from $x_n = g(u_n, x_{n-1})$, then g is said to be SI-invertible.*

Explain more here.

We make special note of a specific driven system. The function

$$g(u, x) = (1 - a)x + a{\cdot}tanh(Au + \alpha aBx) \tag{3}$$

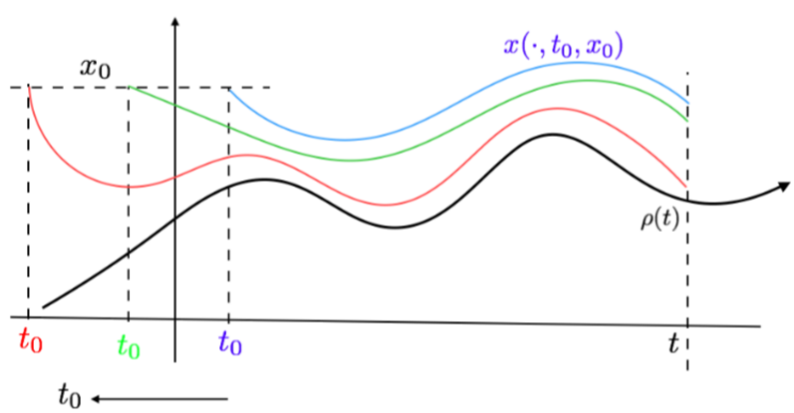
is both SI-invertible and possesses the USP. The proof of these two facts is rather involved and hence the proof, instead of being replicated here, may be found in (cite relevant article) This specific driving function g is used in our implementation and is more completely discussed in chapter 5

Expand: UAP + Embedding → perturbation isn't an issue + we can drop left-infinite sequence.

Despite the ease that one may work with a left-infinite history in the realm of theory, it is impossible to obtain such a sequence in any real-life application. One does not in practice, fortunately, need the entire left-infinite history of the input thanks to the Uniform Attraction Property (UAP). The next definition is stated in an informal manner as the more rigourous definition makes use of processes, a concept which would take some amount of paper to establish and prove - an exercise which would detract from the principal thrust of this paper/project/thesis.

Is it VERY obvious to the reader at this point what the general 'thrust' of the paper IS? Cite articles

Recall that the Unique Attraction Property guarantees that all trajectories *(which ones?)* will converge to the same trajectory in forward time. Incredibly, this permits one to initialise a driven system g with an altogether arbitrary initial value $y_m \in X$ where $m \in \mathbb{Z}$ and the UAP guarantees that the sequence $\{y_{m+1}, y_{m+2}, \ldots\}$ satisfying the relation $y_n = g(u_n, y_{n-1})$ where $k \geq m$ will uniformly approach the elements $\{x_n\}$ of the actual solution. (see [41, Theorem 1] or [11, Eqn. (18)])



One may now appreciate an even more astounding result. g having the USP is equivalent to the UAP and so one need not establish any additional results to ensure that the sequence uniformly approach the unique solution Ψ. This vastly simplifies the work needed to be done in setting up a problem to guarantee that the underlying system will be accurately 'represented' by the conjugate system. **Check this statement again to make sure it's not nonsensical.**

This also solves the problem of perturbations or noise introduced by the observable or measurement function. **(rephrase, expand and explain)**

4.4 The next step in Dynamics

Next we define the relation Y_T induced by (U, T) on $X_U \times X_U$ for a driven system g possessing SI-invertibility. To describe the single-delay lag dynamics formally, we consider a dynamical system $T : U \to U$ and we define a relation on the reachable set X_U, i.e., a subset defined on $X_U \times X_U$ by

$$Y_T := \{(x_{n-1}, x_n) : \{x_k\}_{k\in\mathbb{Z}} \text{ is a solution for some orbit of } T \text{ and } n \in \mathbb{Z}\}.$$

The following theorem establishes the existence of a well-defined map GT describing the single-delay dynamics of the system above.

Theorem 5. *If we let $G_T : Y_T \to Y_T$ be a map defined by the relation $(x_{n-1}, x_n) \mapsto (x_n, x_{n-1})$, then G_T is well-defined (and this results holds even in the absence of g possessing over the USP)*

We're now getting quite close to where we want to be and our results carry more and more weight. Recall that g is only being given inputs from the orbits of T.

Definition 11. *The inverse-limit space $\widehat{U}_T$, a subspace of $\overleftarrow{U}$, is defined by*

$$\widehat{U}_T := \{(\dots, u_{-2}, u_{-1}) : T(u_n) = u_{n+1}\}$$

This space is well-defined since $T : U \to U$ is surjective by assumption. ***Expand why***

Theorem 6. *Graph 1 is exactly the inverse-limit system $(\hat{U}, \hat{T})$.*

Note that H_2 maps an entire left-infinite solution sequence from Ψ to an element in $X \times X$.

We now have the following (compare with 2 above):

graph here causes issues currently

Show that (Y_T, G_T) **is semi-conjugate to** $(\widehat{U}, \widehat{T})$**.**

Summarising the discussing thus far: It is easy to lose onself in the symbols, so we take a moment to review our progress up until this point.

1. We are interested in a dynamical system (U, T) with unknown dynamics for whichever reason.

2. To determine information about this system (U, T), we shift our focus by rather aiming to determining the dynamics of the inverse-limit system $(\widehat{U}, \widehat{T})$ which, given certain assumptions, we can guarantee to be at least semi-conjugate to (U, T).
3. If the driven system g (and $\{u_n\}$ in U is an orbit of T), we have that G_T exists.
4. If, furthermore, g may be shown to have the USP, we showed that (Y_T, G_T) is semi-conjugate to $(\widehat{U}, \widehat{T})$.
5. If T is also (and at this stage this is very much our hope) a homeomorphism, then (Y_T, G_T) topologically conjugate to $(\widehat{U}, \widehat{T})$ in which (U, T) has been embedded!

One can therefore learn the single-delay lag dynamics of the driven system states via G_T with enough data thanks to the USP. This enables us to do at least 2 things: 1) Forecast future values of x_n via iterates of G_T (as G_T can be determined. 2) Forecast future values of u_n.

4.5 A discussion of G_T

By establishing the existence of the map G_T, we've essentially embedded the attractor U **So far we haven't really spoken about an attractor** into the higher dimensional space $X \times X$. In layman's terms, this ensures that there is more "dimensional room" for the underlying system's underlying dynamics to "move". As the dynamics aren't as "squashed", we might therefore hope that the dynamics of G_T are in some sense simpler than that of T. (Taken note of the fact that G_T is a homeomorphism even when T is just continuous)

In *[Koopman]* it is illustrated in an empirical fashion in that the map G_T describes dynamics which are less functionally complex than that of T or of the map $\Phi_{2d,\theta}$ (as discussed above in the section on Takens). This is done by implementing a Recurrent Neural Network (RNN), but not discussed as of yet. *(Add in that we'll be using Pearson coefficient?)*

We opt to learn G_T in an indirect manner by defining a new map Γ. It follows immediately from G_T's existence that Γ also exists.

The reasons for taking this roundabout approach remain to be discussed, but a pat answer may be immediately given: When we have learnt Γ, we can drive the system autonomously and then we know G_T anyway.

Theorem 7. *When x_n and x_{n-1} are successive points on a solution obtained for input-orbit of T $\{u_n\}$, then the map $\Gamma : X \times X \to U$ defined by $(x_{n-1}, x_n) \mapsto u_n$ exists whenever G_T exists [Koopman, Theorem 3c]*

Definition 12. *The projection mapping π_i is defined as follows:*

$$\pi_1 : (a_1, a_2) \to a_1$$

and

$$\pi_2 : (a_1, a_2) \to a_2$$

We now have the graph below:

$$\begin{tikzcd}
\hat{U}_T \arrow[r, "\hat{T}"] \arrow[d, "H_2"'] & \hat{U}_T \arrow[d, "H_2"] \\
Y_T \arrow[r, "G_T"] \arrow[dr, "(\Gamma,\pi_2)"'] & Y_T \\
 & U \times X. \arrow[u, "(\pi_2,g)"']
\end{tikzcd}$$

And a final set of equations - equations that have been constructed from data!

$$u_{k+1} = \pi_1 \circ (\Gamma, \pi_2) \circ (\pi_2, g)(u_k, x_k) \tag{4}$$

$$x_{k+1} = \pi_2 \circ (\Gamma, \pi_2) \circ (\pi_2, g)(u_k, x_k). \tag{5}$$

The problem thus simplifies to the issue of learning the map Γ and combining this with the projection mapping π_2 and the function g, which will be known.

4.6 Advantages of learning Γ

Why are we going such a lengthy roundabout route, one may immediately ask. Why not just learn the map G_T from the get-go? On the surface, this seems to be an arbitrarily chosen path with no real reason for the choice made, so we take a pause again and discuss our reasoning for learning Gamma. There are a number of distinct advantages.

Learning Γ saves computational resources. This follows from the fact that the input u_n at time n lies in a low-dimensional subspace of the high-dimensional subspace $X \times X$ due to the fact that it is zero-padded.

Input- and parameter-related stability is achieved if g has the USP
Define the two types of stabilities
Sentence on u, v, tails, USP and product topology
If we opt to work with $Y - T$, errors can occur. *Expand*
Γ is more stable. *Explain why*

Global dissipativity prevents large numerical errors due to input data. Γ is globally dissipative (*Define*) whereas G_T is not guaranteed to be. IF we were to rather learn G_T, we would risk having

exactly the same problems arise that could arise also for Takens. ***ADD THIS TO ERRORS IN TAKENS: NOT GUARANTEED GLOBALLY DISSIPATIVE***
Expand here ...

Where to add:If g has USP, then any solution to g is a non-autonomous uniform attractor

5 Implementation