Relaxing Conjugacy To Fit Modelling in Dynamical Systems

Joseph D. Skufca\textsuperscript{1,*} and Erik M. Bollt\textsuperscript{1,†}

\textsuperscript{1}Department of Mathematics, Clarkson University, Potsdam New York

(Dated: January 24, 2007)

We address a fundamental modeling issue in science as related to the field of dynamical systems: when is a model of a physical system a good representation? Conjugacy provides a means to determine if two systems are dynamically equivalent. We develop mathematical technology to decide when dynamics of a toy model are like (although not identical to) dynamics of the physical system, since the concept of conjugacy is too rigid for such cases. We contrast the usual methodology where model quality is measured in a Banach space to our dynamically motivated notion of matching orbits as best possible. We highlight our methods with a lower-ordered model of a “noisy” logistic map, and also a simplified model of a Lorenz system such that the usual one-dimensional map model is not exactly justified in the traditional sense.

PACS numbers:

A standard approach to quantifying model accuracy is to measure how “close” the model is to the original system. Often overlooked, however, is that defining “close” depends upon what aspects of the system we are trying to model. In many cases, prediction is our modeling goal, such as when forecasting the weather. The quality of prediction is grounded in numerical analysis on Banach spaces: the “goodness” of short term predictions is based on measurement of residual error. However, in dynamical systems, a model’s quality is typically not based on such error analysis. Continuing, for sake of example, in the field of meteorology, we cite a famous historical example to highlight this long recognized issue. Consider E. Lorenz’s 1965 paper about his 28-variable ODE model of the weather \cite{Lorenz65} which consists of a Galerkin’s projection of a two-level geostrophic model of atmospheric fluid flow. About the matter of choice of parameter values which are initially free in the model, the tuning of which lead to dramatically different dynamic behavior due to a plethora of possible bifurcations, Lorenz says: “Our first choice of constants lead to periodic variations. Subsequent choices yielded irregular....” Expert knowledge as a meteorologist concerning what “reminded” him of realistic weather oscillations guided the model choice. He used a similar modeling approach even by the time of his 1998 paper on a 40-variable model \cite{Lorenz98}. As modeling moves across the sciences, the applied dynamical systems community still tends to choose the “best model” from a model class in an intuitive manner.

In general, when one proposes a mathematical model for some system, the model provides a simplified representation. Modeling can be viewed as the art and science of choosing a “good” representation, where the evaluation of the model’s “goodness” is based on how well it satisfies the purpose for which we are modeling. For example, if we want a simplified representation of the relationship between two variables in a system such that given one, we could predict the other, techniques of approximation theory can find a mathematical model that matches data collected from the system. Choosing an appropriate description of “error” in the problem (least square, square integral, maximum, etc) we can find a well defined “best” model for our system within a model class (linear functions, for example). The field of dynamical systems provides an alternative purpose to the modeling process, focusing not on predicting behavior of the system, but rather, on qualitative characterization of the system. The 3-variable Lorenz system illuminates the richness of behavior that might be created by a convection flow, but is not meant for actually computing that flow. Since the beginnings of the field of dynamical systems by Henri Poincaré \cite{Poincare91}, characterizing a dynamical system asked us to examine topological and geometric features of orbits, rather than focusing on the empirical details of the solution of the dynamical system with respect to a specific coordinate system. One seeks to understand coordinate independent properties, such as the periodic orbit structure — the count and stability of periodic orbits. The question of whether two systems are dynamically the same has evolved into the modern notion of deciding if there is a conjugacy between them \cite{Grebogi81,Young95,Bollt01,Bollt04}.

Given these notions, we often speak of a “toy model” — a dynamical system which is much “like” the “real” system. These subjective evaluations are assertions that the model is satisfactory, but fail to distinguish excellent models from fair ones. We assert that quantifying the quality of a model is an essential problem in science. A primary interest of
this work will be to develop principles and methods to compare dynamical systems when they are not necessarily equivalent (in the sense of conjugacy), but in a manner which respects conjugacy.

**Background.** Given two dynamical systems, \( g_1 : X \rightarrow X \), and \( g_2 : Y \rightarrow Y \), the fundamental departure from a typical measurement of approximation between the two dynamical systems is that we do not directly compare \( g_1 \) and \( g_2 \) under an embedding in a Banach space, (e.g. measuring say \( \| g_1 - g_2 \|_{L^2} \)) because such measurements pay no regard to the central equivalence relationship in dynamical systems — conjugacy. Two systems are conjugate if there is a homeomorphism \( h : X \rightarrow Y \), between the underlying phase space (\( h \) must be 1-1, onto, continuous and \( h^{-1} \) must be continuous), and \( h \) must commute the mappings at each point \( x \in X \), such that \( h \circ g_1 = g_2 \circ h \). Practically, \( h \) represents an “exact” change of coordinates so that the mappings behave exactly the same in either coordinate system. We give the name “commuter” to any function \( f : X \rightarrow Y \) satisfying the commuting relationship,

\[
 f \circ g_1 = g_2 \circ f, \tag{1}
\]

and note that a commuter will be a conjugacy only if it is a homeomorphism. The commuter provides a matching between trajectories for \( g_1 \) and \( g_2 \): over and/or under-representations are reflected as 1-1 and onto problems in \( f \), while trajectories that permit matching only for finite time are related to discontinuities in \( f \). We develop measures of commuters \( f \) that quantify “how much” the \( f \) may fail to be a homeomorphism, which we call homeomorphic defect. Our fundamental contribution is that we note that measurement of defect allows us to quantify the dissimilarity of \( g_1 \) and \( g_2 \) in a manner that is consistent with the dynamical systems approach. We remark on the broad applicability of our approach: if a mathematical model is a simplified description of some system, then the model obviously is not equivalent to the system; therefore, the two are not conjugate. Our research addresses this issue of quantifying the quality in a mathematically grounded manner.

**A prototypical example: Modeling a Noisy Logistic Map.** It is well known that the logistic map \( g_1(x) = rx(1-x) \) is conjugate to the tent map \( g_2(x) = a(1-2|x-\frac{1}{2}|) \) for parameter values \( r = 4 \) and \( a = 2 \). The conjugacy \( h(x) = \frac{1}{2}[1 - \cos(\pi x)] \) provides perhaps the most studied example in the pedagogy of dynamical systems for teaching conjugacy. If \( r \) is perturbed even slightly, the maps fail to be conjugate. Now consider a system with a much stronger perturbation, like a noisy version of the logistic map, which is obviously not conjugate to the tent map. The commuter shown in Fig. 1(R) is not a homeomorphism, since it fails at least one-ness, which we see immediately by direct inspection. The commuter gives the orbit equivalence between the two maps and is the key to understanding the quality of a model. One could argue that the tent map shown is a good candidate to model this “noisy logistic map” because although it simplifies the small scale dynamics it captures the main features of the large scale orbit structure.

\[
\lambda(f) = \alpha_1 \lambda_O(f) + \alpha_2 \lambda_{1-1}(f) + \alpha_3 \lambda_C(f) + \alpha_4 \lambda_{C^{-1}}(f), \tag{2}
\]

with

\[
\lambda_O(f) = \{ \text{amount that } f \text{ is not onto} \}, \quad 
\lambda_{1-1}(f) = \{ \text{amount that } f \text{ is not } 1 \rightarrow 1 \}, \quad 
\lambda_C(f) = \{ \text{amount that } f \text{ is not continuous} \}, \quad 
\lambda_{C^{-1}}(f) = \{ \text{amount that } f^{-1} \text{ is not continuous} \}. 
\]
where we acknowledge that \( f^{-1} \) may not be well defined. Furthermore, we require weights \( 0 \leq \alpha_i \) satisfy \( \sum \alpha_i = 1 \). To define these defects, we assume that there exists measure spaces, \((D_1, A_1, \mu_1)\) and \((D_2, A_2, \mu_2)\), where \( D_1 \subset X \) and \( D_2 \subset Y \), \( A_1 \) and \( A_2 \) are \( \sigma \)-algebras, and \( \mu_1 \) and \( \mu_2 \) are measures. \( D_1 \) and \( D_2 \) are the portions of phase space which are of interest to the modeler, and measures \( \mu_1 \) and \( \mu_2 \) allow the modeler to vary the relative importance of different parts of those sets. We refer to [10] for detailed definitions of \( \lambda_0(f), \lambda_{1,-1}(f), \lambda_C(f), \) and \( \lambda_{C,-1}(f) \).

**Application — modeling a Lorenz system.**

As a concrete example of how defect can be used to find optimal models, consider the flow generated by the Lorenz system with parameters, \( \beta = 8/3, \rho = 60, \) and \( \sigma = 10 \). Although this system is quite different from that studied by [11], we still apply his crucial technique by considering the map of successive maxima. Figure 2 (panel 1) shows the plot of \( z_{n+1} \) vs. \( z_n \) for a long trajectory on the attractor. Because of the strong contraction in the system, the resultant object “appears” to be one-dimensional when viewed from the macroscopic scale. With this motivating observation, we presume the following modeling goal: Find a 1-d map \( g_1 \) from the class of functions of constant slope that best models the dynamic behavior indicated by the successive maxima map, \( g_1 \).

**Note:** \( g_1 \) has no algebraic description. Rather, it reflects a long orbit along the attractor known only as a finite sequence of data, \( \{x_i\} \). Therefore, it serves to illustrate how this technique can be applied in modeling a physical system from time series.

A key step in developing commuters is to assign a partition in each of the spaces, where partitioning can be viewed as equivalent to the symbol dynamics operation of assigning a symbol to each region of phase space [4]. For 1-d maps, a typical partition assigns a symbol to each monotone segment. If we ignore the fine scale structure, the maxima map admits a very natural four-symbol partition, as illustrated by Fig. 2. Observation of a trajectory of the maxima map reveals that points assigned symbol \( d \) always iterate to points assigned symbol \( a \). By requiring that our model have the same allowable symbol transitions as \( g_1 \), we are led to consider candidate models of the form in Fig 3, where a member of the family of models is determined by parameters \( m \) and \( t \), and we denote the resultant map as \( g_2(m,t) \).

Developing a “good” model equates to choosing parameter values \( m \) and \( t \) that create model dynamics that are similar to the dynamics of \( g_1 \). No standard technique exists for making this choice. Note that entropy of any \( g_2(m,t) \) is \( \ln m \). Consequently parameter estimation using entropy methods will be unable to provide means to select \( t \). Rather, we get the best dynamical match by choosing \( m \) and \( t \) so that the resultant commuter is as close as possible to being a conjugacy. For a fixed choice of parameter values, we solve the functional equation

\[
g_2(m,t) \circ f_{m,t} = f_{m,t} \circ g_1. \tag{3}
\]

to find the commuter \( f_{m,t} \), computed as the fixed point of the operator

\[
C(f) = g_2^{-1}(m,n) \circ f \circ g_1. \tag{4}
\]
FIG. 4: Commuter functions. Commuter functions $f_{m,t}$ between the Lorenz system and models $g_{2(m,t)}$ for several choices of parameter values. For ease of visualization, we choose to provide a branched description of the commuter, where the lower branch may be associated with symbol $a$.

Because $g_1$ is known only at each $x_i$, the commuter $f$ is approximated by computing only at these same ordinates, so that $g_1(x_i) = x_{i+1}$ contributes no error to the computation of $g_2^{-1} \circ f \circ g_1$. Figure 4 shows examples to visually assess which might be the best commuter, since homeomorphism would be optimal.

To find the “best” model, we use defect as a cost function to “measure” the difference between the Lorenz system and a model. Because $f_{m,t}$ is computed on a finite set of ordinates, measure based computations can be problematic. To simplify, we use surrogate methods to compute a defect. Let

\[
\tilde{\lambda}(f) = \text{[largest vertical gap in each branch]} + \\
\text{[vertical gap between branches]} + \\
\text{[largest horizontal segment in each branch]}.
\]

This choice of defect ignores fine scale many-to-oneness of the commuter. These small scale variations are consequences of the fact that the maxima map is not a curve, but a thin, fractal structure. Ignoring the fine scale “fuzziness” of the commuter, is consistent with our modeling goal of simplification.

If we use Lebesgue measure of intervals, then evaluation of the defect on a grid of parameter values indicates that a good choice for the parameters is $m = 2.27, t = 0.247$. As an alternative approach, the modeler might decide that it is more important for the model to accurately capture the dynamics which are most frequently observed. Then instead of Lebesgue measure on $D_1$, one would choose the natural measure on the dynamics, which we assume are reflected by the density of points in the long trajectory. To account for this change of emphasis, we measure horizontal segments not by their geometric length but by the fraction of the $x_i$ that fall in that interval. Recomputing the defect over the grid of parameter values, we find that the modeler should choose $m \approx 2.01, t \approx 0.073$. We emphasize that the expert knowledge of the modeler remains a crucial input to this process as the modeler must decide what is “important” for the model to be able to describe. The techniques simply give the modeler a principled way to select the parameter values. Figure 5 illustrates this method of parameter selection.

Here, we have opened discussion of how the art of modeling can be cast in the language of dynamical systems by an appropriate extension of the usual notion of conjugacy to the generalized concept to measure a defect in a commuter function. We have discussed parameter estimation within this context, leading to a well defined concept of quality of simplified toy models representing the nature of the dynamics of a full system. We have noted that expert knowledge of the modeler need not, and even should not, be mathematically removed from the process.
Electronic address: jskufca@clarkson.edu  
† Electronic address: ebollt@clarkson.edu