Timing of transients: quantifying reaching times and transient behavior in complex systems

2017 New J. Phys. 19 083005

(http://iopscience.iop.org/1367-2630/19/8/083005)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
16 November 2016

REVISED
20 June 2017

ACCEPTED FOR PUBLICATION
23 June 2017

PUBLISHED
7 August 2017

PAPER

Timing of transients: quantifying reaching times and transient behavior in complex systems

Tim Kittel[1,2], **Jobst Heitzig**[1], **Kevin Webster**[1] **and Jürgen Kurths**[1,2,3]

[1] Potsdam Institute for Climate Impact Research, Telegrafenberg A31—(PO) Box 60 12 03, D-14412 Potsdam, Germany
[2] Institut für Physik, Humboldt-Universität zu Berlin, Newtonstraße 15, D-12489 Berlin, Germany
[3] Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom

E-mail: Tim.Kittel@pik-potsdam.de

Keywords: early-warning signals, complex systems, nonlinear dynamics, ordinary differential equations, stability against shocks

Supplementary material for this article is available online

Abstract

In dynamical systems, one may ask how long it takes for a trajectory to reach the attractor, i.e. how long it spends in the transient phase. Although for a single trajectory the mathematically precise answer may be infinity, it still makes sense to compare different trajectories and quantify which of them approaches the attractor earlier. In this article, we categorize several problems of quantifying such transient times. To treat them, we propose two metrics, area under distance curve and regularized reaching time, that capture two complementary aspects of transient dynamics. The first, area under distance curve, is the distance of the trajectory to the attractor integrated over time. It measures which trajectories are 'reluctant', i.e. stay distant from the attractor for long, or 'eager' to approach it right away. Regularized reaching time, on the other hand, quantifies the additional time (positive or negative) that a trajectory starting at a chosen initial condition needs to approach the attractor as compared to some reference trajectory. A positive or negative value means that it approaches the attractor by this much 'earlier' or 'later' than the reference, respectively. We demonstrated their substantial potential for application with multiple paradigmatic examples uncovering new features.

1. Introduction

In complex dynamical systems, the importance of a trajectory's transient, i.e. the part of the trajectory distant from the attractor, has been identified in physics research as well as in various other fields. Different phenomena during the process of magnetization for various materials, in particular the domain growth, have been studied extensively [1–3]. In laser physics, it was possible to derive analytical results matching the transient phases of different lasers [4, 5]. In kinetic theory, there has been research on non-equilibrium approaches for more than a century by now [6]. Other modern areas of statistical physics have emphasized the importance of transient dynamics, too, e.g. in social systems [7] and transient phases between jam and free-flow phases in vehicular traffic [8].

Even outside of the direct field of physics, but still within the scope of complex dynamical systems, a focus on transient dynamics has been developed recently. Hastings [9] made a call for more transient analysis of ecological models. An example of this was given by van Geest [10], describing macrophyte-dominated states of lakes as non-equilibrium states. In medicine and biology, epilepsy is seen as a transient phenomenon and much work has been done [11, 12]. In economics, a transient analysis complementing the asymptotic analysis proved to be fruitful, particular supporting with the stability analysis and understanding how to reach the equilibria [13]. Climate change is often seen as a transition to a new situation, i.e. a transient change to a new attractor [14–16]. Closely related, discussions in sustainability sciences are on transient dynamics because they refer to transformations from and to sustainability. Important key topics are the Anthropocene [17–19], particular the great acceleration [20], and planetary boundaries [21, 22].

An important emphasis on *long transients* has been made in [10, 15, 23]. With this term, they refer to trajectories where the relevant and observable phenomena/states, e.g. macrophyte-covered lakes or desert states of the Earth system, are away from the actual attractor, but in the transient phase where a trajectory may stay for a substantial amount of time.

Hastings [9] stressed the importance of different time scales and pointed out how the transient dynamics can be very different and much more interesting than the asymptotic behavior. In addition, he explained how saddles play a central role by inducing long transients. This has been demonstrated in a study by Anderies *et al* [15] in the context of sustainability science. The 'interacting planetary boundary' [21, 22] has been defined by whether states take 'long' to the attractor or not. This idea leads precisely to the main question for this article 'How can we properly quantify the time to reach a system's attractor?', i.e. associate meaningful numbers with it.

This study is meant as a methodological step in direction for applications in real-world systems. So in the following, we focus on being able to do numerical estimations while analytical results are only given to understand general properties.

Often, a trajectory is divided arbitrarily in a transient part and the asymptotics close to the attractor. So we split the main question into two sub-questions: (a) 'What are the problems of these current/intuitive methods to quantify transient time?' and (b) 'How can we mend them?'.

To answer the first question, we work out four essential problems one is confronted with: (I) *infinite reaching time*: the attractor is not reached in finite time for a large class of physically relevant systems; (II) *physical interpretation*: it is unclear how to define precisely 'when the transient is over', so it is ambiguous where to divide between the transient and the asymptotics; (III) *discontinuities*: when having parameter dependence, small changes in the parameter often induce a large (noncontinuous) effect on the measured quantity; and (IV) *non-invariance*: the results depend on the choice of coordinates. Problem (IV) is particularly important, as a result should be a property of the dynamical system and thus independent of the choice of coordinates, i.e. invariant (or correctly transforming) under smooth transformations of the state space (see 'smoothly equivalent' in [24]).

Then, we approach question (b) by formulating two metrics, *area under distance curve* (D) and *regularized reaching time* (T_{RR}). The first one is the integral over the distance to the attractor along the trajectory, and has a physical dimension of time times distance. It measures which trajectories are *reluctant*, i.e. stay distant from the attractor for long, or *eager*, i.e. approach it right away. The second one, T_{RR}, is defined by the difference between the reaching times for the trajectory of interest and a reference trajectory. Thus, it takes a different point of view, actually measuring a time. The idea is that even though the actual reaching times are infinite (problem (I)), their difference is typically finite. So, we can compare trajectories approaching the attractor and define the notions *earlier* and *later*.

We chose four examples to illustrate different features of these metrics. We first use a linear system to understand how the metrics act generally and to observe the divergence of T_{RR} on the strong stable manifold particularly. Also, due to the system's simplicity, analytical solutions are possible. We then use a global carbon cycle model [15] and a model of a generator in the power grid [25] to apply the ideas to some first real world systems. Our final example, the chaotic Rössler oscillator, demonstrates that one can apply these methods to more complex attractors also, in this case a chaotic one. The chosen examples are rather well-understood. So they are good testing cases for the metrics, while their complexity still needs numerical approaches for a proper quantification of reaching times.

Finally, a detailed discussion on how far the two metrics solve the aforementioned problems is given, followed by a summary and an outline of future research.

The remainder of this article is structured as follows. After stating the four essential problems of reaching time definitions in section 2, we illustrate them with a small example. Then, we present the two metrics in section 3 and apply them to examples in section 4. Next, we give a detailed discussion on how far the metrics solve the essential problems in section 5. Finally, we close with a summary and an outlook. Additional information that can be found the supplemental material is available online at stacks.iop.org/NJP/19/083005/mmedia referenced within the article with the prefix 'Suppl. Mat.'.

Assumptions and notations. We assume a general, deterministic and autonomous dynamic system given by the differential equation

$$\dot{x} = f(x) \qquad x \in X \tag{1}$$

with an n-dimensional state space $X = \mathbb{R}^n$ and the right-hand side (rhs) $f(x)$. Usually, we use x to denote an arbitrary state $x \in X$, and refer to specific/fixed states with letters as superscripts, e.g. x^a, x^b, and x^{ref}. The components of a state are written with subscripts, so $x = (x_0, x_1,\ldots,x_{n-1})^\top$ and $x^a = (x_0^a, x_1^a,\ldots,x_{n-1}^a)^\top$. The words 'point' and 'state' are used synonymously for the elements of X. We assume the system (1) to have at least one attractor $\mathcal{A} \subseteq X$ with a basin of attraction $\mathcal{B}_{\mathcal{A}} \subseteq X$. In case the system has more than one attractor, the analysis should be applied to the attractors of interest separately.

For convenience, we will make heavy use of the time-evolution operator φ where $\varphi(t, x)$ is the state after starting at some point x and letting the system evolve for some time $t \geqslant 0$. Hence

$$\varphi(0, x) = x \quad \text{and} \quad \frac{\partial \varphi}{\partial t}(t, x) = f(\varphi(t, x)). \tag{2}$$

When we speak of 'quantifying the transient time' we mean to find a function $X \longrightarrow \mathbb{R}$, a 'metric', that gives a reasonable number for the time a trajectory spent in transient phase for each initial condition $x \in \mathcal{B}_A$.

Additionally, within the article we assume the asymptotics of the system to be understood as we want to focus on the transient only.

2. The problems of reaching time definitions

In this section, we introduce four essential problems. They need to be addressed when aiming to quantify the transient time to reach an attractor $\mathcal{A}$ in a system of type (1). Then, we illustrate them with an example model.

(I) Infinite reaching time. A basic property of a large class of complex systems is that trajectories reach the attractor in infinite time only. That includes even steady states or limit cycles and most systems of ordinary differential equations with smooth rhs functions. This is the fundamental problem why the analysis made in this article is necessary.

(II) Physical interpretation. It is far from being obvious what the terms 'close to the attractor' or 'when the transient is over' means. Often, this is tackled by using some arbitrary threshold ϵ to define what is a 'small distance' to $\mathcal{A}$. But because of problem (I), the time to reach this ϵ-neighborhood typically diverges for $\epsilon \to 0$. So the result depends strongly on the value of ϵ. Note that the focus of this article is to *quantify* the transient time to reach the attractor. So we want to associate meaningful numbers and need to treat this problem.

(III) Discontinuities. When defining a metric to quantify the transient time to reach $\mathcal{A}$ using some parameters e.g. ϵ, the result might depend discontinuously on the parameter. Usually, we want results to change smoothly and, if possible, weakly to changes of the parameter. If there is a discontinuous dependence, then we would expect there to be a corresponding specific property of the system that introduces this behavior.

(IV) Non-invariance. Our focus is on real-world systems. So the transient time should be a general property of the system, and not dependent on the chosen variables or coordinates to represent it. These coordinates correspond to a point of view on the system only. In other terms, invariance under change of coordinates should be given.

Example. While the aforementioned problems are of general nature, we illustrate them next using the example system

$$\dot{x}_0 = 1 - \frac{x_1}{2} - b x_0, \quad \dot{x}_1 = 2(a - x_0^2), \quad a = 2, \qquad b = 0.3. \tag{3}$$

It has a stable focus $x^s = (-\sqrt{a},\, 2(1 + b\sqrt{a}))^\top$ as its only attractor, and a saddle $x^u = (\sqrt{a},\, 2(1 - b\sqrt{a}))^\top$. This has been chosen deliberately simple but is still sufficient to demonstrate all four problems. This way, we do not have to cope with problems inherent to the example system, like high-dimensionality or chaos.

Its flow is shown in figure 1(a). For a chosen trajectory starting at $x^a = (2.8,\, 6.2)^\top$ (see figure 1(a)) the time-dependence of the Euclidean distance to the attractor

$$d_E(\varphi(t, x^a), x^s) = \sqrt{\varphi(t, x^a)^\top \cdot x^s} \tag{4}$$

is depicted in figure 1(a). Two common metrics are the times when an ϵ-neighborhood is entered the first and the last time. So we define the class of sets

$$T^\epsilon(x^a) = \{t \mid \epsilon = d_E(\phi(t, x^a), x^s)\}, \tag{5}$$

that invert the axes of figure 1(b) as depicted in figure 1(c) (blue dotted line). $T^\epsilon(x^a)$ is the set of times when the ϵ-neighborhood is entered or left. Furthermore, the first and last entry times are then $T_F^\epsilon(x^a) = \inf T^\epsilon(x^a)$ and $T_L^\epsilon(x^a) = \sup T^\epsilon(x^a)$ respectively. They are graphed in figure 1(c) also.

The infinite reaching time (problem (I)) is visible in figure 1(c) right away, as $T_F^\epsilon(x^a),\ T_L^\epsilon(x^a) \to \infty$ for $\epsilon \to 0$. By definition, this implies that all elements in $T^\epsilon(x^a)$ will approach ∞ also.

Problem (II): $T_F^\epsilon(x^a)$ and $T_L^\epsilon(x^a)$ depend heavily on the choice of ϵ. So a proper physical interpretation is rather difficult. The notions of 'close to the attractor' or 'when the transient is over' depend strongly on ϵ.

The strong discontinuities (problem (III)) for $T_F^\epsilon(x^a)$ and $T_L^\epsilon(x^a)$ when changing ϵ in figure 1(c) make the choice of a proper ϵ even harder. The discontinuities arise because the trajectory will (for a fixed ϵ) enter and exit the corresponding ϵ-neighborhood several times. This behavior is caused by the complex eigenvalues of the system. It could be circumvented locally by choosing a different distance function, for instance

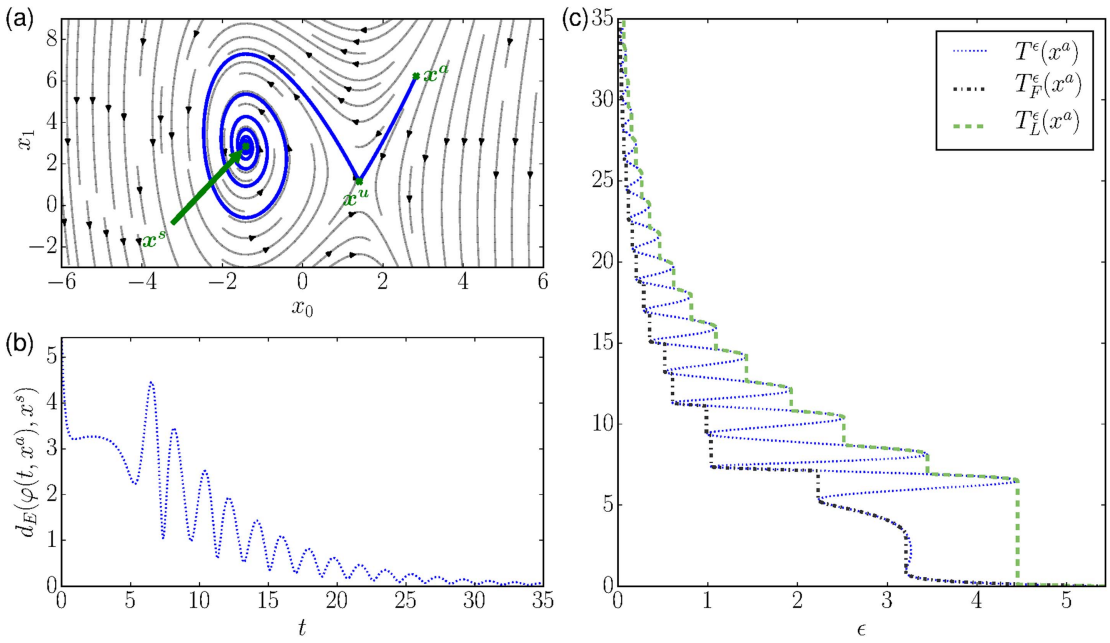

Figure 1. The phase space for the example system equation (3) is depicted in (a). Furthermore, the stable spiraling node x^s and the saddle x^u are added. The trajectory (blue) starting at x^a closely passes by x^u before it finally circulates in to x^s. (b) shows the Euclidean distance d_E (dotted, blue) of this trajectory to its attractor x^s over time t. The first longer dip between $t = 1$ and $t = 5$ is the transient at the (unstable) saddle x^u while the oscillations afterwards are the spiraling around x^s. (c) turns (b) around in order to show the dependence of the time t on some distance $d = \epsilon$ (dotted, blue) of the trajectory to the attractor x^s. Secondly, there are multiple values of ϵ for each t so observables like $t^\epsilon_F(x^a)$ and $t^\epsilon_L(x^a)$ need to be introduced. $t^\epsilon_F(x^a)$ (dash–dotted, black) marks when the first time the ϵ-neighborhood around x^s is entered and $t^\epsilon_L(x^a)$ (dashed, green) the last time. The implications, particularly the arising problems for time definitions, are described in text.

$$d_P(x, x^s) = \|P^{-1} \cdot (x - x^s)\|, \quad P = \begin{pmatrix} 1 & 1 \\ 4\sqrt{a}/\lambda_+ & 4\sqrt{a}/\lambda_- \end{pmatrix}, \tag{6}$$

where $\lambda_\pm = -b/2 \pm \sqrt{b^2/4 - 2\sqrt{a}}$ are the complex eigenvalues of the linearization of (3) around x^s. (This is related to the $\|\cdot\|_P$-norm used in Suppl. Mat. Proposition 3.4.) Unfortunately, this is not so easy for more complex attractors, e.g. the later treated chaotic Rössler system. However, we will present a pragmatic solution to this problem in section 3.2.

Finally, using a different set of coordinates, i.e. smoothly transforming the system, gives different values for $T^\epsilon_F(x^a)$ and $T^\epsilon_L(x^a)$, because the Euclidean distance is not invariant. Hence, the result depends on the set of coordinates chosen for the system and is not invariant under coordinate transformations (problem (IV)). This dependence on the chosen distance is also known to appear in finite-time dynamical systems and their stability [26].

3. Two complementary metrics

To treat the aforementioned problems, we devise two metrics for a general system as equation (1): *area under distance curve* (abbreviated as D) and *regularized reaching time* (T_{RR}). They naturally lead to a transient analysis from separate points of view as explained in the following.

3.1. Area under distance curve

Area under distance curve (D) comes from the idea that a trajectory stays distant from the attractor during the transient while it is close in the asymptotics. A distance function $d(\cdot, \cdot)$ is needed to have notions of 'far' and 'close' and we define D

$$D(x) = \int_0^\infty \mathrm{d}t \; d(\varphi(t, x), \mathcal{A}), \tag{7}$$

where $\mathcal{A}$ is the attractor with the basin $\mathcal{B}_\mathcal{A}$ and φ the time-evolution operator as in equation (2). So we look at the cumulative distance to the attractor and remove the influence of the asymptotics. As (7) is the integral over the distance in time, D is the area below the distance curve. A different point of view is that it is the time weighted by the distance.

As D is defined with the limit of an integral, a note on the convergence is due and can be found in the discussion in section 5.

The distance function d should be between a point in state space and the attractor $\mathcal{A}$. If $\mathcal{A}$ contains more than just one element, it could be the infimum of the distances to all points within. Choosing a tailor-made function $d(\cdot, \cdot)$ allows to adapt the metric to specific research questions, e.g. by letting $d(x, \mathcal{A})$ represent some form of costs or damages due to being away from the attractor. For the idea of D to work, d needs to approach 0 around the attractor and be 0 on it.

Due to the integral representation, D can be estimated numerically directly from the trajectory assuming the attractor is known. The latter was taken as a prerequisite for this article as we want to emphasize the analysis of the transient.

Initial conditions with relatively high values of D are called 'reluctant' and those with low values 'eager'. This terminology is used to emphasize that reluctant states go through large transients distant from the attractor, while eager states approach it directly.

By straightforward differentiation, we can compute the orbital derivative

$$\frac{\partial}{\partial t} D(\varphi(t, x)) = -d(\varphi(t, x), \mathcal{A}), \tag{8}$$

meaning its value strictly decreases along the flow; a property we use later. Furthermore, this shows it to be a Lyapunov function [27]. Furthermore, using equation (8) and adding the condition $D(x) = 0 \ \forall x \in \mathcal{A}$ is an alternative definition for D.

3.2. Regularized reaching time

The second idea, *regularized reaching time* (T_{RR}), is based on time differences between trajectories. It can be interpreted as the additional time (positive or negative) that a trajectory starting at a point of interest needs to approach the attractor after a reference trajectory has already approached it. A positive or negative value means that the trajectory at hand approaches the attractor by this much later or earlier, respectively, than the reference trajectory does.

To formalize this idea, we introduce $t^\epsilon(x)$ as the time a trajectory starting at an arbitrary state x needs in order to reach an ϵ-environment around the attractor. That means for some function $\Delta : X \longrightarrow \mathbb{R}_{\geqslant 0}$ it holds that

$$\epsilon = \Delta(\varphi(t^\epsilon(x), x)), \tag{9}$$

where we want Δ to be 0 on the attractor and $\Delta(\varphi(t, x))$ to be strictly and continuously decreasing in t. This means, in equation (9), Δ has the role of a generalized distance function, measuring how far a point in state space is away from the attractor. Note that $\varphi(t^\epsilon(x), x)$ is the state after starting at x and evolving the system for a time $t^\epsilon(x)$. Hence, equation (9) implicitly defines $t^\epsilon(x)$ to be the time at which an ϵ-environment around the attractor, with respect to the generalized distance function Δ, is entered.

Since the actual reaching times to the attractor are both infinite, T_{RR} is formally described as the limit for $\epsilon \to 0$ of the difference between how long the trajectory starting at some arbitrary state x and the trajectory starting at a chosen (fixed) reference point x^{ref} need to enter the corresponding ϵ-environment

$$T_{\mathrm{RR}}(x; x^{\mathrm{ref}}) = \lim_{\epsilon \to 0} (t^\epsilon(x) - t^\epsilon(x^{\mathrm{ref}})). \tag{10}$$

For hyperbolic fixed points, we prove in Suppl. Mat. section 3 under mild conditions that there exists a class of choices for Δ such that this limit exists for all x within the basin of attraction except the strong stable manifold and the attractor itself. We call the manifold associated to all Lyapunov exponents except the leading one the strong stable manifold. And, if T_{RR} exists, it is unique, i.e. independent of which Δ has been chosen from the class.

Furthermore, we show that T_{RR} is a parametrization of the strong stable foliation. Thus, after a smooth change of coordinates Φ, i.e. a diffeomorphism of the state space, the diffeomorphic image of the strong stable foliation will again parametrize the level sets of T_{RR} in the new variables. Therefore, T_{RR} is invariant under such transformations and it holds that

$$T_{\mathrm{RR}}(\Phi(x); \Phi(x^{\mathrm{ref}})) = T_{\mathrm{RR}}(x; x^{\mathrm{ref}}), \tag{11}$$

where we obviously transformed x^{ref}, too.

T_{RR} represents *the actual time* by how much a trajectory approaches the attractor later or earlier than the one starting at the reference point, so we call states with relatively low T_{RR} 'early' and with high T_{RR} 'late'.

Different choices of x^{ref} (that are not on the strong stable manifold or the attractor) result in additive constants. To be precise, choosing another $x^{\mathrm{ref}\prime}$ yields

$$T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) - T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}\prime}) = T_{\mathrm{RR}}(x^{\mathrm{ref}\prime};\, x^{\mathrm{ref}}). \tag{12}$$

Because the rhs of equation (12) does not depend on x, different choices of x^{ref} do not influence the structure of T_{RR} w.r.t. x. Thus central moments, i.e. ones invariant under shifts, are sensible for analyzing T_{RR} over a distribution of initial conditions in state space; especially the standard deviation proves useful for the examples below. In particular, for any choice of x^{ref} it obviously holds that $T_{\mathrm{RR}}(x^{\mathrm{ref}};\, x^{\mathrm{ref}}) = 0$.

The reference point should not be chosen on the attractor because this gives $t^\epsilon(x^{\mathrm{ref}}) = 0$ for any ϵ, but for $x \in X \backslash \mathcal{A}$ the time $t^\epsilon(x) \to \infty$ for $\epsilon \to 0$. Vice versa, this means when having chosen $x^{\mathrm{ref}} \notin \mathcal{A}$ then $T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) = -\infty \;\; \forall x \in \mathcal{A}$. The same holds for the strong stable manifold.

In order to compute the orbital derivative $\frac{\partial}{\partial t} T_{\mathrm{RR}}(\varphi(t,\, y);\, x^{\mathrm{ref}})$, we use equation (10) and find

$$\frac{\partial}{\partial t} T_{\mathrm{RR}}(\varphi(t,\, y);\, x^{\mathrm{ref}}) = \lim_{\epsilon \to 0} \frac{\partial}{\partial t} t^\epsilon(\varphi(t,\, y)), \tag{13}$$

where $y \in X$ is an arbitrary state and exchangeability of the limit and the derivative has been assumed. Next, we take the derivate with respect to time t in equation (9) for $x = \varphi(t,\, y)$. Sorting the terms appropriately gives

$$0 = \left(\frac{\partial}{\partial t}(\Delta \circ \varphi)\right)(t^\epsilon(\varphi(t,\, y)) + t,\, y) \cdot \left(\frac{\partial}{\partial t} t^\epsilon(\varphi(t,\, y)) + 1\right). \tag{14}$$

$\Delta(\varphi(t,\, x))$ is strictly decreasing in t for any $x \in X$. So its derivative is $\frac{\partial}{\partial t}\Delta(\varphi(t,\, x)) = \left(\frac{\partial}{\partial t}(\Delta \circ \varphi)\right)(t,\, x) < 0$ and in particular non-zero. Hence, $\frac{\partial}{\partial t} t^\epsilon(\varphi(t,\, y)) = -1$, leading finally to the orbital derivative of T_{RR}

$$\frac{\partial}{\partial t} T_{\mathrm{RR}}(\varphi(t,\, y);\, x^{\mathrm{ref}}) = -1. \tag{15}$$

This equation is actually rather natural, as the change of time to approach the attractor along the trajectory should exactly be the time passed. Also, this makes it a Lyapunov function [27].

To use equation (15) as an alternative definition we need another constraint. Because of $T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) = -\infty \;\; \forall x \in \mathcal{A}$, this cannot be done on the attractor (in contrast to D). In case of hyperbolic fixed points, it follows directly from Suppl. Mat. Proposition 3.7 that T_{RR} is a parametrization of the strong stable foliation $\mathcal{F}_{\mathrm{ss}}$, whose definition is recalled in Suppl. Mat. Theorem 3.6. So we can use the constraint that $T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) = 0 \;\; \forall x \in \mathcal{F}_{\mathrm{ss}}^{\mathrm{ref}}$, where we call $\mathcal{F}_{\mathrm{ss}}^{\mathrm{ref}} = \mathcal{F}^{\mathrm{ss}}(x^{\mathrm{ref}})$ the *reference leaf* containing x^{ref}. For more complex attractors, a generalized condition needs to be found and this is part of the outlook.

Suppl. Mat. Proposition 3.4 provides the convergence of T_{RR} in equation (10) for hyperbolic fixed points only. When thinking about more complex attractors that may arise in real-world examples the question of convergence comes up again. A general idea why T_{RR} should converge with a well chosen Δ in this case, too, is that in the asymptotics, trajectories will 'behave similarly' because they are close to the attractor. So, for two very small $\epsilon_1 > \epsilon_2$, the time difference to enter the ϵ_2-environment after entering the one of ϵ_1 should be roughly the same, independent from where a trajectory started. Hence, for two states x and x^{ref} we can assume $t^{\epsilon_2}(x) - t^{\epsilon_1}(x) \approx t^{\epsilon_2}(x^{\mathrm{ref}}) - t^{\epsilon_1}(x^{\mathrm{ref}})$ implying $t^{\epsilon_2}(x) - t^{\epsilon_2}(x^{\mathrm{ref}}) \approx t^{\epsilon_1}(x) - t^{\epsilon_1}(x^{\mathrm{ref}})$. This suggests that the limit in equation (10) might exist. So a crucial problem is to find an appropriate function for Δ in order to get an estimation for T_{RR}.

Estimation of T_{RR}. The first idea for a Δ would be the infimum of the Euclidean distance to the points in the attractor. Basically, this means that t^ϵ should be replaced by T_{F}^ϵ or T_{L}^ϵ from section 2. This would give a very coarse estimation but is probably not the correct choice as the condition of Δ being strictly decreasing along the flow is in general not fulfilled.

A pragmatic choice of Δ is D, the area under distance curve. It fulfills both conditions demanded for Δ (see section 3.1) when using for d the infimum of the Euclidean distance to the attractor points. Hence, we can define $t_D^\epsilon(x)$ as the time until the D (equation (7)) of the trajectory's remainder is ϵ-small

$$\epsilon = D(\varphi(t_D^\epsilon(x),\, x)) = \int_{t_D^\epsilon(x)}^{\infty} \mathrm{d}t \;\, d(\varphi(t,\, x),\, \mathcal{A}). \tag{16}$$

Note that the ideas for D and T_{RR} are generally independent and the usage of D in this case is purely because it fulfills the above mentioned conditions. So it is a good, pragmatic choice.

Using t_D^ϵ defined in equation (16) as the time-function t^ϵ in equation (10) for the estimation of T_{RR}, our numerical results show that this idea is sensible for more complex attractors, e.g. in the Rössler system below.

4. Examples

In order to demonstrate the applicability of the metrics, we selected four examples with differing properties and increasing complexity.

4.1. Linear system with two different time scales

Even though we want to focus on going in the direction of application to real-world systems, understanding some features in a basic linear system proves useful. For general systems, T_{RR} and D can be tackled numerically only. But a linear system can be solved analytically and explicit expressions for both metrics were found. We will first analyze both metrics for a general linear system and then discuss a chosen example.

T_{RR} for a general linear system. For a hyperbolically stable linear system with a (complex-)diagonalizable matrix $A \in \mathbb{R}^{n\times n}$ and the fixed point x^{f} at the origin,

$$\dot{x} = A \cdot x, \tag{17}$$

we decompose $x = \sum_{i=0}^{n-1} \alpha^i v^i$ with coefficients $\alpha^0,...,\alpha^{n-1}$ in the eigenvector basis $v^0,...,v^{n-1}$ with eigenvalues $\lambda^0,...,\lambda^{n-1}$ sorted in descending order by real part. We assume in particular λ^0 to have a strictly larger real part than λ^1 and multiplicity one. Hence we can apply Suppl. Mat. equation (10) derived in the Suppl. Mat. and get

$$T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) = \frac{1}{\lambda_0} \ln \left| \frac{\alpha^{0,\mathrm{ref}}}{\alpha^0} \right|, \tag{18}$$

where $\alpha^{0,\mathrm{ref}}$ is the α^0 coefficient for the reference point x^{ref}. $\alpha^{0,\mathrm{ref}}$ should be non-zero, i.e. x^{ref} should not be on the strong stable manifold.

Note that Suppl. Mat. Proposition 3.4 gives the uniqueness of this result independent of the choice of Δ.

In equation (18), T_{RR} depends only on α^0, meaning the projection of x on the eigenvector corresponding to the least stable eigenvalue λ^0. While this might be counter-intuitive in the beginning, it can be explained: the contributions from all other eigenvalues are vanishing because they decay faster than λ^0 by definition. So for a linear system, only the contribution from λ^0 remains. Also, on the strong stable manifold where $\alpha^0 = 0$, the values for T_{RR} go to $-\infty$ which we mentioned already in section 3.2 for general systems.

D for a general linear system. Taking the system (17) and choosing $d(x,\, \{x^{\mathrm{f}}\}) = d_E(x,\, x^{\mathrm{f}})^2$ the squared Euclidean distance, we calculate D directly by using the definition equation (7)

$$D(x) = \sum_{i,j=0}^{n-1} \frac{-(\alpha^i)^* \alpha^j}{(\lambda^i)^* + \lambda^j} (v^i)^\dagger v^j. \tag{19}$$

Therefore, in case of D, all eigenvalues contribute, contrary to T_{RR}. But they are weighted as can be seen in the denominator. In case of A being symmetric, this formula can be reduced to $D(x) = \frac{1}{2} x^\top A^{-1} x$.

T_{RR} for an example linear system. We choose the $n =$ two-dimensional linear system

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ 4 & -2 \end{pmatrix} \cdot x \tag{20}$$

with a stable and a strong stable eigenvalue and corresponding eigenvectors

$$\lambda^{\mathrm{s}} = -1,\ v^{\mathrm{s}} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \text{ and } \lambda^{\mathrm{ss}} = -2,\ v^{\mathrm{ss}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{21}$$

We choose the reference point to be $x^{\mathrm{ref}} = (1,\ 1)^\top$. Identifying $\lambda^0 = \lambda^{\mathrm{s}}$, $v^0 = v^{\mathrm{s}}$ implies $\alpha^0 = x_0$. Then, using equation (18) gives

$$T_{\mathrm{RR}}(x;\, x^{\mathrm{ref}}) = \ln(|x_0|). \tag{22}$$

This result is also visible in the numerical estimation in figure 3(c); the values of T_{RR} change only in the direction of x_0. The coloring describes the values of the metrics (see the colorbar in the right of the figures) and the green star represents x^{ref}.

In order to get a better feeling for these metrics, we have chosen two exemplary initial conditions, an *early-eager* one and a *late-eager* one, and plotted their trajectories' distance to the attractor over time in figure 2. We see an intuition for T_{RR}: it can be interpreted as the time-shift between the original trajectory and the reference trajectory until the asymptotics match. So we plotted both trajectories shifted to each other using the analytical result for T_{RR} in equation (22).

D for an example linear system. Analyzing D for the example linear system in equation (20) gives

$$D(x) = \frac{11}{6} x_0^2 + \frac{1}{4} x_1^2 + \frac{2}{3} x_0 x_1, \tag{23}$$

where equation (19) has been used. The numerical result in figure 3(a) confirms this.

In figure 2, the blue-shaded area corresponds to the D value which is the same in both cases of our particular choice. This choice was made in order to see how trajectories can have differing T_{RR} values even if the D values match.

The exponential lower bound that comes up in the scatter plot figure 3(b) can be calculated analytically by combining equations (22) and (23)

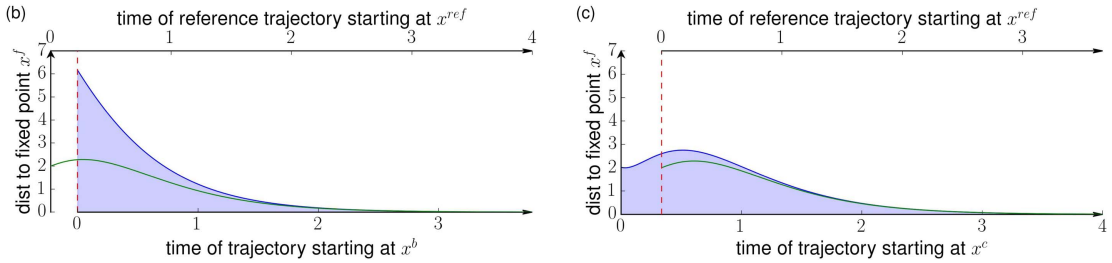

Figure 2. The figure shows for two exemplary initial conditions (a) $x^b = (0.8,\ 2.35)^\top$ and (b) $x^c = (1.4,\ 0.24)^\top$ the distance of the attractor over time (blue curve) in the linear example system of section 4.1. The initial conditions have been chosen such that the D value, which corresponds to the blue-shaded area, is the same for both trajectories, $D(x^b) = D(x^c) = 3.8$. But the trajectory starting at x^b approaches the attractor *earlier* than the reference trajectory (green in (a) and (b)), which in turn is *earlier* than the one from x^c, meaning $T_{\mathrm{RR}}(x^b;\, x^{\mathrm{ref}}) = -0.22 < T_{\mathrm{RR}}(x^{\mathrm{ref}};\, x^{\mathrm{ref}}) = 0 < T_{\mathrm{RR}}(x^c;\, x^{\mathrm{ref}}) = +0.34$. In order to show this, the example trajectories (blue) have been shifted in each plot by the value of T_{RR} with respect to the reference trajectory (green). This demonstrates an intuition behind T_{RR}: it describes by how much one has to shift one trajectory so it matches the asymptotics of the reference trajectory.

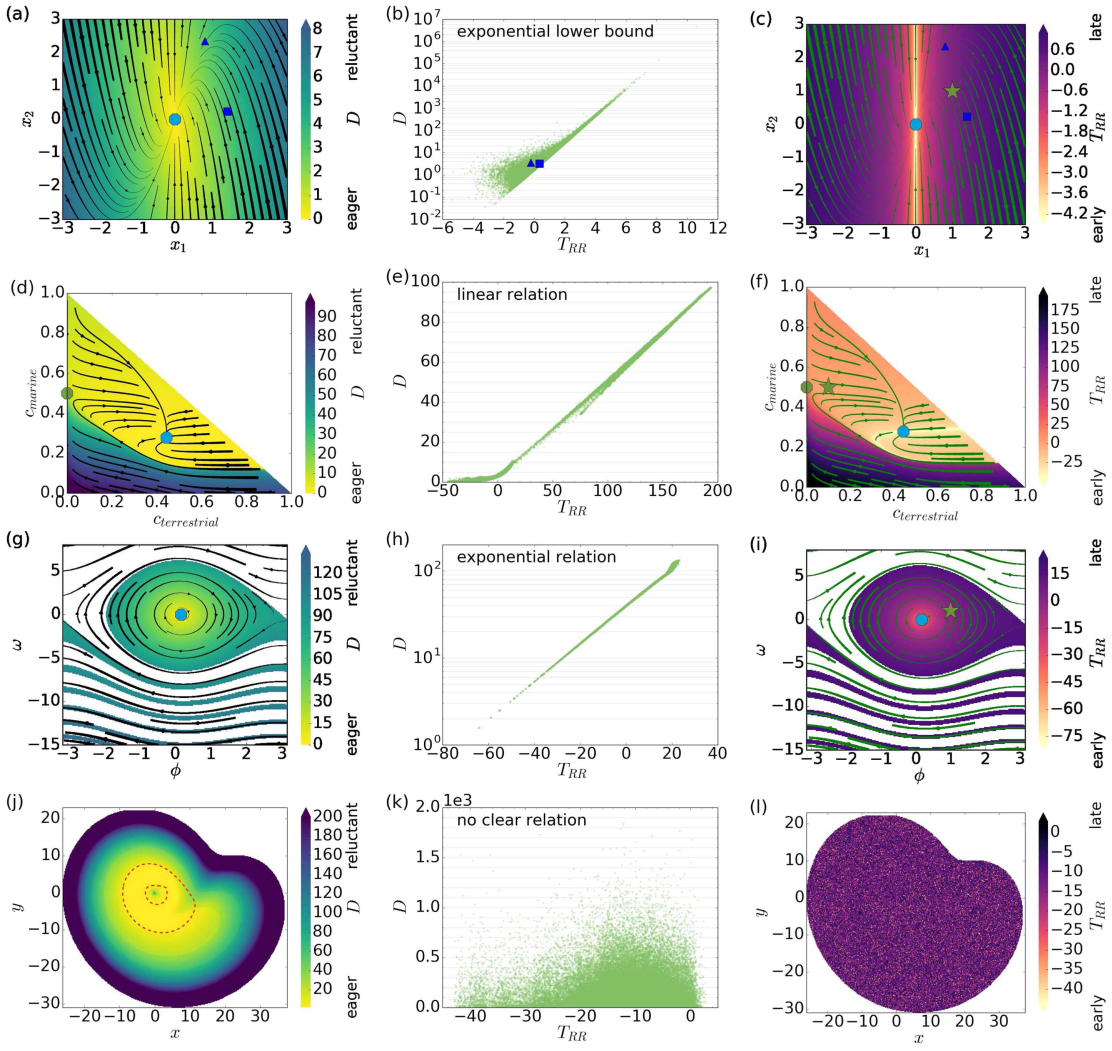

Figure 3. For the presented example systems (top to bottom: linear system, global carbon cycle, generator in a power grid, Rössler system) the two new metrics have been computed for each initial condition in the state space and marked with color, see left column *area under distance curve* (D, D) and right column *regularized reaching time* (T_{RR}). The middle column shows their relations for the particular system. The initial conditions x^b (triangle) and x^c (square) from figure 2 have been marked in (a)–(c), too. As the Rössler system is three-dimensional, the above plot depicts only a slice at fixed $z = 0.6$ where the boundary of the attractor's projection to this plane is shown in dashed red lines. As this is only a projection, D is not 0 for all points within. For comparison, a graphical representation of the full attractor can be found in supplementary material section 2.

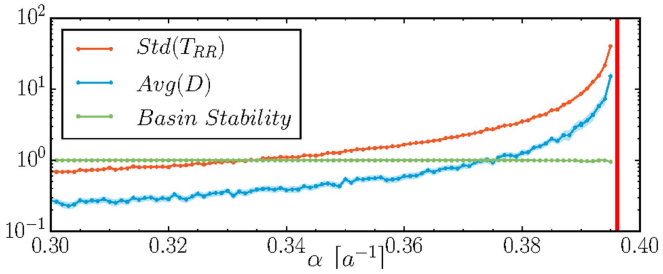

Figure 4. For the *global carbon cycle* in equations (25*a*) and (25*b*), the mean of D and standard deviation of T_{RR} are plotted (with their 5%–95% bootstrap error) and show a divergence before the parameter α (yearly human carbon offtake) reaches the bifurcation value (marked by the red line). For comparison Basin stability is shown, which does *not* show any change because the size of the basin stays constant before the bifurcation.

$$D(x) \geqslant \frac{25}{18} \mathrm{e}^{2T_{\mathrm{RR}}(x;x^{\mathrm{ref}})}. \tag{24}$$

4.2. Global carbon cycle

The second example has been chosen to take a step in the direction of real-world examples. It is a conceptual model of the global carbon cycle proposed by Anderies *et al* [15]. We use the pre-industrialization version. It consists of three dynamical variables, the terrestrial, marine and atmospheric carbon stocks, denoted by $c_{\mathrm{t}} = c_{\mathrm{terrestrial}}$, $c_{\mathrm{m}} = c_{\mathrm{marine}}$ and $c_{\mathrm{a}} = c_{\mathrm{atmospheric}}$ respectively. Furthermore, the conservation of total carbon is formulated in the constraint $C = c_{\mathrm{t}} + c_{\mathrm{m}} + c_{\mathrm{a}} =$ const. Thus, we can reduce the system to 2 state variables c_{t} and c_{m} and rescale the units such that $C = 1$:

$$\dot{c}_{\mathrm{t}} = \mathrm{NEP}(p, r, c_{\mathrm{t}}) - \alpha c_{\mathrm{t}}, \tag{25a}$$

$$\dot{c}_{\mathrm{m}} = I(c_{\mathrm{a}}, c_{\mathrm{m}}), \tag{25b}$$

where NEP is the net Eco-system production, p photosynthesis, r respiration, α harvesting parameter and I diffusion; indirect dependencies have been omitted and more details are in [15, 28]. As the full equations are rather lengthy, we put them in Suppl. Mat. section 1 and refer in the analysis to the flow that is drawn in figures 3(d) and (f) and the α parameter stated above. The whole phase space of equations (25*a*) and (25*b*) is the basin of the attraction of the fixed point in the middle marked by a blue dot; the dynamics is drawn as streams. The trajectories starting in the lower part have to pass by a 'desert-like' saddle (with $c_{\mathrm{t}} = 0$) at the left (green dot).

The color in figure 3(f) depicts T_{RR} and the first finding is the splitting of the basin of attraction. The strong stable manifold of the stable node becomes visible as a light beige line due to its low values of T_{RR}, i.e. as very *early* states because $T_{\mathrm{RR}} \to -\infty$. So it is the separatrix for the observed splitting. Also, the expected smooth increase of the return times when distancing (along the trajectories) from the attractor can be observed.

Still, the splitting of the basin of attraction is visible for values of $c_{\mathrm{terrestrial}} < 0.3$, where it is only due to quantitatively different behavior and the visible boundary is actually a rather sharp but still continuous transition. (The latter statement follows right from Suppl. Mat. Theorem 3.6 and Suppl. Mat. Proposition 3.7.) Looking at figure 3(f) one can also see that the boundary becomes more and more fuzzy for even smaller values of $c_{\mathrm{terrestrial}}$, demonstrating that there is really a need for a quantitative analysis.

When applying D to this model (figure 3(d)), the splitting of the basin can be observed again. In contrast to T_{RR}, the strong stable manifold of the stable node is not visible because D can be seen as a (by distance) weighted time and the contributions from the asymptotic part where the difference in the Lyapunov spectrum matters are negligible.

Furthermore, we see a clear linear correlation of both metrics in figure 3(e) because all trajectories starting in the lower part have to pass by at the saddle on the left and spend a long time there.

Both metrics work as *early-warning signals* [14, 29], too. When increasing α, corresponding to the harvest of terrestrial carbon, the system passes through a subcritical pitchfork bifurcation where the saddle becomes stable and the lower-left part of the phase space splits off. The divergences of the two metrics' statistics as seen in figure 4 prove their prebifurcational sensitivity, while other systemic indicators like basin stability [30] do not change (up to numerical fluctuations, see figure 4). Note that in this example, a Lyapunov exponent analysis of the saddle would be able to predict the bifurcation due to the simplicity of the saddle also. However, in case of a more complex saddle, this would become arbitrarily difficult while this numerical estimation would still be possible for both metrics.

4.3. Generator in a power grid

As the next example, we chose the swing equation in equation (26), a basic model describing the dynamics of a single generator connected to a large power grid [31]. It consists of two dynamical variables, the phase θ and angular frequency ω, both in a reference frame rotating at the grid's rated frequency. The parameters of the system correspond to the net power production $P = 1$ (at the node), the capacity of the transmission line $K = 6$ and dampening $\alpha = 0.1$.

$$\dot{\phi} = \omega, \quad \dot{\omega} = 2P - \alpha\omega - 2K \sin\phi. \tag{26}$$

In this form, which is used in electrical engineering [25, 32], it is formally equivalent to a pendulum with constant driving and damping.

The stable fixed point at $\omega^s = 0$, $\phi^s = \arcsin\frac{P}{K}$ describes a state of synchronization. For the chosen set of parameters, the system exhibits another attractor: a limit cycle at larger positive values of ω. For negative values, the two basins of attraction are interleaved. A more detailed introduction and analysis can be found in [25, 31, 33].

Calculating T_{RR} inside the basin of the stable fixed pointed (ω^s, θ^s) yields figure 3(i). There is basically no color change away from the attractor, so we can see that a trajectory barely spends any time in the transient and goes quickly to the attractor. Analogously, figure 3(g) for D leads to the same conclusion as T_{RR}.

Comparing both metrics in figure 3(h) shows that they are closely linked. Note that this time D is presented on a logarithmic scale, so the relation is exponential and what we see here is actually the influence of the linearized part of the system. The accumulation in the upper right corresponds to the initial conditions with lower values of ω. This means, they only go through a very short transient and spend most of their time in the part where the linearization holds.

The white parts in the phase spaces figures 3(g) and (i) correspond to the basin of attraction of the limit cycle. As this means the system is away from synchrony, the generators would usually switch off before reaching it. So we did not include it in the analysis.

4.4. Chaotic Rössler oscillator

Although we have proven the convergence of T_{RR} for fixed points only, we show with the chaotic Rössler system [34, 35] that both metrics are applicable to higher-dimensional and more complex attractors, too. The equations are

$$\dot{x} = -y - z, \quad \dot{y} = x + ay, \quad \dot{z} = b + z(x - c), \tag{27}$$

where x, y and z are the coordinates in state space. While this naming convention is not in line with the rest of the article, it has been chosen as it is standard for these equations.

Figure 3(l) shows a slice of the phase space with the standard parameters $a = 0.2, b = 0.2, c = 5.7$ for T_{RR} and the expected sensitivity to initial conditions for chaos is observed: *early* and *late* trajectories lie closely together and the metric T_{RR} has low spatial correlation.

In contrast, D shows in figure 3(j) surprisingly smooth changes of an embryo-like shape. Because the focus of this article is on transient dynamics a new feature of the chaotic Rössler system is uncovered: while the attractor is chaotic, the basin of attraction is very regular. D focuses on the initial transient and the chaotic asymptotics is filtered out. For comparison, the boundaries of the attractor's projection have been added with dashed red lines in figure 3(j) and depictions of the attractor are in Suppl. Mat. section 2.

Furthermore, T_{RR} can be applied as an early-warning signal in this case, too. In order to demonstrate this, we chose to vary a as it has a crucial influence on the system's dynamics (see the bifurcation diagram in figure 5 (green)). For values of $a < 0.006$ (see [36]) there is only a single stable fixed point. At $a \approx 0.006$ a limit cycle emerges due to a Hopf bifurcation [36]. For $a > 0.11$, several period doublings are observed, finally leading to chaos for $a > 0.155$. Even in the chaotic regime, further bifurcations can be observed.

In figure 5, the standard deviation of the T_{RR} distribution from randomly chosen initial conditions inside the basin of attraction is given. Due to the sensitive dependence on initial conditions, the reference value varies a lot and hence introduce shifts in the distribution that do not describe actual changes in the system's dynamics. To remove this effect, it is crucial to use central moments like the standard deviation.

T_{RR} is strongly sensitive to any qualitative changes in the dynamics of the system, incl. even chaos–chaos transitions. Closely observing figure 5 uncovers that there is a base-line with little fluctuations at $\mathrm{Std}(T_{RR}) \approx 10$ complemented with strong peaks. In the chaotic regime, the peaks correspond directly to qualitative changes. Also, we observe sensible changes during the period-doubling phase and a strong increase before the Hopf bifurcation at $a \approx 0.006$, proving the usefulness as an *early-warning signal*.

The abrupt downward peak at $a \approx 0.11$ is unexpected and more details are needed to clarify it. The other peaks correspond well with the transitions visible in the bifurcation diagram.

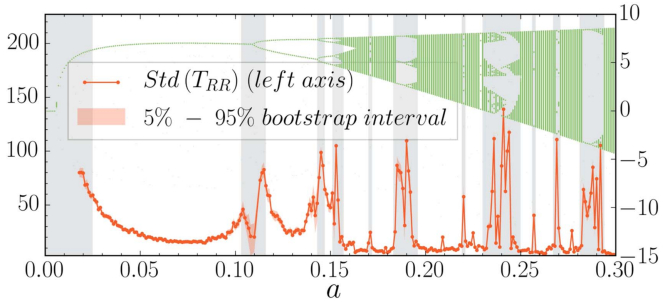

Figure 5. The bifurcation diagram (green) of the Rössler system for varying the parameter a in equation (27) was computed from the local maxima in z of the attractor and T_{RR} (orange) shows a strong sensitivity to these qualitative changes. The gray background is used so the reader can more easily connect the peaks in T_{RR} to the corresponding parts in the bifurcation diagram.

5. Discussion

In order to see how far the two proposed metrics answer the question 'How can we properly quantify the time to reach a system's attractor?' we will go along the four essential problems that have been worked out in section 2 for this discussion: (I) infinite reaching time, (II) physical interpretation (III) discontinuities and (IV) non-invariance.

Area under distance curve (D) has been defined as the cumulative distance to the attractor over time in order to emphasize the idea that a trajectory stays 'far' from the attractor in the transient while being close in the asymptotics. The distance d is not necessarily meant in the mathematical sense [37], but it only needs to approach 0 around the attractor and be 0 on it. In that way, it is possible to choose the appropriate d for different research questions, e.g. asking about costs or damages. Even in these interpretations D is a metric capturing the transient time, because there are only contributions when the trajectory is distant from the attractor, i.e. still in its transient phase. Another point of view is to see D as the time to reach the attractor weighted by the distance.

We understand Problem (I), infinite reaching time, as solved. For hyperbolic attractors and d being a mathematical distance function, the integral in equation (7) does converge. Trajectories approach the attractor exponentially in the asymptotics and the integral over the exponential envelope is finite.

While this covers most systems relevant for real-world applications, in some very specific cases, D might be infinite. The asymptotic tail of the integral might not converge, i.e. the trajectory does not approach it 'fast enough'. This means, either this is the wanted result or d has not been chosen appropriately. In the first case, it could be for example that D was computed for an initial condition that is not economically feasible, so the cost diverges. Furthermore, this would imply that even though the attractor is systemically stable, it is not economically feasible to cope with small perturbations.

From a technical perspective a divergence in D can be understood as indicating that d has not been chosen matching to the system. E.g. using the Euclidean distance and $\dot{x} = -\frac{1}{2}x^3$ where the solutions are $\varphi(t,\, x) = \mathrm{sign}(x)\frac{1}{\sqrt{t+|\,x\,|^{-2}}}$, D does not converge. Another example is to take a linear system $\dot{x} = -x$ with $x < 1$. Using $d(x,\, \{0\}) = -\frac{1}{\ln|\,x\,|}$ with $d(0,\, \{0\}) = 0$ gives $D \to \infty$.

This can usually be solved by choosing an appropriate d. E.g. choosing for $\dot{x} = -\frac{1}{2}x^3$ using $d(x,\, \{0\}) = \exp(-|x|^{-1})$ and $d(0,\, \{0\}) = 0$ gives finite values for D.

Problem (II) is solved because there is no direct parameter. Still, as there is the indirect dependence on d a discussion is necessary and given in comparison to the first and last entry time to an ϵ-environment $T_{\mathrm{F}}^{\epsilon}(x)$ and $T_{\mathrm{L}}^{\epsilon}(x)$ respectively. For them, a small change in ϵ will have a huge impact on the measured times because for $\epsilon \to 0$ both values go to infinity. Furthermore, if one would locally change the way how the distance to the attractor is measured, the values for $T_{\mathrm{F}}^{\epsilon}(x)$ and $T_{\mathrm{L}}^{\epsilon}(x)$ would change drastically, too.

Because D is defined as the cumulative d over time, a local change in d will have only minor effects on the exact value, so even estimated functions for d with some uncertainty can be used.

Problem (III), discontinuities, have been avoided in D by using the integral representation. Hence the function is even differentiable along the flow (see equation (8)).

We see Problem (IV), non-invariance, as solved, if d has been chosen with some meaning, e.g. economic damages. Then one can simply represent the economic damage function in the changed coordinates, because the meaning is independent of the coordinates. This reasoning is not mathematical but context-dependent. From a purely mathematical point of view, if d is just any distance function, generally the result is not invariant under

change of coordinates as it depends on geometric features of the system. But as we want to go in the direction of real-world systems, a model-specific choice of d is compulsory anyway.

Regularized reaching time (T_{RR}) has been defined as the difference in time to approach the attractor.

Problem (I), infinite reaching time, does not appear for all states in the basin of attraction except the attractor and the strong stable manifold. In case of the attractor, the trajectory will stay on it while trajectories from other points approach the attractor only; and, by definition, points on the strong stable manifold approach the attractor a lot faster, also in the asymptotics. So these infinities are actually reasonable results. Also, both are usually of a smaller dimension than the state space. Hence coincidently being there is unlikely, and these cases are rather irrelevant for real-world applications.

Problems (II) and (III) are intrinsically solved by avoiding parameters. The necessary choice of x^{ref} introduces a constant shift only while not changing the structure of the function. When looking at central moments of T_{RR}, i.e. ones invariant under shifts, this dependence on the choice of x^{ref} disappears completely as they would only shift the mean. So an analysis by changing the system's parameters is possible. This has been done in the examples for the global carbon cycle and Rössler system and T_{RR} has been confirmed as an early-warning signal. This analysis can be seen as a systemic approach to the concept of *critical slowing down* (CSD) [14, 29, 38] after a shock, i.e. an instantaneous and non-infinitesimal perturbation, uncovering prebifurcational changes in the transient behavior. In contrast, CSD is usually done with (local) noise only. The usage of shocks has been developed in the context of Basin stability [30, 31] and its extensions [23, 39–42].

Problem (IV), non-invariance, is proven to be solved for hyperbolic fixed points. In case of more complex attractors, we can currently only define an estimation of T_{RR} which depends on geometric properties. So invariance might not be given and more research due in that direction. An important step in that direction has been done by writing down the properties of Δ which imply that the necessary way of measuring how a trajectory approaches might not be local (except fixed-points). The used pragmatic choice of $\Delta = D$ demonstrates this as it basically says that the remainder of the trajectory should have an ϵ small value of D only.

An assumption that has been made during the proof of invariance of T_{RR} for hyperbolic fixed points is: the eigenvalue of the RHS's Jacobian with the largest real-part is either unique and with multiplicity 1 or there are two that are complex conjugated to each other. However, this condition is not really constraining because we assume most real world systems fulfill it.

Comparison. The metrics have been applied to several examples and we will discuss a comparison between both metrics here. They are depicted in figures 3(b), (e), (h) and (k) and show different relations, as stated in the figures. The exponential lower bound and the exponential relation for the linear system and swing equation respectively come mostly from the asymptotic behavior, in particular as the linear system does (by definition) not have any nonlinearities. Still the relations are different as the asymptotic behavior differs slightly, too, one being a node the other a focus. This shows that even though we clearly focus on the transient, it is actually important to be aware of the asymptotic behavior, too. And one cannot analyze the former without knowing about the latter.

In contrast, the linear relation for the global carbon cycle really points to the transient behavior only. It is due to the states passing by the 'desert-like' saddle. Finally, there seems to be no clear relation between both metrics for the Rössler example, pointing to the chaotic behavior. Still, both metrics have separately been useful, D demonstrating the smoothness of the basin of attraction and the standard deviation of T_{RR} being sensitive to qualitative changes of the system.

Other methods. When developing this research on measuring times to approach the attractor, we had the impression that there are two more common ideas, additionally to the first and last entry time. We do not intend to have a complete overview of all methods but would like to discuss these two shortly here. This part refers to a general system in the sense of (1).

The first idea is to develop metrics based on characteristic times. These are usually defined as the time until a quantity is reduced to $1/\mathrm{e}$ of its original value [43]. This quantity could be a distance to the attractor or a coordinate. From this definition it already follows that they are subject to problem (IV). Also, even if the quantity is at $1/\mathrm{e}$ of its initial value, the trajectory might still be far away from the attractor and in its transient dynamics. Lastly, taking a one-dimensional linear system and assuming the quantity is the coordinate, the characteristic time is constant for all initial conditions. This is counter-intuitive when thinking about a time to approach the attractor.

The second idea for general systems is to use Lyapunov exponents [44]. They have units of inverse time and are invariant under changes of coordinates. However, they are actually a property of the attractor. So they do not capture the transient but only the asymptotics closely around and at the attractor.

6. Summary and outlook

In this article, we have treated the question: 'How can we properly quantify the time to reach a system's attractor?'

First, we have worked out the four essential problems of quantifying the timing of transients in order to develop two new metrics, *area under distance curve* D and *regularized reaching time* T_{RR}. As the focus of this work is meant to be on making a first step to real-world systems, we have applied the metrics numerically to four chosen examples systems, observing different features. Finally, we have discussed in detail how far the metrics treat the four essential problems.

With this approach, interesting features of the examples have been uncovered. Using the global carbon cycle, we have demonstrated the importance of the transient analysis, as the desert state is only a saddle but nevertheless passing by there would lead to an extinction of humanity. The splitting of the basin of attraction is partially due to the strong stable manifold of the attractor but it continues for lower values of $c_{\mathrm{terrestrial}}$ where it is only due to quantitatively different behavior demonstrating the need for quantitative methods. Particularly interesting is how the (central) statistics of our metrics are a systemic approach to the concept of CSD leading to an interpretation as early-warning signals, which we have demonstrated also. The independence of the choice of reference points has been achieved by the usage of central moments. In case of the generator in a power grid, most of the relevant dynamics seems to be dominated by the linearization of the equations around the focus.

In order to prove the applicability to more complex dynamics, we have used our metrics on the Rössler system, too, and found the smoothness of the attractor's basin with D. As the attractor itself is chaotic, this smoothness is surprising. T_{RR} reacts strongly to the sensitivity to initial conditions of the chaotic system and one might want to ask whether there is a relation to winding numbers when approaching the attractor. Still, its worth is displayed when varying the a parameter. This parameter has strong influence on the Rössler system's dynamics and T_{RR} reacts strongly to the different bifurcations and even the chaos–chaos transitions, proving again its worth as *early-warning-signal*.

We have not performed any comparative analysis with the mentioned first- and last-entry-time approaches because these behave inconsistently and their quantitative results are arbitrary, as discussed at length in section 2.

The detailed discussion on the two metrics have showed that, while they do treat the four essential problems, they do not fully solve them and further investigation is needed. Also, they come from two very different basic ideas so the comparison showed that they really measure independent features but can improve the understanding of a system by combining them. For both metrics, we have showed that they are Lyapunov-functions. While some properties have already been used in the article, these definitions in terms of orbital derivatives may be a rich groundwork for the next steps.

Four directions of immediate future research are due:

(1) Working on the definition of T_{RR} using the Lyapunov function properties. This step is crucial in order to further the understanding of transient analysis and needs to take the attractor into account as well. Hence, the analysis of more complex attractors and basin shapes, e.g. riddled basins, is part of this.

(2) Applying the current definition of the metrics, in particular using the estimation of T_{RR} with $\Delta = D$, to understand the implications and the precise use cases better. Furthermore, their relations to topological structures, e.g. in complex networks [45], need to be worked out in detail. This part, even though complementary, should be done in accordance with the results in (1).

(3) On the numerical side, it is important to introduce more sophisticated methods of Lyapunov function estimations, where a starting point is the work by Giesl and Hafstein [46]. The curse of dimensionality is going to be a problem for network systems, hence methods for estimation of these metrics' statistics in such kinds of systems induce a need for developing specific algorithms.

(4) Comparison of the timing of transients in model output and observation data as the new observable time is now available.

Acknowledgments

The authors thank the anonymous referees for their detailed and constructive feedback.

This paper was developed within the scope of the IRTG 1740/TRP 2011/50151-0, funded by the DFG/FAPESP. This work was conducted in the framework of PIKs flagship project on coevolutionary pathways (COPAN). The authors thank CoNDyNet (FKZ 03SF0472A) for their cooperation. The authors gratefully acknowledge the European Regional Development Fund (ERDF), the German Federal Ministry of Education and Research and the Land Brandenburg for supporting this project by providing resources on the high performance computer system at the Potsdam Institute for Climate Impact Research. The authors thank the developers of the used software: Python [47], Numerical Python [48] and Scientific Python [49].

The authors thank Sabine Auer, Karsten Bölts, Catrin Ciemer, Jonathan Donges, Reik Donner, Jasper Franke, Frank Hellmann, Jakob Kolb, Chiranjit Mitra, Finn Müller-Hansen, Jan Nitzbon, Anton Plietzsch Stefan Ruschel, Tiago Pereira da Silva, Francisco A Rodrigues, Paul Schultz, and Lyubov Tupikina for helpful discussions and comments.

References

[1] Barkema G, Marko J and De Boer J 1994 *Europhys. Lett.* **26** 653
[2] Gaulin B D and Spooner S 1987 *Phys. Rev. Lett.* **58** 668–71
[3] Chou Y C and Goldburg W I 1979 *Phys. Rev.* A **20** 2105
[4] Fiutak J and Mizerski J 1980 *Z. Phys.* B **39** 347–52
[5] Tang C, Telle J and Ghizoni C 1975 *Appl. Phys. Lett.* **26** 534–7
[6] Krapivsky P L, Redner S and Ben-Naim E 2010 *A Kinetic View of Statistical Physics* (Cambridge: Cambridge University Press)
[7] Castellano C, Fortunato S and Loreto V 2009 *Rev. Mod. Phys.* **81** 591
[8] Chowdhury D, Santen L and Schadschneider A 2000 *Phys. Rep.* **329** 199–329
[9] Hastings A 2004 *Trends. Ecol. Evol.* **19** 39–45
[10] Van Geest G, Coops H, Scheffer M and van Nes E 2007 *Ecosystems* **10** 37–47
[11] Schaffer W M, Kendall B, Tidd C W and Olsen L F 1993 *Math. Med. Biol.* **10** 227–47
[12] Fisher R S, Boas W v E, Blume W, Elger C, Genton P, Lee P and Engel J 2005 *Epilepsia* **46** 470–2
[13] Fisher F M 1989 *Disequilibrium Foundations of Equilibrium Economics* (Cambridge: Cambridge University Press)
[14] Lenton T M 2011 *Nat. Clim. Change* **1** 201–9
[15] Anderies J M, Carpenter S R, Steffen W and Rockström J 2013 *Environ. Res. Lett.* **8** 044048
[16] Heitzig J, Kittel T, Donges J F and Molkenthin N 2016 *Earth Syst. Dyn.* **7** 21–50
[17] Crutzen P J 2002 *Nature* **415** 23
[18] Steffen W *et al* 2011 *Ambio* **40** 739–61
[19] Waters C N *et al* 2016 *Science* **351** aad2622
[20] Steffen W, Broadgate W, Deutsch L, Gaffney O and Ludwig C 2015 *Anthropocene Rev.* **2** 81–98
[21] Rockström J *et al* 2009 *Ecol. Soc.* **14** 32
[22] Steffen W *et al* 2015 *Science* **347** 1259855–1–10
[23] van Kan A, Jegminat J, Donges J F and Kurths J 2016 *Phys. Rev.* E **93** 042205–1–7
[24] Kuznetsov Y A 2013 *Elements of Applied Bifurcation Theory* vol 112 (New York: Springer)
[25] Weckesser T, Jóhannsson H and Østergaard J 2013 Impact of model detail of synchronous machines on real-time transient stability assessment *Bulk Power System Dynamics and Control-IX Optimization, Security and Control of the Emerging Power Grid, (IREP) Symp.* (IEEE) pp 1–9
[26] Bhat S P and Bernstein D S 2000 *SIAM J. Control Optim.* **38** 751–66
[27] Giesl P 2007 *Construction of Global Lyapunov Functions using Radial Basis Functions* (Berlin: Springer)
[28] Heck V, Donges J F and Lucht W 2016 *Earth Syst. Dyn.* **7** 783–96
[29] Scheffer M, Bascompte J, Brock W A, Brovkin V, Carpenter S R, Dakos V, Held H, Van Nes E H, Rietkerk M and Sugihara G 2009 *Nature* **461** 53–9
[30] Menck P J, Heitzig J, Marwan N and Kurths J 2013 *Nat. Phys.* **9** 89–92
[31] Menck P J, Heitzig J, Kurths J and Joachim S H 2014 *Nat. Commun.* **5** 1–8
[32] Yuan Y, Kubokawa J and Sasaki H 2003 *IEEE Trans. Power Syst.* **18** 1094–102
[33] Schultz P, Heitzig J and Kurths J 2014 *New J. Phys.* **16** 125001
[34] Rössler O E 1976 *Phys. Lett.* A **57** 397–8
[35] Zgliczynski P 1997 *Nonlinearity* **10** 243
[36] Barrio R, Blesa F, Dena A and Serrano S 2011 *Comput. Math. Appl.* **62** 4140–50
[37] Heitzig J 2002 Mappings between distance sets or spaces *PhD Thesis* Universität Hannover
[38] Scheffer M *et al* 2012 *Science* **338** 344–8
[39] Klinshov V V, Nekorkin V I and Kurths J 2015 *New J. Phys.* **18** 013004
[40] Hellmann F, Schultz P, Grabow C, Heitzig J and Kurths J 2016 *Sci. Rep.* **6** 1–12
[41] Mitra C, Kurths J and Donner R V 2015 *Sci. Rep.* **5** 1–10
[42] Mitra C, Choudhary A, Sinha S, Kurths J and Donner R V 2017 *Phys. Rev.* E **95** 032317
[43] Clark M M 2011 *Transport Modeling for Environmental Engineers and Scientists* (New York: Wiley)
[44] Cvitanović P, Artuso R, Mainieri R, Tanner G and Vattay G 2016 *Chaos: Classical and Quantum* (Copenhagen: Niels Bohr Inst.)
[45] Havlin S *et al* 2012 *Eur. Phys. J. Spec. Top.* **214** 273–93
[46] Giesl P and Hafstein S 2015 *Discrete Continuous Dyn. Syst.* B **20** 2291–331
[47] Van Rossum G and Drake F L Jr 1995 Python Reference Manual (Centrum voor Wiskunde en Informatica Amsterdam)
[48] Ascher D *et al* 2001 *Numerical Python*
[49] Jones E *et al* 2001 SciPy: Open Source Scientific Tools for Python