



HAL
open science

Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space

Tim Kittel, Finn Müller-Hansen, Rebekka Koch, Jobst Heitzig, Guillaume Deffuant, Jean-Denis Mathias, Jürgen Kurths

► **To cite this version:**

Tim Kittel, Finn Müller-Hansen, Rebekka Koch, Jobst Heitzig, Guillaume Deffuant, et al.. Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space. 2020. hal-02961471

HAL Id: hal-02961471

<https://hal.science/hal-02961471>

Preprint submitted on 8 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Operationalization of Topology of Sustainable Management to Estimate Qualitatively Different Regions in State Space

Tim Kittel¹, Finn Müller-Hansen^{1,2,a}, Rebekka Koch¹, Jobst Heitzig¹, Guillaume Deffuant³, Jean-Denis Mathias³, and Jürgen Kurths

¹ Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, FutureLab on Game Theory and Networks of Interacting Agents, P.O. Box 60 12 03, D-14412 Potsdam, Germany

² Mercator Research Institute on Global Commons and Climate Change (MCC), EUREF Campus 19, Torgauer Straße 12-15, 10829 Berlin, Germany

³ Irstea, UR LISC Laboratoire d'ingénierie des systèmes complexes, 24 avenue des Landais, BP 20085 Aubière, F-63172, France

Abstract To apply the framework *Topology of Sustainable Management* (TSM) by Heitzig et al. (2016) to dynamical models, we connect it to *viability theory* via a variant definition of the former. This enables us to use the Saint-Pierre algorithm to estimate the main partition of TSM. Furthermore, we present an extension of the algorithm to compute implicitly defined capture basins, regions of state space as defined in viability theory. We use a low-complexity model coupling environmental and socio-economic dynamics to demonstrate the applicability of this approach. With this example, we highlight how common problems in estimations of these regions like (i) an unbounded state space and (ii) highly varying time scales can be circumvented by introducing appropriate coordinate transformations. The article thus shows how algorithmic approaches from viability theory can be used to get a better understanding of the state space of manageable dynamical systems.

[TODO Jobst: incorporate newer refs in intro, ref to Felix, go through earlier reviewer comments and response.]

[TODO J-D: update refs to viability papers?]

[TODO all: update contact details, thanks and funding info in acknowledgements]

1 Introduction

When charting the pathways to a sustainable future for humanity, different constraints we should abide by have been proposed. The idea of *Planetary Boundaries* which has been developed by Rockström et al. [58] and extended by Steffen et al. [66] is to define a set of biophysical boundaries chosen in order to ensure human development if they

^a Correspondence to Finn Müller-Hansen, e-mail: mueller-hansen@mcc-berlin.net

are respected. While they provide safety by defining a *safe operating space*, Raworth [56] pointed out that a human future needs to concern justness as well, hence she introduced the *Social Foundations*. Combined, they define the *safe and just operating space*, often referred to as SAJOS.

While there is much research on refining the current definitions, e.g. for freshwater [27] and phosphorus [17], extending them, e.g. terrestrial net primary plant production [62], and downscaling them [28], another focus has emerged: their interaction due to the system’s intrinsic dynamics [2; 29].

On a formal level, many similar problems on different scales can be found, e.g. quotas for fishing [22; 19], water management [57] or language competition [20].

One approach to treat these kind of problems was developed by Heitzig et al. [30] with the framework on *Topology of Sustainable Management* (TSM). It has already shed a new light on the implications of boundaries when taking the dynamics and possible management into account. Complex structures in state space, that we call regions, corresponding to a hierarchy of safety levels, may occur naturally from this concept. These regions in state space are a *qualitative* classification and differ in how secure they are and how much management they need to either stay within the desirable region or reach it. Combined, they give a partition that we want to identify. As pointed out in [30], it might be important to have this kind of qualitative analysis before performing a quantitative optimization. They have carefully chosen the terms “default dynamics” and “management” in order to emphasize that usually only a slight influence on the dynamics might be possible but not a complete control, e.g. in the global climate system. On a formal level however, a relation to *viability theory* (VT), a subfield of control theory, has already been sketched out. The analysis of the models in [30] was possible to do without an automated classification as the model systems were all two-dimensional. The need to apply this framework to more complex models leads directly to the question governing this article: “How to properly operationalize the TSM-framework? And how to automatize the identification of the regions in state space?”

Viability theory was developed by Aubin and his collaborators [8; 4; 9; 7] to address the problem of maintaining a dynamical system within a set of desirable states, e.g. in our case delimited by planetary boundaries. The main concept, the *viability kernel*, is defined as the set of states from where it is possible to keep its trajectory within the desirable set indefinitely. Therefore, it is possible to avoid the transgression of the boundaries. The second main concept is the *capture basin*. It is defined with respect to a *target set*, i.e. a chosen set of states that one wants to reach. Then, the capture basin is the set of states from which the target set can be reached. Hence this concept is very close to the one of reachability in control theory. Viability Theory has been applied in many domains, e.g. economics [5], fisheries [22; 19], wireless sensor node [37], forest [44], language competition [20], social networks [46], sustainability management [25; 59; 60], resilience modeling [43; 23]. Mathias et al. [45] recently used viability theory for showing our rapidly shrinking capacity to comply with the planetary boundaries on climate change. Several algorithms for the estimation of viability kernels have been developed, among them the Saint-Pierre algorithm [63] with extension to machine-learning-based classification methods [24], reachability based viability [42] or optimal-control-problem-based ideas [13; 25; 59].

In this article, we will present a variant definition of TSM based on the main ideas of VT, which differs from the original definition by using control theory’s notion of reachability instead of “safe reachability” (as defined in [30]). Also, we require a target set be reached in finite but arbitrary large time, instead of infinite time because it simplifies the computation and is more realistic. But for computationally supported estimations of specific models these differences are usually not relevant. This new approach to TSM allows us to access the tools from the established field of VT, in

particular the Saint-Pierre Algorithm, leading to an operationalization of the TSM framework.

Furthermore, we develop a *nonlinear, local time homogenization* solving the problem of vastly differing time scales when estimating viability kernels and capture basins using the Saint-Pierre algorithm. It fully homogenizes the time scale of dynamics while keeping the major properties of the system invariant, e.g. fixed points, other attractors, μ -Lipschitz continuity, C^∞ , etc. if they were present in the original system. Even though the approach is straightforward, it is essential for the analysis of the example model used in this article.

Commonly, the state space of a model can be unbounded in at least one variable, e.g. economic production, but the Saint-Pierre algorithm is for bounded sets only. To tackle the problem of unbounded state spaces, we propose a nonlinear coordinate transformation to a bounded state space. In particular, this transformation has the property that one can choose a scale to be “resolved best”.

Certain regions, the so-called *eddies*, of the TSM-partition in state space are shown to be naturally defined as implicit target sets. To identify these numerically, we also have developed an extension of the Saint-Pierre algorithm.

Having a definition in terms of viability theory enables us to operationalize the TSM-framework. We demonstrate it with a three dimensional example model coupling environmental and socio-economic dynamics [65]. Similar to other low-complexity models [71; 10; 40; 33; 2; 14; 55; 36] it focuses on the interaction between the *ecosphere* and the *anthroposphere* [38]. The former is represented by an atmospheric carbon stock and the latter by the (global) economic output and a knowledge stock on renewable energy production, possibly leading to a technological change. The model’s complexity has been held low so it can be used as an appropriate first example where the TSM partition is estimated automatically. Two common policies, we call managements, can be implemented in a conceptual manner: (i) *low growth* and (ii) climate mitigation by inducing an *energy transformation*.

In such a model, there are two naturally arising boundaries. The first one is the *planetary boundary on climate change* (PB-CC) [58; 66] which limits the amount of atmospheric carbon. The second one is a wealth-related social foundation prescribing a threshold below that the yearly economic output should not fall.

The article is structured as follows. First we give details on software availability in Section 2. In Section 3, we recall the necessary notions from viability theory in order to introduce the variant definition of TSM in Section 4. In Section 5, we show our means to deal with an unbounded state space and then introduce the nonlinear, local time-homogenization. Next, we shortly recall the Saint-Pierre algorithm and present its extension for implicitly defined target sets. In Section 6 we introduce the example model with its different managements and analyze it using the tools developed before. Finally, we close with a summary and an outlook for necessary future work in Section 7.

2 Software availability

3 Viability theory

In this section, we shortly introduce the notions of *viability theory* (VT) needed in this article, following [4; 6]. We start with a time-continuous ($t \in \mathbb{T} := \mathbb{R}_+$) dynamical control system

$$\dot{x} = f(x, u) \tag{1}$$

with $x \in \mathcal{X} = \mathbb{R}^n$, the state space, and $u \in \mathcal{U}$, the set of all possible values for the control parameter u . We call f the right-hand side function (RHS). Note that no dependency of \mathcal{U} on x is assumed for notational simplicity but the idea could be extended in a straightforward manner.

A function $q_{x_0}: \mathbb{T} \rightarrow \mathcal{X}$ is called a *solution* for an (arbitrary) initial condition $x_0 \in \mathcal{X}$ if there exists a measurable function of time $\pi: \mathbb{T} \rightarrow \mathcal{U}$, called *policy*, such that for any time $t \in \mathbb{T}$ the condition $\frac{dq}{dt}(t) = f(q(t), \pi(t))$ is fulfilled and $q_{x_0}(0) = x_0$.

A *viable set* \mathcal{V} of a *constraint set* $\mathcal{Y} \subseteq \mathcal{X}$ is then defined as a set of initial conditions x for which there exists a *viable solution* q_x that stays within \mathcal{Y} forever:

$$\forall x \in \mathcal{V} \exists q_x \forall t \in \mathbb{T}: q_x(t) \in \mathcal{Y}. \quad (2)$$

The largest viable set of \mathcal{Y} is called the *viability kernel* $\text{Viab}_{\mathcal{U}}(\mathcal{Y})$. The set of possible controls \mathcal{U} is given as a subscript as we will distinguish different controls later. If one assumes only some constant control $u_f \in \mathcal{U}$, i.e. any solution has a policy $\pi(t) = u_f$, the corresponding viability kernel is called a *viability niche*

$$\text{VN}_{u_f}(\mathcal{Y}) := \text{Viab}_{\{u_f\}}(\mathcal{Y}). \quad (3)$$

The *capture basin* of a *target set* $\mathcal{Z} \subseteq \mathcal{X}$ is the subset of the state space in \mathcal{X} for which there exists a solution of Eq. (1) that reaches \mathcal{Z} in finite time

$$\text{Capt}_{\mathcal{U}}^{\mathcal{Y}}(\mathcal{Z}) = \{x_0 \in \mathcal{X} \mid \exists \text{ solution } q_{x_0} \exists T \in \mathbb{T}: \\ q_{x_0}(T) \in \mathcal{Z} \wedge \forall t < T: q_{x_0}(t) \in \mathcal{Y}\}. \quad (4)$$

In case no constraint set is given, the whole state space is assumed, i.e. $\text{Capt}_{\mathcal{U}}(\mathcal{Z}) := \text{Capt}_{\mathcal{U}}^{\mathcal{X}}(\mathcal{Z})$.

4 A variant of the topology of sustainable management based on viability theory

In this section, we present a variant definition of the TSM-partition based on VT. From the point of view of VT it might be seen as an extension of the theory when considering a qualitative distinction of controls into default and management, leading to a partition of the state space and multiple dilemmas, whereof we introduce only the “lake dilemma”.

We elucidate some central notions using a metaphorical example of ducklings in Fig. 1. The water region represents the state space and the streamlines represent the dynamics. The ducklings can either swim with the flow (default dynamics) or struggle and swim against it (management). However, it is not possible to swim up a waterfall once they have dropped down. The desirable region, a safe environment providing enough food and nesting places for the ducklings, is on the left and the undesirable region full of predators on the right. In the following, we will introduce the general concept for each region first and then explain where it comes up in the ducklings example.

The definition is based on a general control system as Eq. (1) where we additionally require that a *default control* $u_0 \in \mathcal{U}$ is separated out from *all possible controls* \mathcal{U} . $\mathcal{U}_m = \mathcal{U} - \{u_0\}$ are the *manageable controls*.¹ Hence we call $f(x, u_0)$ the *default flow* giving rise to the *default dynamics* and the dynamics corresponding to the manageable controls is called *management options*. Furthermore, we require a division of the state space into *desirable* $\mathcal{X}^+ \subseteq \mathcal{X}$ and *undesirable* $\mathcal{X}^- := \mathcal{X} - \mathcal{X}^+$.

¹ Here and in the following we use the lax difference and union notation with “−” and “+” for sets.

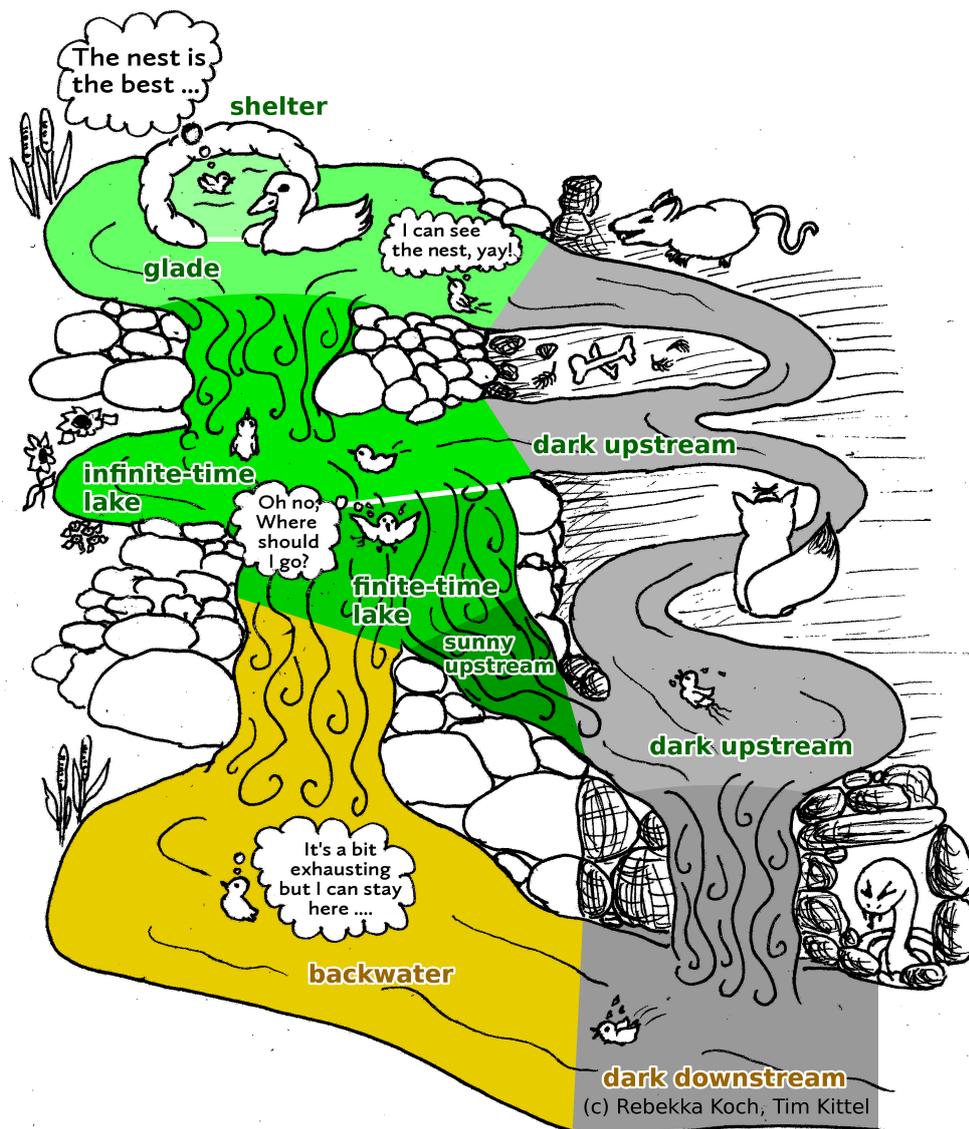


Figure 1: Metaphorical example for the concepts introduced in Section 4, which is a variant definition of the Topology of Sustainable Management (TSM) framework developed by Heitzig et al. [30]. It depicts a river flowing downwards with ducklings on it that may swim through the desirable (left) and undesirable parts (right). In the nest, which corresponds to the *shelter*, the ducklings do not have to swim and can stay there forever without effort. Outside they will slowly drift down or can swim, i.e. *manage*, against the downward-flowing stream, if the stream lines are not curly, but long. In areas with curly stream lines, the waterfalls, the flow is so strong that the ducklings move with the flow anyway. The example gives rise to a number of qualitatively different regions from the TSM-framework.

As already sketched in [30], when approaching the idea of TSM from VT's point of view, we can find it on two basic quantities, the *shelters* \mathcal{S} and the *manageable region* \mathcal{M}

$$\mathcal{S} := \text{VN}_{u_0}(\mathcal{X}^+) \quad (5)$$

$$\mathcal{M} := \text{Viab}_{\mathcal{U}}(\mathcal{X}^+). \quad (6)$$

Shelters are the viability niche of \mathcal{X}^+ with respect to the default control u_0 only, meaning the system will stay in the desirable region forever *without management*, thus being the safest regions in the state space. For the ducklings the shelter is the nest, a place in which they can stay in a safe environment forever without swimming against the stream.

Inside the manageable region, one can also stay in the desirable region forever, but may need *all* possible dynamics. In the case of the ducklings-model it corresponds to the waters in the desirable region in Fig. 1 (including the shelter). This is a special case only as we did not want to make the picture too complex and we refer the reader to the detailed explanations in [30] for the full framework.

The set $\text{Capt}_{\mathcal{U}}(\mathcal{M})$ from where the manageable region can be reached is naturally divided into the *upstream*

$$\mathcal{U} := \text{Capt}_{\mathcal{U}}(\mathcal{S}) \quad (7)$$

from where the shelter is reachable and the rest, called *downstream*

$$\mathcal{D} := \text{Capt}_{\mathcal{U}}(\mathcal{M}) - \mathcal{U}. \quad (8)$$

As both sets, the upstream and the downstream, may have qualitatively very different dynamics inside, we introduce a finer partition. From the *glades*

$$\mathcal{G} = \text{Capt}_{\mathcal{U}}^{\mathcal{X}^+}(\mathcal{S}) - \mathcal{S} \quad (9)$$

the shelter can be reached through the sun. In our ducklings example, a glade is formed by the water around the nest. There, the ducklings need to swim against the flow to get back inside the nest and meanwhile they stay inside the desirable region.

Another region inside the *upstream* are the *lakes*

$$\mathcal{L} = \mathcal{U} \cap \mathcal{M} - \mathcal{S} - \mathcal{G}. \quad (10)$$

There, the shelters are reachable through the undesirable region only. However, one can decide to stay within the desirable region forever but will always need the management options. This leads to a qualitative choice, the *lake dilemma*: Inside of a lake one has to choose between eventual safety and uninterrupted desirability. Deciding for the first requires crossing the undesired region with all its consequences. But deciding for the second implies that there will always be a need for management options.

Furthermore, there are two kinds of lakes, *time-limited lakes* \mathcal{L}_l and *time-unlimited lakes* \mathcal{L}_u

$$\mathcal{L}_u = \text{Viab}_{\mathcal{U}}(\mathcal{L}), \quad (11)$$

$$\mathcal{L}_l = \mathcal{L} - \mathcal{L}_u. \quad (12)$$

Because time-unlimited lakes are the viability kernel of the lakes, there is no time pressure for a decision on the dilemma. In contrast, there is a fixed deadline in time-limited lakes. This leads to a stronger form of the dilemma called the *pressing lake dilemma*. Within the metaphorical ducklings model in Fig. 1 the time-unlimited lake is

the waterfall starting from the glade and the subsequent calmer waters. The ducklings can swim in the calmer waters as long as they want to. The time-limited lake is the following waterfall that splits into two streams. Here, the ducklings drop down the waterfall and they have only a moment to decide for the left or the right.

The rest of the upstream is split into the *remaining sunny* and the *dark upstream*

$$\mathcal{U}^{(+)/-} = (\mathcal{U} - \mathcal{M}) \cap \mathcal{X}^{+/-} \quad (13)$$

depending on whether the actual state is inside the desirable or undesirable region. Within the ducklings model, both, the remaining sunny and the dark upstream, are present.

The *Backwaters*

$$\mathcal{W} = \mathcal{M} - \mathcal{U} \quad (14)$$

belong to the manageable region as well as to the downstream. It is possible to stay in the desirable region by managing but the shelters are not reachable. The calm waters in the left lower corner of Fig. 1 is a backwater because the ducklings can swim against the stream and stay inside the backwater but due to the waterfalls, they cannot reach the nest.

Analogously to the upstream, the rest of the downstream is divided into the *remaining sunny* and the *dark downstream*

$$\mathcal{D}^{(+)/-} = (\text{Capt}_{\mathcal{U}}(\mathcal{W}) - \mathcal{U} - \mathcal{W}) \cap \mathcal{X}^{+/-}. \quad (15)$$

In Fig. 1, the part of the waters that belongs to the undesirable region of the downstream is the dark downstream.

If the desirable region can be reached over and over again one is inside the *eddies* \mathcal{E} that are divided into *sunny eddies* \mathcal{E}^+ and *dark eddies* \mathcal{E}^- . The metaphorical image behind the naming is that of a circular flow where one part is in the desirable and the other part in the undesirable region. They are the maximal pair of sets fulfilling

$$\mathcal{E}^{+/-} \subseteq \mathcal{X}^{+/-} - \mathcal{U} - \mathcal{D}, \quad (16a)$$

$$\mathcal{E}^{+/-} \subseteq \text{Capt}_{\mathcal{U}}(\mathcal{E}^{-/+}), \quad (16b)$$

$$\mathcal{E} = \mathcal{E}^+ + \mathcal{E}^-. \quad (16c)$$

The worst regions are *trenches*

$$\Theta = \mathcal{X} - \text{Capt}_{\mathcal{U}}(\mathcal{X}^+) \quad (17)$$

because once inside one cannot reach the desirable region ever again. Inside the *abysses*

$$\mathcal{Y} = \mathcal{X} - \mathcal{U} - \mathcal{D} - \mathcal{E} - \Theta \quad (18)$$

one can reach the desirable region a finite number of times only, and again it is distinguished between sunny and dark abysses $\mathcal{Y}^{+/-} = \mathcal{Y} \cap \mathcal{X}^{+/-}$. This completes the main partition of the TSM framework.

5 Estimation

In order to estimate the TSM-partition for a chosen model, several ingredients might be necessary. The Saint-Pierre algorithm, that we want to use, is based on finitely

discretizing the state space and then using local linear approximations of the dynamics. Hence, it is applicable to bounded state spaces only. In case the relevant part of the state space is unbounded, we need to map it to a bounded space first. Also, vastly differing time scales might be problematic for the linear approximations, so there is a need to homogenize the time scales.

Then, we will sketch the Saint-Pierre algorithm and show how it can even be used to estimate implicitly defined capture basins, e.g. the eddies of the TSM-partition.

5.1 Dealing with an unbounded state space

There are multiple ways to map an unbounded state space to a bounded one, depending on the specific need for the system. In case of the example system analyzed later, each coordinate is bounded from below and unbounded from above. This is rather common in socio-economic models, in particular due to continuous economic growth. Hence, we propose a solution that maps each coordinate separately.

We assume a general dynamic system given by a set of ordinary differential equations

$$\dot{x} = f(x) \quad (19)$$

with $x \in \mathbb{R}_{\geq 0}^n$. In contrast to Eq. (1), there is no control parameter, because the dependence on the control parameter is irrelevant here and the mapping can be done for general ordinary differential equations. We imply an extension to control systems by considering fixed controls. Then we propose the coordinate transformation

$$\begin{aligned} \Phi : \mathbb{R}_{\geq 0}^n &\longrightarrow [0, 1]^n \\ (x_i) &\longmapsto \left(\frac{x_i}{x_{i,mid} + x_i} \right), \end{aligned} \quad (20)$$

where $x_{i,mid} \in \mathbb{R}_{\geq 0}$ are parameters. Applying this transformation on the dynamics leads to a new set of ordinary differential equations

$$\dot{y} = F(y) := ((D\Phi \cdot f) \circ \Phi^{-1})(y) \quad (21)$$

$$= \frac{(1 - y_i)^2}{x_{i,mid}} f \left(\frac{y_i x_{i,mid}}{1 - y_i} \right) \quad (22)$$

for $y \in [0, 1]^n$, i.e. inside a bounded space, where $D\Phi$ is the Jacobian of Φ and \circ is the symbol for function composition. The parameters $x_{i,mid}$, summarized to the vector x_{mid} , are precisely the scales for each coordinate that is “resolved best” because $\Phi(x_{mid}) = (\frac{1}{2}, \frac{1}{2}, \dots)^T$. So they should be taken to be around the main region of interest.

5.2 Nonlinear, local time-homogenization

A problem during the estimation of viability kernels from Section 3 is a possibly inhomogeneous time scale, i.e. that the (norm of the) RHS function of the control system (1) can have values through *several orders of magnitude*. For instance, models like Eq. (22) often lead to divergences at the upper boundary of a coordinate.

This problem can actually be addressed by rescaling the time of the system in a nonlinear way. The used definitions of viability theory depend only qualitatively but not quantitatively on time. Hence viability kernels and capture basins in the rescaled system are equivalent to the original ones.

As the control parameter is not necessary for the rescaling of the system, we use a differential equation

$$\dot{\mathbf{y}} = F(\mathbf{y}). \quad (23)$$

We propose to use the new system

$$\dot{\mathbf{y}} = \tilde{F}(\mathbf{y}) := \frac{F(\mathbf{y})}{\|F(\mathbf{y})\| + \epsilon}. \quad (24)$$

Assuming ϵ is small enough, this new system generally fulfills three criteria:

1. The systems (23) and (24) are *orbitally equivalent* (cf. [39, p. 42], Definition 2.4), i.e. the trajectories of solutions with the same initial conditions follow the same path. In other words, only the time has been rescaled.
2. Everywhere away from fixed points $\|F(\mathbf{y})\| \gg \epsilon$ holds and hence the time scale is properly homogenized

$$\|\tilde{F}(\mathbf{y})\| = \frac{\|F(\mathbf{y})\|}{\|F(\mathbf{y})\| + \epsilon} \approx 1. \quad (25)$$

3. At fixed points of the original system, the function goes to zero with the same properties as f at that point (e.g. μ -Lipschitz or C^∞ , same Lyapunov-Exponents etc.) because within a small enough environment of the fixed point $\|F(\mathbf{y})\| \ll \epsilon$ holds, thus

$$\tilde{F}(\mathbf{y}) = \frac{F(\mathbf{y})}{\|F(\mathbf{y})\| + \epsilon} \approx \frac{1}{\epsilon} F(\mathbf{y}). \quad (26)$$

Because the units of the coordinates of \mathbf{y} might be different from each other, there is no real physical interpretation of \tilde{F} . But that is not necessary either as it is only an auxiliary system for the estimation with the Saint-Pierre Algorithm.

5.3 Sketch of the Saint-Pierre algorithm

The Saint-Pierre algorithm [63] was developed in order to estimate the viability kernel of a control system Eq. (1).

It starts with a discretization \mathcal{Y}_h of the constraint set \mathcal{Y} where a point $x \in \mathcal{Y}$ is at most at a distance h of a point $\mathbf{y} \in \mathcal{Y}_h$. Furthermore, a small time step $\Delta t > 0$ is chosen and it supposes that the set of controls \mathcal{U} is discrete (if not, it is also discretized). It supposes that f is l -Lipschitz and there exists an upper bound M of \mathcal{Y} .

The algorithm starts by computing, for each point $x \in \mathcal{Y}_h$ and for each control $u \in \mathcal{U}$, the successors $S(x, u)$ of x when applying control u , for a linearized, extended dynamics defined from f . Note that the linearization is done in the x variable only. It is extended in the sense that the successors include all the points located in a ball around $x + f(x, u) \cdot \Delta t$. The successors $S(x, u)$ of x when applying control u are given by

$$S(x, u) = \left\{ \mathbf{y} \in \mathcal{Y}_h \mid \|\mathbf{y} - (x + f(x, u) \cdot \Delta t)\| \leq h + \frac{Ml}{2} \cdot (\Delta t)^2 \right\}. \quad (27)$$

² Often, a regular grid with resolution h is chosen for this discretization, but this is by no means a necessity.

This extension of the dynamics guarantees that the algorithm described below converges to the actual viability kernel when Δt and the resolution of the grid decrease to 0. Computing and storing all the successors for each point of the grid rapidly becomes computationally heavy when the dimensionality of the the state space is large and the grid resolution is small (this is an example of the famous curse of dimensionality).

Then, the algorithm builds a series of discrete sets (subsets of the grid \mathcal{Y}_h) $K_0 = \mathcal{Y}_h, K_1, \dots, K_n$ such that $K_{i+1} \subset K_i$, defined as follows:

$$K_{i+1} = \{x \in K_i \mid \exists u \in \mathcal{U}: S(x, u) \cap K_i \neq \emptyset\} \quad (28)$$

After a finite number of steps, the algorithm reaches a fixed point, i.e. $K_{n+1} = K_n$. The set K_n is the viability kernel of \mathcal{Y}_h for the linearized, extended discrete dynamics. Saint-Pierre [63] shows that this set converges to the viability kernel of the continuous time dynamics when Δt and h tend to 0 appropriately. Note that the approximations are done from the exterior of the viability kernel: generally, the approximation includes points that do not belong to the actual viability kernel (their proportion decreases when the resolution of the grid decreases).

This algorithm has been extended by Deffuant et al. [24] for using continuous sets K_i , using a machine learning algorithm that takes as input the points of the grid that belong to K_i and the ones that do not, and derives an approximation of its boundary. This opens the possibility to represent continuous viability kernels that are defined more conveniently than a huge set of points.

A slight modification of the of the algorithm described above enables us to approximate the capture basin. We start with a discretization of the state space \mathcal{X}_h , analogously to \mathcal{Y}_h above, and define the discretized target set $\mathcal{X}_h = \mathcal{X} \cap \mathcal{X}_h$ for a target set \mathcal{X} . Again, we create a series of discrete sets K'_i with $K'_0 = \mathcal{X}_h$ and where the successors of all elements in K'_{i+1} are in K'_i

$$K'_{i+1} = \{x \in \mathcal{X}_h \mid \exists u \in \mathcal{U}: S(x, u) \cap K'_i \neq \emptyset\}. \quad (29)$$

Again, after a finite number of steps, the algorithm reaches a fixed point, i.e. $K'_{n+1} = K'_n$, and K'_n is the capture basin of \mathcal{X}_h in \mathcal{X}_h for the linearized extended discrete dynamics. In contrast to the viability estimation, this is an approximation from the interior.

Improvements and extensions to this algorithm are currently under intensive research. Relations to dynamical programming [26] and other extensions [11; 64] can provide the minimal time to reach the target set. Also, one can even find controllers that drive the system to the target set [16; 41; 18].

5.4 Estimation of implicitly defined capture basins: eddies

For capture basins we only care about entering a target set at least once. However, eddies are defined by being able to reach the sunny part over and over again, so the most natural definition is an implicit one as in Eqs. (16a) and (16b). In order to estimate them, we find an alternate definition in terms of a limit process.

We start by defining the largest sets that could contain eddies

$$\mathcal{E}^+_0 = \mathcal{X}^+ - \mathcal{U} - \mathcal{D}, \quad (30a)$$

$$\mathcal{E}^-_0 = \mathcal{X}^- - \mathcal{U} - \mathcal{D}, \quad (30b)$$

and then use the iteration step

$$\mathcal{E}^-_i = \text{Capt}_{\mathcal{U}}(\mathcal{E}^+_{i-1}) \cap \mathcal{E}^-_{i-1}, \quad (31a)$$

$$\mathcal{E}^+_i = \text{Capt}_{\mathcal{U}}(\mathcal{E}^-_i) \cap \mathcal{E}^+_{i-1} \quad (31b)$$

for $i = 1, 2, \dots$. Note that \mathcal{E}^-_i used in Eq. (31b) is already computed in Eq. (31a). So this can really be seen as a step-by-step prescription. Thus, the eddies can be recovered as

$$\mathcal{E}^+ = \lim_{i \rightarrow \infty} \mathcal{E}^+_i, \quad (32a)$$

$$\mathcal{E}^- = \lim_{i \rightarrow \infty} \mathcal{E}^-_i. \quad (32b)$$

The limit exists because both sequences are monotone and nonincreasing. The Saint-Pierre algorithm works on a discretized state space with finitely many elements. Hence, the existence of the limit provides that there exists an $k \in \mathbb{N}$ such that $\mathcal{E}^+ = \mathcal{E}^+_k = \mathcal{E}^+_{k-1}$ and $\mathcal{E}^- = \mathcal{E}^-_k = \mathcal{E}^-_{k-1}$ and the algorithm converges after a finite number of steps.

This iteration process follows the idea of being able to visit the sunny part over and over again and is an algorithmic description for the estimation of eddies. Other similarly implicitly given sets can be estimated by adjusting this basic idea.

6 Example: the AYS low-complexity model of climate change, wealth, and energy transformation

We demonstrate the operationalization of the TSM-framework using a three dimensional example model.

To develop a low-complexity model already incorporating climate change, welfare growth and energy transformation, we took inspiration from Kellie-Smith and Cox [33] and added a renewable energy sector with a learning-by-doing dynamics. Its structure is depicted in Fig. 2.

Our model has only three dynamical variables. The first is the *excess atmospheric carbon stock* A [GtC = giga tons of carbon], measured w.r.t. a pre-industrial level $A_0 \approx 600$ GtC. It increases with current CO₂ emissions E [GtC/a = GtC per year]. Taking A_0 as an estimate for the long-term no-emissions equilibrium value, we assume A approaches zero if $E = 0$, due to carbon uptake by oceans, plants and soil. To keep the complexity of the model as low as possible, we do not explicitly model a carbon cycle as in Anderies et al. [2] but simply assume the carbon uptake leads to an exponential relaxation towards equilibrium on a characteristic time scale of $\tau_A \approx 50$ a [a = years]. Hence our first model equation is

$$\frac{dA}{dt} = E - A/\tau_A, \quad (33)$$

where E will be derived below from economic assumptions.

The second variable is *economic output / production* Y [US\$/a] representing the relation to wealth of a society, using the gross world product as its indicator as usual. We assume the economy to have a positive basic growth rate $\beta \approx 3\%$ [1/a] and additional climate impacts as in [33]. As a proxy for temperature we simply use A , effectively assuming an infinitely fast greenhouse effect. Hence this terms is represented by $-\theta AY$ where $\theta \approx 8.57 \cdot 10^{-5}/(\text{GtC a})$ is a temperature sensitivity parameter chosen such that the total growth rate $\beta - \theta A$ becomes negative when A exceeds the level corresponding to a global warming of $+2^\circ\text{C}$. This gives

$$\frac{dY}{dt} = \beta Y - \theta AY. \quad (34)$$

The third dynamical variable is the *renewable energy knowledge stock* S that indicates how much knowledge is available for the production of renewable energy R

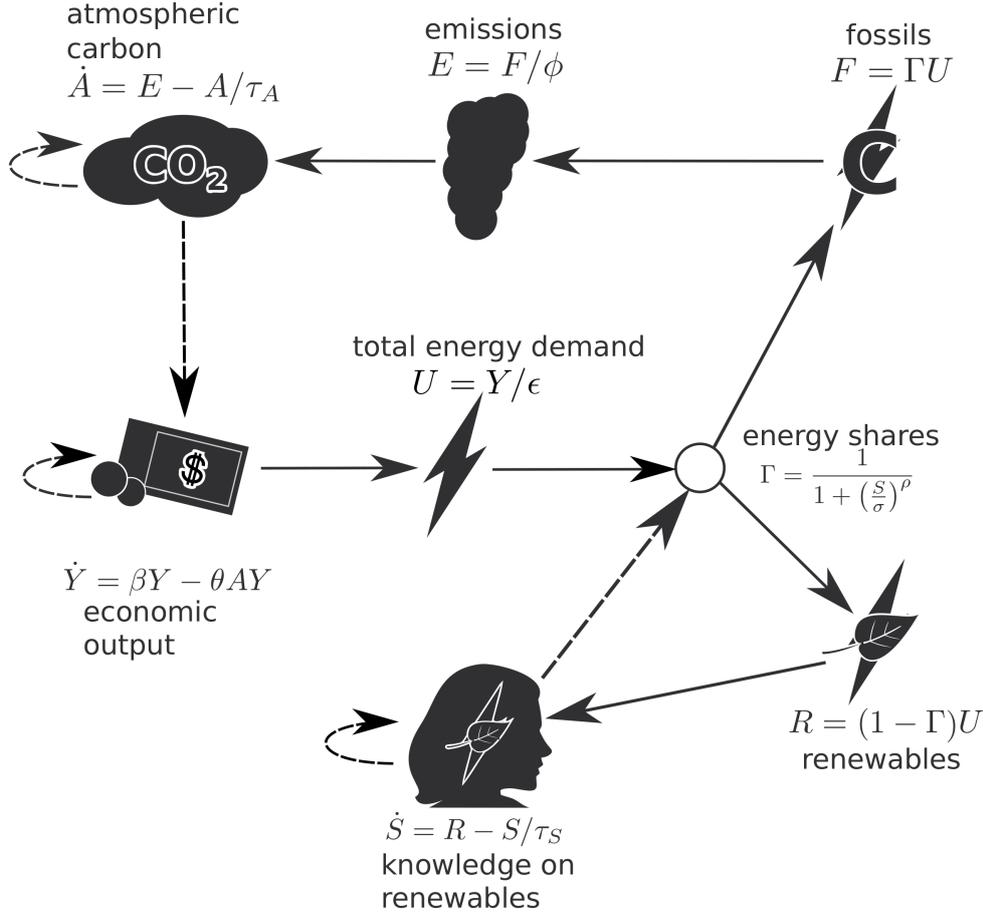


Figure 2: The interplay of the three dynamical variables *excess atmospheric carbon stock* A , *economic production* Y and *renewable energy knowledge stock* S and the five dependent variables *energy demand* U , *fossil (or renewable) energy flow* F (or R), *emissions* E and the *share of the fossil sector* Γ .

[GJ/a = giga joule per year]. In accordance with Wright's law (e.g., [53]) of learning-by-doing, we basically identify S with the past cumulative production of renewables and thus measure it in units of [GJ]. To account for the human capital component, we additionally assume that knowledge depreciates on a characteristic time scale of $\tau_S \approx 50$ a. Cumulation and depreciation then give

$$\frac{dS}{dt} = R - S/\tau_S, \quad (35)$$

where R will be derived below.

Finally, to determine E and R , we use the following simplistic economic assumptions. The energy demand U [GJ/a] is proportional to the economic output

$$U = Y/\epsilon, \quad (36)$$

where $\epsilon \approx 147$ US\$/GJ is an energy efficiency parameter. This demand is satisfied by a mix of fossil and renewable energy which are assumed to be perfect substitutes

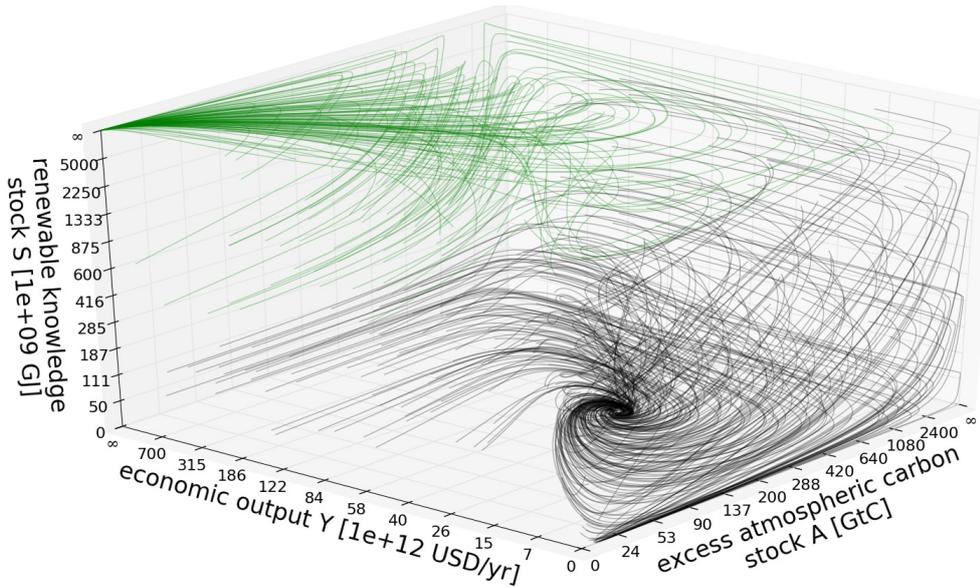


Figure 3: (color online) The default flow of the AWS model is sampled with trajectories from randomly distributed initial conditions on nonlinearly scaled axes so the full states space $X = \mathbb{R}_{\geq 0}^3$ is displayed (more details in Section 6.5). Green trajectories end up at the green attractor x_g and black ones at x_b which are analyzed in Section 6.1.

(and ignoring other energy sources such as agriculture and other bioenergy). Their respective shares are determined by a price equilibrium. We assume convex monomial cost functions and unit costs of renewable energy that show a power-law decay with growing S [53]. This implies that the fossil sector has a share given by the sigmoidal function

$$\Gamma = \frac{1}{1 + \left(\frac{S}{\sigma}\right)^\rho}, \quad (37)$$

where $\sigma \approx 4 \cdot 10^{12} \text{ GJ}$ is the break-even knowledge level at which renewable and fossil extraction costs become equal, and $\rho \approx 2$ is a dimensionless parameter determined from the cost convexity and learning rate. Γ approaches unity (no renewables) as $S \rightarrow 0$ and zero (no fossils) as $S \rightarrow \infty$. Fossil and renewable energy flows and emissions are then

$$F = \Gamma U, \quad R = (1 - \Gamma) U, \quad E = F/\phi, \quad (38)$$

where our final parameter $\phi \approx 4.7 \cdot 10^{10} \text{ GJ/GtC}$ is the fossil fuel combustion efficiency. This completes the model equations. Appendix B contains details on how we estimated the parameters.

The 3 dynamical variables A , Y , and S are interrelated due to the various connecting equations and the nonlinearities arise particularly due to Eq. (37) and Eq. (34). The resulting flow is depicted in Fig. 3 where the basins of the two attractors (discussed in Section 6.1) are already colored differently.

In Section 6.4, we introduce management options for β and σ , hence the control parameter u is two-dimensional and the particular control u_0 for the default flow is

$$u_0 = \begin{pmatrix} \beta \\ \sigma \end{pmatrix}. \quad (39)$$

6.1 Attractors

With the above parameter values, the dynamics has two fixed points. The “black fixed point” x_b at

$$x_b = \begin{pmatrix} A_b \\ Y_b \\ S_b \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\theta} \\ \frac{\phi \epsilon \beta}{\theta \tau_A} \\ 0 \end{pmatrix} = \begin{pmatrix} 350 \text{ GtC} \\ 4.84 \cdot 10^{13} \text{ US\$} \\ 0 \end{pmatrix} \quad (40)$$

corresponding to a carbon based economy without renewable energy use, reduced economic output, and constant climate damages. And there is the “green fixed point” x_g at

$$x_g = \begin{pmatrix} A_g \\ Y_g \\ S_g \end{pmatrix} = \begin{pmatrix} 0 \\ +\infty \\ +\infty \end{pmatrix} \quad (41)$$

corresponding to eventually unbounded exponential growth of economic output and renewable knowledge. Also, there is an exponential decline of fossil usage and emissions towards zero. The mathematical meaning of “ $+\infty$ ” is made clear in Section 6.5.

Both attractors are rather extreme cases. We find this acceptable because this model is a first example and we want to focus on the transients. We understand the asymptotics to be conceptual.

6.2 Current state

The current state $x_c = (A_c, Y_c, S_c)$ can be estimated. A_c is currently around 240 GtC, corresponding to a concentration of 400 ppm [12], and the world gross product of 2015 is around 70 Trillion US\$ [68]. S_c is estimated on the basis of the total past renewable energy consumption of roughly $2 \cdot 10^{12}$ GJ [52]. Since this figure has accumulated over roughly the same time as the characteristic knowledge depreciation time, $\tau_S = 50$ a, we assume roughly half of it has already depreciated, leaving 10^{12} GJ. Because of the large error margins involved in estimating this figure and because it contains hydroelectricity whose growth potential is somewhat problematic, we aim at staying on the conservative side with our estimate and again take only half of this value, giving $5 \cdot 10^{11}$ GJ.

$$x_c = \begin{pmatrix} A_c \\ Y_c \\ S_c \end{pmatrix} = \begin{pmatrix} 240 \text{ GtC} \\ 7 \cdot 10^{13} \text{ US\$} \\ 5 \cdot 10^{11} \text{ GJ} \end{pmatrix} \quad (42)$$

6.3 Desirable states

Steffen et al. [66] set the the planetary boundary for climate change (PB-CC) to 350 ppm [ppm = parts per million] with an uncertainty zone until 450 ppm. We chose the desirable states to be where $A > A_{PB} = 345$ GtC (above the pre-industrial level A_0), corresponding to the looser 450ppm boundary (see Appendix B for the conversion).

Raworth [56] demands that the planetary boundaries are to be complemented with social boundaries that she calls *social foundations*. Combining both ideas is important

to have not only a safe operating space but also a just operating space. Her *social foundations* shaped only illustrative indicators so we have to choose a suitable one for this case. A pragmatic choice is a lower boundary of $Y_{SF} = 4 \cdot 10^{13}$ US\$ (SF-Y), the economic production of the year 2000. The exact value is open for discussion, but as our model has rather low complexity this choice seems reasonable. When trying to refine the number, one should model the distributions in order to include distributive justice and tackle inequality.

Note that for the default dynamics, the green fixed point x_g is not violating either of the boundaries, while the black one x_b violates the PB-CC as can be seen in Fig. 4a.

6.4 Management options

The above parameter values define what we consider the *default dynamics* since they represent a “business-as-usual” case. This means humanity applies no specific management that would alter “the way things usually go”.

In addition to the default dynamics, we study some *management options* representing possible policy choices that may be combined in any way, leading to more or less shifted trajectories.

(i) The option of *low growth* (LG) reduces the basic growth rate β to half its value $\beta_{LG} = 1.5\%/a$, leading to the control

$$u_{LG} = \begin{pmatrix} \beta_{LG} \\ \sigma \end{pmatrix}. \quad (43)$$

This moves the black fixed point to $x_{b,LG} = (A, Y, S) = (175 \text{ GtC}, 2.42 \cdot 10^{13} \text{ US\$/a}, 0)$, no longer violating the PB-CC (see Fig. 4b) but now violating the SF-Y.

(ii) Climate mitigation by inducing an *energy transformation* (ET), e.g. via taxing fossils and/or subsidizing renewable resource use. These policy instruments shift the relative costs of fossil and renewable energy, which according to Eq. (37) can be effected in our model by a reduction of σ . Hence, we represent this option by reducing σ to approx. $(1/2)^{1/\rho} = 1/\sqrt{2}$ of its default value, i.e. to $\sigma_{ET} = 2.83 \cdot 10^{12}$ GJ, corresponding to dividing the renewable to fossil cost ratio by half and leading to the control

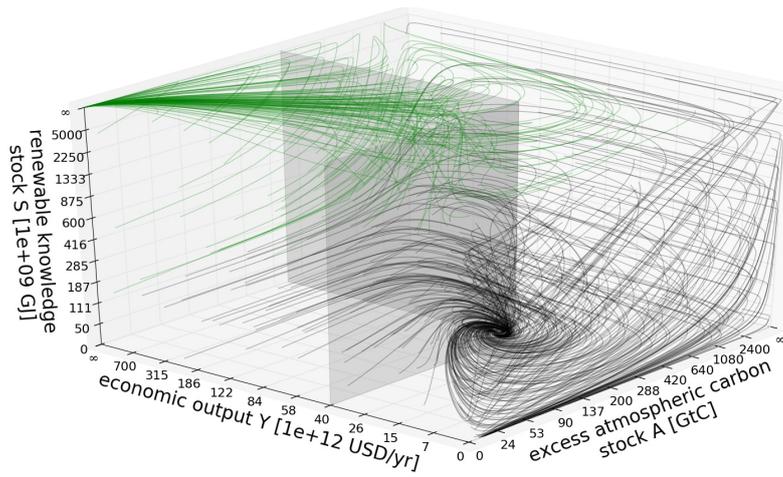
$$u_{ET} = \begin{pmatrix} \beta \\ \sigma_{ET} \end{pmatrix}. \quad (44)$$

This does not affect the location of the two attractors. But, more important, it changes the shape of the basins of attractions. When carefully inspecting Fig. 4c, one can see that the volume of the green fixed point’s the basin of attraction is enlarged in comparison to the default flow in Fig. 4a. Within the concept of Basin Stability [47; 48] the volume of the basin of attraction has been found to be an important indicator for an attractor’s stability hence we will use a similar approach for the bifurcation analysis in Section 6.7.

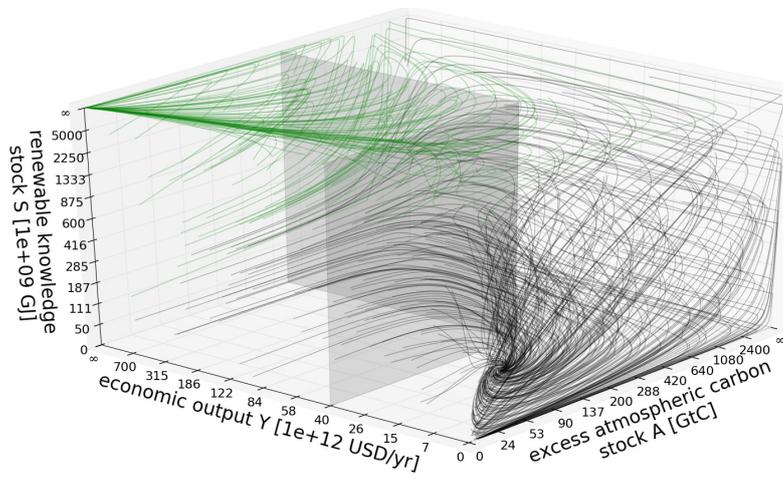
With these two management options, the set of possible managements is given by $\mathcal{U}_m = \{u_{LG}, u_{ET}\}$.

6.5 Dealing with the unbounded state space

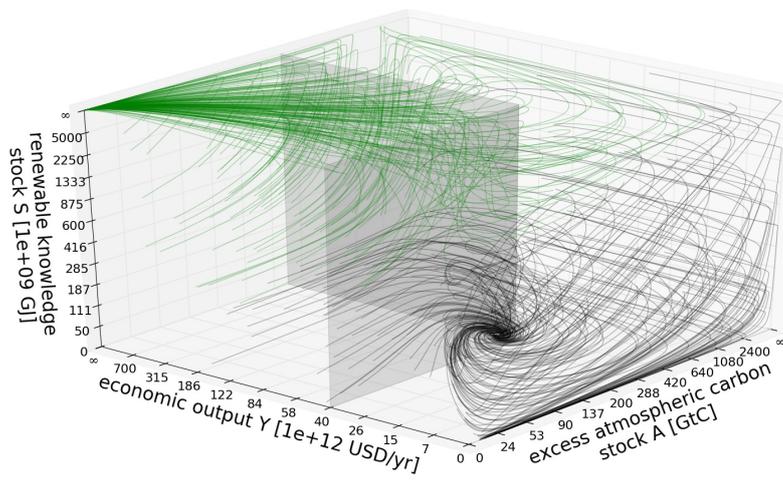
The model (33)-(38) in Section 6 is defined on an unbounded state space and in this section we map it to a bounded one for two reasons: the attractor at “ $x_g = (0, +\infty, +\infty)$ ” and the need for a bounded state space in order to apply the Saint-Pierre algorithm.



(a) Default dynamics



(b) Low growth (LG)



(c) Energy transformation (ET)

Figure 4: (color online) The flows of the AWS model for (a) the default dynamics, (b) the low growth (LG) and (c) the energy transformation management management (ET) option including the combination of both boundary. Note how the black attractor x_b changes its position as discussed in Section 6.4.

In order to make this mathematically sound, we map $\mathcal{X} = \mathbb{R}_{\geq 0}^3$ (parameterized by the variables $x = (A, Y, S)$) to a bounded space $\mathcal{Y} = [0, 1]^3$ (parameterized by transformed coordinates $y = (a, y, s)$) and then add the point $y_g = (0, 1, 1)$ which is the equivalent of x_g in these new coordinates. So we perform a change of coordinates

$$\dot{x} = f(x) \quad \longrightarrow \quad \dot{y} = F(y), \quad (45)$$

where we switch from the old right-hand side (RHS) f to the new RHS F . Following the explanations in Section 5.1, we use the transformation

$$\begin{aligned} \Phi : \mathcal{X} = [0, \infty)^3 &\longrightarrow \mathcal{Y} = [0, 1]^3 \\ A &\longmapsto a = \frac{A}{A_{mid} + A} \\ Y &\longmapsto y = \frac{Y}{Y_{mid} + Y} \\ S &\longmapsto s = \frac{S}{S_{mid} + S}, \end{aligned} \quad (46)$$

where the parameters $x_{mid} = (A_{mid}, W_{mid}, S_{mid})^T$ are such that $\Phi(x_{mid}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^T$. They can be understood as the scale where the ‘‘resolution is the best’’. Hence, changing this value does not qualitatively influence the result, but a good choice can make them clearer.

This is exactly the transformation that has been used to create Figs. 3 and 4a to 4c and will be used for all the following figures, too. As we care most about the current state of the world, we choose $x_{mid} = x_c$.

Using Eq. (22), we get a new set of ODEs with $y = (a, y, s)$ as coordinates

$$\dot{a} = \frac{W_{mid}}{\phi \epsilon A_{mid}} \gamma (1-a)^2 \frac{y}{1-y} - \frac{a(1-a)}{\tau_A}, \quad (47a)$$

$$\dot{y} = y(1-y) \left(\beta - \theta A_{mid} \frac{a}{1-a} \right), \quad (47b)$$

$$\dot{s} = (1-\gamma) \frac{W_{mid}}{\epsilon S_{mid}} (1-s)^2 \frac{y}{1-y} - \frac{s(1-s)}{\tau_S}, \quad (47c)$$

$$\gamma = \frac{(1-s)^\rho}{(1-s)^\rho + \left(\frac{S_{mid}s}{\sigma} \right)^\rho}, \quad (47d)$$

where γ is the equivalent of Γ in Eq. (37) but in the y -coordinates. The fixed points in the new y -coordinates are

$$y_g = \begin{pmatrix} a_g \\ y_g \\ s_g \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \longleftrightarrow \quad x_g = \begin{pmatrix} A_g \\ Y_g \\ S_g \end{pmatrix} = \begin{pmatrix} 0 \\ \infty \\ \infty \end{pmatrix} \quad (48a)$$

$$y_b = \begin{pmatrix} a_b \\ y_b \\ s_b \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\beta + \theta A_{mid}} \\ \frac{\phi \epsilon \beta}{\phi \epsilon \beta + Y_{mid} \theta \tau_A} \\ 0 \end{pmatrix} \quad \longleftrightarrow \quad x_b = \begin{pmatrix} A_b \\ Y_b \\ S_b \end{pmatrix} = \begin{pmatrix} \frac{\beta}{\theta} \\ \frac{\phi \epsilon \beta}{\theta \tau_A} \\ 0 \end{pmatrix}. \quad (48b)$$

Now, we formally extend the dynamics such that $F(y_g) = 0$.

6.6 Results

The model introduced in Section 6 has been analyzed in its compactified form Eqs. (47a) to (47d). For that, we used the nonlinear, local time-homogenization from Section 5.2

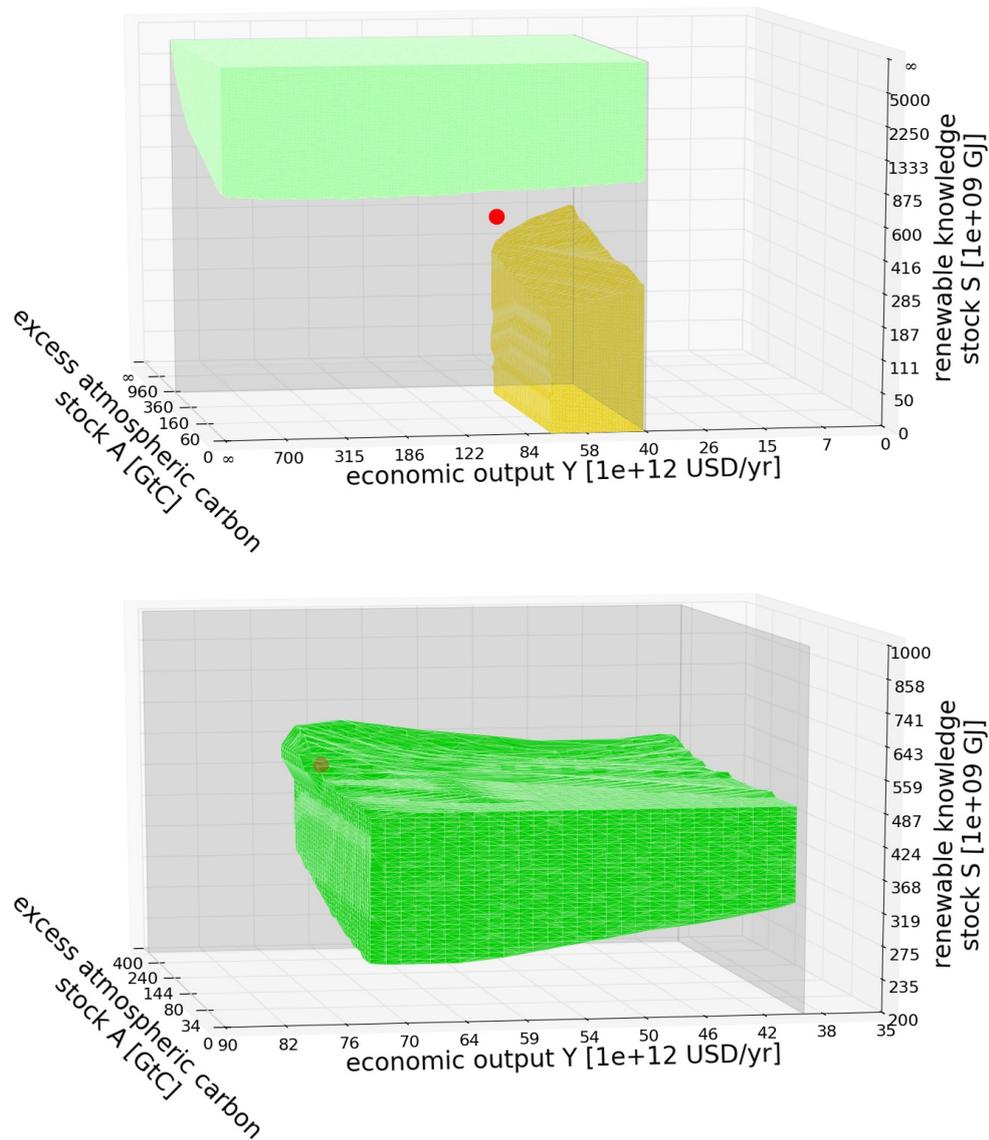


Figure 5: (color online) The safest region, the Shelter (left, light green), is the part where enough knowledge on renewable resource use has already been accumulated and hence the relative price of the renewable resource is low enough such that even the exponentially growing total energy demand which is proportional to the Welfare can be fully compensated by the renewable energy resource so the planetary boundaries are not transgressed. The Backwater (left, yellow) corresponds to the region where one can stay using the *low growth* management option leading finally to a *zero-growth economy* but still staying within the boundaries.

and the Saint-Pierre algorithm that was sketched in Section 5.3. We do not write out the equations for the time-homogenized version as they are lengthy, their calculations straight-forward and they do not give much insight.

The most important identified regions are depicted in Figs. 5a and 5b and we use them for the following discussion.

The first regions to note are the *shelters* and *backwaters*, depicted in Fig. 5a. The former, where one can stay without management forever in the sun and which is the safest region thus, is in our model the invariant kernel of the green fixed point's basin of attraction when restricting to the desirable states \mathcal{X}^+ only. When being in \mathcal{X}^+ and having accumulated already enough knowledge for the energy production with renewable resources, they become so cheap that there is basically no need for fossil fuels anymore. So the remaining (excess) CO_2 (above long-term equilibrium) is removed over time due to the carbon uptake, leading the system to the green fixed point. The *glades*, where one can reach the shelter through the desirable region, are just a thin layer under the shelter so they have not been included in Fig. 5a.

The *backwater*, where one can stay in the sun forever but needs to apply management over and over again, is the part of the desirable region where the growth of economic output and hence of emissions can be restricted. That way, the atmospheric carbon concentration can be kept within the planetary boundary. Also, within the backwater the decarbonization of the economy is impossible, since the given maximal carbon tax and renewable subsidy policy are too weak to make renewables competitive with fossil fuel. Instead, one can manage to stay in a state that corresponds to a carbon-based economy where the low background economic growth is compensated by the climate impacts. The atmospheric carbon level is relatively high but still within the boundary and in equilibrium with the emissions. Hence, using the low growth option properly, one can stay within the desirable region but cannot reach the green fixed point. Note that for simplicity, we did not include the option to choose a value of the base growth rate lying between the two options β and β_{LG} . So formally, the management strategy required to stay in the desirable state described above involves a fast switching between β and β_{LG} , since either of these two extreme values alone leads to a black fixed point in the undesirable region. Still, it is easy to see that this management strategy is equivalent to using a constant intermediate value of β_m (e.g. $\beta_m = 2.7\%/a$) instead. This dynamics has a black fixed point that lies in the desirable region. So this real-world option is implicitly included in the model. In other words, since the TSM-framework allows arbitrarily fast and frequent switches between the management options, one only needs to model the ‘‘corners’’ of the option space explicitly and gets all intermediate options (= all convex combinations). Still, replacing the value of β_{LG} with β_m would introduce further changes, since the maximal management in the transient would be restricted and thus the size of the backwaters reduced. We show this in Section 6.7. Here, it becomes obvious that the TSM-framework incorporates the asymptotics and the transient of a model.

The current state x_c (as estimated in Section 6.2) seems to lie between the two regions discussed above. It is in a region we call a *time-limited lake* as can be seen in the zoom in Fig. 5b. This means humanity appears to be currently facing the so-called lake dilemma where we have to make the qualitative decision between staying within the desirable region uninterrupted but being in need of management forever or going through the undesirable region to finally end up in the shelter. In the model, the choice is between using the energy-transformation option (ET) in order to speed up the knowledge accumulation on renewable resource use and finally reach the green fixed point or the low growth option (LG) in order to restrict the total energy demand and reach the black fixed point. Note that even with a combined usage of the different options (convex combinations) it cannot be avoided to transgress the boundary when going for the first choice.

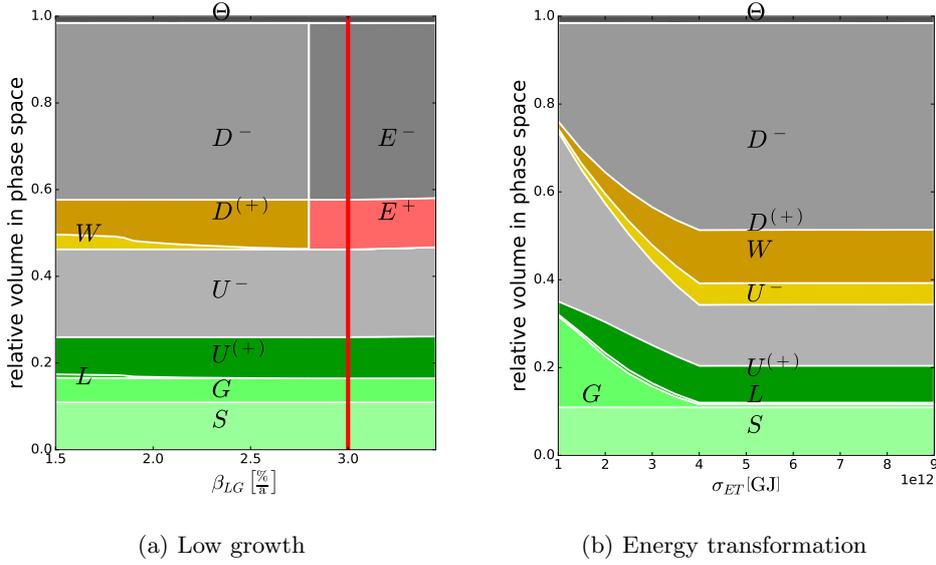


Figure 6: The bifurcation diagrams when varying the parameters governing the two management options from their maximally deviating value to their corresponding value of the default dynamics (except (a) where the default dynamics value is marked by a red vertical line). In (a) a *Downstream-Eddies-Bifurcation* occurs when varying β_{LG} due to the stable focus of the low growth option crossing the boundary between desirable and undesirable states. In (b), we observe that for small σ_{ET} , i.e. strong management, the glade and dark upstream increase in size drastically because the meaning of the energy transformation option is to make the transition to the renewable resource use easier.

6.7 Bifurcation analysis

Within the example model from Section 6, varying the parameters corresponding to the two managements options may lead to bifurcations because of possible changes in the topological structure of the state space with respect to TSM. As an indicator for the bifurcations we use the relative volume of each region, motivated by the concept of Basin Stability [47; 48] and its extensions [31; 35; 49; 69; 34; 50; 51]. Because we use uniformly distributed points in state space for the Saint-Pierre algorithm, we estimate the relative volume of one region with the number of points associated to this region over the total number.

When varying β_{LG} in Fig. 6a corresponding to the *low growth* management option (LG) from 1.5%/a to 3.5%/a, a *downstream-eddies bifurcation* occurs. Until the fixed point of the LG flow crosses the planetary boundary at the critical value of $\beta_{LG} = \beta_{PB} \approx 2.95\%/a$ there is always a *backwater*. Beyond, there are only *eddies* left. The eddies occur because the focus of the default flow and the one of the log-growth flow are both in the undesirable region. One can (in this case) switch between the two flows in a smart way such that one circles far around both foci and can reach the desirable region over and over again while having to pass through the undesirable region in between. Due to the discretization in state space and time during the estimation, the bifurcation seems to occur already at $\approx 2.8\%/a$ in Fig. 6a.

As discussed in Section 6.6 the *backwater* for $\beta_{LG} = 1.5\%/a$ occurs because there exists an in-between value $\beta_m = 2.7\%/a$ where the focus lies in the desirable region. In Fig. 6a, the volume of the backwater for $\beta'_{LG} = \beta_m = 2.7\%/a$ is smaller than for $\beta_{LG} = 1.5\%/a$, because the maximal management in the transient is restricted.

In Fig. 6b, the change in σ_{ET} from the *energy-transformation* management option is depicted and a notable change for small σ_{ET} corresponding to strong management can be observed. As the name implies, this management speeds up the transformation to renewable resource use and thus the glade increases in size, because from more initial conditions it is possible to reach the green fixed point without transgressing the boundaries.

7 Summary & Outlook

To study the operationalization of the topology of sustainable management (TSM), we have introduced a variant definition of it in terms of viability theory (VT).

Using this connection, we have been able to apply the Saint-Pierre algorithm in order to operationalize the analysis of the TSM partition. Because the algorithm works on bounded spaces only, we have introduced a coordinate transformation to a bounded space. The transformation parameters are chosen to fix the scale of “highest resolution”. Furthermore, we solved the problem that time scales may vary through orders of magnitude by introducing a nonlinear, local time homogenization.

These novel concepts were applied to an example system combining climate change, welfare and energy transformation. While the system was kept minimalistic, a rich topology was found. The current state of the world was estimated within the scope of the model; particularly interesting was that it seems to be inside a finite-time lake. So, humanity is facing a pressing lake-dilemma where it has to make a choice between two qualitatively different options. Furthermore, we performed a bifurcation analysis under change of the management parameters and found a downstream-eddies bifurcation.

The example system had a rather low-complexity so we can focus on the operationalization of TSM. But for that reason, relevant parts, e.g. a carbon cycle, an economic cycle and different energy productions, were represented simplistically only, so a more complex model is necessary. The one developed in [54; 55] seems to be a good candidate. The complexity is higher than the one used in this article but the analysis might still be possible.

Increasing the dimensionality of the system induces the need for improved algorithms also. While there exist some [13; 42; 15; 1], they need to be adjusted and extended to fit to the computation of the TSM partition.

This article 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 PIK’s 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 [70], Numerical Python [3] and Scientific Python [32].

The authors thank Sabine Auer, Wolfram Barfuss, Karsten Bölts, Catrin Ciemer, Eduardo Costa, Jonathan Donges, Reik Donner, Jaap Eldering, Jasper Franke, Roberto Guelleri, Frank Hellmann, Jakob Kolb, Julien Korinman, Till Koster, Chiranjit Mitra, Jan Nitzbon,

Ilona Otto, Tiago Pereira, Camille Pognard, Francisco Rodrigues, Edmilson Roque dos Santos, Stefan Ruschel, Paul Schultz, Fabiano Berardo de Sousa, Lyubov Tupikina, and Kilian Zimmerer for helpful discussions and comments.

References

References

1. I. Alvarez, R. Reuillon, and R. D. Aldama. Viabilitree: a ld-tree framework for viability-based decision. *archives-ouvertes.fr*, 2016.
2. J. M. Anderies, S. R. Carpenter, W. Steffen, and J. Rockström. The topology of non-linear global carbon dynamics: from tipping points to planetary boundaries. *Environmental Research Letters*, 8(4):044048, dec 2013. doi: 10.1088/1748-9326/8/4/044048.
3. D. Ascher, P. F. Dubois, K. Hinsen, J. Hugunin, T. Oliphant, et al. Numerical python, 2001.
4. J. Aubin. *Viability theory*. Birkhäuser, 1991.
5. J. P. Aubin. *Dynamic economic theory: a viability approach*, volume 5. Springer Verlag, 1997.
6. J. P. Aubin. Viability kernels and capture basins of sets under differential inclusions. *SIAM Journal on Control and Optimization*, 40(3):853–881, 2001. doi: 10.1137/S036301290036968X.
7. J.-P. Aubin and A. Cellina. *Differential inclusions: set-valued maps and viability theory*, volume 264. Springer Science & Business Media, 2012.
8. J.-P. Aubin and G. Da Prato. Stochastic viability and invariance. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 17(4):595–613, 1990.
9. J.-P. Aubin, A. Bayen, and P. Saint-Pierre. *Viability Theory: New Directions*. Springer, 2011.
10. W. Barfuss, J. F. Donges, M. Wiedermann, and W. Lucht. Sustainable use of renewable resources in a stylized social-ecological network model under heterogeneous resource distribution. *Earth System Dynamics Discussions*, 2016:1–13, 2016. doi: 10.5194/esd-2016-15.
11. Bayen, Crück, and Tomlin. Guaranteed overapproximations of unsafe sets for continuous and hybrid systems: Solving the hamilton-jacobi equation using viability techniques. *Lecture Notes In Computer Science*, 2002.
12. R. A. Betts, C. D. Jones, J. R. Knight, R. F. Keeling, and J. J. Kennedy. El nino and a record co2 rise. *Nature Climate Change*, 2016.
13. O. Bokanowski, S. Martin, R. Munos, and H. Zidani. An anti-diffusive scheme for viability problems. *Applied Numerical Mathematics*, 56(9):1147–1162, 2006.
14. J. A. Brander and S. M. Taylor. The simple economics of Easter Island: A Ricardo-Malthus model of renewable resource use. *The American Economic Review*, 88(1):119–138, 1988.
15. A. Brias, J. Mathias, and G. Deffuant. Accelerating viability kernel computation with cuda architecture: application to bycatch fishery management. *Computational Management Science*, 13(3):371–391, 2016.
16. P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre. Set-valued numerical analysis for optimal control and differential games. *Annals of the International Society of Dynamic Games.*, 1998.
17. S. R. Carpenter and E. M. Bennett. Reconsideration of the planetary boundary for phosphorus. *Environmental Research Letters*, 6(1):014009, 2011.
18. L. Chapel and G. Deffuant. Svm approximation of value function contours in target hitting problems. In *Informatics in control, automation and robotics. 8th*

- International Conference, ICINCO 2011*, volume 174 of *Lecture Notes in Electrical Engineering*, pages 37–48. Springer, 2011.
19. L. Chapel, G. Deffuant, S. Martin, and C. Mullon. Defining yield policies in a viability approach. *Ecological Modelling*, 212(1):10–15, 2008.
 20. L. Chapel, X. Castelló, C. Bernard, G. Deffuant, V. M. Eguíluz, S. Martin, and M. San Miguel. Viability and resilience of languages in competition. *Plos one*, 5(1):e8681, 2010.
 21. P. Ciais, C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. L. Quéré, R. Myneni, S. Piao, and P. Thornton. The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. *Change, IPCC Climate*, pages 465–570, 2013. doi: 10.1017/CBO9781107415324.015.
 22. P. M. Cury, C. Mullon, S. M. Garcia, and L. J. Shannon. Viability theory for an ecosystem approach to fisheries. *ICES Journal of Marine Science: Journal du Conseil*, 62(3):577–584, 2005.
 23. G. Deffuant and N. Gilbert, editors. *Viability and Resilience of Complex Systems: Concepts, Methods and Case Studies from Ecology and Society*. Springer, 2011.
 24. G. Deffuant, L. Chapel, and S. Martin. Approximating viability kernels with support vector machines. *IEEE transactions on automatic control*, 52(5):933–937, 2007.
 25. M. Delara and L. Doyen. *Sustainable Management of Natural Resources. Mathematical Models and Methods*. Springer, 2008.
 26. H. Frankowska. Optimal trajectories associated with a solution of the contingent hamilton-jacobi equation. *Applied Mathematics and Optimization*, pages 291–311, 1989.
 27. D. Gerten, H. Hoff, J. Rockström, J. Jägermeyr, M. Kummu, and A. V. Pastor. Towards a revised planetary boundary for consumptive freshwater use: role of environmental flow requirements. *Current Opinion in Environmental Sustainability*, 5(6):551–558, 2013.
 28. T. Häyhä, P. L. Lucas, D. P. van Vuuren, S. E. Cornell, and H. Hoff. From Planetary Boundaries to national fair shares of the global safe operating space - How can the scales be bridged? *Global Environmental Change*, 40:60–72, 2016. doi: 10.1016/j.gloenvcha.2016.06.008.
 29. V. Heck, J. F. Donges, and W. Lucht. Collateral transgression of planetary boundaries due to climate engineering by terrestrial carbon dioxide removal. *Earth System Dynamics*, 7(4):783–796, 2016. doi: 10.5194/esd-7-783-2016.
 30. J. Heitzig, T. Kittel, J. F. Donges, and N. Molkenhain. Topology of sustainable management of dynamical systems with desirable states: from defining planetary boundaries to safe operating spaces in the Earth system. *Earth System Dynamics*, 7(1):21–50, 2016. doi: 10.5194/esd-7-21-2016.
 31. F. Hellmann, P. Schultz, C. Grabow, J. Heitzig, and J. Kurths. Survivability of Deterministic Dynamical Systems. *Scientific reports*, 6:29654(29654):1–12, 2016. doi: 10.1038/srep29654.
 32. E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001. URL <http://www.scipy.org/>. [Online; accessed 2016-05-10].
 33. O. Kellie-Smith and P. M. Cox. Emergent dynamics of the climate-economy system in the Anthropocene. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 369(1938):868–86, 2011. doi: 10.1098/rsta.2010.0305.
 34. T. Kittel, J. Heitzig, K. Webster, and J. Kurths. Timing of transients: Quantifying reaching times and transient behavior in complex systems. *arXiv preprint arXiv:1611.07565*, 2016.

35. V. V. Klinshov, V. I. Nekorkin, and J. Kurths. Stability threshold approach for complex dynamical systems. *New Journal of Physics*, 18(1):013004, 2015.
36. J. J. Kolb and J. Heitzig. Social tipping as opinion dynamics of bounded rational agents. approaches on modeling divestment, (in prep.).
37. C. T. Kone, G. DeSousa, and J. D. Mathias. Adaptive management of energy consumption, reliability and delay of wireless sensor node: application to ieee 802.15.4 wireless sensor node. *Plos One*, page Under revisions, 2017.
38. A. Kuhn and T. Heckelei. Anthroposphere. In *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*, pages 282–341. Springer, 2010.
39. Y. A. Kuznetsov. *Elements of applied bifurcation theory, Second Edition*, volume 112. Springer Science & Business Media, 1998.
40. S. J. Lade, S. Niiranen, J. Hentati-Sundberg, T. Blenckner, W. J. Boonstra, K. Orach, M. F. Quaas, H. Österblom, and M. Schlüter. An empirical model of the baltic sea reveals the importance of social dynamics for ecological regime shifts. *Proceedings of the National Academy of Sciences*, 112(35):11120–11125, 2015.
41. Lhommeau, Jaulin, and Hardouin. Inner and outer approximation of capture basin using interval analysis. In *4th International Conference on Informatics in Control, Automation and Robotics (ICINCO2007)*, pages 1–5, 2007.
42. J. N. Maidens, S. Kaynama, I. M. Mitchell, M. M. Oishi, and G. A. Dumont. Lagrangian methods for approximating the viability kernel in high-dimensional systems. *Automatica*, 49(7):2017 – 2029, 2013. doi: <http://dx.doi.org/10.1016/j.automatica.2013.03.020>.
43. S. Martin. The cost of restoration as a way of defining resilience: a viability approach applied to a model of lake eutrophication. *Ecology and Society*, 9(2), 2004.
44. J. Mathias, B. Bonté, T. Cordonnier, and F. DeMorogues. Using the viability theory for assessing flexibility of forest managers under ecological intensification. *Environmental Management*, 56:1170–1183, 2015.
45. J. Mathias, J. Anderies, and M. Janssen. On our rapidly shrinking capacity to comply with the planetary boundaries on climate change. *Scientific Reports - Nature*, 7(42061):1–7, 2017. doi: 10.1038/srep42061.
46. J.-D. Mathias, S. Lade, and V. Galaz. Multi-level policies and adaptive social networks - a conceptual modeling study for maintaining a polycentric governance. *International Journal of the Commons*, 11(1):220–247, 2017. doi: 10.18352/ijc.695.
47. P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths. How basin stability complements the linear-stability paradigm. *Nature Physics*, 9(2):89–92, 2013.
48. P. J. Menck, J. Heitzig, J. Kurths, and H. Joachim Schellnhuber. How dead ends undermine power grid stability. *Nature Communications*, 5:3969:1–8, jun 2014. doi: 10.1038/ncomms4969.
49. C. Mitra, J. Kurths, and R. V. Donner. An integrative quantifier of multistability in complex systems based on ecological resilience. *Scientific reports*, 5:16196 (16196):1–10, 2015. doi: 10.1038/srep16196.
50. C. Mitra, A. Choudhary, S. Sinha, J. Kurths, and R. V. Donner. Multi-node basin stability in complex dynamical networks. *arXiv preprint arXiv:1612.06015*, 2016.
51. C. Mitra, T. Kittel, A. Choudhary, J. Kurths, and R. V. Donner. Recovery time after localized perturbations in complex dynamical networks. *arXiv preprint arXiv:1612.06015*, 2017.
52. Mongabay Organization. Carbon dioxide emissions charts, 2005. URL http://rainforests.mongabay.com/09-carbon_emissions.htm.
53. B. Nagy, J. D. Farmer, Q. M. Bui, and J. E. Trancik. Statistical Basis for Predicting Technological Progress. *PLoS ONE*, 8(2):1–7, 2013. doi: 10.1371/jour-

- nal.pone.0052669.
54. J. Nitzbon. Bifurcation analysis and parameter estimation in low-dimensional models of global human-nature coevolution. Master's thesis, Georg-August-Universität Göttingen, 2016.
 55. J. Nitzbon, J. Heitzig, and U. Parlitz. Sustainability, collapse and oscillations of global climate, population and economy in a simple world-earth model, (in prep.).
 56. K. Raworth. A Safe and Just Space For Humanity: Can we live within the Doughnut? *Nature*, 461:1–26, 2012. doi: 10.5822/978-1-61091-458-1.
 57. B. D. Richter, R. Mathews, D. L. Harrison, and R. Wigington. Ecologically sustainable water management: managing river flows for ecological integrity. *Ecological Applications*, 13(1):206–224, 2003. doi: 10.1890/1051-0761(2003)013[0206:ESWMMR]2.0.CO;2.
 58. J. Rockström, W. L. Steffen, K. Noone, A. Persson, F. S. Chapin III, E. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, and Others. Planetary boundaries: exploring the safe operating space for humanity. *Ecology and Society*, 14(2), 2009.
 59. C. Rougé, J.-D. Mathias, and G. Deffuant. Extending the viability theory framework of resilience to uncertain dynamics, and application to lake eutrophication. *Ecological Indicators*, 29:420 – 433, 2013. doi: <http://dx.doi.org/10.1016/j.ecolind.2012.12.032>.
 60. C. Rougé, J.-D. Mathias, and G. Deffuant. Relevance of control theory to design and maintenance problems in time-variant reliability: The case of stochastic viability. *Reliability Engineering and System Safety*, 132:250 – 260, 2014.
 61. E. S. Rubin, I. M. L. Azevedo, P. Jaramillo, and S. Yeh. A review of learning rates for electricity supply technologies. *Energy Policy*, 86:198–218, 2015. doi: 10.1016/j.enpol.2015.06.011.
 62. S. W. Running. A measurable planetary boundary for the biosphere. *science*, 337(6101):1458–1459, 2012.
 63. P. Saint-Pierre. Approximation of the viability kernel. *Applied Mathematics and Optimization*, 29(2):187–209, 1994.
 64. P. Saint-Pierre. Approximation of viability kernels and capture basins for hybrid systems. In *Control Conference (ECC), 2001 European*, pages 2776–2783. IEEE, 2001.
 65. H. J. Schellnhuber. Discourse: Earth system analysis—the scope of the challenge. In *Earth System Analysis*, pages 3–195. Springer, 1998.
 66. W. Steffen, K. Richardson, J. Rockstrom, S. E. Cornell, I. Fetzer, E. M. Bennett, R. Biggs, S. R. Carpenter, W. de Vries, C. A. de Wit, C. Folke, D. Gerten, J. Heinke, G. M. Mace, L. M. Persson, V. Ramanathan, B. Reyers, and S. Sorlin. Planetary boundaries: Guiding human development on a changing planet. *Science*, 347(6223):1259855–1259855, feb 2015. doi: 10.1126/science.1259855.
 67. The World Bank Group. Energy intensity level of primary energy, 2016. URL <http://data.worldbank.org/indicator/EG.EGY.PRIM.PP.KD>.
 68. The World Bank Group. Gdp, 2016. URL <http://data.worldbank.org/indicator/NY.GDP.MKTP.CD>.
 69. A. van Kan, J. Jegminat, J. F. Donges, and J. Kurths. Constrained basin stability for studying transient phenomena in dynamical systems. *Phys. Rev. E*, 93:042205, Apr 2016. doi: 10.1103/PhysRevE.93.042205.
 70. G. Van Rossum and F. L. Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.
 71. M. Wiedermann, J. F. Donges, J. Heitzig, W. Lucht, and J. Kurths. Macroscopic description of complex adaptive networks coevolving with dynamic node states. *Physical Review E*, 91(5):052801, 2015.

Appendix

A Existence of eddies

Eddie-like We call a pair of sets $\mathcal{A}^{+/-} \subseteq \mathcal{X}$ *eddie-like* iff they fulfill (i) $\mathcal{A}^{+/-} \subseteq \mathcal{X}^{+/-} - \mathcal{U} - \mathcal{D}$ and (ii) $\mathcal{A}^{+/-} \subseteq \text{Capt}(A^{-/+})$. Note the inverted order of the signs in the last term.

Union of two Eddie-like pair sets are also Eddie-like We claim that for two eddie-like pairs of sets $\mathcal{E}^{+/-}_1$ and $\mathcal{E}^{+/-}_2$ the union pair $\mathcal{E}^{+/-}_3 = \mathcal{E}^{+/-}_1 \cup \mathcal{E}^{+/-}_2$ is eddie-like, too.

Proof The first condition is trivially fulfilled and the second one follows straight away from $\text{Capt}(\mathcal{A}) \cup \text{Capt}(\mathcal{B}) = \text{Capt}(\mathcal{A} \cup \mathcal{B})$ for two state sets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$. Hence, the union of all eddie-like pairs of sets is maximal and *eddies* exist. \square

B Parameter estimation

To get a roughly realistic setting, we estimated the parameters of the model using several publicly available data sources.

A_0 was taken from [21] and slightly rounded. τ_A and β were taken from [33]. ϕ was based on the ton oil equivalent of various fossil fuels and a typical mass share of 90% carbon in fossil fuels, as described in [54].

Assuming that two degrees warming correspond to a carbon concentration of 450 ppm and thus to a carbon stock of 950 GtC (both being 1.6 times their pre-industrial value), we require that the total growth rate $\beta - \theta A_1$ becomes zero for $A_1 \approx 950 \text{ GtC} - A_0 = 350 \text{ GtC}$, hence θ was taken to be $\beta/A_1 \approx 8.57 \cdot 10^{-5}/(\text{GtC a})$.

ϵ was estimated from the World Bank's primary energy intensity data [67].

For τ_S , the characteristic depreciation time of renewable energy knowledge, no reliable source was found, so we made a very coarse guess by setting it roughly to the length of an average working life of 50 a.

The break-even knowledge level σ was also estimated very coarsely. According to past cumulative world consumption of renewable energy is $\approx 2 \cdot 10^{18} \text{ Btu} \approx 2 \cdot 10^{12} \text{ GJ}$ or roughly 20 years of world energy consumption. To be on the conservative side and avoid overestimating the potential of renewables, we took σ to be two times that value.

ρ was set as follows. We assume fossil and renewable energy production costs of $C_F \propto F^{1+\gamma}$ and $C_R \propto R^{1+\gamma}/S^\lambda$, where $\gamma > 0$ is a convexity parameter and $\lambda > 0$ is a learning exponent. Then energy prices are $\pi_F \propto \partial C_F / \partial F \propto F^\gamma$ and $\pi_R \propto \partial C_R / \partial R \propto R^\gamma / S^\lambda$. In the price equilibrium, $\pi_F = \pi_R$, hence $R/F \propto S^{\lambda/\gamma}$, and thus $\rho = \lambda/\gamma$. According to [61], the learning rate $LR = 1 - 2^{-\lambda}$ of several renewables is around 1/8, hence $\lambda \approx \log_2(8/7) \approx 0.2$. Assuming a mild convexity of $\gamma \approx 0.1$, we get $\rho \approx 2$.