Earth Syst. Dynam. Discuss., 6, 435–488, 2015
www.earth-syst-dynam-discuss.net/6/435/2015/
doi:10.5194/esdd-6-435-2015

Earth System Dynamics Discussions
Open Access

Topology of sustainable management in dynamical Earth system models with desirable states

J. Heitzig[1] and T. Kittel[1,2]

[1]Potsdam Institute for Climate Impact Research, P.O. Box 60 12 13, 14412 Potsdam, Germany
[2]Humboldt-Universität zu Berlin, Department of Physics, Newtonstr. 15, 12489 Berlin, Germany

Received: 13 February 2015 – Accepted: 24 February 2015 – Published: 12 March 2015

Correspondence to: J. Heitzig (heitzig@pik-potsdam.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

To keep the Earth system in a desirable region of its state space, such as the recently suggested "tolerable environment and development window", "planetary boundaries", or "safe (and just) operating space", one not only needs to understand the quantitative internal dynamics of the system and the available options for influencing it (management), but also the structure of the system's state space with regard to certain qualitative differences. Important questions are: which state space regions can be reached from which others with or without leaving the desirable region? Which regions are in a variety of senses "safe" to stay in when management options might break away, and which qualitative decision problems may occur as a consequence of this topological structure?

In this article, as a complement to the existing literature on optimal control which is more focussed on quantitative optimization and is much applied in both the engineering and the integrated assessment literature, we develop a mathematical theory of the qualitative topology of the state space of a dynamical system with management options and desirable states. We suggest a certain terminology for the various resulting regions of the state space and perform a detailed formal classification of the possible states with respect to the possibility of avoiding or leaving the undesired region. Our results indicate that before performing some form of quantitative optimization, the sustainable management of the Earth system may require decisions of a more discrete type that come in the form of several dilemmata, e.g., choosing between eventual safety and uninterrupted desirability, or between uninterrupted safety and increasing flexibility.

We illustrate the concepts and dilemmata with conceptual models from classical mechanics, climate science, ecology, economics, and coevolutionary Earth system modelling and discuss their potential relevance for the climate and sustainability debate.

1 Introduction

The sustainable management of systems mainly governed by an internal dynamics for which one desires to stay in a certain region of their state space, such as a "tolerable environment and development (E&D) window" in a model of the Earth system (Schellnhuber, 1998), requires first and foremost an understanding of the *topology* of the system's state space in terms of what regions are in some sense "safe" to stay in, and to what qualitative degree, and which of these regions can be reached with some degree of safety from which others regions, either by the internal ("default") dynamics or by some alternative dynamics influenced by some form of management. In the Earth system context, it may be very hard to define meaningful "planetary boundaries" (Rockström et al., 2009a; Steffen et al., 2015) and relate them to sustainable development goals without such an analysis.

Also the question whether it suffices to influence the system by active management for only a limited time to reach a safe region or whether it might be necessary to repeat active management indefinitely or even continue it uninterruptedly in order to avoid undesired state space regions seems quite relevant in view of urgent problems such as the climate policy debate. E.g., if suitable mitigation policies such as certain forms of energy market regulation can transform the economic system in a way that allows one to eventually deregulate the market again, then for how long can one delay mitigation until this feature is lost and only permanent regulation can help? Or, if certain adaptation or geoengineering options might be cheaper than mitigation but require an uninterrupted management or lead to a less well-known region of state space (Kleidon and Renner, 2013), which of these qualitatively different properties is preferable?

We will see that such questions about a "safe" or "safe and just operating space" (Rockström et al., 2009b; Raworth, 2012) may lead to decision dilemmata that cannot easily be analysed in a purely optimization-based framework. The paradigm of optimal control, which is much applied in both the engineering and the integrated assessment literature, on the one hand does not provide sufficient concepts for such a qualitative

ESDD
6, 435–488, 2015

analysis and on the other hand typically requires quite a lot of additional knowledge, in particular, some or other form of *quantitative* evaluation of states.

In this article, we will derive in a purely topological way a thorough and precise *qualitative* classification of the possible states of a system with respect to the possibility of avoiding or leaving some given undesired region by means of some given management options. Our results indicate that in addition to (or maybe rather before) performing some form of quantitative optimization, the sustainable management of a system may require decisions of a more discrete type, e.g., choosing between eventual safety and permanent desirability, or between permanent safety and increasing future options, etc. This appears even more so in the presence of strong nonlinearities, multistable regimes, bifurcations, and tipping elements (Lenton et al., 2008; Schellnhuber, 2009), where small state changes due to random perturbations or deliberate management may not only have large consequences but can lead to qualitative and possibly irreversible changes. The resulting dilemmata that we will discuss are summarized in Table 1.

To indicate the wide scope of applicability of our concepts in various subdisciplines of Earth system Science, we illustrate the concepts and dilemmata with conceptual models from classical mechanics, climate science, ecology, economics, and coevolutionary Earth system modelling.

In contrast to the somewhat related but more formal approach of sequential decision problems in discrete-time systems (Botta et al., 2015), we focus on the more easily applicable class of *continuous-time* systems and their models here. Our classification is based on a distinction between default and alternative trajectories of a system, and suitably adapted *reachability* concepts from control theory and the important but vast field of viability theory (Aubin, 2009; Aubin et al., 2011; Aubin and Saint-Pierre, 2007). Since physical models of global-scale processes or other macroscopic systems are usually of a statistical physics nature in the sense that they represent the aggregate effects of many micro-scale processes by suitable approximations, their proper interpretation typically requires one to expect small (actually or seemingly) random perturbations. We

ESDD
6, 435–488, 2015

J. Heitzig and T. Kittel

take this into account here by strengthening the usual notion of reachability to one of *stable reachability,* and by requiring the featured subsets of state space to be topologically open (instead of closed) sets, so that infinitesimal perturbations cannot kick the system out of them.

Naturally, we will have to introduce some formal notation but try to keep it to a minimum in the main text, reserving a more detailed formal treatment and a comparison of our concepts with those of viability theory for the Appendix.

1.1 Formal framework

Let us assume a *manageable dynamical system with desirable states,* given by the following components:

(i) A dynamical system with a *state space* $X \neq \varnothing$, a *default dynamics* represented by a family of continuous *default (forward) trajectories* $\tau_x(t)$ (with $x \in X$, $t \geqslant 0$, and $\tau_x(0) = x$), and a state space *topology* given by a system of open sets $\mathcal{T} \subseteq 2^X$. (Typically, $X \subseteq \mathbb{R}^n$ and $\mathcal{T}$ is the Euclidean topology on $\mathbb{R}^n$ restricted to X.)

(ii) A notion of "desirable states" represented by an open set $X^+ \in \mathcal{T}$, called the *sunny region,* e.g. defined via some "tolerable E&D window" (Schellnhuber, 1998), whose complement $X^- = X - X^+$ we call the *dark (region)*. (For readability, we use a minus sign instead of a backslash to denote subtraction of sets in this paper.)

(iii) A notion of "management options" represented by a (finite or infinite) family $\mathcal{M}_x$ of continuous *admissible trajectories* μ for each $x \in X$ (with $\mu(0) = x$, $\mu(t) \in X$ for all $t \geqslant 0$).

We require that $\tau_x \in \mathcal{M}_x$ and assume that one can switch immediately to any trajectory $\mu \in \mathcal{M}_x$ whenever in state x. Extending the established common metaphor of a boat on water, we say the system "floats" when it follows a default trajectory, and that we may "row" the system along any other admissible trajectory.

ESDD
6, 435–488, 2015

J. Heitzig and T. Kittel

Note that although, formally, we consider deterministic autonomous systems only, non-deterministic systems can be incorporated by considering probability distributions as states, time-delay systems can be treated similarly, and externally driven or otherwise explicitly time-dependent systems can be covered by including time t as a variable with $\dot{t} = 1$ into the state vector. Also, if management involves some form of inertia, e.g., if not the propelling vector $\boldsymbol{v}$ of a boat but only its acceleration $\dot{\boldsymbol{v}}$ can be changed discontinuously, the proper way to model this in our framework would be to treat $\boldsymbol{v}$ as part of the state.

2 Qualitative distinction of regions w.r.t. sustainable manageability of desirability

The main idea of the coarsest of our classifications of states is to first identify (i) a safe region where management is unnecessary, called the *shelters* S, and (ii) a less safe but larger *manageable region* M where one can permanently avoid the dark at least by management. Then we classify all states with regard to whether and how X^+, S, and M can be stably reached from the current state by management. For each state, we ask: (iii) Can S be stably reached, and if so, can the dark be avoided on the way? (iv) If not, can M be stably reached? (v) If not, can we stably reach X^+ over and over again, or at least once again? We will see that these criteria lead to a partition of state space into a “cascade” consisting of five main regions that we will call the *upstream* U, the *downstream* D, the *eddies* E, the *abysses* Υ, and the *trenches* Θ. Each of these will then be split up further into sets such as the *glades* G, *lakes* L, and *backwaters* W, etc., by asking further qualitative questions. In choosing these figurative terms, we try to avoid a too technically-sounding language and rather extend the useful and common metaphor of “flows” and “basins” in a natural way that is consistent with the basic picture of a river, an ocean, and different regions therein, without trying to match their common-language meanings too accurately.

ESDD
6, 435–488, 2015

To acknowledge the fact that all real-world dynamics and management will be subject to at least infinitesimal noise and errors, we base the formal definition of these state space regions on certain notions of *invariant open kernel, sustainability,* and *stable reachability,* whose symbolic mathematical definitions and algebraic properties are detailed in the Appendix.

2.1 Shelters, manageable region, upstream and downstream

The *invariant open kernel* of a set $A \subseteq X$, denoted $A^{\iota\circ}$, is the largest open subset of A that contains the default trajectories of all its own points (its existence and uniqueness is nontrivial but proved in the Appendix). Note that $A^{\iota\circ}$ may be empty. Now, we define the *shelters* as the invariant open kernel of the sunny region,

$$S = (X^+)^{\iota\circ}. \tag{1}$$

Each (topologically) connected component of S is called an individual *shelter.* S contains all sunny states whose default trajectories stay in the sunny region X^+ forever without any management even when infinitesimal (or small enough) perturbations occur. In other words, when inside S, one *will* "stably" stay in X^+ *by default.*

We call an open set $A \in \mathcal{T}$ *sustainable* (in the basic sense of the word, simply meaning that it can be sustained) iff it contains an admissible trajectory for each of its points, i.e., iff for all $x \in A$, there is $\mu \in \mathcal{M}_x$ with $\mu(t) \in A$ for all $t \geqslant 0$. Again, the openness requirement ensures a minimal form of stability against small perturbations. The *sustainable kernel* of a set $A \subseteq X$, denoted $A^{\mathcal{S}}$, is the largest sustainable open subset of A (again, see Appendix for existence and uniqueness; in Viability Theory (Aubin, 2001) this corresponds to the "viability kernel" of A). Also $A^{\mathcal{S}}$ may be empty. Let us call the sustainable kernel of the sunny region the *manageable region:*

$$M = (X^+)^{\mathcal{S}} \supseteq S. \tag{2}$$

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

In other words, when inside M, one *can* stably stay in X^+ *by management.* Note that in the (rather uninteresting) case where $\mathcal{M}_x = \{\tau_x\}$ ("no management"), sustainability is the same as open invariance, and $M = S$.

Next, we introduce a suitable notion of stable reachability to overcome two problems with the classical notion of (plain) reachability known from control theory, where a state y is reachable from another state x iff (this abbreviation meaning "if and only if") it lies on some admissible trajectory starting at x (Sontag, 1998).

First, we want a stable fixed point y of the default dynamics to be counted as stably reachable from a (sufficiently small) neighbourhood of itself although one might only get arbitrarily close to y instead of getting to y in finite time. Second, we want stable reachability to imply that small perturbations along the way can't render the target unreachable. To solve this conceptual task in a mathematically convenient way, we define stable reachability here via the following binary relation between sets. We call an open set $C \in \mathcal{T}$ a *forecourt* for some set $Y \subseteq X$, denoted $C \rightsquigarrow Y$, iff one can approach Y arbitrarily closely from everywhere in C without leaving C, or, more precisely, iff for all $x \in C$, there is $\mu \in \mathcal{M}_x$ so that for all open sets $Z \in \mathcal{T}$ with $Z \supseteq Y$, there is $t > 0$ with $\mu(t) \in Z$ and $\mu(t') \in C$ for all $t' \in [0, t]$. Now, for a state $x \in X$ and some set $A \subseteq X$, we say that another state $y \in X$ or another set $Y \subseteq X$ are *stably reachable from x through A*, denoted $x \rightsquigarrow_A y$ or $x \rightsquigarrow_A Y$, iff x is in some subset of A that is a forecourt for $\{y\}$ or Y, respectively. The set of states from where Y can be stably reached through A is denoted $(\rightsquigarrow_A Y)$. (This is a stable version of what Aubin, 2001 would call a "capture basin" of Y.) Note that in these definitions, the order in which the logical quantifiers "for all" and "there exists" appear is critical for some of the resulting properties. If Y is open, the definitions can be somewhat simplified (see the Appendix).

Using this notion of stable reachability, we can now define the *upstream* U of the manageable system as the set of states from where the shelters S can be stably reached. Likewise, the *downstream* D consists of all states from which the manageable region

ESDD
6, 435–488, 2015

M but not the shelters can be stably reached:

$$U = (\rightsquigarrow_X S) \supseteq S, \tag{3}$$

$$D = (\rightsquigarrow_X M) - (\rightsquigarrow_X S) = (\rightsquigarrow_X M) - U \supseteq M - U. \tag{4}$$

Note that in the *no-management case,* U is basically (i.e., up to boundary effects due to our openness requirement) the basin of attraction of S, and $D = \varnothing$.

Example 1

The simple, one-dimensional toy example depicted in Fig. 1 illustrates the above four sets (and most of those we will introduce later). It has a default dynamic along the blue line downwards at a speed proportional to slope, but management is able to move upwards instead whenever the slope is small enough (on the thin blue lines in Fig. 1). The chosen undesirable region is indicated in grey. The shelter consists of the two segments just left of point a and it can be stably reached from everywhere properly left of a, hence that whole region constitutes the upstream. The manageable region is the union of the shelter with the regions denoted "lake", "glade", and "backwater" (to be explained in Sect. 2.3), and it can be stably reached from everywhere properly left of point b, hence the downstream is the right-open interval from a to b. Note that while one can stay in the lake forever by management, one cannot reach the safety of the shelter from there without crossing the dark; this kind of states is the topic of the lake dilemma discussed in Sect. 2.3.

2.2 Trenches, abysses, eddies, and the main cascade

On the other, dark end of what we will call the main cascade, we first define the *trenches* Θ as that region in the dark from which one cannot stably reach the sunny region even once,

$$\Theta = X - (\rightsquigarrow_X X^+) \tag{5}$$

ESDD
6, 435–488, 2015

(this concept approximately corresponds to the "catastrophe domains" of Schellnhuber, 1998).

Now we turn to the region from where one cannot avoid ending up in the trenches if one has to fear infinitesimal perturbations. We define the *abysses* Υ (uppercase greek letter Upsilon) as the closure of this region, minus the trenches:

$$\Upsilon = \overline{\{x \in X \,|\, \forall \mu \in \mathcal{M}_x \, \exists t \geqslant 0 : \mu(t) \in \Theta\}} - \Theta. \tag{6}$$

The closure is taken since already an infinitesimally small perturbation from a point in this closure can make the trenches unavoidable.

Finally, the *eddies* E are the remainder of X, i.e., the part from where the manageable region cannot be stably reached but the trenches can be avoided:

$$\begin{aligned} E &= X - U - D - \Upsilon - \Theta \\ &= (X - (\rightsquigarrow_X M)) \cap (X - (\Upsilon + \Theta)). \end{aligned} \tag{7}$$

Thus, when in the eddies, even though one can reach the sunny part over and over again, one cannot stay there forever but has to visit the dark repeatedly.

A connected component of Θ, Υ, or E will be called an individual *trench, abyss,* or *eddy,* and the latter two typically have sunny and dark parts.

In the *no-management case,* the trenches basically equal the invariant kernel of X^-, the abysses basically equal the rest of the basin of attraction of the trenches, and the eddies is basically the union of those trajectories that will forever alternate between X^+ and X^- (all up to subtleties on the boundaries related to stability).

The system $C = \{U, D, E, \Upsilon, \Theta\}$ is a partition of X which we call the *main cascade* because of the following mutual reachability restrictions:

$$U \not\leftrightsquigarrow_X D \not\leftrightsquigarrow_X E \not\leftrightsquigarrow_X \Upsilon \not\leftrightsquigarrow_X \Theta. \tag{8}$$

In other words, one might at best be able to go in the "downstream" direction by default or by management, from upstream to downstream to the eddies to the abysses to the trenches, but not in the other, "upstream" direction (see also Fig. 3).

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

In Example 1, Θ and Υ are nonempty, as indicated in Fig. 1, while E is empty, which is typical for systems without circular flows and with a sufficiently simply shaped X^+.

Example 2: gravity pendulum fun-ride

This simple example is based on a classical mechanical system and is designed in particular to illustrate the eddies and to introduce the lake dilemma. People sit in a fun ride resembling a gravity pendulum with angle θ and angular velocity ϖ and default dynamics given by $\dot{\theta} = \varpi$ and $\dot{\varpi} = -\sin\theta$. An optional additional clockwise acceleration of the pendulum of magnitude $a > 0$ leads to alternative admissible trajectories on which for some time interval(s) one has $\dot{\varpi} = -\sin\theta - a$. The sunny region is where $|\varpi| < \ell$, for some $\ell > 0$ representing a safety speed limit above which people might get sick.

The unique shelter is delimited by the default trajectory leading through the points $\theta = 2k\pi, \varpi = \pm\ell$, see Fig. 2. For all states on a default trajectory that has $\varpi > 0$ (counterclockwise pendulum motion) at least some of the time, there is an admissible trajectory leading into the shelter by "braking" whenever $\varpi > 0$. Hence the upstream U equals the region strictly above the default trajectory with $\varpi < 0$ that connects the unstable saddle point at $\theta = (2k+1)\pi, \varpi = 0$ with itself. This shows how the boundaries of the shelters and other regions may be found by identifying tangential or otherwise significant points and backtracing the default and alternative trajectories leading to them.

Close to that saddle point, at θ slightly larger than $(2k+1)\pi$ and $\varpi \ll 1$, there is also an admissible trajectory that stays close to there (by braking repeatedly for short intervals), hence that point is in the manageable region M. This is a typical example of how a region close to a saddle point may become manageable due to an alternative feasible trajectory with a slightly *shifted saddle point.*

However, for choices such as $a = 0.6$ and $\ell = 0.75$ (Fig. 2), there is no admissible trajectory leading from $\theta \approx (2k+1)\pi, \varpi \approx 0$ into the shelters without entering the region with $|\varpi| > \ell$. This is another instance of the lake dilemma (see Sect. 2.3).

ESDD
6, 435–488, 2015

The region below and including the default trajectory that touches the line $\varpi = -\ell$ from below is the trenches, and the region between it and the upstream is the eddies. Downstream and abysses are empty in this example.

Example 3: complete main cascade

We include this synthetic example (without figure) to show that all of the above sets may be nonempty at once. It has a circularly symmetric default dynamics in 2-D polar coordinates r, ϕ:

$$\dot{r} = f(r) = -\frac{r(r-2)(r-3)(r-5)(r-6)(r-8)(r-11)}{(9+r)^3},$$
$$\dot{\phi} = g(r) = r(r-5.5)(r-8)(r-8.5)(r-11)/100.$$

It has a stable fixed point at $r = 0$, stable limit cycles at $r \in \{3, 6, 11\}$, unstable ones at $r \in \{2, 5, 8\}$, and changes in rotational direction at $r \in \{5.5, 8.5\}$ (between limit cycles) and on the stable limit cycles at $r \in \{8, 11\}$.

We assume the management options are so that the admissible trajectories are those with $\dot{r} \in [f(r) - 1/5, f(r) + 1/5]$ and $\dot{\phi} = g(r)$, i.e., one can row only radially, with a relative speed of at most $1/5$ and arbitrarily large acceleration. For r in the intervals $R_1 \approx [.01, 1.8]$, $R_2 \approx [3.65, 4.05]$, $R_3 \approx [6.7, 7.7]$, and $R_4 \approx [11.05, \infty)$, we have $f(r) < -1/5$ so that no stopping or rowing "outwards" is possible in the corresponding rings, while rowing "inwards" is always possible. If we choose the sunny region to be the (not circularly symmetric) half-plane $X^+ = \{y = r \sin\phi > 1\}$, then the upstream U is the interior of the region outside R_3, with approx. $r > 7.7$; the downstream D is the half-open ring between the outer bounds of R_2 and R_3, with approx. $r \in (4.05, 7.7]$; the unique trench is slightly larger than the disc $r \leqslant 1$; the unique abyss is approx. the ring with $r \in (1, 1.8)$ including R_1; and the unique eddy is approx. the ring with $r \in [1.8, 4.05]$ including R_2.

ESDD
6, 435–488, 2015

2.3 Glades, the lake dilemma, backwaters, and the manageable partition

As seen in Examples 1 and 2 above, some of the states in the manageable region M may be in $U = (\rightsquigarrow_X S)$ but not in $(\rightsquigarrow_{X^+} S)$. This motivates the definition of two subsets of M via the relation of *sunny stable reachability*, $\rightsquigarrow_{X^+}$, namely (i) the *glades* G, from where the shelters can be stably reached through the sun, and (ii) the *lakes* L, from where the shelters can be stably reached only through the dark:

$$G = (\rightsquigarrow_{X^+} S) - S, \tag{9}$$

$$L = M \cap U - (\rightsquigarrow_{X^+} S) = M \cap U - S - G. \tag{10}$$

The lakes are a particularly interesting region since there one has a qualitative decision problem which we call the *lake dilemma:* shall one stay in the sun by rowing over and over again, but risking to float into the dark if the paddle breaks, or shall one move into the shelters, accepting a temporary passage through the dark, to be able to recline in safety eventually? In other words, the lake dilemma is a choice between uninterrupted desirability and eventual safety. Below we will encounter more qualitative dilemmata of this and other types.

While $\{S, G, L\}$ is a partition of $M \cap U$, also the downstream D may contain a manageable part, the *backwaters* W. This is the region where one may stay in the sun forever by rowing over and over again, but where one may not stably reach the shelters at all, not even through the dark:

$$W = M \cap D = M - U. \tag{11}$$

This completes the *manageable partition*

$$M = S + G + L + W. \tag{12}$$

Also, both U and D may contain points outside M, which we call the *dark upstream/downstream,*

$$U^- = U \cap X^-, \quad D^- = D \cap X^-, \tag{13}$$

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ESDD

6, 435–488, 2015

J. Heitzig and T. Kittel

and the *remaining sunny upstream/downstream,*

$$U^{(+)} = (U \cap X^{+}) - M, \quad D^{(+)} = (D \cap X^{+}) - M, \tag{14}$$

leading to the *upstream* and *downstream partitions*

$$U = S + G + L + U^{(+)} + U^{-}, \quad D = W + D^{(+)} + D^{-}; \tag{15}$$

see Figs. 1 and 5 for illustrations. Finally, one can divide the eddies and abysses into sunny and dark parts:

$$E^{\pm} = E \cap X^{\pm}, \quad \Upsilon^{\pm} = \Upsilon \cap X^{\pm}. \tag{16}$$

Figure 3 summarizes all the regions introduced so far.

In the (pathological) *no-management case,* some of the regions will coincide and others will be empty, and one can represent their relationship also by means of *symbolic dynamics* (beim Graben and Kurths, 2003): Assign each state x a symbolic sequence representing the sequence of its trajectory's transitions between the sunny (+) and dark (−) regions, and use the wildcard $*$ to denote repetitions of zero or more symbols. Then (up to peculiarities that may occur for boundary states) $S = M = (+)$, $U^{-} = (-)(+-)^{*}(+)$, $U^{(+)} = (+-)(+-)^{*}(+)$, $G = L = D = \varnothing$, $E^{+} = (+-)^{\infty}$, $E^{-} = (-+)^{\infty}$. $\Upsilon^{+} = (+)(-+)^{*}(-)$, $\Upsilon^{-} = (-+)(-+)^{*}(-)$, and $\Theta = (-)$.

Example 4: carbon cycle and planetary boundaries

Anderies et al. (2013) proposed a conceptual model of the global carbon cycle capturing its main features while keeping the model sufficiently low-dimensional to be able to discuss the planetary boundaries concept with it. We use their model for pre-industrial times, which has three dynamical variables c_m, c_t and $c_a = 1 - c_m - c_t$ representing the maritime, terrestrial, and atmospheric shares of the fixed global carbon stock. The dynamics is of the form

$$\dot{c}_m = a_m(c_a - \beta c_m), \quad \dot{c}_t = f(c_a, c_t) - \alpha c_t,$$

ESDD
6, 435–488, 2015

where a_m, β are diffusion parameters, f is a function representing photosynthesis and respiration, and α governs the relative human offtake from the terrestrial carbon stock.

Since the latter can be considered the natural human management option for this system, we assume the default flow has $\alpha_+ = 0.5$, while management can reduce it by half to $\alpha_- = 0.25$, which results in the trajectories shown in Fig. 4. Both have a unique stable fixed point in the interior of the state space which is globally attractive for all states with $c_t > 0$.

In order to roughly represent the planetary boundaries relating to climate change, biosphere integrity, and ocean acidification (Rockström et al., 2009b; Steffen et al., 2015), we require a "sunny" state to have sufficiently low atmospheric carbon, at least a minimum value of terrestrial carbon, and not too large maritime carbon, leading to a dark region of the shape shown in Fig. 4. If the unmanaged fixed point is sunny, one obtains an upstream situation with a shelter, a glade, and a remaining sunny upstream $U^{(+)}$ as shown in the figure. For our (quite arbitrarily) chosen parameter values, a trajectory starting in the sunny upstream is likely to first cross the climate boundary and then the biosphere boundary before getting back into the sunny region, whereas it seems quite unlikely to cross the acidification boundary.

Example 5: alternative plant types and multistability

In this fictitious example typical of a managed multistable system, two plant types 1,2 compete for some fixed patch of land, modify the soil, and are harvested. Their growth follows a logistic-type dynamics, with land cover proportions $x_{1,2} \in [0,1]$ following the equations

$$\dot{x}_1 = x_1(K_1(x_{1,2}) - x_1) - h_1 x_1,$$
$$\dot{x}_2 = r x_2(K_2(x_{1,2}) - x_2) - h_2 x_2.$$

In this, $r > 1$ is a constant productivity quotient, $h_{1,2}$ are the harvest rates, and the two dynamic capacities $K_1(x_{1,2}) = \sqrt{x_1(1-x_2)} \leqslant 1$ and $K_2(x_{1,2}) = \sqrt{x_2(1-x_1)} \leqslant 1$ represent

ESDD
6, 435–488, 2015

the fact that each type modifies the soil quickly to its own benefit but to the other type's disadvantage. On the default trajectories, both harvest rates $h_{1,2}$ equal some rather high value h_+, leading to low equilibrium harvests.

We assume management can repeatedly choose between the default and two types of alternative trajectories. One type has $h_{1,2} = h_- < h_+$, representing management by restricting harvests politically in order to yield higher long-term harvests, but without aiming to change the plant mix, as depicted in Fig. 5 (left). The other has $h_2 = 0$ and $h_1 = 2h_+$, representing management by temporarily protecting type 2 in order to change the plant mix to the higher productivity plant; we assume that this moratorium results in more intense harvesting of type 1, as depicted in Fig. 5 (right). We assume that both options exist simultaneously at all times (the separate plots of Fig. 5 are only for better discernability of the trajectories). We set the desirable region to where $x_1 + x_2 > \ell$, in order to ensure some minimum harvests.

For the choice $r = 2$, $h_+ = 0.2$, $h_- = 0.1$, $\ell = 0.65$ of the figure, the shelter S around the desirable high productivity stable fixed point of the default dynamics at $\approx (0, 0.79)$ is delimited by the default trajectory that meets the boundary to the undesired region tangentially. It can be stably reached from all states with $x_2 > 0$, hence the upstream is $U = \{(x_1, x_2) | x_2 > 0\}$. The border of the glade G can be found by backtracking the "widest" admissible trajectory that meets the boundary to the undesired region tangentially; this turns out to be a type-2 management trajectory as seen in Fig. 5 (right).

From the undesirable lower productivity stable fixed point of the default dynamics (with $h_{1,2} = h_+$) at $\approx (0.52, 0)$, one can also (and faster) get to the higher productivity stable fixed point of the first type of *managed* dynamics with $h_{1,2} = h_-$, at $\approx (0, 0.79)$, and stay there as long as management holds, so that this region is part of the manageable region M. The exact boundary of this region (which soon turns out to be a lake, L) is the "widest" admissible trajectory that meets the boundary to the undesired region tangentially; in this case, this trajectory turns out to be a type-1 management trajectory as seen in Fig. 5 (left). To get from this type 1-dominated region to the type 2-dominated shelter S via the other management option of protecting type 2, one has to cross the

ESDD
6, 435–488, 2015

undesired middle region in which both types coexist at a low level due to soil conditions that are suboptimal for both types. Hence this region is a lake.

This form of the lake dilemma can also occur in other multistable systems when one of the attractors is in the dark but sufficiently close to the sunny region so that constant management can sustain the system in a sunny place near that attractor, and when other management options may push the system towards another, sunny attractor after crossing the dark.

This example also shows that the more management options exist, the less trivial it is to find the boundaries between regions even in two-dimensional systems. For higher dimensions, one will usually have to rely on specializes numerical algorithms such as the Viability Kernel Algorithm of Frankowska and Quincampoix (1990) from viability theory.

Example 6: bifurcations of a directly manageable flow

If a system passes through a bifurcation, the classification of states by the criteria outlined above will typically change. Let us examine some typical cases that can occur in the exemplary case where management can directly affect the flow by changing the default derivative $\dot{x} = F(x)$ of a one-dimensional system by at most one unit, so that the admissible trajectories are those with $\dot{x} \in [F(x) - 1, F(x) + 1]$. (See Example 9 for the case where management is via changing a parameter instead.)

Assume $X^+ = \{|x| < \ell\}$ for some $\ell \gg 1$, and the default flow has a *subcritical pitchfork bifurcation,* say $F(x) = x^3 - rx$, where for $r > 0$ the stable fixed point $x = 0$ is surrounded by two unstable ones at $x = \pm\sqrt{r}$ and becomes unstable itself for $r \leqslant 0$, as depicted by the solid and dotted pale blue lines in Fig. 6 (top-left). Then for $r > 0$, we have a shelter-and-glade situation with a shelter $S = (-\sqrt{r}, \sqrt{r})$ and two glades $G = (-g(r), -\sqrt{r}] + [\sqrt{r}, g(r))$ where $g(r) > \sqrt{r}$ is the upper solution to $F(g(r)) - 1 = 0$. But for $r \leqslant 0$, the shelter disappears and the glades merge and are converted into a backwater $W = (-g(r), g(r))$. In both cases, this is surrounded by two sunny abysses

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

$\Upsilon^+ = (-\ell, -g(r)] + [g(r), \ell)$ and two trenches $\Theta = (-\infty, \ell] + [\ell, \infty)$. One may call this transition a *backwater/glade bifurcation.*

An interesting case is a *saddle-node bifurcation* such as the one in Fig. 6 (top-right), with $F(x) = -r - x^2$ and a critical parameter $r = 0$ at which the stable and unstable fixed points at $x = \pm\sqrt{-r}$ collide. First, at the critical point, the shelter caused by the stable fixed point and its glade are transformed into a backwater. Then, somewhat later (at $r = 1$), the maximally achievable value of $\dot{x}$ becomes negative and the backwater ceases to exist and only the sunny abyss remains. One may call this a *glade/backwater/abyss transition.*

If a stable fixed point approaches and eventually enters deeply into the dark region, this may also be called a form of "bifurcation" that causes a similar transition in the classification of states. If $F(x) = -r - x$ and $X^+ = \{x > 0\}$, as in Fig. 6 (bottom-left), then again two changes occur: At $r = 0$, the shelter-and-upstream situation of $r < 0$, with $S = (0, \infty)$ and $U^- = (-\infty, 0]$, converts into a backwater-and-downstream situation with $W = (0, \infty)$ and $D^- = (-\infty, 0]$. Then at $r = 1$, this further converts into an abyss-and-trench situation of $r \geqslant 1$ with $\Upsilon^+ = (0, \infty)$ and $\Theta = (-\infty, 0]$. One could thus call this a *shelter/backwater/abyss transition.*

Finally, a transition with three steps is caused if the fixed point passes through a narrower strip of dark, as in Fig. 6 (bottom-right), where again $F(x) = -r - x$ but now $X^+ = \{|x| > 1/4\}$. Here the shelter is again first transformed into a backwater at $r = -1/4$, but then into a lake L when the fixed point leaves the dark again at $r = +1/4$, and even later into a remaining sunny upstream $U^{(+)}$ once the maximally achievable value of $\dot{x}$ at the upper boundary of the dark, i.e., at $x = 1/4$, becomes negative. We suggest to call this a *shelter/backwater/lake/upstream transition.*

Example 7: combined population and resource dynamics

This model was used in Brander and Taylor (1998) to explain the rise and fall of the native civilization on Rapa Nui (aka Easter Island) before western contact. It is derived

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

from simple economic principles and leads to a a modified Lotka-Volterra model with a finite resource. The human population x is preying on the island's forest stock y,

$$\dot{x} = \delta x + \phi\gamma xy, \quad \dot{y} = ry(1 - y/\kappa) - \gamma xy$$

for some parameters $\gamma, \delta, \kappa, \phi, r$ representing growth and harvest rates and the stock's capacity.

We assume management will either reduce the default harvest factor γ_0 to some smaller value $\gamma_1 < \gamma_0$ to avoid overexploitation of the resource, or increase it to a larger value $\gamma_1 > \gamma_0$ to avoid famine. Our choice of the sunny region relies on two principles. The absolute population should not drop below a threshold x_{min} and the relative decline in population under the default dynamics, $-\dot{x}/x$, should not exceed a value of ℓ. Hence $X^+ = \{x > x_{min}$ and $y > y_{min} = -(\ell + \delta)/\phi\gamma_0\}$.

The resulting partition is depicted in Fig. 7 for $\phi = 4, r = 0.04$ and different choices of $\gamma_0, \gamma_1, \delta, \kappa, x_{min}, y_{min}$. The system not only displays similar transitions as discussed in Example 6 above but can also produce nonempty eddies (Fig. 7, bottom-left): one can repeatedly visit the sunny region by suitably switching between a low default harvest factor and a managed higher harvest factor, but one cannot avoid getting back into the undesired region of a low or fast declining population. Such an "optimal" management strategy would then lead to nearly periodic cycles.

Example 8: substitution of a dirty technology

In this purely economic example, a certain good (e.g., cars or electric energy) comes in two perfectly substitutable types using two different technologies, one "dirty" and one "clean" (e.g., fossil-fuel-driven and electric cars, or conventional and renewable energy), whose convex production costs C_1, C_2 decrease via learning-by-doing similar to Wright's law (Nagy et al., 2013) as $C_i(y_i) = \gamma_i y_i^{1+\sigma_i}/(1+\sigma_i)x_i^{\alpha_i}$. In this, y_i is the amount produced per-time, x_i is cumulative past production (with $\dot{x}_i = y_i$), γ_i are cost factors, $\sigma_i > 0$ are convexity parameters, and $\alpha_i > 0$ are learning exponents. We assume that

ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

demand D depends linearly on price, $D(p) = D_0 - \delta p$, $\delta > 0$, that demand equals production, $D = y_1 + y_2$, and that price equals marginal costs, $p = \partial C/\partial y_i = \gamma_i y_i^{\sigma_i}/x_i^{\alpha_i}$. One can then uniquely solve for the produced amounts y_i, getting some formula $y_i = f_i(x_1, x_2)$. This results in a two-dimensional system with state variables x_1, x_2 and equations $\dot{x}_i = f_i(x_1, x_2)$.

Let us put $D_0 = 1$, $\delta = 1$, $\sigma_i \equiv 1/5$, $\alpha_i \equiv 1/2$, and assume that the default dynamics has $\gamma_i \equiv 1$, so that the long-term default behaviour is $p(t) \to 0$, $D(t) \to 1$, and, if the dirty technology 1 is the traditional one, so that $x_1(0) > x_2(0)$, we have $x_1(t) \to \infty$, $x_2(t) \to \hat{x}_2 < \infty$, $y_1(t) \to 1$, and $y_2(t) \to 0$, i.e., usage of the clean technology 2 will die out. To depict the diverging behaviour, we used the transformation $z_i = x_i/(1 + x_i)$ in Fig. 8. Now, consider the management option of lowering γ_2 to a value of, say, 1/2 by subsidising the clean technology 2 to induce a technological change (Jaffe et al., 2002; Kalkuhl et al., 2012).

This leads to the alternative dynamics depicted in Fig. 8, showing that for some initial states with $x_1 > x_2$ one can now get $x_2(t) \to \infty$ and $y_1(t) \to 0$. The goal of keeping the usage of the dirty technology below some limit, $y_1 < \ell < 1$, corresponds to a desirable region in terms of x_1, x_2, whose border can be computed as $x_2 = x_1(1/\ell - 1 - 1/\ell^{4/5}\sqrt{x_1})^{2/5}$, see Fig. 8. That goal is automatically fulfilled in the top-left shelter region, can also be sustained by management (subsidies) in the glade region below it, and can at least be reached eventually from the remaining sunny upstream below the glade and from the dark upstream which is delimited by the management trajectory that meets the upper right corner. The region below the latter trajectory has a large dark part equalling the unique trench and a smaller sunny part in the bottom left corner that equals the unique abyss. There are no downstream or eddies in this example.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ESDD

6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

3 Finer distinction of regions w.r.t. mutual reachability of different type

If we choose a slightly larger sunny region in Fig. 5 by lowering ℓ to $\ell = 0.45$, the unmanaged fixed point with $y = 0$ gets into X^+ and the former lake around it now becomes a second shelter, which might be called a *shelter/lake transition.* But from this shelter the other, more desirable shelter can still only be reached through the dark. We will see in this section that this is an instance of what we will call the *harbour dilemma* below. An even more profound dilemma can occur when one has to choose between two shelters that cannot be reached from each other at all, not even through the dark, which we will call a *port dilemma* below. Both relate to the question of which parts or subregions of the above introduced regions may be stably reached from which other parts through the shelters, or at least through the sunny region, or at all. In order to study these questions, we'll introduce three additional, successively finer partitions derived from the stable reachability relations $\rightsquigarrow_X$, $\rightsquigarrow_{X^+}$, and $\rightsquigarrow_S$.

3.1 The ports and rapids partition and network

While from each state in U, one can stably reach some part of S, one cannot in general navigate freely inside S or U or any other member of the main cascade C. To get a formal notion of "free navigation", we say that a set $P \subseteq X$ is *portish* iff it has $x \rightsquigarrow_X y$ for all $x, y \in P$, is topologically connected, and does not intersect two different eddies, abysses, or trenches. A maximal portish set is called a *port.* We show in the Appendix that each two ports are disjoint, each port is completely contained in one of the sets $U, D, E, \Upsilon^-, \Theta$, none can intersect Υ^+, each *returnable* state (i.e., an x with $x \rightsquigarrow_X x$) is in a port, but no *transitional* state (x with $x \not\rightsquigarrow_X x$) is. In the pendulum example of Fig. 2, the returnable points are those in $U + D$ because of the periodic frictionless default flow and the possibility of counteracting small perturbations by braking or accellerating at some later point of the perturbed trajectory. In the eddies and below, this is not possible after an accelerating perturbation, hence those regions are transitional. In the plant types example of Fig. 5, there are also transitional regions, e.g. to the top and right where all

ESDD
6, 435–488, 2015

admissible trajectories lead down and left; and in the technological change example of Fig. 8 all points are transitional because of the positive growth of the knowledge stocks.

To extend the system $\mathcal{P}$ of all ports into a partition of all of X that is finer than $\mathcal{C}$, it seems natural to distinguish the non-port states w.r.t. which ports they can reach and from which ports they can be reached: we thus say that two non-port states x,y are *port-equivalent* iff they are in the same member of $\mathcal{C}$, do not lie in two different eddies, abysses, or trenches, and if $x \rightsquigarrow_X P \Leftrightarrow y \rightsquigarrow_X P$ and $P \rightsquigarrow_X x \Leftrightarrow P \rightsquigarrow_X y$ for all $P \in \mathcal{P}$. Each maximal topologically connected set of port-equivalent states is now called a *rapid*. This ensures that not only U and D are partitioned into ports and rapids, but so is each individual eddy, abyss, and trench (again, see the Appendix for a more formal treatment). Figure 3 (top) shows rapids on the left ends of the downstream and the eddy, and another one that equals the abyss.

In Example 1 (Fig. 1), the two ports are the two closed intervals where the default flow is slow, one in the upstream and one in the downstream. There is a rapid to the left of the left port and another between the left port and point a, and these two rapids are port-equivalent since they can reach the left but not the right port. Similarly, the right port is surrounded by two port-equivalent rapids. Finally, there is a singleton rapid consisting only of the point a and a last one formed by point b and all that is to the right of it; from these two port-equivalent rapids, no port can be stably (!) reached.

The ports and rapids together form the *ports and rapids partition*, $\mathcal{PR}$, which is finer than $\mathcal{C}$. It can also be interpreted as a (directed) *ports and rapids network* whose nodes are the ports and rapids and whose links represent the remaining stable reachabilities between them, thus concisely summarizing the overall structure of all management options. When there are only finitely many ports and rapids, one may drop indirect links in the network representation since reachability is transitive (if $A \rightsquigarrow B \rightsquigarrow C$, also $A \rightsquigarrow C$). Because of this transitivity, $\mathcal{P}$ is mathematically also a *quasi-ordered* set.

ESDD
6, 435–488, 2015

Example 6 revisited

Let us return to one case discussed in Example 6 and see how reachability may change under pitchfork bifurcations.

Assume the default flow has a *supercritical pitchfork bifurcation*, say $F(x) = rx - x^3$, so that at $r = 0$ the stable fixed point $x_0 = 0$ splits in two at $x_\pm = \pm\sqrt{r}$ separated by an unstable one at $x = 0$, as depicted in Fig. 6 (top-right). Then the port surrounding the stable fixed point $x = 0$, $P_0 = (-g(r), g(r))$, where $g(r)$ is the solution to $F(g(r)) + 1 = 0$, eventually also splits in three ports P_0 and $P_\pm$, separated by two rapids $R_\pm$; their borders are depicted by the dashed red lines in Fig. 6 (top-left). But this happens only at a larger value of r, namely at $r = 3/\sqrt[3]{4} \approx 1.9$, since at this point the maximum of $F(x)$ exceeds unity and the two stable fixed points $x_\pm$ can no longer be reached from each other. The corresponding ports and rapids network has these arrows: $P_- \rightsquigarrow_X R_- \rightsquigarrow_X P_0 \leftrightsquigarrow_X R_+ \leftrightsquigarrow_X P_+$. One may call this transition a *port pitchfork bifurcation*.

For the *subcritical pitchfork bifurcation* of Example 6, we get the same port pitchfork bifurcation, only that now the reachability arrows are reversed: $P_- \leftrightsquigarrow_X R_- \leftrightsquigarrow_X P_0 \rightsquigarrow_X R_+ \rightsquigarrow_X P_+$.

Example 9: bifurcations with manageable parameter

This example system is designed to illustrate the relationship of reachability and bifurcations of a dynamical system that can only be managed through its parameter. It has a two-dimensional state space $X = \{(r, y)\}$ where $y \in \mathbb{R}$ has a default dynamics

$$\dot{y} = h(y|r) = -(4 + r^2)^3 y^3 + (2r^2 - 1)(4 + r^2)y + e^r - 10$$

that cannot be managed directly, and $r \in \mathbb{R}$ is a variable with no default dynamics ($\dot{r} = 0$) which however can be managed with velocity at most 100 and arbitrarily large acceleration, leading to admissible trajectories with $\dot{r} \in [-100, 100]$ and $\dot{y} = h(y|r)$.

If r is instead interpreted as a parameter of the one-dimensional system $\dot{y} = h(y|r)$, X can be interpreted as its bifurcation space in which one can plot the loci of stable (solid

ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

lines) and unstable (dotted lines) fixed points as in Fig. 9. As one can see, there are three saddle-node bifurcations at $r_1 \approx -2.2$, $r_2 \approx 1.735$, and $r_3 \approx 4.9$ with monostable regions $r_1 < r < r_2$ and $r > r_3$, and bistable regions $r < r_1$ and $r_2 < r < r_3$. Individual and paired saddle-node bifurcations (with often result from fold bifurcations) occur frequently in bistable Earth system components such as the Atlantic Meridional Overturning Circulation (AMOC) or other tipping elements (Schellnhuber, 2009). The resulting network of ports and rapids is depicted in Fig. 10 (top).

3.2 The port dilemma

The ports and rapids partition is helpful in the discussion of a certain type of dilemma that results from two different objectives which may not be easily balanced: (i) the objective of being in or reaching a state with high *intrinsic desirability,* e.g., as measured by some qualitative preference relation finer than the mere distinction between "desirable" ("sunny") and "undesirable" ("dark"), or even by some quantitative evaluation such as a welfare function; or (ii) the objective of retaining an amount of *flexibility* as large as possible by being in or reaching a state from which a large part of state space (e.g., as measured by Lebesgue measure) is reachable. Flexibility may be important in particular in situations in which there is some uncertainty about future management options and/or future preferences (Kreps, 1979).

Fig. 10 (top) illustrates that these objectives can conflict if we assume that in this example a larger value of y has more intrinsic desirability. If one is in the left port or in one of the two rapids from which both ports can be reached, one has to decide whether to move to the right port that contains the state with the largest y, or to the left port from where one can reach the larger part of state space later on. This type of dilemma, which we call a *port dilemma* here, does not occur when one is already in the right port or in some rapid from which at most one port can be reached.

ESDD
6, 435–488, 2015

3.3 The harbours and channels partition and network

Since it does not take into account the definition of the desirable region X^+ at all, the ports and rapids partition $\mathcal{PR}$ is not fine enough to be compatible to the manageable partition $\mathcal{M}$, e.g., the backwater W in Fig. 1 contains only part of the downstream port. However, we can easily partition X^+ in a similar way, based on sunny stable reachability instead of (plain) stable reachability.

A set $H \subseteq X$ is *harbourish* iff it has has $x \rightsquigarrow_{X^+} y$ for all $x, y \in H$, is topologically connected, does not intersect two different lakes, eddies, or abysses, and does not intersect two different connected components of $S+G$. A maximal harbourish set is called a *harbour.* Let $\mathcal{H}$ be the system of all harbours.

There may be sunny states in a port which are not in any harbour. Also those states can be similarly partitioned as above: two non-harbour states $x, y \in X^+$ are *harbour-equivalent* iff they are in the same member of $\{S+G, L, U^{(+)}, W, D^{(+)}, E^+, \Upsilon^+\}$, do not lie in two different lakes, eddies, or abysses, do not lie in two different connected components of $S+G$, and if $x \rightsquigarrow_{X^+} H \Leftrightarrow y \rightsquigarrow_{X^+} H$ and $H \rightsquigarrow_{X^+} x \Leftrightarrow H \rightsquigarrow_{X^+} y$ for all $H \in \mathcal{H}$. Each maximal topologically connected set of harbour-equivalent states is now called a *channel* and lies completely in either one port or one rapid (see Appendix for proofs).

The resulting *harbours and channels partition* of X^+, $\mathcal{HC}$, is finer than $\mathcal{PR}$ and can be interpreted as another network whose links represent sunny stable reachability.

In Fig. 1, the lake consists of one harbour only, the unique upstream port contains a second harbour in $S+G$, and the sunny part of the unique downstream port forms the third harbour, while the sunny part of each rapid is a channel. For the bifurcation example, choosing $X^+ = \{y > -1/3\}$, the harbours and channels network and its relation to $\mathcal{PR}$ is depicted in Fig. 10 (middle). Note that on the leftward dashed management trajectory in the middle of the bifurcation diagram, there is a leftmost point from where one can still "turn around" and reach (if only unstably) the right part without entering the dark region; this point is a corner of the right harbour (but not belonging to it, for stability reasons), and below it is a channel leading to another harbour in the bottom-left.

ESDD
6, 435–488, 2015

J. Heitzig and T. Kittel

3.4 The harbour dilemma

The harbours and channels partition is helpful in the discussion of dilemmata involving (i) the objective of *staying* in a desirable state and (ii) the objective of eventually reaching a state with higher desirability or flexibility, called a *harbour dilemma* here.

Figure 10 (middle) illustrates such a conflict. If one is in the bottom-left harbour or in one of the two channels that can only reach that harbour, one has to decide whether to stay there, ensuring permanent desirability, or move through the dark to the eventually more desirable (and more flexible) right harbour.

The lake dilemma of Example 5 is converted to a harbour dilemma if one lowers $\ell = 0.65$ to $\ell = 0.45$, which moves the stable equilibrium at $(0.5, 0)$ into the sunny part and converts the lake into a second shelter.

3.5 The docks and fairways partition and network

Note that although $\mathcal{HC}$ is finer than $\mathcal{PR}$ and is compatible with most regions from the manageable partition, there is still one important region that can have a nontrivial intersection with harbours and channels, namely the shelters S. In order to complete our hierarchy of partitions and networks of regions, we therefore introduce a third and finest partition level, restricted to S, based on *safe stable reachability,* $\rightsquigarrow_S$.

A set $D \subseteq X$ is *dockish* iff it has $x \rightsquigarrow_S y$ for all $x, y \in D$, is topologically connected and does not intersect two different shelters. A maximal dockish set is called a *dock.* Let $\mathcal{D}$ be the system of all docks. Again, there may be sheltered states in a harbour which are not in any dock. Hence two non-dock states $x, y \in S$ are called *dock-equivalent* iff they belong to the same shelter and $x \rightsquigarrow_S D \Leftrightarrow y \rightsquigarrow_S D$ and $D \rightsquigarrow_S x \Leftrightarrow D \rightsquigarrow_S y$ for all $D \in \mathcal{D}$. Each maximal topologically connected set of dock-equivalent states is now called a *fairway* and lies completely in either one harbour or one channel.

The resulting *docks and fairways partition* of S, $\mathcal{DF}$, is finer than $\mathcal{HC}$ and can be interpreted as still another network whose links represent safe stable reachability.

Figures 1 and 10 (bottom) show examples.

ESDD
6, 435–488, 2015

3.6 The dock dilemma

Finally, the docks and fairways partition is helpful in the discussion of dilemmata involving (i) the objective of staying in a *safe* state (i.e., in the shelters) and (ii) the objective of eventually reaching a state with higher desirability or flexibility, called a *dock dilemma* here.

Figure 10 (right) illustrates such a conflict. If one is in the top-left dock (containing the top-left stable branch) or a fairway that can only reach that dock, one has to decide whether to stay there, ensuring the permanent safety of the shelter, or move through the sunny but unsafe part to an eventually more desirable or more flexible dock in the right harbour.

4 Summary of the introduced hierarchy of partitions and networks

Summarizing, we have now a hierarchy of ever-finer partitions of the system's state space at our hands. We began with the main cascade $C = \{U, D, E, \Upsilon, \Theta\}$, its refinement into the partition $\{S, G, L, U^{(+)}, U^-, W, D^{(+)}, D^-, E^+, E^-, \Upsilon^+, \Upsilon^-, \Theta\}$ (see Fig. 3), and the further refinement by topological connectedness into individual shelters, glades, lakes, backwaters, eddies, abysses, and trenches. These partitions represent the qualitative differences in stable reachability of the shelters or the manageable set, thus allow for a first classification of states w.r.t. the possibilities of sustainable management, and may reveal decision problems of the type of "lake dilemma" (e.g., Figs. 1, 2, 5), where one has to choose between *uninterrupted desirability and eventual safety.*

A different refinement of C into the ports-and-rapids network $\mathcal{PR}$ is still based on stable reachability alone but contains other details suitable for the identification and discussion of possible "port dilemmata" that involve a choice between *higher desirability and higher flexibility.* Inside the desirable region X^+, $\mathcal{PR}$ can be refined into the harbours-and-channels network $\mathcal{HC}$ suitable for the discussion of "harbour dilemmata" that involve a choice between *uninterrupted desirability and eventually higher desir-*

ESDD
6, 435–488, 2015

ability or flexibility, and further into the docks-and-fairways network $\mathcal{DF}$ suitable for the discussion of "dock dilemmata" that involve a choice between *uninterrupted safety and eventually higher desirability or flexibility* (Table 1, Fig. 10).

These three networks may also be interpreted as a three-level "network of networks" with "nodes" representing state space regions of different quality and size. A network-theoretic analysis of it using methods such as the node-weighted measures of Heitzig et al. (2011) may especially be interesting in the context of varying system parameters and bifurcations such as those in Fig. 6, but is beyond the scope of this article.

5 Conclusions

We have presented a formal classification of the possible states of a dynamical system into regions of state space which differ qualitatively in their safety, the possibilities of reaching a safe state, the possibilities of avoiding undesired states, and in the amount of flexibility for future management.

Based on an assumed main division of the systems states into only two classes, desirable ("sunny") and undesirable ("dark"), we have constructed a hierarchy of partitions of a system's state space, whose member regions we suggested to name by metaphorical names either corresponding to the general image of a boat floating or rowing on a complex water system, such as "upstream", "downstream", "eddy", "abyss", "trench", "lake", and "backwater", or corresponding to the image of a "shelter" surrounded by a "glade". To capture the nature of and relationships between the different regions, we have introduced the notion of stable reachability and the corresponding three-level reachability network of "ports", "harbours", "docks", "rapids", "channels", and "fairways", and illustrated our concepts with conceptual example models from classical mechanics, climate science, ecology, economics, and coevolutionary Earth system modelling. Most of the different regions can readily be found in most models for either most or at least selected parameter settings. A notable exception are the "eddies" which, due to their circular feature, can be expected to occur much more rarely in real-world, nonconservative systems, especially when thermodynamic or otherwise irreversible processes are involved, such as soil degradation. Example 7 however illustrates how eddies may occur in coevolutionary systems and might incentivize management cycles that lead to undampened periodic ups and downs. It must remain an open question here whether this effect might be an additional explanation for empirically observable cycles such as business or resource cycles when management is involved.

The introduced concepts have then been used to point out a number of qualitatively different decision problems, the "lake, port, harbour, and dock dilemmata". In our opinion, one particularly nasty form of decision problem is the lake dilemma, where one has to choose between uninterrupted desirability and eventual safety, and Example 5 indicates that this dilemma may easily occur at least in ecological systems or other multistable systems with a sunny attractor and another one slightly in the dark. Since the transformation of socio-metabolic processes or complex industrial production systems may resemble the soil transformation of Example 5, one may also expect the lake dilemma to occur in the socio-metabolic and economic subsystems of the Earth. The form of lake seen near the saddle point in the pendulum Example 2 can also occur in other nonlinear oscillators, e.g. the Duffing oscillator or models of glacial cycles that resemble it such as Saltzman et al. (1982) and Nicolis (1987), when a management option exists that has a slightly shifted saddle point. This indicates that the lake dilemma may also occur in purely physical subsystems of the Earth system.

We believe that our concepts may be especially useful in the context of the current debate about planetary boundaries, a possible safe and just operating space for humanity, and the necessary socio-economic transitions to reach it or stay in it. We suggest that the region delimited by some identified set of planetary boundaries in the sense of Rockström et al. (2009a) and Steffen et al. (2015) and some similar socio-economic limits, e.g. those relating to the Millenium sustainable development goals (Raworth, 2012), should be interpreted in our framework as a natural choice for the desirable region X^+, although their definitions already contain some reasoning about the further dynamical consequences of violating the boundaries. From these bound-

ESDD
6, 435–488, 2015

aries, one may then try to identify one or more smaller shelter regions S that can be considered a safe operating space in the sense that, once there, no further large-scale management in the form of global policies is necessary to stay within the limits for all times (or at least for a sufficiently long planning horizon).

If it turns out that the current state of the Earth is outside the shelters, one should then aim next at trying to decide whether it is in the upstream. If so, knowledge about whether it is in a glade or lake or not, and which safe docks can be stably reached will be necessary in order to choose a management path. In the lake case, this choice would first require a decision about whether a temporary violation of the limits can be justified by the eventual safety. In addition, a port dilemma may necessitate a decision between higher desirability and higher flexibility at this point. Only after these qualitative decisions it seems advisable to optimize the chosen type of management pathway by means of more traditional control and optimization theory, hopefully using accurate enough quantitative estimates of the involved options, costs, and benefits. Once in the shelters, one may start caring about improving the state further by moving between docks to either improve desirability or flexibility, but this may require a risky temporary passage through a sunny but unsafe region (which poses a dock dilemma) or even a passage through the dark (which poses a harbour dilemma).

If we are not in the "upstream" of the Earth system, prospects are worse. Violating the limits can then only be avoided by management, either eventually forever (if in the downstream), or only repeatedly but with repeated violations occuring (if in the eddies), or even only for a limited time with an ultimate descent into the undesired region (if in the abysses or already in the trench). In these last cases, one may either attempt at repeating the analysis with a less ambitious, "second best" definition of the desirable region X^+, or simply revert to quantitative optimization, e.g., to minimize some damage function.

We hope that the theoretical considerations outlined here may be of some help to sharpen the important debate of how a transition to a safe desirable state of the Earth system can be managed.

ESDD
6, 435–488, 2015

Appendix A: Assumptions, notation

We use sloppy set theoretic notation when no confusion arises: union $A+B=A\cup B$, difference $A-B=A\setminus B$, power set $2^A=\{B\subseteq A\}$. For a more formal treatment than in the main text, we assume the following:

A *state space* $X\neq 0$ with some Hausdorff topology $\mathcal{T}\subseteq 2^X$ (i.e. a system of open sets that separate each two points) on it whose elements we call *states* or *points* (e.g. $X\subseteq\mathbb{R}^n$ with Euclidean topology). X may be compact or unbounded, finite- or infinite-dimensional, etc.

A flow (= deterministic continuous-time autonomous dynamical system) on X (e.g. a model of human-nature coevolution or any other Earth system model) given by a family of continuous ("business-as-usual" or) *default trajectories* $\tau_x:[0,\infty)\to X$ with $\tau_x(0)=x$ and $\tau_{\tau_x(t)}(t')=\tau_x(t+t')$ for all initial conditions $x\in X$ and all relative time points $t,t'\geqslant 0$. We don't require further smoothness properties of the flow, like differentiability, to avoid having to assume a richer topological structure for X than just a general topological space, and to avoid unnecessarily complicated notions and familiarity with, e.g., differential geometry. Although flows are often represented by ordinary differential equations, their solutions are sometimes not unique, hence our notion of flow is in terms of trajectories instead, to allow us to distinguish, e.g., a 1-D flow with $\dot{x}=\sqrt{x}$ and $\tau_0(t)\equiv 0$ from the flow that has also $\dot{x}=\sqrt{x}$ but $\tau_0(t)=t^2/4$.

An open nonempty set $X^+\in\mathcal{T}$ of desirable states, called the *sunny region,* e.g. defined by means of some notion of "tolerable E&D window" (Schellnhuber, 1998). We call the complement $X^-=X-X^+\neq 0$ the *dark (region).* We require openness for convenience so that infinitesimal perturbations can't lead from sunny to dark part, and trajectories cannot touch the sunny region without entering it for a strictly positive amount of time. Although in most of our examples, X^+ is a simply shaped, connected, convex, and often compact set, none of these properties is required for the theory presented here except openness.

ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

A family of nonempty sets $\mathcal{M}_x$ of admissible trajectories from each $x \in X$ that includes τ_x and is closed under switching, i.e., if $\mu \in \mathcal{M}_x$, $t > 0$, $x' = \mu(t)$, and $\mu' \in \mathcal{M}_{x'}$, then the trajectory defined by $\mu''(t'') = \mu(t)$ for $t'' \leqslant t$ and $\mu''(t'') = \mu'(t'' - t)$ for $t'' > t$ is also in $\mathcal{M}_x$. This requirement corresponds to the so-called "semigroup" axiom of mathematical control theory (Sontag, 1998). Note that we do not allow any explicit time dependency of flow or management, but such dependencies can as usual be encoded by including time as a state variable. Also, if management can change a parameter of the model, that parameter has to be transformed to a (slow) state variable with "zero" default dynamics of its own to meet our framework.

Appendix B: Further properties and proofs

The proofs only require an understanding of general topological spaces, in particular of openness and continuity, but not of any higher-level concepts from differential topology or the like.

Proposition 1 (Existence and uniqueness)
For all $A \subseteq X$:

1. There is a unique largest (default-trajectory-) invariant and open subset $A^{\iota\circ} \subseteq A$, containing all other such sets.

2. Every invariant and open set is sustainable. In particular, S is.

3. There is a unique largest sustainable subset $A^S \subseteq A$, and $A^S \supseteq A^{\iota\circ}$ containing all other such sets.

Proof.

1. Let $\mathcal{I}(A)$ be the system of all open subsets $B \subseteq A$ for which $\tau_x(t) \in B$ for all $x \in B, t > 0$. The proposition is proved by showing that $\mathcal{I}(A)$ is a *kernel system,* i.e.,

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

contains the empty set (which is trivial) and contains the union $\bigcup \mathcal{B}$ of any of its subsets $\mathcal{B} \subseteq \mathcal{I}(A)$. The latter follows from the fact that the system of all open sets, $\mathcal{T}$, is a kernel system by definition, and if $x \in \bigcup \mathcal{B}$, then $x \in B \in \mathcal{B}$, hence $\tau_x(t) \in B \subseteq \bigcup \mathcal{B}$ for all $t > 0$. Now $A^{\iota\circ} = \bigcup \mathcal{I}(A) \in \mathcal{I}(A)$.

2. Since $\tau_x \in \mathcal{M}_x$.

3. Similarly, the system $\mathcal{S}(A)$ of all sustainable subsets $B \subseteq A$ is a kernel system: if $x \in \bigcup \mathcal{B}$, then $x \in B \in \mathcal{B}$, hence there is $\mu \in \mathcal{M}_x$ with $\mu(t) \in B \subseteq \bigcup \mathcal{B}$ for all $t > 0$. Now $A^{\mathcal{S}} = \bigcup \mathcal{S}(A) \in \mathcal{S}(A)$. 2. implies $A^{\mathcal{S}} \supseteq A^{\iota\circ}$.

Q.E.D.

Proposition 2 (Stable reachability)
For all $A, A', C, Y, Z \subseteq X$ and $x, y, z \in X$:

1. *If Y is open, then (i) $C \rightsquigarrow Y$ iff for all $x \in C$, there is $\mu \in \mathcal{M}_x$ so that there is $t > 0$ with $\mu(t) \in Y$ and $\mu(t') \in C$ for all $t' \in [0, t]$; and (ii) $x \rightsquigarrow_A Y$ iff there is and open $C \subseteq A$ with $x \in C$ and for all $x' \in C$, there is $\mu \in \mathcal{M}_{x'}$ so that there is $t > 0$ with $\mu(t) \in Y$ and $\mu(t') \in C$ for all $t' \in [0, t]$.*

2. *If $x \rightsquigarrow_A Y$, then x is in the interior (= largest open subset) of A, A°, and there is an open set $B \ni x$ with $x' \rightsquigarrow_A Y$ for all $x' \in B$. Hence, each $(\rightsquigarrow_A Y)$ is open.*

3. *Transitivity:*

$$x \rightsquigarrow_A y \rightsquigarrow_{A'} Z \implies x \rightsquigarrow_{A+A'} Z,$$
$$x \rightsquigarrow_A y \rightsquigarrow_{A'} z \implies x \rightsquigarrow_{A+A'} z.$$

   In particular, $\rightsquigarrow_A$ is a transitive (but not necessarily reflexive) relation.

4. *If A is open, it is stably reachable from each of its elements. In particular, since $S = (X^+)^{\iota\circ} \subseteq (X^+)^{\mathcal{S}} = M$ is open, S is also included in $U = (\rightsquigarrow_X S)$.*

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

Proof.

1. (i) Assume $C \rightsquigarrow Y \in \mathcal{T}$ and let $x \in C$. Then, by definition of forecourts, there is $\mu \in \mathcal{M}_x$ so that for all open sets $Z \in \mathcal{T}$ with $Z \supseteq Y$, there is $t > 0$ with $\mu(t) \in Z$ and $\mu(t') \in C$ for all $t' \in [0,t]$. Since Y is open, it is such a Z, proving the first direction.

   For the other direction, assume that for all $x \in C$, there is $\mu \in \mathcal{M}_x$ so that there is $t > 0$ with $\mu(t) \in Y$ and $\mu(t') \in C$ for all $t' \in [0,t]$. Let $x \in C$, choose such a $\mu \in \mathcal{M}_x$ and $t > 0$, and let $Z \in \mathcal{T}$ with $Z \supseteq Y$ be an open set. Then $\mu(t) \in Y \subseteq Z$ as required.

   (ii) By definition of stable reachability, $x \rightsquigarrow_A Y$ iff there is an open $B \subseteq A$ with $x \in B \rightsquigarrow Y$. By (i), $B \rightsquigarrow Y$ iff for all $x' \in B$, there is $\mu \in \mathcal{M}_{x'}$ so that there is $t > 0$ with $\mu(t) \in Y$ and $\mu(t') \in B$ for all $t' \in [0,t]$.

2. Assume $x \rightsquigarrow_A Y$. Then $x \in X$ for some open $B \subseteq A$, hence $x \in B \subseteq A^\circ$. Also, $B \rightsquigarrow Y$ hence $x' \rightsquigarrow_A Y$ for all $x' \in B$. Hence $(\rightsquigarrow_A Y)$ contains an open neighbourhood of each of its points and is thus open itself.

3. We show this by concatenating suitably chosen admissible trajectories between points close to x, y, Z. Let $x \rightsquigarrow_A y \rightsquigarrow_{A'} Z$, choose open sets $B \subseteq A$, $B' \subseteq A'$ with $x \in B \rightsquigarrow \{y\}$ and $y \in B' \rightsquigarrow Z$, and put $B'' = B + B' \subseteq A + A'$, then $x \in B''$ and B'' is open. To show that $B'' \rightsquigarrow Z$, we let $x'' \in B''$ and show that there is $\mu \in \mathcal{M}_{x''}$ so that for all open $W'' \supseteq Z$, there is $t > 0$ with $\mu(t) \in W''$ and $\mu(t') \in B''$ for all $t' \in [0,t]$.

   If $x'' \in B'$, there is such a μ with $\mu(t') \in B' \subseteq B''$ for all $t' \in [0,t]$ since $B' \rightsquigarrow Z$.

   If $x'' \notin B'$ instead, $x'' \in B \rightsquigarrow \{y\}$, hence we find $\nu \in \mathcal{M}_{x''}$ so that for all open $W \supseteq \{y\}$, there is $t > 0$ with $\nu(t) \in W$ and $\nu(t') \in B$ for all $t' \in [0,t]$. Since B' is such a W, we find $t'' > 0$ with $\nu(t'') \in B'$ and $\nu(t') \in B$ for all $t' \in [0,t'']$. For $y' = \nu(t'') \in B' \rightsquigarrow Z$, we then find $\nu' \in \mathcal{M}_{x''}$ so that for all open $W'' \supseteq Z$, there is $t > 0$ with $\nu'(t) \in W''$ and $\nu'(t') \in B'$ for all $t' \in [0,t]$. Now define μ by putting $\mu(t') = \nu(t')$ for $t' \in [0,t'']$ and $\mu(t') = \nu'(t'-t'')$ for $t' \geq t''$. Then $\mu \in \mathcal{M}_{x''}$ because of our assumptions on

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ESDD

6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

$\mathcal{M}$, and for all open $W'' \supseteq Z$, there is $t > 0$ with $v'(t) \in W''$ and $v'(t') \in B + B' = B''$ for all $t' \in [0,t]$, as required.

The z case follows from putting $Z = \{z\}$. Transitivity is the special case of $A' = A$.

4. For $x \in A \in \mathcal{T}$, we show $x \rightsquigarrow_A A$ by showing $A \rightsquigarrow A$. Let $x' \in A$. By 1., we have to find $\mu \in \mathcal{M}_{x'}$ and $t > 0$ with $\mu(t') \in A$ for all $t' \in [0,t]$. Since A is open and $\tau_{x'}$ is continuous, $\tau_{x'}$ is such a μ.

Q.E.D.

Proposition 3 (Main cascade)

1. *$U = (\rightsquigarrow_X S)$ and $D + U = (\rightsquigarrow_X M)$ are open, $\Theta = X - (\rightsquigarrow_X X^+)$ and $\mathrm{Y} + \Theta$ are closed, $E + D + U = X - \mathrm{Y} - \Theta$ is open, and $\{U, D, E, \mathrm{Y}, \Theta\}$ form a partition of X.*
2. *For all $u \in U, d \in D, e \in E, y \in \mathrm{Y}, \theta \in \Theta$, we have $u \not\leftrightsquigarrow_X d \not\leftrightsquigarrow_X e \not\leftrightsquigarrow_X y \not\leftrightsquigarrow_X \theta$.*
3. *If $W = \varnothing$, also $D = \varnothing$.*

Proof.

1. Openness follows from Prop. 2, the partition covers X by definition of E, and the only nontrivial disjointness is that between the open set $D + U = (\rightsquigarrow_X M)$ and the closed set $\mathrm{Y} + \Theta = \overline{\{x \in X \,|\, \forall \mu \in \mathcal{M}_x \exists t \geqslant 0 : \mu(t) \in \Theta\}}$. If x is in both sets, there is also $x' \in (\rightsquigarrow_X M) \cap \{x \in X \,|\, \forall \mu \in \mathcal{M}_x \exists t \geqslant 0 : \mu(t) \in \Theta\}$, but then there is $\mu'_x \in \mathcal{M}_x$, $t' > 0$ with $\mu'_x(t') \in M$, and by definition of M there is then also some $\mu \in \mathcal{M}_x$ with $\mu(t) \in X^+$ for all $t \geqslant t'$. But, by assumption, there is $t \geq 0$ with $\mu(t) \in \Theta$. Since $\Theta \cap X^+ = 0$, we have $t < t'$, but by definition of Θ, this contradicts $\mu(t') \in X^+$. Hence such an x cannot exist.
2. Because of trasitivity and 1., $d \rightsquigarrow_X u \in U = (\rightsquigarrow_X S)$ would imply $d \rightsquigarrow_X S$ and thus $d \in U \cap D = \varnothing$; $e \rightsquigarrow_X d \in D = (\rightsquigarrow_X M) - U$ would imply $e \rightsquigarrow_X M$ and thus $e \in (U +$

ESDD
6, 435–488, 2015

$D) \cap E = \varnothing$. If one could reach the eddies from the abysses, one could avoid the trenches: Assume $y \rightsquigarrow_X e \notin \Upsilon + \Theta = \overline{\{x \in X \,|\, \forall \mu \in \mathcal{M}_x \, \exists t \geqslant 0 : \mu(t) \in \Theta\}}$. Since the latter is closed, its complement is open, so there is $\mu \in \mathcal{M}_y$ and $t > 0$ with $\mu(t) \notin \Upsilon + \Theta$. For $x = \mu(t)$, we find $\mu' \in \mathcal{M}_x$ and $t'' > 0$ with $\mu'(t') \notin \Theta$ for all $t' > t''$. Concatenating μ with μ' gives a similar member of $\mathcal{M}_y$, in contradiction to $y \in \Theta$. Finally, if $\theta \rightsquigarrow_X y$ and $\theta \in \Theta$, then $y \in \Theta$ by definition of Θ, hence $y \notin \Upsilon$.

3. This follows from $D = (\rightsquigarrow_X M) - U = D = (\rightsquigarrow_X W)$.

Q.E.D.

Proposition 4 (Ports, rapids, harbours, etc.)

1. *Each two ports (or harbours or docks) are disjoint.*
2. *Each port lies completely in one of $U, D, E, \Upsilon^-, \Theta$, no port intersects Υ^+.*
3. *Each harbour (or dock) lies completely in one port (or harbour).*
4. *Each channel (or fairway) lies completely in one member of $\mathcal{PR}$ (or $\mathcal{HC}$).*
5. *These partitions are successive refinements of each other: $\mathcal{C}, \mathcal{PR}, \mathcal{HC}, \mathcal{DF}$.*
6. *If a harbour H intersects some of the regions $S+G$, L, U^+, W, or D^+, it is already completely contained in that region.*

ESDD
6, 435–488, 2015

Proof.

1. Assume $y \in A \cap A'$ for two different maximal portish (or harbourish or dockish) sets A, A' and put $B = A + A'$. But then B is itself portish (or harbourish or dockish) because stable reachability is transitive. This contradicts the maximality of A and A'.

2. By Prop. 3 if $x \rightsquigarrow_P y \rightsquigarrow_P x$ then x and y they must belong to the same member of $\mathcal{C}$, hence each port lies completely in one of them.

   To show that a port $P \subseteq \mathrm{Y}$ is already in Y^-, assume $x \in P \cap \mathrm{Y}^+ \subseteq X^+ \in \mathcal{T}$. We will now construct a contradiction by constructing an admissible trajectory from x that avoids Θ forever. Since $x \rightsquigarrow_X x$ and X^+ is open, there is an open set $A \subseteq X^+$ with $y \rightsquigarrow_X x$ for all $y \in A$. Since τ_x is continuous and A open, we find $t_0 > 0$ with $\tau_x(t) \in A$ for all $t \in [0, t_0]$. Let $y = \tau_x(t_0)$ and pick a $\mu \in \mathcal{M}_y$ that returns arbitrarily closely to x. Let $\mathcal{A}$ be the set of all open $A \subseteq X^+$ with $x \in A$, and choose a $t_A > 0$ with $\mu(t_A) \in A$ for all $A \in \mathcal{A}$ (this requires the Axiom of Choice which we will assume here). Let $t_1 = \inf_{A \in \mathcal{A}} \sup_{B \in \mathcal{A}, B \subseteq A} t_B \geqslant 0$. Since $y \in \mathrm{Y} + \Theta$, there is $t' > 0$ with $\mu(t'') \in \Theta$ for all $t'' > t'$, hence $t_A \leqslant t'$ for all $A \in \mathcal{A}$ and thus $t_1 \leqslant t'$. Next we show that $\mu(t_1) = x$. If $\mu(t_1) = y \neq x$, one can choose $A \in \mathcal{A}$ and $C \in \mathcal{T}$ with $y \in C$ and $A \cap C = \varnothing$ (this is the only point where we need the Hausdorff property). Since μ is continuous, there are $t_l < t_1$ and $t_u > t_1$ with $\mu(t') \in C$ for all $t' \in [t_l, t_u]$. By definition of t_1, there is $A' \in \mathcal{A}$ with $\sup_{B \in \mathcal{A}, B \subseteq A'} t_B \in [t_1, t_u]$. Putting $A'' = A \cap A' \in \mathcal{A}$, we then also have $\sup_{B \in \mathcal{A}, B \subseteq A''} t_B \in [t_1, t_u]$, hence there is $B \in \mathcal{A}$ with $B \subseteq A'' \subseteq A$ and $t_B \geqslant t_l$ and hence $\mu(t_B) \in C$ by choice of t_l. But $\mu(t_B) \in B \subseteq A$ by choice of t_B. Hence $\mu(t_B) \in A \cap C = \varnothing$, a contradiction. So $\mu(t_1) = x$ after all. Finally we concatenate $\tau_x[0, t_0]$ and $\mu[0, t_1]$ infinitely many times and get an admissible trajectory from x that avoids Θ forever.

3. Since $\rightsquigarrow_S$ refines $\rightsquigarrow_{X^+}$, which refines $\rightsquigarrow_X$.

ESDD
6, 435–488, 2015

J. Heitzig and T. Kittel

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

4. Since dock-equivalence refines harbour-equivalence, which refines port-equivalence.

5. Follows from 2.–4.

6. This follows directly from the definitions of $S+G, L, U^+, W$, and D^+ by means of $\rightsquigarrow_X$ and $\rightsquigarrow_{X^+}$ and the transitivity of those relations.

Q.E.D.

Remarks:

- In general, $A^{\iota\circ}$ may be properly smaller than both the interior $(A^\iota)^\circ$ of the largest invariant subset A^ι of A and the largest invariant subset of A°, $(A^\circ)^\iota$. The three sets can only be shown to be equal under additional smoothness assumptions on τ and $\mu \in \mathcal{M}_x$.

- The set of all states that are stably reachable from x need not be closed or open and need not contain any of the intermediate states that lie on the trajectories $\mu \in \mathcal{M}_x$ used in stable reachability.

- $x \rightsquigarrow_A Y$ does not imply $x \rightsquigarrow_A y$ for any $y \in Y$, since, after a perturbation, other points in Y may be reachable than before.

- For two points x, y in the same port, harbour, or dock A, one may still not have $x \rightsquigarrow_A y$ since the intermediate states on the trajectories from x to y may not be *stably* reachable from x and thus may not belong to A. In other words, perturbations may still push the system temporarily out of a port, harbour, or dock (as can be seen in Example 9), but one can then return to the same port, harbour, or dock. For this reason, the directed reachability network is typically acyclic but may contain reachability cycles in pathological situations.

ESDD
6, 435–488, 2015

– The scope of possible connection topologies that may occur as the reachability network of a managed system contains at least all acyclic finite or countably infinite directed graphs, as can be seen by the following construction: given an acyclic directed graph, one can construct a topologically equivalent network of water bowls which are connected by water tubes leading from a dedicated "drain" at the bottom of the source ball to a common entrance at the top of the target ball. Let water flow into all balls without incoming tubes and out of all outgoing tubes through grilles, determining the default dynamics of a small submarine floating in the water. Then assume the submarine can be propelled strongly enough to move freely inside each ball and to each drain, but not strongly enough to leave the ball through the entrance at the top, against the direction of the water flow. By making parts of the balls and tubes opaque and moving some of the drains from the bottom to the sides of the ball, the construction can be extended to show that also all internally consistent three-level acyclic networks can occur as the three-level network of ports, harbours, and docks.

Appendix C: Relationship to viability theory

The vast mathematical literature on viability theory (VP), summarized in (Aubin, 2009; Aubin et al., 2011), also treats the question of which regions of state space can be reached from which others when a system's dynamics has some additional degrees of freedom that may represent unknown internal components such as human behaviour, or unknown external drivers, or options for management or control.

Its fundamental concepts of *viable domain, viability kernel,* and *capture basin* correspond to our notions of *sustainable set, sustainable kernel,* and sets of the form $\rightsquigarrow_K A$, but the concepts differ in that we require these sets to be topologically open, to account for possible infinitesimal perturbations. In VP, these and other sets are usually required to be closed instead, and while this has some advantages for proving deep results such as certain convergence properties, it also requires VP to focus on a more restric-

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

tive class of systems (differential inclusions and/or Marchaud maps, vector spaces as state spaces) than we do. While our purely topological existence proof only relies on the fact that the sustainable sets form a kernel system, the proof that a viability kernel exists is harder and requires additional smoothness assumptions on the system of possible trajectories.

On the other hand, we have added the distinction between default and alternative trajectories here to be able to talk about the consequences of having to manage a system only temporarily or repeatedly. Consequently, our notion of *shelter* has no counterpart in standard VP, and our notion of *invariance* differs from theirs since it refers to the default dynamics only.

Similarly, our notion of *stable reachability* differs in two important ways from VP's reachability: on the one hand, we require it to be "safe" against infinitesimal perturbations, on the other, we allow a trajectory to need infinite time to reach a target exactly (which does not count as reachable in VP) if it can reach arbitrarily small neighbourhoods of the target in finite time, so that in our theory, asymptotically stable fixed points are reachable via the default dynamics. This difference can easily be seen in a slightly changed version of Fig. 6 (top-right): Assume $\dot{x} = -r - x^2$ and $\dot{r} \in [-1,0]$, i.e., management can only move to the left. While in our theory, the stable branch is stably reachable from below, it is not so in VP since that takes infinite time.

Despite these differences, algorithms such as the tangent method and the viability kernel algorithm by Frankowska and Quincampoix (1990) are quite helpful in our context, too, and we have the following approximate correspondences: $U \approx$ capture basin of S; $M \approx$ viability kernel of X^+; $U + D \approx$ capture basin of M; $E^+ + Y^+ \approx$ the "shadow" of X^+; and $\Theta \approx$ "invariance kernel" of X^-. In the reachability network of networks, the union of ports and rapids "between" two given ports P, P' (and similarly for harbours and docks) corresponds to what is called a "connection basin" between P and P' in VP.

Acknowledgements. This work was conducted in the framework of PIK's flagship project on Coevolutionary Pathways (COPAN). This paper was developed within the scope of the IRTG 1740/TRP 2011/50151-0, funded by the DFG/FAPESP. We thank Wolfram Barfuss,

ESDD
6, 435–488, 2015

Nicola Botta, Jonathan Donges, Vera Heck, Ulrike Kornek, Jürgen Kurths, Wolfgang Lucht, Finn Müller-Hansen, John Schellnhuber, Ricarda Winkelmann and the COPAN team for fruitful discussions and helpful comments.

References

Anderies, J. M., Carpenter, S. R., Steffen, W., and Rockström, J.: The topology of non-linear global carbon dynamics: from tipping points to planetary boundaries, Environ. Res. Lett., 8, 044048, doi:10.1088/1748-9326/8/4/044048, 2013. 448, 482

Aubin, J.-P.: Viability kernels and capture basins of sets under differential inclusions, SIAM J.Control Optim., 40, 853–881, doi:10.1137/S036301290036968X, 2001. 441, 442

Aubin, J.-P.: Viability Theory, Birkhäuser, Boston, 2009. 438, 473

Aubin, J.-P. and Saint-Pierre, P.: An introduction to viability theory and management of renewable resources, in: Advanced Methods for Decision Making and Risk Management in Sustainability Science, edited by: Kropp, J. and Scheffran, J., chap. 2, Nova Science Publishers, New York, 43–80, 2007. 438

Aubin, J.-P., Bayen, A., and Saint-Pierre, P.: Viability Theory. New Directions, Springer-Verlag Berlin Heidelberg, 2011. 438, 473

beim Graben, P. and Kurths, J.: Detecting subthreshold events in noisy data by symbolic dynamics, Phys. Rev. Lett., 90, 100602, doi:10.1103/PhysRevLett.90.100602, 2003. 448

Botta, N., Jansson, P., and Ionescu, C.: A computational theory of policy advice and avoidability, in preparation, 2015. 438

Brander, J. A. and Taylor, M. S.: The simple economics of Easter Island: a Ricardo-Malthus model of renewable resource use, Am. Econ. Rev., 88, 119–138, 1998. 452

Frankowska, H. and Quincampoix, M.: Viability kernels of differential inclusions with constraints: algorithms and applications, International Institute for Applied Systems Analysis – Working Paper, 1990. 451, 474

Heitzig, J., Donges, J. F., Zou, Y., Marwan, N., and Kurths, J.: Node-weighted measures for complex networks with spatially embedded, sampled, or differently sized nodes, Eur. Phys. J. B, 85, 1–22, doi:10.1140/epjb/e2011-20678-7, 2011. 462

Jaffe, A., Newell, R., and Stavins, R.: Environmental policy and technological change, Environ. Resour. Econom., 22, 41–69, 2002. 454

ESDD
6, 435–488, 2015
Topology of sustainable management
J. Heitzig and T. Kittel

Kalkuhl, M., Edenhofer, O., and Lessmann, K.: Learning or lock-in: optimal technology policies to support mitigation, Resour. Energy Econ., 34, 1–23, 2012. 454

Kleidon, A. and Renner, M.: A simple explanation for the sensitivity of the hydrologic cycle to surface temperature and solar radiation and its implications for global climate change, Earth Syst. Dynam., 4, 455–465, doi:10.5194/esd-4-455-2013, 2013. 437

Kreps, D. M.: A representation theorem for "Preference for flexibility", Econometrica, 47, 565–577, doi:10.2307/1910406, 1979. 458

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, P. Natl. Acad. Sci. USA, 105, 1786–1793, 2008. 438

Nagy, B., Farmer, J. D., Bui, Q. M., and Trancik, J. E.: Statistical basis for predicting technological progress, PLoS ONE, 8, 1–7, doi:10.1371/journal.pone.0052669, 2013. 453

Nicolis, C.: Long-term climatic variability and chaotic dynamics, Tellus A, 39, 1–9, doi:10.3402/tellusa.v39i1.11734, 1987. 463

Raworth, K.: A safe and just space for humanity: can we live within the doughnut?, Oxfam Policy and Practice: Climate Change and Resilience, 8, 1–26, doi:10.5822/978-1-61091-458-1, 2012. 437, 463

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, III, F. S., Lambin, E., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H., Nykvist, B., De Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J.: Planetary boundaries: exploring the safe operating space for humanity, Ecol. Soc., 14, 32, available at: http://www.ecologyandsociety.org/vol14/iss2/art32/, 2009a. 437, 463

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J. A.: A safe operating space for humanity, Nature, 461, 472–475, doi:10.1038/461472a, 2009b. 437, 449

Saltzman, B., Sutera, A., and Hansen, A. R.: A possible marine mechanism for internally generated long-period climate cycles, J. Atmos. Sci., 39, 2634–2637, doi:10.1175/1520-0469(1982)039<2634:APMMFI>2.0.CO;2, 1982. 463

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

Schellnhuber, H. J.: Discourse: Earth system analysis – the scope of the challenge, in: Earth System Analysis: Integrating Science for Sustainability, edited by: Schellnhuber, H. J. and Wenzel, V., chap. 1, Springer, Berlin/Heidelberg, 3–195, doi:10.1007/978-3-642-52354-0_1, 1998. 437, 439, 444, 465

Schellnhuber, H. J.: Tipping elements in the Earth system, P. Natl. Acad. Sci. USA, 106, 20561, doi:10.1073/pnas.0705414105, 2009. 438, 458

Sontag, E. D.: Mathematical Control Theory: Deterministic Finite Dimensional Systems, 2nd Edn., Springer Science + Business Media New York, 1998. 442, 466

Steffen, W., Richardson, K., Rockström, J., Cornell, S., Fetzer, I., Bennett, E., Biggs, R., Carpenter, S. R., de Wit, C. A., Folke, C., Mace, G., Persson, L. M., Veerabhadran, R., Reyers, B., and Sörlin, S.: Planetary boundaries: guiding human development on a changing planet, Science, 347, 6223, doi:10.1126/science.1259855, 2015. 437, 449, 463

Table 1. Preview of dilemma types discussed in the article.

Name	Option 1	Option 2
“Lake” dilemma	uninterrupted desirability	eventual safety
“Port” dilemma	higher desirability	higher flexibility
“Harbour” dilemma	uninterrupted desirability	eventually higher desirability or flexibility
“Dock” dilemma	uninterrupted safety	eventually higher desirability or flexibility

ESDD
6, 435–488, 2015

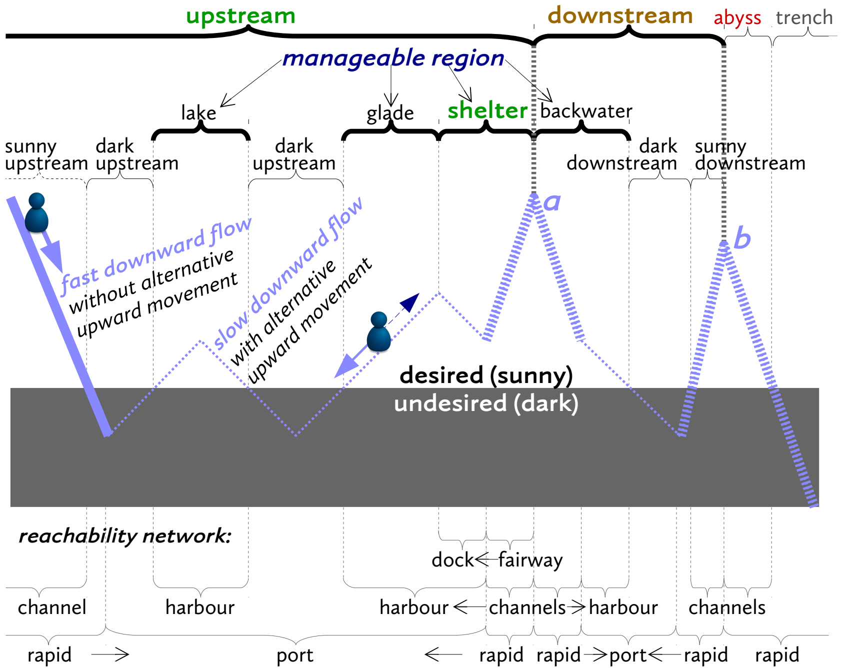

Figure 1. Introductory Example 1. A system moves along the blue line: downward by default (pale blue arrows), but in some regions management can move it in the opposite direction (dark blue arrow) in order to avoid the undesired "dark" region. *Shelters, manageable region, upstream and downstream* (boldface, Sect. 2.1) and other regions from the *main cascade* (top line, Sect. 2.2). Regions from the finer *manageable partition* (below, Sect. 2.3). See Fig. 3 for a systematic summary of these concepts. Bottom: three-level *reachability network* (Sect. 3).

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

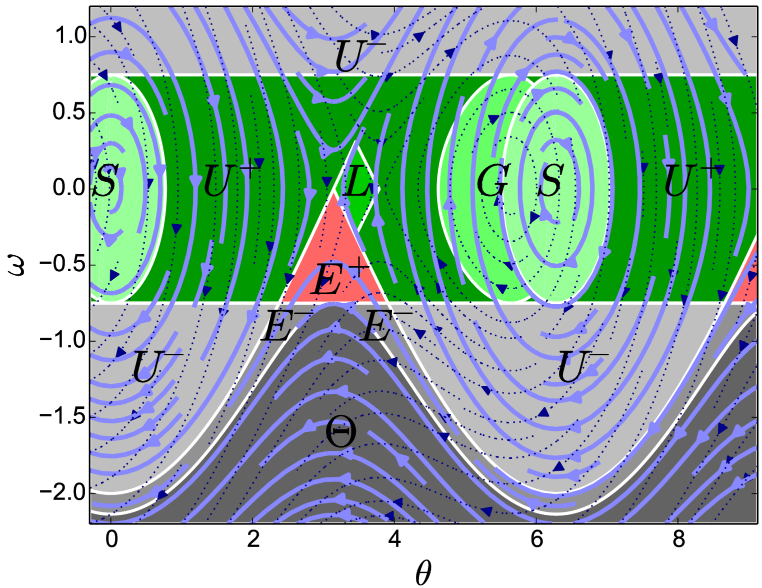

Figure 2. Example 2, phase portrait of a gravity pendulum fun-ride with management by one-sided acceleration and undesirable fast rotations. Thick pale blue lines: default dynamics. Dotted dark blue lines: alternative management trajectories. Grey strips: undesired ("dark") region. Colored areas and labels: derived state space partition (see text), colors as defined in Fig. 3. White: region borders. Note the nonempty eddies E and the lake dilemma (L).

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

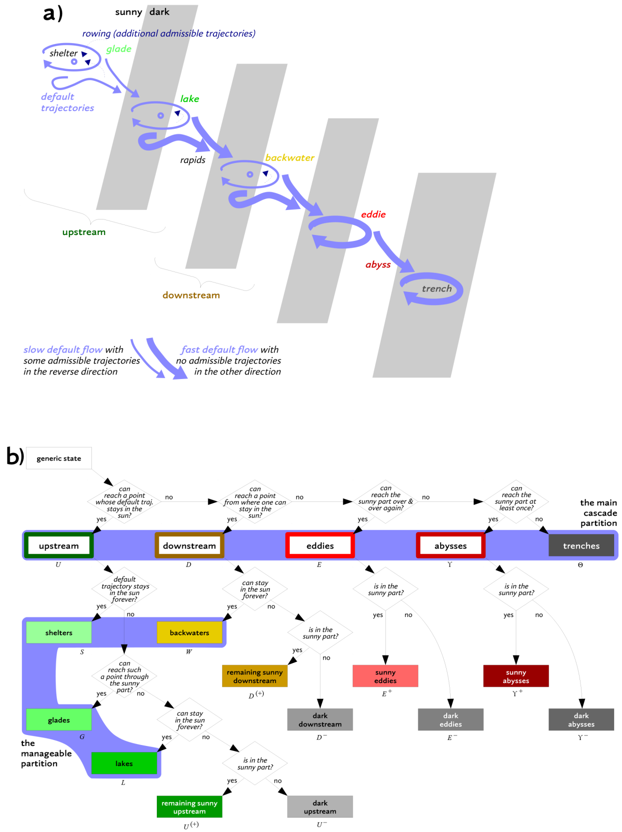

Figure 3. Top: Schematic illustration of a system in which all regions defined so far are nonempty. Bottom: Decision tree summarizing the partition of a manageable dynamical system's state space w.r.t. stable reachability of the desired region or the shelters (main cascade), and the finer partition of the manageable region (Sect. 2.3). The color scheme (grey undesired regions, green upstream regions, yellow downstream regions, red eddies and abysses, lighter meaning better) is also used in the remaining figures.

ESDD
6, 435–488, 2015

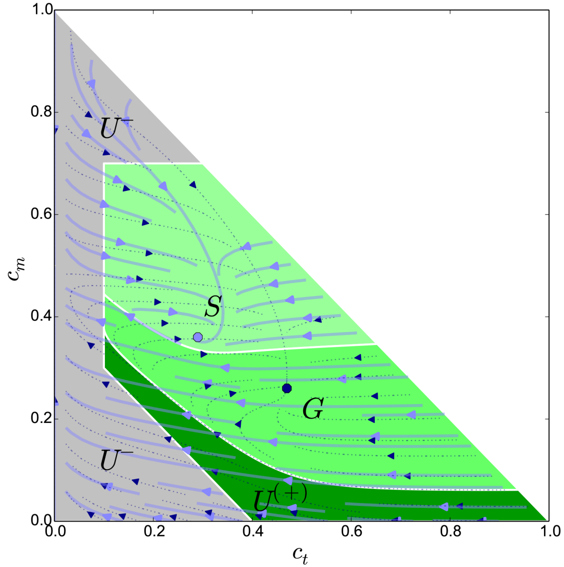

Figure 4. Example 4, carbon cycle model of Anderies et al. (2013). Grey area: undesired region defined by (i) upper bounds for maritime carbon c_m (white horizontal line, representing a planetary boundary related to ocean acidification) and atmospheric carbon $1-c_\mathrm{t}-c_\mathrm{m}$ (white diagonal line, related to a climate change boundary) and a lower bound for terrestrial carbon c_t (white vertical line, representing an ecosystem services planetary boundary). Assuming a management option of reducing the human offtake rate by half leads to a division of the remaining state space into a shelter S around the globally stable fixed point of the default dynamics, a glade G from where S can be reached by management without violating the bounds, and a remainining sunny upstream $U^{(+)}$ from where one cannot avoid violating the bounds temporarily.

ESDD
6, 435–488, 2015

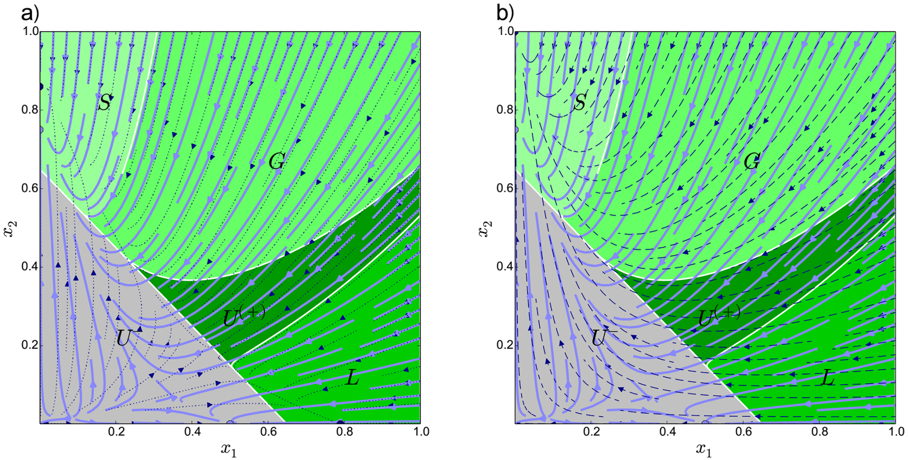

Figure 5. Example 5, showing all upstream regions and illustrating the lake dilemma. A bistable system of two competing plant types with two simultaneous management options (depicted in separate plots only for discernability). Management by a general harvesting quota (trajectories shown left) can ensure desirable long-term harvests of the less productive type x_1 ("lake" L). Management by temporary protection of the more productive type x_2 (right) can cause a transition to the desirable fixed point (in the "shelter" S), but only through the undesired region of low harvests (gray). The state space partition boundaries resulting from both options together (white curves) and a desirable minimum harvest boundary (white diagonal) follow one of the three admissible trajectory types at each point.

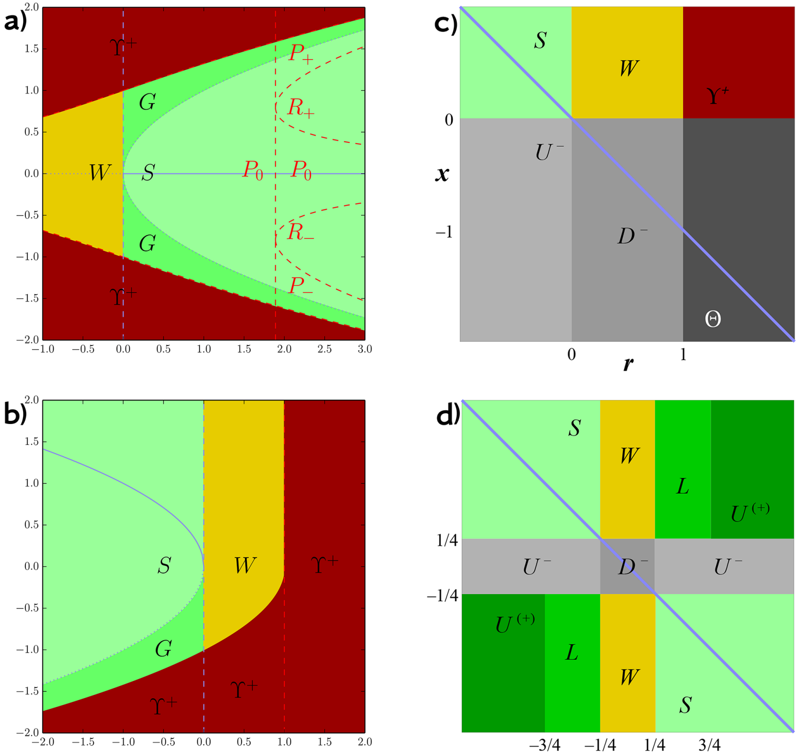

Figure 6. Example 6, showing how parameter changes can change the quality of states due to bifurcations. Top-left: backwater/glade bifurcation and later port pitchfork bifurcation caused by a subcritical pitchfork bifurcation of the default flow (a similar port pitchfork bifurcation is caused by a supercritical pitchfork bifurcation). Top-right: glade/backwater/abyss transition caused by a saddle-node bifurcation, with the second critical value marked in red. Bottom-left: shelter/backwater/abyss transition caused by the transition of a stable fixed point into the deep dark. Bottom-right: shelter/backwater/lake/upstream transition caused by the transition of a stable fixed point through a dark strip.

ESDD
6, 435–488, 2015

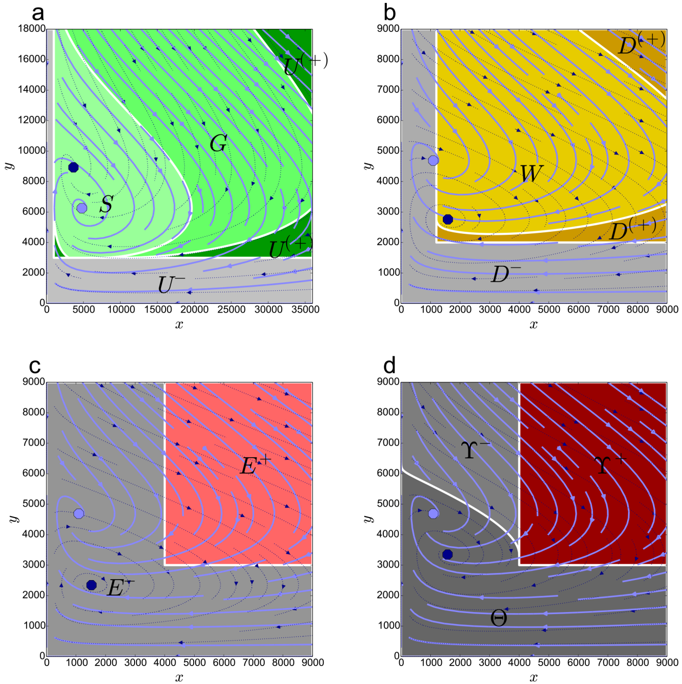

Figure 7. Example 7, coevolution of a population x and a resource stock y. In all cases, $\phi = 4, r = 0.04$. When the globally stable fixed point of the default dynamics (pale blue) falls into X^+, only upstream regions occur (top-left, $\gamma_0 = 4 \times 10^{-6} > \gamma_1 = 2.8 \times 10^{-6}, \delta = -0.1, \kappa = 12\,000, x_{\min} = 1000, y_{\min} = 3000$). When it falls into X^- instead, but the stable fixed point of the alternative management trajectory (dotted dark blue) is in X^+, then only downstream regions occur (top-right, $\gamma_0 = 8 \times 10^{-6} < \gamma_1 = 13.6 \times 10^{-6}, \delta = -0.15, \kappa = 6000, x_{\min} = 1200, y_{\min} = 2000$). Otherwise (bottom, $\gamma_0 = 8 \times 10^{-6} < \gamma_1, \delta = -0.15, \kappa = 6000, x_{\min} = 4000, y_{\min} = 3000$), the analysis depends on whether one can repeatedly reach X^+ by switching between default and alternative trajectories: For $\gamma_1 = 16 \times 10^{-6}$ (bottom-left), only eddies occur, while for $\gamma_1 = 11.2 \times 10^{-6}$ (bottom-right), only abysses and trenches occur.

ESDD
6, 435–488, 2015

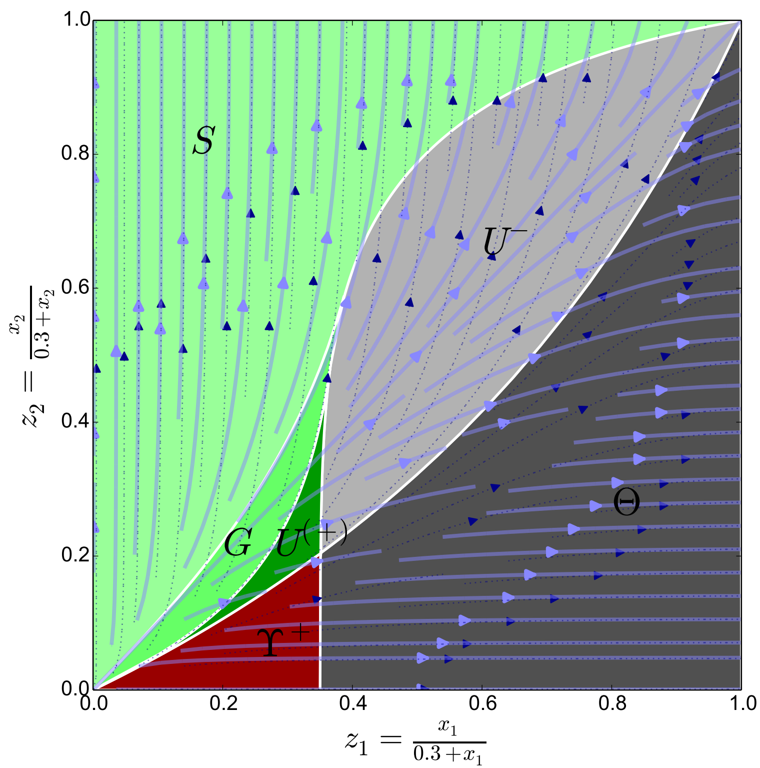

Figure 8. Example 8, coevolution of past cumulative production of a dirty technology x_1 and a clean one (x_2) without (pale blue curves) and with (dotted dark blue curves) a subsidy for the clean technology. Undesired region with too high future usage of the dirty technology colored in grey. Knowledge stocks $x_{1,2}$ were transformed to $z_{1,2} = x_{1,2}/(0.3 + x_{1,2})$ in order to capture their divergence to $+\infty$.

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel

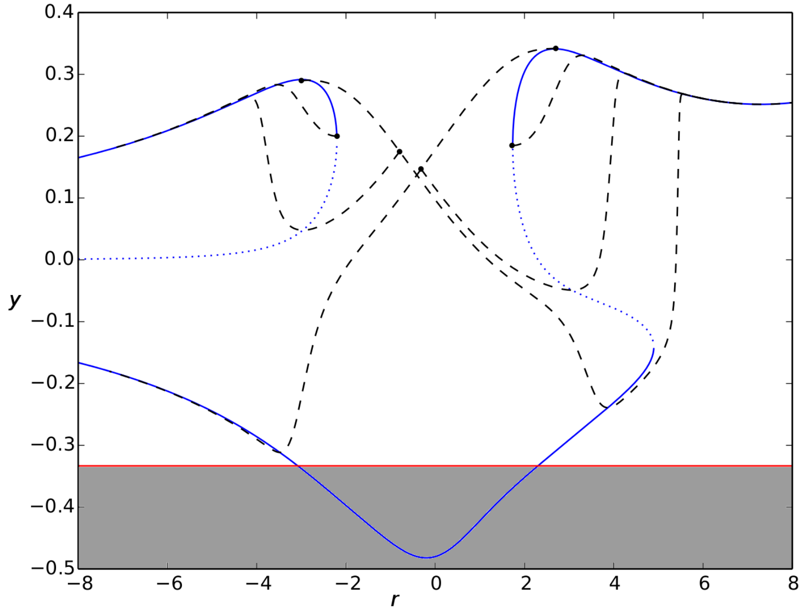

Figure 9. Example 9, bifurcations with manageable parameter. Loci of stable (solid blue) and unstable (dotted blue) fixed points of $\dot{y} = -(4+r^2)^3y^3+(2r^2-1)(4+r^2)y+e^r-10$. Extreme admissible management trajectories (dashed black lines) and their starting points (black dots). Border (red line) between sunny region $y > -1/3$ and the dark (grey). See Fig. 10 for an analysis.

ESDD
6, 435–488, 2015

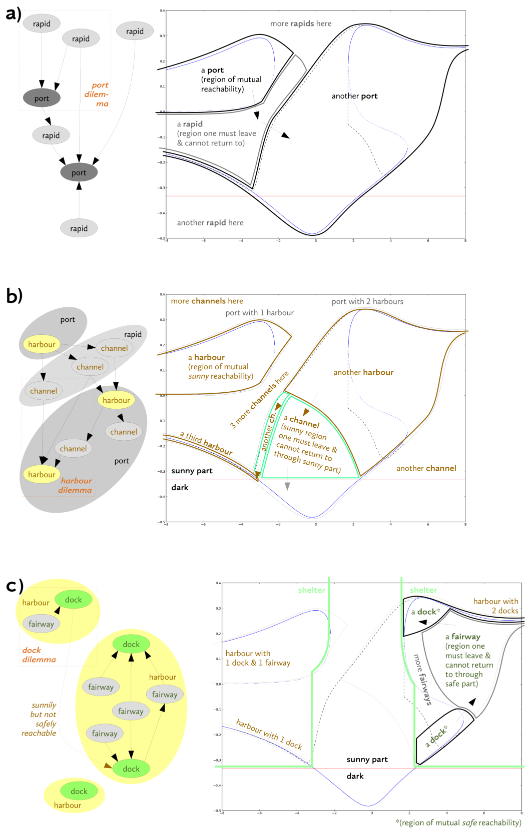

Figure 10. Three-level stable reachability network of ports and rapids (top), harbours and channels (middle), and docks and fairways (bottom), and related dilemmata in the bifurcation example.

ESDD
6, 435–488, 2015

Topology of sustainable management

J. Heitzig and T. Kittel