Chapter 2

Qualitative Reasoning with Model Ensembles

This chapter introduces qualitative differential equations (QDEs) from a new perspective and recalls basic results from differential inclusions and viability theory. In the first section, I develop a novel formal framework which includes QDEs, differential inclusions, causal loop diagrams and other modelling approaches common in sustainability science as special cases. These approaches are highly relevant to account for challenges of uncertainty and generality (cf. Chapter 1). Called framework of model ensembles, common methods like ensemble runs in climate change research are generalised. More importantly, the advantages of QDEs as a modelling tool for syndrome research and natural resource management are clarified. The framework is a guiding principle for the whole thesis. In contrast to traditional approaches to QDEs, I explicitly use graph theoretical concepts to simplify definitions and proofs. Differential inclusions, recalled in the third section, are a generalisation of ordinary differential equations (ODEs). They take account of uncertainties in a way which is complementary to QDEs, and they are the basis for introducing viability theory in the fourth section. This inverse method is, in particular, well-suited to assess model ensembles and to design control strategies. It thus integrates issues of normativity transparently into the analysis. Together with the graph theoretical approach to QDEs, differential inclusions and viability theory will be fundamental for the new techniques I develop in Chapter 3.

2.1 Model Ensembles and Set-Valued Solution Operators

Rather than considering a single model, e.g. an ODE system, the basic idea of the framework of model ensembles I define in this thesis is to consider a whole set of models systematically. This is necessary to investigate real-world systems which cannot be formally described in a unique way due to uncertainty or generality. Although not formalised from the systematic perspective I develop here, such a style of reasoning is common (not only) in sustainability science, e.g. for parameter variation (e.g. Stainforth et al. 2005), model comparison (e.g. Gregory et al. 2005), scenario development (e.g. Nakićenović et al. 2000; Millennium Ecosystem Assessement 2005; Swart et al. 2004), and to characterise patterns of global environmental change (e.g. Schellnhuber et al. 2002). Also, causal loop diagrams as introduced

in Chapter 1 can be seen as a general description of a broad range of different systems. The concise framework proposed in this thesis generalises these ideas.

A **model ensemble** $\mathcal{M}$ is defined as a set of functions $f : X \times \mathbb{R}_+ \to \mathbb{R}^n$ on a **state space** $X \subseteq \mathbb{R}^n$. These functions are called **models**, each describing a possible configuration of a real-world system under investigation or one example of the pattern to be analysed. The set $\mathcal{E}$ contains functions $x(\cdot) : \mathbb{R}_+ \to X$, being the space of **admissible trajectories** of the systems, e.g. $\mathcal{E} = C^1(\mathbb{R}_+, X)$. Each model $f \in \mathcal{M}$ defines a family of initial value problems

$$\begin{aligned} \dot{x} &= f(x,t), \\ x(0) &= x_0, \end{aligned}$$

with $x_0 \in X$. We call the set of all initial value problems given by a model ensemble the **systems of the model ensemble** $\mathcal{M}$. It is also possible to consider model ensembles which only contain autonomous models.

Of course, the systems of the model ensemble have (in general) different solutions. Thus, a *set* of trajectories must be assigned to each initial value x_0. The set-valued **solution operator** $\mathcal{S}_\mathcal{M}(\cdot) : X \to \mathcal{P}(\mathcal{E})$ (of a model ensemble $\mathcal{M}$ with respect to a state space X and admissible trajectories $\mathcal{E}$), assigning to an initial state a subset of $\mathcal{E}$, is defined by

$$\mathcal{S}_\mathcal{M}(x_0) := \{x(\cdot) \in \mathcal{E} \mid x(0) = x_0, \exists f \in \mathcal{M} \forall t \in \mathbb{R}_+ : \dot{x}(t) = f(x(t), t)\}.$$

Depending on $\mathcal{E}$ it may be sufficient that the ODE only holds almost everywhere. We call the elements of $\mathcal{S}_\mathcal{M}(X)$ the **solutions of the model ensemble** $\mathcal{M}$ (with respect to a state space X and admissible trajectories $\mathcal{E}$). If an application requires a very general model, $\mathcal{M}$ is the collection of all cases which have to be analysed. Similarly, in the case of uncertainties, $\mathcal{M}$ is defined to subsume all systems which must be considered. These are given by the part of the knowledge base which is certain to a high degree, while variation is admitted for uncertain parameters, functions or processes. The solution operator is closely related to the concept of an evolutionary system as defined by Aubin (2001). The main challenge in reasoning with model ensembles is to find relevant structure in $\mathcal{S}_\mathcal{M}(X)$. This includes

1. Representing a model ensemble in a way which is adequate to the modeller and allows for a formal treatment,
2. Efficient algorithms to determine $\mathcal{S}_\mathcal{M}(X)$ from a (possibly infinite) model ensemble $\mathcal{M}$,
3. Detecting structural features of the solutions of the model ensemble.

An example for the latter is to introduce normative settings in a formalised way by a partition $X = A \cup B$ of the state space into a preferable region A and a problematic region B. If for *all* $x(\cdot) \in \mathcal{S}_\mathcal{M}(x_0)$ and for all $t \geq 0$ the relation $x(t) \in B$ holds such that the system is "locked in" B due to its intrinsic dynamic interactions, it may be said that "a catastrophic outcome is unavoidable". Or if $\exists x(\cdot) \in \mathcal{S}_\mathcal{M}(x_0), t \geq 0 : x(t) \in A$, then "it is possible to sustain preferable conditions". Such features are very robust in that they hold for a whole model ensemble and not just one model. They are introduced more formally in section 2.4 (p. 45). We now provide some examples for model ensembles.

EXAMPLE 1: Let $\mathcal{M}$ contain only one function $f : X \times \mathbb{R}_+ \to \mathbb{R}^n$ which is Lipschitz on X, and let the admissible trajectories be $\mathcal{E} = C^1(\mathbb{R}_+, X)$. Then, $\mathcal{S}_\mathcal{M}(x_0)$ contains the usual solutions of the initial value problem with $x(0) = x_0$ which exist on $\mathbb{R}_+$. □

EXAMPLE 2: Given a function $f' : X \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n, (x, t; p) \mapsto f'(x, t; p)$, depending on a parameter vector p, and a finite set P of possible parameterisations, define the finite model ensemble

$$\mathcal{M} := \{f \in C(X \times \mathbb{R}_+, \mathbb{R}^n) \mid f(x, t) = f'(x, t; p), p \in P\}.$$

Then, the solution operator with respect to a set of admissible trajectories provides all "scenario runs" for the different parameterisations. □

EXAMPLE 3: For a given autonomous measurable function $f' : X \times U \to \mathbb{R}^n, (x, u) \mapsto f'(x, u)$, where $U \subseteq \mathbb{R}$ is a given interval of control values, we define the infinite model ensemble

$$\mathcal{M} := \{f : X \times \mathbb{R}_+ \to \mathbb{R}^n \text{ measurable } \mid f(x, t) = f'(x, u(t)), u(t) \in U\}.$$

Taking absolutely continuous functions as admissible trajectories, the solution operator $\mathcal{S}_\mathcal{M}(x_0)$ describes all trajectories starting from x_0 which result from any measurable open-loop control $u(\cdot) : \mathbb{R}_+ \to U$. □

EXAMPLE 4: Another example (which will be formalised in detail in section 2.2 and section 2.3, p. 42) are causal loop diagrams. I adopt a straightforward way to interpret a causal loop diagram (cf. Richardson 1986). Suppose that the diagram is given by a directed graph where edges are marked with signs. The vertices represent real variables, say $x_1, \ldots, x_n$. An ODE $\dot{x} = f(x)$ with state vector $x = (x_1 \ldots x_n)^t \in \mathbb{R}^n$ is said to be consistent with the diagram if the signs of the partial derivatives $\operatorname{sgn}(D_j f_i(x))$ correspond to the signs of the edges. If there is no edge between two variables, the partial derivative vanishes. For example, if there is a positive edge from profit expectations x_1 to effort x_2 (cf. Fig. 1.1, p. 13), then $D_1 f_2(x) > 0$, i.e. high profit expectations increase the change of effort. Since such conditions are fulfilled by multiple functions, this provides a collection of ODEs which are all consistent with the influence diagram, and thus a model ensemble $\mathcal{M}$. □

Within the framework of model ensembles, further questions can be posed which are highly relevant for sustainability science. If it is not possible find relevant features common to all solutions of a model ensemble $\mathcal{M}$ we can try to identify subsets $\mathcal{M}' \subseteq \mathcal{M}$ for which such robust properties can be determined. The characterisation of $\mathcal{M}'$ is associated with the discovery of structural features which bring about problematic or desirable system behaviour. In other terms, conditions under which certain (sub)pattern evolve are found. If $\mathcal{M}$ is partially determined by certain control measures imposed on the system, and $\mathcal{M}'$ by alternative control measures, the differences between the solution operators $\mathcal{S}_\mathcal{M}(X)$ and $\mathcal{S}_{\mathcal{M}'}(X)$ are of interest. This kind of reasoning is used intensively in Chapter 4. Moreover, there are cases where solutions of a model ensemble are artifacts from the assumptions the modeller made. Then it is important to restrict $\mathcal{M}$ so that the artifacts are eliminated. Very "unlikely" or "irrelevant" cases which cannot be refuted on base of the original model ensemble are further

reasons to restrict $\mathcal{M}$ or even $\mathcal{E}$. Generally, a **restriction** means a restriction of the model ensemble to some $\mathcal{M}' \subseteq \mathcal{M}$, of the admissible trajectories to some $\mathcal{E}' \subseteq \mathcal{E}$, or of the state space to some $X' \subseteq X$. This is the organising principle for the new techniques developed in Chapter 3, where $\mathcal{E}$ and $\mathcal{M}$ are increasingly restricted step by step.

2.2 Qualitative Differential Equations

In this section I reformulate the concept of qualitative differential equations (QDEs) from a new graph theoretical perspective which differs from the algorithmic view in the original work of Kuipers (1994). This allows for a straightforward introduction, for some new propositions, and paves the road for the development of advanced techniques in Chapter 3. Moreover, I present QDEs as an example for model ensembles, making the introduction more systematic. This embeds the theory into the broader framework developed in section 2.1 and provides a common ground for the comparison to other methods, e.g. differential inclusions (cf. section 2.3, p. 42). The core idea of QDEs is to scan the state and velocity space of the systems of a model ensemble, i.e. a set of ordinary differential equations which share common monotonicity properties: we assume that the dynamics of the system are governed by an ODE $\dot{x} = f(x), f : \mathbb{R}^n \to \mathbb{R}^n$, where only the signs of the partial derivatives $D_i f_j, i, j = 1, \ldots, n$ are known. This approach, called qualitative reasoning, was developed in the 1980s in the field of artificial intelligence. It was made to resemble commonsense reasoning, to model with incomplete knowledge, and to increase the efficiency of automatic reasoning algorithms by abstracting from details which are not necessary to perform a given task. In the context of sustainability science QDEs are very valuable to analyse causal loop diagrams and to deal with uncertainty, generality and non-quantitative knowledge (cf. Chapter 1). As information about the signs of the partial derivatives is not sufficient to set up a unique ODE, we have to deal with a model ensemble and its set-valued solution operator. The latter contains a broad variety of solutions, which nevertheless share some common properties. These are derived by deducing all possible sequences of sign vectors $\operatorname{sgn}(\dot{x})$ just from the signs of the partial derivatives – without solving any system of the model ensemble. The full theory of QDEs permits restricting the model ensemble by requiring more than prescribed signs, e.g. algebraic relations on the state space, which will be introduced below. However, already the signs of the partial derivatives offer an interesting way to interpret a causal loop diagram.

The theory of QDEs is a prominent approach to qualitative reasoning. Other related approaches are confluences (de Kleer and Brown 1984) and qualitative process theory (Forbus 1984). Confluences are equations on the domain of signs. Thus, positive and negative relations between variables of a model can be formulated, and conclusions about their signs can be drawn. The challenge here – and also in the other approaches – are the ambiguities of sign algebra (Williams 1991). In a very simple example we do not know the sign of $x + y$ if $x > 0$ and $y < 0$. To close sign algebra, in addition to $\{[+], 0, [-]\}$, an element $[?]$, denoting an unspecified sign, has to be introduced. This gives rise to different interpretations of confluences, e.g. whether $x = y$ is valid for $x = [+]$ and $y = [?]$ or not.

Qualitative process theory (QPT) explicitly characterises dependencies of variables by four basic types: positive/negative influence of one variable x on the *change* $\dot{y}$ of another,

and positive/negative relationships between variables x, y. Thus, dynamics come into the reasoning process: we can infer under which conditions a variable increases or decreases. In this area, much work was done on building models automatically from libraries of model fragments (Falkenhainer and Forbus 1991).

QDEs extend this approach by introducing so called landmarks, i.e. critical values of the state variables where signs of partial derivatives change (Kuipers 1984). Meanwhile, several other extensions have been made (Kuipers 1994; Kuipers 2001), some of which will be reviewed below in section 2.2.4 (p. 36). In parallel, a broad field of applications has emerged, e.g. in finance (Benaroch and Dhar 1995), epidemiology and genetics (Heidtke and Schulze-Kremer 1998; Trelease et al. 1999), chemistry (Juniora and Martin 2000; Syed et al. 2002), ecology (Guerrin and Dumas 2001; Bredeweg and Salles 2003), the automotive industry (Price and Snooke 1997; Sachenbacher 2001) and sustainability science (Petschel-Held et al. 1999; Petschel-Held and Lüdeke 2001; Eisenack and Kropp 2001; Kropp et al. 2002; Sietz et al. 2005; Eisenack et al. 2006).

In the domain of mathematics, the works of Dordan (1992, Dordan (1995), Aubin (1996), and Hüllermeyer (1997) introduce the concept of a monotonic cell, consisting of all states x such that $f(x)$ has a given sign vector. A trajectory can be described qualitatively by the sequence of monotonic cells it visits. By imposing additional restrictions on f, they investigate the issue of the existence of solutions of a QDE more seriously than in the literature from computer science. They also generalise the approach to other partitions of the state space than by signs, called qualitative frames. However, the approach is more restrictive in that only single ODEs are considered. This is interesting in itself, but not sufficient for our purposes where generality of models and uncertainties have to be taken into account.

In the first subsection the basic concepts of QDEs will be reformulated using graph theoretical concepts and the framework of model ensembles. Then, the basic concepts are extended by introducing landmarks. In the third subsection, the QSIM algorithm is outlined. Finally, I give an overview of advanced QDE techniques to tackle large solutions of QDEs.

2.2.1 Basic Qualitative Differential Equations

For the sake of simplicity I start with a version of QDEs which is closer to QPT. It omits algebraic constraints, e.g. on the signs of the right-hand sides of ODEs, and only considers influences on the change of variables; the velocity, but not the state space is investigated. This will be extended in the next subsection. There is one difference to the original work of Kuipers (1994): the focus is on qualitative states which persist for time *intervals* – but intermediate states occurring only for a point in time (so called time-point states) are not represented explicitly. This has the advantage that solutions of QDEs can be displayed in a much more accessible form (Eisenack and Petschel-Held 2002).

At first we specify the kind of model ensemble which constitutes a QDE (cf. section 2.1, p. 17). For this some elementary sign algebra is needed (Williams 1991). By $\mathcal{A} := \{[+], 0, [-]\}$ we denote the domain of signs, and by $\mathcal{A}_* := \{[+], 0, [-], [?]\}$ the domain of extended signs, where $[?] = \{[+], 0, [-]\}$ denotes an unknown sign. A tolerance relation, denoted here by $\approx$, is defined for $\sigma_1, \sigma_2 \in \mathcal{A}_*$ by

$$\sigma_1 \approx \sigma_2 \text{ iff } \sigma_1 = \sigma_2 \text{ or } \sigma_1 \in \sigma_2 \text{ or } \sigma_2 \in \sigma_1.$$

addition	[+]	0	[-]	[?]	multiplication	[+]	0	[-]	[?]
[+]	[+]	[+]	[?]	[?]	[+]	[+]	0	[-]	[?]
0	[+]	0	[-]	[?]	0	0	0	0	0
[-]	[?]	[-]	[-]	[?]	[-]	[-]	0	[+]	[?]
[?]	[?]	[?]	[?]	[?]	[?]	[?]	0	[?]	[?]

Table 2.1: Addition and multiplication on the domain of extended signs $\mathcal{A}_*$.

We say that σ_1 and σ_2 are consistent if $\sigma_1 \approx \sigma_2$. All signs are consistent with $[?]$. Addition $+$ and multiplication $\cdot$ of signs are defined in Tab. 2.1. For $\sigma_1, \sigma_2 \in \mathcal{A}_*$ and $0 \neq \sigma_3 \in \mathcal{A}$, rules such as the multiplicative cancellation law

$$\sigma_1 = \sigma_2 \cdot \sigma_3 \text{ iff } \sigma_1 \cdot \sigma_3 = \sigma_2 \tag{2.1}$$

are valid. Additionally, we will use tuples and matrices of (extended) signs ($\mathcal{A}^n$, $\mathcal{A}_*^{n\times n}$). The sign operator $[\cdot] := \operatorname{sgn}(\cdot)$ is extended component wise to vectors and matrices, as well as the relation $\approx$. Now a model ensemble can be defined:

DEFINITION 1: *For a given $n \times n$ matrix of signs $\Sigma = (\sigma_{i,j})_{i,j=1,\dots,n}$, $\sigma_{i,j} \in \mathcal{A}_*$, and a state space $X \subseteq \mathbb{R}^n$ we define the* ***monotonic ensemble***

$$\mathcal{M}(\Sigma, X) := \{f \in C^1(X, \mathbb{R}^n) \mid \forall x \in X : \left[\mathcal{J}(f)(x)\right] \approx \Sigma\},$$

where $\mathcal{J}(f)$ denotes the Jacobian of f. We call a function $x(\cdot) \in C^1([0,T], \mathbb{R}^n)$, possibly $T = \infty$, ***reasonable****, if there is only a finite set of points t with $\dot{x}(t) = 0$ on any bounded interval, and define the space of admissible trajectories $\mathcal{E}$ by all reasonable functions with values in X. We call the systems of the model ensemble $\mathcal{M}(\Sigma, X)$ a* ***QDE****.*

A monotonic ensemble $\mathcal{M}(\Sigma, X)$ is a model ensemble which only contains autonomous models. In most cases the state space of the model is clear from the context, and we simply write $\mathcal{M}(\Sigma)$. Although a set of ODE systems is not an equation we use this designation in analogy to Kuipers (1994). The reason for this is that a QDE can be "solved" by considering a constraint satisfaction problem, i.e. a relational equation over a finite set, which is explained in more detail in section 2.2.3 (p. 32).

Based on DEF. 1, a set-valued solution operator $\mathcal{S}_{\mathcal{M}(\Sigma)}(\cdot)$ is defined (with respect to the state space X and the space of reasonable functions as admissible trajectories $\mathcal{E}$). The set of solutions of the monotonic ensemble $\mathcal{S}_{\mathcal{M}(\Sigma)}(X)$ contains all reasonable solutions of all ODE systems contained in the QDE. It should be noted that the properties of the monotonic ensemble are not sufficient to guarantee a global solution for every $f \in \mathcal{M}(\Sigma)$.

As outlined in the previous section, we are interested in analysing the solution operator. Of course, solving all ODEs is not an option. Therefore, the basic idea of the theory of qualitative differential equations is to determine a certain discretisation of the solutions of the monotonic ensemble directly from Σ. Reasonable functions are chosen as admissible trajectories to allow for a discretisation which tracks the sign vectors $[\dot{x}(t)]$ for each solution. Let $x(\cdot)$ be reasonable. If the relation $\left[\dot{x}_i(t_1)\right] = -\left[\dot{x}_i(t_2)\right]$ holds for one $i \in \{1, \dots, n\}$ and some $t_1 < t_2, t_1, t_2 \in \mathbb{R}_+$, there is a point $t \in (t_1, t_2)$ with $\dot{x}_i(t) = 0$ due to continuity.

If there is one $\tau \in \mathbb{R}_+$ with $\dot{x}(\tau) \neq 0$, there is a maximal open interval J containing τ, such that $\left[\dot{x}(\cdot)\right]$ remains constant over J. Being reasonable guarantees that the set of times t where the sign of $\dot{x}(t)$ changes is discrete and countable. Therefore, the sign vectors of the velocity vector of a solution with increasing t can be written as a sequence.

DEFINITION 2: *For a given reasonable function $x(\cdot)$ on $[0,T]$ we have an ordered sequence of sign jump points (t_j) with $t_0 = 0$ which subsequently contains all boundary points of the closures of all sets $\{t \in [0,T] | [\dot{x}(t)] = v\}$ with $v \in \{[-],[+]\}^n$. We construct a sequence of sign vectors $\tilde{x} = (\tilde{x}_j) := \left(\left[\dot{x}(\tau_j)\right]\right)$, where we arbitrarily choose $\tau_j \in (t_j, t_{j+1})$. If the sequence (t_j) is finite with m elements, we choose $\tau_m \in (t_m, T)$. The sequence $\tilde{x}$ is called* ***abstraction*** *of $x(\cdot)$.*

Note that the abstraction $\tilde{x}$ does not depend the concrete values $\tau_j \in (t_j, t_{j+1}), j \in \mathbb{N}$, since the sign vector $\left[\dot{x}(t)\right]$ is constant on any interval (t_j, t_{j+1}). The sequence (t_j) is chosen from the *boundary points* to have maximal intervals with constant sign vector. They are taken from the *closure* of $\{t \in [0,T] | [\dot{x}(t)] = v\}$ to ignore saddle points where components of $\dot{x}$ vanish for one point in time, but afterwards attain the same sign as before. If there is a sign jump at t, there is a $i \in \{1, \ldots, n\}$ such that the following conditions hold:

$$\left.\begin{array}{ll} (i) & \dot{x}_i(t) = 0, \\ (ii) & \exists \epsilon : \forall t^- \in [t-\epsilon, t) : \ \dot{x}_i(t^-) \neq 0 \text{ or } \forall t^+ \in (t, t+\epsilon] : \ \dot{x}_i(t^+) \neq 0, \\ (iii) & \exists \epsilon : \forall t^- \in [t-\epsilon, t),\ t^+ \in (t, t+\epsilon] : [\dot{x}_i(t^-)] \neq [\dot{x}_i(t^+)]. \end{array}\right\} \tag{2.2}$$

The set of the abstractions of all solutions of a monotonic ensemble is (in general) still infinite. However, they are entailed by a finite graph in the following way:

DEFINITION 3: *Let $\mathcal{M}(\Sigma)$ be a monotonic ensemble, $\mathcal{E}$ the set of reasonable trajectories and $\mathcal{S}_{\mathcal{M}(\Sigma)}(\cdot)$ the corresponding solution operator. We denote the set of the abstractions of the solutions by*

$$\tilde{\mathcal{S}}_{\mathcal{M}(\Sigma)} := \{\tilde{x} \mid \exists x_0 \in X, x(\cdot) \in \mathcal{S}_{\mathcal{M}(\Sigma)}(x_0) : \tilde{x} \text{ is the abstraction of } x(\cdot)\}.$$

Then, the directed ***state-transition graph*** *G of the monotonic ensemble is defined by the vertices*

$$V(G) := \{v \in \mathcal{A}^n \mid \exists\, \tilde{x} \in \tilde{\mathcal{S}}_{\mathcal{M}(\Sigma)}, j \in \mathbb{N} : \ \tilde{x}_j = v\},$$

called ***qualitative states****, and the edges*

$$E(G) := \{(v,w) \mid \ \exists\, \tilde{x} \in \tilde{\mathcal{S}}_{\mathcal{M}(\Sigma)}, j \in \mathbb{N} : \ \tilde{x}_j = v \text{ and } \tilde{x}_{j+1} = w\},$$

called ***qualitative transitions****.*

For convenience, the state-transition graph of a monotonic ensemble is also called the state-transition graph of a QDE. Thus, we have defined a directed graph G such that all sequences of $\tilde{\mathcal{S}}_{\mathcal{M}(\Sigma)}$ describe a path in G, i.e. the graph "covers" all reasonable solutions of initial value problems

$$\begin{aligned} \dot{x} &= f(x) \\ x(0) &= x_0 \end{aligned} \tag{2.3}$$

with $f \in \mathcal{M}(\Sigma)$. Note that G is loop free, since subsequent coefficients of the abstraction of a reasonable function are different. The state-transition graph is the basic tool to investigate the structure of $\mathcal{S}_{\mathcal{M}(\Sigma)}(X)$. We are interested in the existence of vertices and edges of G which depends on Σ, and in further properties of G. The state-transition graph of a QDE can be efficiently computed by the so-called QSIM algorithm, which will be outlined in section 2.2.3 (p. 32). Some features of state-transition graphs can already be shown directly.

Which vertices occur in a state-transition graph? Most basically, $\{[-],[+]\}^n \subseteq V(G)$ due to the following reasons: by chain rule $\ddot{x} = \mathcal{J}(f)(x) \cdot \dot{x}$, such that for assumptions about the sign matrix $\left[\mathcal{J}(f)(x)\right]$ not all sign vectors $[\dot{x}]$ are consistent with all sign vectors $[\ddot{x}]$. However, since no claims about $[\ddot{x}]$ are made, no $[\dot{x}] \in \{[-],[+]\}^n$ can be excluded from being a vertex. The situation is more complicated if some $\dot{x}_i \equiv 0$ on (t_j, t_{j+1}), which implies that also $\ddot{x}_i \equiv 0$ on the same interval.

I now present a new necessary criterion for such a vertex to exist. For this, we need the set $Z_0(v) := \{i = 1, \ldots, n | v_i = 0\}$, which assigns to a sign vector $v \in \mathcal{A}^n$ the indices of vanishing components.

PROPOSITION 1: *If $v \in V(G)$, then for all $i \in Z_0(v)$*

$$\exists j, k \notin Z_0(v), j \neq k : 0 \neq \sigma_{i,j} v_j \approx -\sigma_{i,k} v_k \neq 0$$

or $\forall j \notin Z_0(v) : \sigma_{i,j} = 0$.

PROOF: If $v \in V(G)$ there is an $f \in M(\Sigma)$ and $x_0 \in X$ such that for a solution $x(\cdot)$ to Eq. (2.3) there exists an interval (t_1, t_2) such that $\forall t \in (t_1, t_2) : [\dot{x}(t)] = v$. Thus,

$$\forall i \in Z_0(v) : 0 = \dot{x}_i(t) = \ddot{x}_i(t) = \sum_{j=1,\ldots,n} D_j f_i \cdot \dot{x}_j. \tag{2.4}$$

If $\forall j \notin Z_0(v) : \sigma_{i,j} = 0$, then

$$\begin{aligned} \ddot{x}_i(t) &= \sum_{j \in Z_0(v)} D_j f_i \cdot \dot{x}_j + \sum_{j \notin Z_0(v)} D_j f_i \cdot \dot{x}_j \\ &= \sum_{j \in Z_0(v)} D_j f_i \cdot 0 + \sum_{j \notin Z_0(v)} 0 \cdot \dot{x}_j = 0. \end{aligned}$$

If this is not the case, there must be a $j \notin Z_0(v)$ such that $D_j f_i \cdot \dot{x}_j \neq 0$, i.e. $\sigma_{i,j} v_j \neq 0$. Then, Eq. (2.4) can only be true if there is also a $k \notin Z_0(v)$ with $[D_k f_i \cdot \dot{x}_k] = -[D_j f_i \cdot \dot{x}_j]$, i.e. $\sigma_{i,j} v_j \approx -\sigma_{i,k} v_k \neq 0$. □

Additionally, every state-transition graph contains the vertex 0, representing the equilibria of systems of the monotonic ensemble. A sufficient criterion for the existence of vertices is omitted here, since it will be argued in section 2.2.4 (p. 36) that vertices with vanishing components have only limited relevance for applications.

Now, I will prove a characterisation for the existence of edges in the state-transition graph G which cannot be found in the previous literature. It is simplified by considering only vertices with non-vanishing components. When two qualitative states v, w differ only

in one component i, there must be a solution of the monotonic ensemble $x(\cdot)$, defined by a model f, which transgresses the main isocline $f_i(x) = 0$ at some time, because this isocline separates the regions of the phase space where $[f(x)] = v$ and $[f(x)] = w$, respectively (the so called monotonic cells). A necessary condition for such a transgression is an appropriate sign of $\ddot{x}_i$ on the main isocline, e.g. if $v_i = [-]$ and $w_i = [+]$, then $[\ddot{x}_i] \approx [+]$ is needed. We define the **intermediate state** $v \wedge w$ for $v, w \in \mathcal{A}$ by

$$(v \wedge w)_i := \begin{cases} v_i & \text{if } v_i = w_i, \\ 0 & \text{if } v_i \neq w_i, \end{cases}$$

Thus, $Z_0(v \wedge w)$ are the indices of the components which change from v to w (or which are constant in one or both states).

PROPOSITION 2: *Let $v, w \in V(G)$, $v \neq w$, and $Z_0(v) = Z_0(w) = \varnothing$. Then, $(v, w) \in E(G)$ iff*

$$\forall i \in Z_0(v \wedge w)\, \exists j \notin Z_0(v \wedge w) : w_i \cdot (v \wedge w)_j \approx \sigma_{i,j}.$$

PROOF: At first, we prove that the existence of an edge $(v, w) \in E(G)$ implies the condition. By DEF. 3 (p. 23), there is only an edge (v, w) if

$$\left.\begin{array}{l} \exists f \in M(\Sigma), x' \in X, t_0 < t_1 < t_2 \in \mathbb{R}_+, \\ \qquad \text{and a solution to Eq. (2.3) with } x_0 = x' \text{ such that} \\ \qquad \forall t \in [t_0, t_1) : [\dot{x}(t)] = v \text{ and} \\ \qquad \forall t \in (t_1, t_2] : [\dot{x}(t)] = w. \end{array}\right\} \tag{2.5}$$

Moreover, $[\dot{x}(t_1)] = (v \wedge w)$ due to continuity of $\dot{x}$. Consider an $i \in Z_0(v \wedge w)$. If $v_i = [+]$ and $w_i = [-]$, the velocity x_i obtains a maximum at t_1. Thus, by continuity, $[\ddot{x}_i(t_1)] = [-] = w_i$. In the case of a minimum, $[\ddot{x}_i(t_1)] = [+] = w_i$. Other cases are not possible since $Z_0(v) = Z_0(w) = \varnothing$. The equality $w_i = [\ddot{x}_i(t_1)] = [(D_1 f_i \ \ldots \ D_n f_i) \cdot \dot{x}(t_1)]$ can only hold if $\exists j \in \{1, \ldots, n\}$ such that $w_i = [D_j f_i(x(t_1)) \cdot \dot{x}_j(t_1)]$. It follows that $\sigma_{i,j} \cdot (v \wedge w)_j \approx w_i$. The index j cannot be an element of $Z_0(v \wedge w)$, because otherwise the sign $\sigma_{i,j} \cdot 0 = 0$ would be not consistent with the sign $w_i \neq 0$. From Eq. (2.1), we conclude that $w_i \cdot (v \wedge w)_j \approx \sigma_{i,j}$.

Now we show that the condition implies the existence of an edge $(v, w) \in E(G)$. Suppose that Σ and $v, w \in V(G)$ are given and the condition is satisfied. Since $Z_0(v) = Z_0(w) = \varnothing$, it holds for all $i \in Z_0(v \wedge w)$ that $w_i = -v_i \neq 0$. It has to be shown that Eq. (2.5) holds. We construct appropriate f, x', t_0, t_1 and t_2.
Choose $c \in \mathbb{R}^n$ such that $\forall i \in Z_0(v \wedge w) : c_i = 0$ and $\forall i \notin Z_0(v \wedge w) : c_i = (v \wedge w)_i \cdot 1$. Let $k(i) \notin Z_0(v \wedge w)$ be an index (depending on i) such that $w_i \cdot (v \wedge w)_{k(i)} \approx \sigma_{i,k(i)}$ for all $i \in Z_0(v \wedge w)$, which exists due to the condition. Define the matrix $A = (a_{i,j}) \in \mathbb{R}^{n \times n}$ by $a_{i,k(i)} := n \cdot w_i \cdot (v \wedge w)_{k(i)}$ if $i \in Z_0(v \wedge w)$, and $a_{i,j} := 1 \cdot w_i \cdot (v \wedge w)_j$ if $i \notin Z_0(v \wedge w)$ or $j \neq k(i)$. We set $f(x) := Ax + c$. Clearly, $f \in \mathcal{M}(\Sigma)$ since $[\mathcal{J}(f)(x)] = [A] \approx \Sigma$.
Now choose an arbitrary $t_1 \in \mathbb{R}_+$ and an open interval J_0 containing t_1 such that $x(\cdot) : J_0 \to \mathbb{R}^n$ is a solution of the ODE $\dot{x} = f(x), x(t_1) = 0$. Let $i \notin Z_0(v \wedge w)$. Since $\dot{x}(t_1) = c$ and

$c_i \neq 0$, and due to continuity of $x(\cdot)$, there is also an open interval J_i containing t_1 such that $\forall t \in J_i : [\dot{x}_i(t)] = (v \wedge w)_i = w_i$.

Such an interval also exists for $i \in Z_0(v \wedge w)$:
Since $\forall t \in J_0 : \ddot{x}(t) = A\dot{x}(t)$ and $(v \wedge w)_{k(i)} \neq 0$,

$$\begin{aligned}\ddot{x}_i(t_1) &= \sum_{j=1,\ldots,n} a_{i,j}\dot{x}_j(t_1) \\ &= a_{i,k(i)}c_{k(i)} + \sum_{\substack{j \notin Z_0(v\wedge w),\\ j \neq k(i)}} a_{i,j}c_j \\ &= n \cdot w_i \cdot (v \wedge w)_{k(i)} \cdot 1 \cdot (v \wedge w)_{k(i)} + \sum_{\substack{j \notin Z_0(v\wedge w),\\ j \neq k(i)}} a_{i,j}c_j.\end{aligned}$$

Observe that

$$n \cdot w_i \cdot (v \wedge w)_{k(i)} \cdot 1 \cdot (v \wedge w)_{k(i)} = n \cdot w_i,$$

and

$$|\sum_{\substack{j \notin Z_0(v\wedge w),\\ j \neq k(i)}} a_{i,j}c_j| = |\sum_{\substack{j \notin Z_0(v\wedge w),\\ j \neq k(i)}} 1 \cdot w_i \cdot (v \wedge w)_j \cdot 1 \cdot (v \wedge w)_j| \leq n - 1.$$

Consequently $[\ddot{x}_i(t_1)] = w_i$, and, due to continuity, there is an open interval J_i containing t_1 such that $\forall t \in J_i : [\ddot{x}_i(t)] = w_i$. Now choose $t_0 < t_1 < t_2$ such that $t_0, t_2 \in \bigcap_{i=0,\ldots,n} J_i$ and define x' component-wise by

$$x'_i := x_i(t_1) - \int_{t_0}^{t_1} \dot{x}_i(\tau)d\tau,$$

such that $x(\cdot)$ is a solution of $\dot{x} = f(x), x(t_0) = x'$. Then, for all $t \in [t_0, t_1), i = 1, \ldots, n$:

$$[\dot{x}_i(t)] = \left[\dot{x}_i(t_1) - \int_t^{t_1} \ddot{x}_i(\tau)d\tau\right] = -w_i = v_i,$$

because $\dot{x}_i(t_1) = 0$ and $\forall \tau \in [t, t_1] : [\ddot{x}_i(\tau)] = w_i$.
By analogy for all $t \in (t_1, t_2) : [\dot{x}(t)] = w$. □

This proposition shows that every relevant edge actually corresponds to a solution of the monotonic ensemble. We now provide a simple example of a QDE and its state-transition graph.

EXAMPLE 5: Suppose we have a sign matrix

$$\Sigma = \begin{pmatrix} 0 & + \\ - & 0 \end{pmatrix},$$

and take $X = \mathbb{R}^2$ as state space. The monotonic ensemble $\mathcal{M}(\Sigma)$ contains all continuously differentiable functions $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\forall x \in \mathbb{R}^2 : D_2 f_1 > 0, D_1 f_2 < 0, D_1 f_1 =$

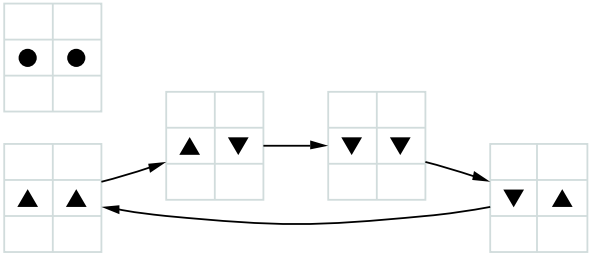

Figure 2.1: State-transition graph of the QDE of $\mathcal{M}(\Sigma)$ (computer-generated output, produced by the QSIM algorithm explained below in section 2.2.3, p. 32). The vertices represent qualitative states. The columns in each vertex correspond to the variables x_1, x_2. The triangles indicate whether $\dot{x}_i$ is positive, negative or vanishes.

$D_2 f_2 = 0$, and the QDE contains all initial value problems Eq. (2.3) with $f \in \mathcal{M}(\Sigma)$ and $x_0 \in \mathbb{R}^2$. The solution operator $\mathcal{S}_{\mathcal{M}(\Sigma)}(x_0)$ contains all reasonable solutions of these initial value problems which start in x_0, and $\tilde{\mathcal{S}}_{\mathcal{M}(\Sigma)}$ contains all abstractions, i.e. sequences of "trends" $[\dot{x}(\cdot)]$ of these trajectories which can be traced as paths in the state-transition graph G of the monotonic ensemble.

We infer the structure of G. As remarked above, $\{[-],[+]\}^2 \subseteq V(G)$. It follows from PROP. 1 (p. 24) that there is no other vertex in G except equilibrium: If $v = (0\ v_2)^t$, and consequently $Z_0(v) = \{1\}$, both necessary conditions are violated – there are no two different indices $j, k \notin Z_0(v), j \neq k$. Only $1 \in Z_0(v)$ and $2 \notin Z_0 = (v)$, but $\sigma_{1,2} = [+] \neq 0$. The argument for $v = (v_1\ 0)^t$ is analogous.

In this situation all edges can be found by using PROP. 2. For $v = ([+]\ [+])^t$ and $w = ([+]\ [-])^t$, it holds that $(v \wedge w) = ([+]\ 0)^t$ and $Z_0(v \wedge w) = \{2\}$. We have $w_2 \cdot (v \wedge w)_1 = [-] = \sigma_{2,1}$, such that $(v, w) \in E(G)$. But since $v_2 \cdot (v \wedge w)_1 = [+] \neq [-] = \sigma_{2,1}$, there is no edge from w to v. The arguments are analogous for all pairs of vertices (see Fig. 2.1 for the result). □

2.2.2 Qualitative Differential Equations with Landmarks

This section extends the basic definitions, e.g. to prescribe zeros of the models of a monotonic ensemble and algebraic constraints. By introducing so called landmarks into the state space, the latter can be investigated more closely. Basically, landmarks are introduced (i) to exclude models from a monotonic ensemble where main isoclines do not pass through qualitatively prescribed points of the state space, and (ii) to allow regions of phase space with different sign matrices $[\mathcal{J}(f)(x)]$ which are separated by landmarks. If they are only used in the first way, landmarks provide a restriction of monotonic ensembles as defined in section 2.1 (p. 17). Again, the following definitions are different from the original work of Kuipers (1994): all landmarks are introduced by the modeller and not generated during qualitative simulation. Hence, the following theory is related to what Kuipers terms the "envisionment representation".

Although the situation is almost analogue to DEF. 1, 2 and 3, we need new concepts, mainly to prepare more elaborate model ensembles and abstractions of reasonable functions. For every variable x_i of the model, we specify an ordered set of symbolic **landmarks** $\langle \lambda_{i,1}, \ldots, \lambda_{i,k_i} \rangle$. For every $i = 1, \ldots, n$ there is an ordered set

$$Q_i := \langle \lambda_{i,1}, \{\lambda_{i,1}, \lambda_{i,2}\}, \lambda_{i,2}, \ldots, \{\lambda_{i,k_i-1}, \lambda_{i,k_i}\}, \lambda_{i,k_i} \rangle,$$

where single landmarks and pairs of consecutive landmarks alternate. The cross product $Q := \times_{i=1,\ldots,n} Q_i$ is called the **quantity space**, and $S := \times_{i=1,\ldots,n}(Q_i \times \mathcal{A})$ is the **qualitative state space** of the model. To extract specific components of a **qualitative state** $v \in S$ by projection $\pi_i : \times_{i=1,\ldots,n} X_i \to X_i$, the standard notations in qualitative reasoning are:

$$\mathrm{qval}_i(v) = \pi_i(v) \in Q_i \times \mathcal{A}, \tag{2.6}$$
$$\mathrm{qmag}_i(v) = \pi_1(\mathrm{qval}_i(v)) \in Q_i, \tag{2.7}$$
$$\mathrm{qmag}(v) = \times_{i=1,\ldots,n} \mathrm{qmag}_i(v) \in Q, \tag{2.8}$$
$$\mathrm{qdir}_i(v) = \pi_2(\mathrm{qval}_i(v)) \in \mathcal{A}, \tag{2.9}$$
$$\mathrm{qdir}(v) = \times_{i=1,\ldots,n} \mathrm{qdir}_i(v) \in \mathcal{A}^n. \tag{2.10}$$

They are called **qualitative value** (Eq. 2.6), **qualitative magnitude** (Eqs. 2.7 and 2.8) or **qualitative direction** (Eqs. 2.9 and 2.10), respectively. Symbolic landmarks are interpreted as numbers in $\mathbb{R}$ and the pairs of landmarks as open intervals, such that a qualitative state can be assigned to each point of the state space X: A vector

$$\Lambda = (\lambda_{1,1} \ \ldots \ \lambda_{1,k_1} \\ \vdots \\ \lambda_{n,1} \ \ldots \ \lambda_{n,k_n})^t,$$

such that all $\lambda_{i,j} \in \mathbb{R}$, is called **landmark vector** if $\forall i = 1, \ldots, n : \lambda_{i,1} < \lambda_{i,2} < \ldots < \lambda_{i,k_i}$. For a given landmark vector, the **state abstraction** of a joint state and velocity vector $(x\, \dot{x})^t \in X \times \mathbb{R}^n$ is obtained by the mapping $a_\Lambda : X \times \mathbb{R}^n \to S$ such that

$$\mathrm{qdir}_i\big(a_\Lambda(x, \dot{x})\big) = [\dot{x}_i],$$
$$\mathrm{qmag}_i\big(a_\Lambda(x, \dot{x})\big) = \begin{cases} \lambda_{i,j} & \text{if } x_i = \lambda_{i,j} \\ \{\lambda_{i,j}, \lambda_{i,j+1}\} & \text{if } x_i \in (\lambda_{i,j}, \lambda_{i,j+1}). \end{cases}$$

However, we will not consider a specific landmark vector Λ, but only assume the existence of landmark vectors with prescribed qualitative properties. This is why landmarks are introduced as *symbolic* values. One exception is the landmark "`0`", which is always associated with $0 \in \mathbb{R}$. This is necessary to extract the sign of a qualitative magnitude: If $0 \in Q_i$ and $q \in Q_i$, we define $[q] = [+]$ for $q > 0$, $[q] = [-]$ for $q < 0$, and $[q] = 0$ for $q = 0$; more generally

$$[q]_\lambda := \begin{cases} [+] & \text{if } q > \lambda, \\ 0 & \text{if } q = \lambda, \\ [-] & \text{if } q < \lambda, \end{cases}$$

where $\lambda \in Q_i$ is a landmark. We are now ready to proceed with further definitions extending the concepts from QDEs without landmarks.

DEFINITION 4: *For a state space X, a quantity space Q and the associated state space S, let $\mu : Q \to \mathcal{A}_*^{n\times n}$ be a mapping which assigns a matrix of (extended) signs $\mu(q)$ to each element q of the quantity space. Additionally, we take a family of relations $C := \{C_1, \ldots, C_m\}$ on the state space, $C_j \subseteq S, j = 1, \ldots, m$, called* ***constraints****, as given. Then, we obtain a* ***monotonic landmark ensemble***

$$\begin{aligned}\mathcal{M}(\mu, C) := \{f \in & C^1(X, \mathbb{R}^n) \mid \\ & \exists \textit{ landmark vector } \Lambda\, \forall x \in X : \\ & \big[\mathcal{J}(f)(x)\big] \approx \mu\big(\mathrm{qmag}(a_\Lambda(x, f(x)))\big) \\ & \textit{and} \\ & \forall j = 1, \ldots, m : a_\Lambda(x, f(x)) \in C_j\}.\end{aligned}$$

The systems of a monotonic landmark ensemble $\mathcal{M}(\mu, C)$ are called a ***QDE****.*

It must be noted that μ and C can be chosen such that $\mathcal{M}(\mu, C) = \varnothing$ in a non-trivial way. In section 2.2.3 (p. 32) there are additional remarks on the consistency of μ and C. Taking the reasonable functions with values in X as space of admissible trajectories $\mathcal{E}$, we obtain a set-valued solution operator $\mathcal{S}_{\mathcal{M}(\mu,C)}(\cdot) : X \to \mathcal{P}(\mathcal{E})$. As before, not the solutions of the monotonic landmark ensemble themselves will be investigated, but their abstraction:

DEFINITION 5: *For a given reasonable function $x(\cdot)$ on $[0, T]$, possibly $T = \infty$, and a landmark vector Λ, there is an ordered sequence of jump points (t_j), with $t_0 = 0$ and subsequently containing all boundary points of the closures of all sets $\{t \in [0, T] | a_\Lambda(x(t), \dot{x}(t)) = v\}$ with $v \in S$ such that $\forall i = 1, \ldots, n : \mathrm{qdir}_i(v) \neq 0$ and $\mathrm{qmag}_i(v)$ is a pair of consecutive landmarks. The* ***landmark abstraction*** *of $x(\cdot)$ with respect to Λ is the sequence of qualitative states $\tilde{x} = (\tilde{x}_j) := \big(a_\Lambda(x(\tau_j), \dot{x}(\tau_j))\big)$ with arbitrarily chosen $\tau_j \in (t_j, t_{j+1})$. If (t_j) is finite with m elements, we choose $\tau_m \in (t_m, T)$.*

Again, the landmark abstraction does not depend on actual selection of $\tau_j, j \in \mathbb{N}$. By denoting $v(t) := a_\Lambda(x(t), \dot{x}(t))$, the definition implies that if t is a jump point, there exists an $i = \{1, \ldots, n\}$ such that the following three properties hold:

$$\left.\begin{array}{ll} (i) & \mathrm{qdir}_i(v(t)) = 0 \text{ or } \exists l \in \{1, \ldots, k_i\} : \mathrm{qmag}_i(v(t)) = \lambda_{i,l} \\ (ii) & \exists \epsilon\, \forall t^- \in [t-\epsilon, t) : \ \mathrm{qval}_i(v(t^-)) \neq \mathrm{qval}_i(v(t)) \\ & \qquad \text{or } \forall t^+ \in (t, t+\epsilon] : \ \mathrm{qval}_i(v(t^+)) \neq \mathrm{qval}_i(v(t)), \\ (iii) & \exists \epsilon\, \forall t^- \in [t-\epsilon, t),\ t^+ \in (t, t+\epsilon] : \mathrm{qval}_i(s(t^-)) \neq \mathrm{qval}_i(s(t^+)). \end{array}\right\} \tag{2.11}$$

The landmark abstraction of the solutions of a monotonic landmark ensemble are traced as paths of a directed graph as in the previous section.

DEFINITION 6: *Let $\mathcal{M}(\mu, C)$ be a monotonic landmark ensemble, $\mathcal{E}$ the space of reasonable trajectories and $\mathcal{S}_{\mathcal{M}(\mu,C)}(\cdot)$ the corresponding solution operator. We denote the set of the landmark abstractions of the solutions by*

$$\tilde{\mathcal{S}}_{\mathcal{M}(\mu,C)} := \{\tilde{x} \mid \exists x(\cdot) \in \mathcal{S}_{\mathcal{M}(\mu,C)}(x_0), x_0 \in X : \tilde{x} \textit{ is the abstraction of } x(\cdot)\}.$$

Then, the directed ***state-transition graph*** G *of the monotonic landmark ensemble is defined by the vertices*

$$V(G) := \{v \in S \mid \exists\, \tilde{x} \in \tilde{\mathcal{S}}_{\mathcal{M}(\mu,C)}, j \in \mathbb{N} :\ \tilde{x}_j = v\},$$

called ***qualitative states****, and the edges*

$$E(G) := \{(v,w) \mid\ \exists\, \tilde{x} \in \tilde{\mathcal{S}}_{\mathcal{M}(\mu,C)}, j \in \mathbb{N} :\ \tilde{x}_j = v \textit{ and } \tilde{x}_{j+1} = w\},$$

called ***qualitative transitions****.*

Again, G is always loop free. Although criteria for the existence of an edge in a state-transition graph are, in principle, a simple extension of PROP. 2 (p. 25), the situation becomes more complex since a larger number of cases has to be considered. Only a sketch of the various cases is provided here because they offer limited new insight. For this discussion (and later in Chapter 3), the extension of the notion of the intermediate state $v \wedge w \in S$ is helpful, which is defined component wise by

$$\mathrm{qdir}_i(v \wedge w) := \begin{cases} \mathrm{qdir}_i(v) & \text{if } \mathrm{qdir}_i(v) = \mathrm{qdir}_i(w), \\ 0 & \text{if } \mathrm{qdir}_i(v) \neq \mathrm{qdir}_i(w), \end{cases} \tag{2.12}$$

$$\mathrm{qmag}_i(v \wedge w) := \begin{cases} \mathrm{qmag}_i(v) & \text{if } \mathrm{qmag}_i(v) = \mathrm{qmag}_i(w), \\ \lambda_{i,j} & \text{if } \mathrm{qmag}_i(v) \neq \mathrm{qmag}_i(w) \\ & \text{and } \lambda_{i,j} \in \mathrm{qmag}_i(v) \cap \mathrm{qmag}_i(w) \neq \varnothing. \end{cases}$$

It should be noted that this operation is not defined for all pairs of states. In this case, $(v,w) \notin E(G)$, since their qualitative magnitudes are not adjacent. Due to continuity, it is not possible that a trajectory $x(\cdot)$ jumps from $x_i(t_1) < \lambda_{i,j}$ to $x_i(t_2) > \lambda_{i,j+1}$, $t_1 < t_2$, if there is not $t \in (t_1, t_2)$ such that $\lambda_{i,j} < x_i(t) < \lambda_{i,j+1}$.

An obvious necessary criterion for the existence of an edge (v,w) is that the intermediate state has to satisfy the relations in C, i.e. $\forall j = 1,\ldots,m : (v \wedge w) \in C_j$. For another criterion suppose that (i) $v \wedge w$ is defined, (ii) $Z_\lambda(v \wedge w) := \{i = 1,\ldots,n | \mathrm{qmag}_i(v) \neq \mathrm{qmag}_i(w)\} \neq \varnothing$ (i.e. some qualitative magnitudes change), (iii) for all $i \in Z_\lambda(v \wedge w)$ the qualitative magnitudes $\mathrm{qmag}_i(v)$ and $\mathrm{qmag}_i(w)$ are pairs of landmarks, and (iv) $Z_0(v \wedge w) = \varnothing$ (i.e. all qualitative directions remain unchanged). Then, if there is an $i \in Z_\lambda(v \wedge w) : \mathrm{qmag}_i(w) > \mathrm{qmag}_i(v)$, there can only be an edge (v,w) if $\mathrm{qdir}_i(v \wedge w) = \mathrm{qdir}_i(v) = [+]$. Otherwise, if $\mathrm{qmag}_i(w) < \mathrm{qmag}_i(v)$, the ith component has to decrease, $\mathrm{qdir}_i(v) = [-]$.

It can be seen that such criteria become very complicated because there are situations where only qualitative directions change (as investigated in PROP. 2, p. 25), where one or more qualitative magnitudes change, where some qualitative values remain constant on a landmark, or where multiple of such events occur in parallel. All these cases are incorporated into the QSIM algorithm which is presented in the next section. However, some of them are of limited practical relevance. Exclusion of solutions of the monotonic landmark ensemble where one component is constant is discussed in section 2.2.4 (p. 36). In section 3.2 (p. 62) I develop a method to exclude cases where several qualitative magnitudes and several qualitative directions change at the same time and which have limited relevance: both techniques restrict the set of admissible trajectories to a subset of the reasonable functions.

EXAMPLE 6: Suppose we have a system defined by an ODE of the form $\dot{x} = f(x)$ on the state space $X = \mathbb{R}_+$ with $f \in C^1(\mathbb{R}, \mathbb{R})$, but we only the following properties of f: There exist three real numbers $0 < \lambda_1 < \lambda_2 < \lambda_3$ such that λ_1, λ_3 are the only zeros of f, i.e.

$$f(\lambda_1) = f(\lambda_3) = 0.$$

Moreover, the function is increasing below λ_2 and decreasing above:

$$\begin{aligned} \forall x < \lambda_2 : D_x f(x) > 0, \\ \forall x > \lambda_2 : D_x f(x) < 0, \\ D_x f(\lambda_2) = 0. \end{aligned}$$

Quantitative values are not known for $\lambda_1, \lambda_2, \lambda_3$. We want to describe this situation by a monotonic landmark ensemble to find all trajectories which may be produced by such a system. At first, the qualitative state space has to be constructed on the basis of adequate landmarks:

$$Q = Q_1 := \big\langle 0, \{0, \lambda_1\}, \lambda_1, \{\lambda_1, \lambda_2\}, \lambda_2, \{\lambda_2, \lambda_3\}, \lambda_3, \{\lambda_3, \infty\}, \infty \big\rangle,$$

yielding $S = Q \times \mathcal{A}$ and a state abstraction $a_\Lambda : \mathbb{R}_+ \times \mathbb{R} \to S$ for every landmark vector $\Lambda = (\lambda_0 \ \lambda_1 \ \lambda_2 \ \lambda_3 \ \lambda_4)^t$ with $\lambda_0 = 0, \lambda_4 = \infty$, which describes a possible quantitative configuration of the the landmarks. Secondly, we need the map μ and the set of constraints C. We define

$$\mu : Q \to \mathcal{A}_*, q \mapsto \begin{cases} [+] & \text{if } x < \lambda_2, \\ 0 & \text{if } x = \lambda_2, \\ [-] & \text{if } x > \lambda_2, \end{cases}$$

to express the monotonicity properties of f and $C := \{C_1, C_2\}, C_1, C_2 \subseteq S$ with

$$\begin{aligned} C_1 &:= \{v \in S \mid \mathrm{qmag}_1(v) \leq \lambda_2 \Rightarrow \mathrm{qdir}_1(v) = [\mathrm{qmag}_1(v)]_{\lambda_1}\}, \\ C_2 &:= \{v \in S \mid \mathrm{qmag}_1(v) \geq \lambda_2 \Rightarrow \mathrm{qdir}_1(v) = -[\mathrm{qmag}_1(v)]_{\lambda_3}\}, \end{aligned}$$

to express the zeros of f. All together, we obtain the monotonic landmark ensemble

$$\begin{aligned} \mathcal{M}(\mu, C) := \{f \in & C^1(\mathbb{R}_+, \mathbb{R}) \mid \\ & \exists \text{ landmark vector } \Lambda \, \forall x \in \mathbb{R}_+ : \\ & \big[D_x f(x)\big] = \mu\big(\mathrm{qmag}(a_\Lambda(x, f(x)))\big) \\ & \text{and} \\ & \forall j = 1, \ldots, m : a_\Lambda(x, f(x)) \in C_j\}, \end{aligned}$$

and a QDE consisting of all initial value problems given by functions $f \in \mathcal{M}(\mu, C)$ and initial values $x_0 \in \mathbb{R}_+$. This defines a set-valued solution operator $\mathcal{S}_{\mathcal{M}(\mu,C)}(\cdot) : \mathbb{R}_+ \to \mathcal{P}(\mathcal{E})$ with the reasonable functions $x(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ as space of admissible trajectories $\mathcal{E}$. It assigns all reasonable solutions to an initial value.

Based on the state abstraction, every solution $x(\cdot) \in \mathcal{S}_{\mathcal{M}(\mu,C)}(x_0)$ can be discretised by landmark abstraction, yielding a sequence $(\tilde{x}_j), \tilde{x}_j \in S$. All these sequences are paths in the

state-transition graph, which can be computed with the QSIM algorithm (see next subsection). The result is presented in Fig. 2.2. Note that only qualitative states which persist over time intervals are represented as vertices in the graph. There is, e.g. no qualitative state v with $\mathrm{qmag}(v) = 0$ and $\mathrm{qdir}(v) = [-]$. □

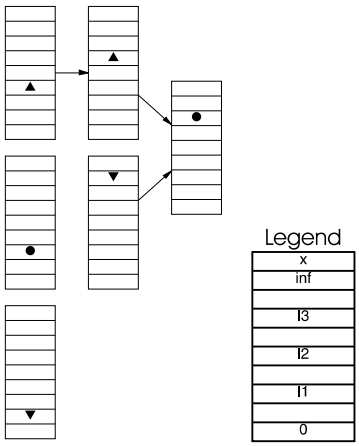

Figure 2.2: State-transition graph of the QDE defined in EX. 6 (computer-generated output). Each vertex represents a qualitative state, each edge a qualitative transition. The small boxes inside each vertex v represent all possible qualitative magnitudes where singular landmarks and intervals defined by pairs of consecutive landmarks alternate (see legend). The box which contains a symbol denotes $\mathrm{qmag}(v)$, at the same time indicating whether $\mathrm{qdir}(v) = [+]$, $\mathrm{qdir}(v) = 0$ or $\mathrm{qdir}(v) = [-]$.

Several larger QDEs will be presented and analysed in Chapter 4. QDEs with landmarks can describe much more complex settings than the causal loop diagrams introduced in Chapter 1. In addition to negative and positive influences on *change rates of variables*, there can be *direct* influences on the variables, represented by the constraints C. Moreover, whether an influence is negative or positive can depend on the (qualitative) value of another (or the same) variables via the mapping μ.

2.2.3 The QSIM Algorithm

Basic Ideas

The QSIM algorithm is a procedure to compute the state-transition graph of a QDE. This subsection outlines its basic ideas and the script language needed in Chapter 4 and the Appendix. For the complete formal specification we refer to the original work of Kuipers (1994). The core of the QSIM algorithm is a specialised solver for a so called constraint satisfaction problem for which various methods exist (Mackworth 1977; Mackworth 1987; Tsang 1993).

DEFINITION 7: *A **constraint satisfaction problem** is given by a set of n variables $V = \{x_1, \ldots, x_n\}$ and a set of n associated finite domains $D = \{D_1, \ldots, D_n\}$, where the variables take their values. Furthermore, a set of m **constraints** $C = \{C_1, \ldots, C_m\}$, which are*

subsets of the state space $S = \times_{i=1,\ldots,n} D_i$, has to be given. The ***solution*** *of a constraint satisfaction problem is the maximal subset $L \subseteq S$ such that $\forall i = 1, \ldots, m : L \subseteq C_i$.*

Computing the solution can be complicated if $C_1, \ldots, C_m$ are represented in a way such that $\cap_{i=1,\ldots,m} C_i$ cannot be compute directly. For a monotonic landmark ensemble $\mathcal{M}(\mu, C)$, the basic idea is to take $D_i = Q_i \times \mathcal{A}, i = 1, \ldots, n$, and the constraints C. Additional constraints are implied by the sign information given by μ. The solution L includes all qualitative states which are consistent with the model assumptions. This is a superset of the vertices of the state-transition-graph G, because it also contains states which cannot be obtained for a time interval (which is required by DEF. 2, p. 23 or DEF. 5, p. 29). Excluded are qualitative states v where for an $i \in \{1, \ldots, n\} : \mathrm{qdir}_i(v) \neq 0$ and $\exists j \in \{1, \ldots, k_i\} : \mathrm{qmag}_i(v) = \lambda_{i,j}$, i.e. where a component of a solution $x_i(\cdot)$ of a system of $\mathcal{M}(\mu, C)$ is at a landmark and $\dot{x}_i \neq 0$ – this is only possible for a *point* in time, i.e. for intermediate states $(v \wedge w)$, and not on a time interval. However, these intermediate states must also satisfy the constraints C. It is possible that the solution L does not contain every combination of qualitative directions (contrary to section 2.2.1, p. 21). The edges $(v, w) \in E(G)$ are found by considering the sign matrices $\mu(v), \mu(w)$, because they indicate under which conditions the qualitative direction or the qualitative magnitude of a state can change (cf. PROP. 2, p. 25).

In practice, the algorithm takes account of several technical details, mainly to cover all special cases for edges in G as discussed above. An example for another technicality is that the state space is augmented by the velocity space to allow for a more flexible formulation of constraints: The qualitative state space S is supplied with the components $i' = n+1, \ldots, 2n$, $Q_{i'} := \langle \{-\infty, 0\}, 0, \{0, +\infty\} \rangle$ with $\dot{x}_i = x_{i'}$. This makes the state abstractions of vectors $(x\ \dot{x}\ \ddot{x})^t$ available via

$$\mathrm{qmag}_{i'}(a_\Lambda(x, \dot{x})) = \begin{cases} \{-\infty, 0\} & \text{if } \dot{x}_{i'} < 0, \\ 0 & \text{if } \dot{x}_{i'} = 0, \\ \{0, +\infty\} & \text{if } \dot{x}_{i'} > 0, \end{cases}$$

and $\mathrm{qdir}_{i'}(a_\Lambda(x, \dot{x})) = [\ddot{x}_i]$. The link of $\dot{x}_i$ to $x_{i'}$ is expressed by a constraint of the form

$$\{v \in S \mid \forall i = 1, \ldots, n : \mathrm{qdir}_i(v) = [\mathrm{qmag}_{n+i}(v)]\}. \tag{2.13}$$

Altogether, the constraints describing $\mathcal{M}(\mu, C)$ can be specified using a script language as outlined below. Then, the QSIM algorithm applies a constructive filtering technique. It begins with one user-specified qualitative state v which is an element of the solution L of the constraint satisfaction problem. Then, all *potential* successor states of v are generated, i.e. states where qualitative magnitudes change in accordance with current qualitative directions, or where qualitative directions change in accordance with the signs of the Jacobian. This is as an application of PROP. 2 (p. 25) and its full extension to monotonic landmark ensembles. Each potential successor is checked to establish whether it belongs to L: the procedure efficiently evaluates all constraints $C_1, \ldots, C_m$ for each potential successor. If one constraint is violated, the state is "filtered out". With this procedure, sequences of "surviving" *actual* states are generated. In these sequences state which can occur for a time interval

and for time points (so called I-states and P-states) alternate. Only the former are relevant for the state-transition graph, but all special cases like extrema of solutions occurring exactly on a landmark or concurrent changes of qmag and qdir can be covered by considering P-states. Thus, if $v \wedge w$ is a P-state succeeding v, and $w \in \cap_{i=1,\ldots,m} C_i$ is a new I-state succeeding $v \wedge w$, then w is introduced as a vertex and (v, w) as an edge in G. After this, new potential successors of w are generated and "filtered" iteratively, until no new valid successor is found.

One more subtlety in the filtering procedure arises from the fact that the number of potential successors is usually too large for efficient computation. In practice, a more complicated algorithm is applied. In a first step only the possible successors of some components of a qualitative state are generated (so-called partial assignments) and discarded if they violate a single constraint $C_1, \ldots, C_m$. Then, all surviving values for one component are combined with all surviving values for other components. This has to be checked against all constraints iteratively until all complete, consistent states are constructed. The algorithm is designed such that the guaranteed coverage theorem holds (Kuipers 1994, p. 118, adapted to the terminology of this thesis):

PROPOSITION 3: *Let $x(\cdot)$ be a reasonable solution of the monotonic landmark ensemble $\mathcal{M}(\mu, C)$, and Λ a landmark vector. Then, the landmark abstraction $\tilde{x}$ of $x(\cdot)$ is a path in the graph computed by the QSIM algorithm.*

It is important to note that there are also paths in the graph which are not an abstraction of a solution of the monotonic landmark ensemble. Such a path is called **spurious behaviour**. The QSIM algorithm is complete, meaning that the abstractions of all solutions are represented, but not sound, i.e. there is spurious behaviour – a fact that can generally be proven (Say and Akin 2002). In PROP. 2 (p. 25) soundness was only shown for edges (v, w), describing the short but important paths v, w. Thus, the information contained in the state-transition graph can be seen as negative: If there is no edge from a vertex v to a vertex w, we can be sure that there is no solution of the monotonic landmark ensemble for which the landmark abstraction attains w directly after v. Therefore, QDEs are complementary to common numerical simulation. While the latter strives to use sound but incomplete algorithms (every computed trajectory approximates a solution, but it is not possible to compute an infinite ensemble quantitatively), the QSIM algorithm is unsound but complete.

The QSIM Script Language

There are several implementations of the QSIM algorithm, e.g. in Lisp (University of Texas at Austin 1991) or in C (Dvorak 1998). The latter is the base for an extended C++ version developed for this thesis (which was already used for the examples in the previous subsections). We now introduce the basic syntax and semantics of the script language to formulate a QDE in these implementations. The variables of the model together with their quantity spaces are defined by a block

```
(quantity-spaces
  <(var ( <lm> ) ''text'')>
)
```

where <lm> is the sequence of landmarks associated with variable var, i.e. a component of the (qualitative) state space. Angular brackets < > indicate that a specification can be repeated multiple times. The description text comments on the variable. The most basic constraint requires that the derivative of one variable x is determined by a variable y (cf. Eq. 2.13).

```
((d//dt x y))
```

If i is the index of variable x and j the index of y this constraint is valid for qualitative states v with $\mathrm{qdir}_i(v) = [\mathrm{qmag}_j(v)]$. Simple monotonic relationships between two variables are expressed by

```
((M+ x y) <(lx ly)>)
```

or

```
((M- x y) <(lx ly)>)
```

with a (possibly empty) sequence of m pairs (lx ly), called **corresponding values**, where lx represents a landmark of the variable x (with index i) and ly a landmark of variable y (with index j), denoting sequences of landmarks $\lambda_{i,1}, \ldots, \lambda_{i,m}$ and $\lambda_{j,1}, \ldots, \lambda_{j,m}$. A qualitative state v satisfies the M+ constraint if

$$\begin{aligned} &\mathrm{qdir}_i(v) = \mathrm{qdir}_j(v), \\ &\text{and } \forall k = 0, \ldots, m : \; [\mathrm{qmag}_i(v)]_{\lambda_{i,k}} = [\mathrm{qmag}_j(v)]_{\lambda_{j,k}}, \end{aligned}$$

while for the M- constraint

$$\begin{aligned} &\mathrm{qdir}_i(v) = -\mathrm{qdir}_j(v), \\ &\text{and } \forall k = 0, \ldots, m : \; [\mathrm{qmag}_i(v)]_{\lambda_{i,k}} = -[\mathrm{qmag}_j(v)]_{\lambda_{j,k}}. \end{aligned}$$

Rows in the sign matrix Σ (without a coefficient $[?]$) are expressed by multivariate monotonic function constraints

```
(((M <s>) <var>))
```

where <s> denotes a sequence of m signs $\sigma_1, \ldots, \sigma_m \in \mathcal{A}$ (which are different from 0), and <var> denotes $m+1$ variables with indices $i_1, \ldots, i_{m+1}$. Then, a qualitative state v respects this constraint if one of the following two conditions holds:

$$\begin{aligned} \mathrm{qdir}_{i_{m+1}}(v) = 0 \Rightarrow &\exists j, k \in \{1, \ldots, m\}, j \neq k : \\ &\sigma_j \cdot \mathrm{qdir}_{i_j}(v) = -\sigma_k \cdot \mathrm{qdir}_{i_k}(v) \neq 0, \\ \mathrm{qdir}_{i_{m+1}}(v) \neq 0 \Rightarrow &\exists j \in \{1, \ldots, m\} : \sigma_j \cdot \mathrm{qdir}_{i_j}(v) = \mathrm{qdir}_{i_{m+1}}(v). \end{aligned}$$

This constraint can also be extended to corresponding values

```
(((M <s>) <var>) <(tuple)> )
```

with `<(tuple)>` being a sequence of $(m+1)$-tuples of landmarks with length k, denoted by $\lambda_{i_1,1}, \ldots, \lambda_{i_1,k}$, $\lambda_{i_2,1}, \ldots, \lambda_{i_2,k}$, …, $\lambda_{i_{m+1},1}, \ldots, \lambda_{i_{m+1},k}$, and additionally requiring for a qualitative state v and for all $l = 1, \ldots, k$ that

$$\left(\forall j = 1, \ldots, m : \sigma_j \cdot [\mathrm{qmag}_{i_j}(v)]_{\lambda_{i_j,l}} = \sigma_{m+1}\right) \Rightarrow [\mathrm{qmag}_{i_{m+1}}(v)]_{\lambda_{i_{m+1},l}} = \sigma_{m+1}.$$

Other important constraints are defined in a similar way, e.g. `U-` and `U+`, capturing the monotonicity properties of functions $f(x), x \in \mathbb{R}$ which increase on one side of a landmark and decrease on the other (cf. Ex. 6, p. 31), and `((mult x y z))`, called qualitative multiplication, generalising the monotonicity properties of an algebraic equation like $x_k(t) = x_i(t) \cdot x_j(t)$. For a full specification of this and further constraints I refer to Kuipers (1994).

2.2.4 Advanced Techniques

I open this subsection with a short discussion of the strengths and weaknesses of QDEs. The latter motivates various extensions of the basic approach. Here, established extensions are presented, while in Chapter 3 I introduce new abstraction and restriction techniques. Using the framework of model ensembles developed in section 2.1 (p. 17) and the graph theoretical description introduced in section 2.2 (p. 20), I can formulate and systematise most of these techniques in a new and consistent way. The concept of a restriction technique is directly related to the restriction of a model ensemble, and is the guiding principle for most parts of Chapter 3. The generic definition of abstraction techniques I provide subsumes various known abstraction techniques. It will be used again and complemented with viability theory in Chapter 3.

The main advantages of QDEs concern generality, uncertainty and non-quantitative knowledge. As discussed in Chapter 1, problems of sustainability science show such properties. Based on the framework of model ensembles, QDEs clearly account for the first two of them (cf. section 2.1, p. 17): Each QDE represents a broad variety of systems with some common (monotonicity) properties. This can subsume a typical pattern comprising many instances, or a set of potentially valid model formulations between which the modeller cannot discriminate. QDEs also account for non-quantitative knowledge, since the characteristics of the system are only expressed in terms of trends (signs), influences (signs) and thresholds (landmarks). Hence, it is not necessary to measure variables of a real-world system quantitatively, making it much more easy to introduce variables which are difficult to operationalise (e.g. profit expectations, well-being, political power).

On the other hand, the core of the approach is deterministic: a monotonic landmark ensemble contains only autonomous models. We cannot take variability of parameters into account. Secondly, as discussed in the previous subsection, we have the problem of spurious behaviour: Not every path in the state-transition graph corresponds to a solution of the monotonic landmark ensemble. From the practical perspective, a main challenge is posed by very large state-transition graphs resulting from QDEs with many variables. The number of potential qualitative states increases exponentially with the dimension of the qualitative state space, termed the state explosion problem (Valmari 1998). A deeper obstacle is the tendency of larger QDEs to produce weakly structured state-transition graphs, indicated by a typically

increasing average degree of vertices. This makes it less "predictable" which successor state w will be observed when a system is in a given qualitative state v. This is not an artifact of the method, but simply a consequence of the generality (or uncertainty) of the model assumptions contained in an influence diagram: If many partial derivatives of a component function $f_i, f \in \mathcal{M}(\Sigma)$ do not vanish, it is more likely that there exists a $j \notin Z_0(v \wedge w)$ such that $w_i \cdot (v \wedge w)_j \approx \sigma_{i,j}$, resulting in multiple successors (cf. PROP. 2, p. 25). All these problems are obstacles to drawing conclusions for the management of an environmental system. This is because only very limited knowledge is expressed by a monotonic landmark ensemble.

Most techniques to handle a large or weakly structured state-transition graph G can be classified by the following scheme:

1. Abstraction: simplify the representation of G to foster meaningful analysis.
2. Restriction: reduce the number of vertices and edges of G by introducing additional types of assumptions about the system. This is mainly by restricting the state space, the monotonic landmark ensemble, or the space of admissible trajectories.
3. Analysis: consider G as a database which is not displayed but queried for interesting model properties.

Here, the concept of "abstraction" is formally different from section 2.2.1 (p. 21) and section 2.2.2 (p. 27), but the use of both meanings is standard in the qualitative reasoning literature.

Abstraction techniques

There are two types of abstraction procedures: they can either be guided solely by the structure of state-transition graph G itself (automated abstraction), or can depend on user-specified criteria (user-guided abstraction). Both are useful, and thus they are very often combined. All abstraction techniques perform some kind of clustering of the state-transition graph, resulting in an abstracted state-transition graph which disregards some features of (and therefore abstracts from) the original state-transition graph (Clancy and Kuipers 1993; Mallory et al. 1996). There is also related work on the abstraction of finite state machines (Oikonomou 1996). We propose the following generic definition of abstraction, which is also fundamental for section 3.1 (p. 52).

DEFINITION 8: *Let G be a state-transition graph and $G_i, i = 1, \ldots, m$ a partition of disjoined subgraphs of G with $\bigcup_{i=1,\ldots,m} V(G_i) = V(G)$. The* ***abstracted state-transition graph*** *G' is then given by*

$$V(G') := \bigcup_{i=1,\ldots,m} \{V(G_i)\}$$

and

$$E(G') := \{(v', w') \mid v' \neq w' \text{ and } \exists v \in v', w \in w' : (v, w) \in E(G)\}.$$

The new loop-free graph G' contains a vertex for every subgraph G_i (defined by the union of the qualitative states in this subgraph). The edges of G' are "inherited" from the edges of G. Different abstraction procedures are distinguished by the chosen subgraphs.

One major breakthrough in (automated) abstraction was so called chatter-box abstraction (Clancy and Kuipers 1993; Clancy 1997), since it reduces both the number of states and the computation effort of the QSIM algorithm drastically. The idea is to cluster qualitative states between which the system can arbitrarily float back and forth (called chattering). Based on the new DEF. 8 this technique can be described in a simple and concise way. For a graph G, define the bi-directed subgraph $G^- := G \cap G^{-1}$, which contains all vertices but only the bi-directed edges of G.

DEFINITION 9: *A **chatter-box** is a subgraph of G which is induced by a strongly connected component of the bi-directed subgraph G^-.*

Chatter-boxes can be used for abstraction since

PROPOSITION 4: *Chatter-boxes are disjoined subgraphs of G.*

PROOF: Due to maximality, the strongly connected components of G^- are disjoined. Thus, also the subgraphs of G induced by these components are disjoined. □

DEFINITION 10: *Let $G_i, i = 1, \ldots, l$ be the family of all chatter-boxes of a state transition graph G, and $G_i, i = l+1, \ldots, m$ the trivial subgraphs of G which have only one vertex which is not a member of a chatter-box. The abstracted state-transition graph G' based on this family is called the **chatter-box abstraction** of G.*

The most common version of the chatter-box algorithm detects only chatter-boxes where all qualitative states in the subgraph have the same qualitative magnitudes. If a chatter-box contains different qualitative magnitudes, it is split to different disjoined subgraphs: this **simple chatter-box abstraction** ignores so called landmark chatter. The advantage is that this version can be easily integrated into the QSIM algorithm and no post processing is needed. Since this reduces the number of qualitative states which have to be considered by the filtering algorithm, its efficiency increases. We will use simple chatter-box abstraction for most of the applications in Chapter 4.

Another very powerful simplification procedure is projection, also called generation of a variable focus (Mallory et al. 1996; Clancy et al. 1997). For this user-guided procedure, a set of variables of interest has to be specified by the modeller. The idea is to observe only differences in these variables and to disregard the values of all other variables. For example, merely technical auxiliary variables can be ignored or hypotheses about a restricted set of state variables can be verified by this technique. The variables not to be ignored are described as a index set $I = \{j_1, \ldots, j_p\} \subseteq \{1, \ldots, n\}$ (denoting components of the qualitative state space). We consider two types of projection: simple and faithful. The first abstracts the state-transition graph taking as subgraph states which have identical qualitative values in all variables of interest. The second type of projection additionally splits these subgraphs into weakly connected components, thus maintaining connectivity properties.

By $\pi_I : V(G) \to \times_{j \in I}(Q_j \times \mathcal{A})$, we denote the usual projection $v \mapsto (v_{j_1} \ \ldots \ v_{j_p})^t$. Two vertices $v, w \in V(G)$ are called **simple projection equivalent** with respect to I if $\pi_I(v) = \pi_I(w)$. Define the undirected supergraph $G^+ := G \cup G^{-1}$ of G, which contains

a bi-directed edge for every edge in G. Two vertices v, w are called **faithful projection equivalent** if, additionally, in G^+ there is a path $v_0, \ldots, v_k$, $v_0 = v, v_k = w$ such that $\pi_I(v_0) = \pi_I(v_1) = \ldots = \pi_I(v_k)$ or if $v = w$. Both relations are obviously equivalence relations, and we obtain equivalence classes on $V(G)$ which induce disjoined subgraphs of G.

DEFINITION 11: *Let $G_i, i = 1, \ldots, m$ be the family of subgraphs of G induced by a simple projection equivalence defined by I. The resulting abstracted state-transition graph $G' =: \pi_I(G)$ is called the **simple projection** of G with respect to I. If $G_i, i = 1, \ldots, m$ are given by a faithful projection equivalence defined by I, $G' =: \bar{\pi}_I(G)$ is called the **faithful projection** of G with respect to I.*

Besides producing a simplified graph, these projections have some useful properties. For a path $v_0, \ldots, v_k$ in G, we can construct a sequence of vertices $v'_0, \ldots, v'_l$, in $\pi_I(G)$, $l \leq k$ by taking the sequence $\pi_I(v_0), \ldots, \pi_I(v_k)$ and removing all elements which are identical to their predecessor. Then $l = 1$ or $v'_0, \ldots, v'_l$ is a path in $\pi_I(G)$. The same holds true for $\bar{\pi}_I$. As a consequence, structures are preserved under $\bar{\pi}_I$ and π_I. A subgraph H of G is weakly connected, if for all $v, w \in V(H)$, $v \neq w$, there is a path $v, \ldots, w$ in G^+. If a subgraph $H \subseteq G$ is weakly or strongly connected, the same applies to $\pi_I(H) \subseteq \pi_I(G)$ (unless $|\pi_I(H)| = 1$) and to $\bar{\pi}_I(H) \subseteq \bar{\pi}_I(G)$ (unless $|\bar{\pi}_I(H)| = 1$). For the inverse case, we only have a weaker property:

PROPOSITION 5: *(1) If $v'_1, \ldots, v'_k$ is a path in $\bar{\pi}_I(G)^+$, then for all $v, w \in V(G)$ such that $\bar{\pi}_I(v) = v'_1$ and $\bar{\pi}_I(w) = v'_k$ there exists a path $v, \ldots, w$ in G^+.*
(2) If $H \subseteq \bar{\pi}_I(G)$ is weakly connected, so is $\bar{\pi}_I^{-1}(H) \subseteq G$.

PROOF: (1) Since for all $v'_i, v'_{i+1}, i = 1, \ldots, k-1$, there is always an edge (v'_i, v'_{i+1}) in $\bar{\pi}_I(G)^+$, there must also be an edge $(v_i, v_{i+1}) \in E(G^+)$ with $\bar{\pi}(v_i) = v'_i$ and $\bar{\pi}(v_{i+1}) = v'_{i+1}$. Since $\bar{\pi}_I$ is faithful, all $u \in V(G)$ with $\bar{\pi}(u) = v'_i$ are strongly connected in G^+. Thus, there is a path from v to w in G^+.
(2) Take $v, w \in V(\bar{\pi}_I^{-1}(H))$ and choose $v', w' \in V(H)$ such that $\bar{\pi}(v) = v'$ and $\bar{\pi}(w) = w'$. Since H is weakly connected, there is a path $v', \ldots, w'$ in $\bar{\pi}_I(G)^+$. Consequently, part (1) of the proposition guarantees the existence of a path $v, \ldots, w$ in G^+, i.e. $\bar{\pi}_I^{-1}(H)$ is weakly connected. □

Simple projection will be necessary to display some results in Chapter 4.

A third abstraction technique is called state-based (Fouché and Kuipers 1991; Clancy and Kuipers 1993), and can be used if the quantity space contains a larger number of landmarks. The modeller chooses some landmarks which are of limited interest. All states with a qualitative magnitude adjacent to these landmarks are regarded as equivalent if they also have the same qualitative direction. By this equivalence relation, disjoined subgraphs are defined and an abstracted state-transition graph can be computed.

Restriction Techniques

One simple restriction method is the analytical function constraint (Kuipers 1994). In principle, it is possible that a component of a solution of a monotonic landmark ensemble is constant over a time interval. If we restrict the set of admissible trajectories to functions which are analytical (at least in some components), this is only possible if the component remains constant forever. From the modelling perspective this restriction is reasonable if a constant evolution of one variable over some time must be regarded as a very improbable marginal case in the real world. If an analytical function constraint with a prescribed index set $I \subseteq \{1, \ldots, n\}$ is applied to a state-transition graph, all vertices v are for which at least one $i \in I$ exists such that $\mathrm{qdir}_i(v) = 0$ are eliminated. The effect of this restriction technique can be increased if it is integrated into the QSIM algorithm to reduce the number of states which have to be checked at each filtering step. The analytical function constraint is used for some of the applications in Chapter 4.

Phase plane constraints are a path-dependent technique (Lee and Kuipers 1988). Since a QDE contains only autonomous systems, non-constant trajectories cannot intersect. For each tuple of landmarks from the quantity space Q and a landmark vector a specific state in X is given. The set of admissible trajectories is restricted to the reasonable functions which always pass through such states in the same direction. However, this is not very restrictive for high dimensional systems.

The integration of quantitative knowledge to QDEs, so called semi-qualitative reasoning, is still a challenge (Berleant and Kuipers 1992; Kuipers 1994; Berleant and Kuipers 1998; Kay 1998; Moldenhauer et al. 1999). If we define a qualitative model, we possibly disregard quantitative knowledge which may be (partially) available: Quantitative knowledge cannot be expressed by a monotonic landmark ensemble. If we had full quantitative knowledge, we could restrict the model to a single ODE. In the more interesting case, we come up with a "hybrid" model ensemble by restricting a monotonic landmark ensemble (partially) with quantitative constraints. For such a semi-qualitative reasoning three types of quantitative knowledge are considered: (i) quantitative landmark intervals, (ii) functional envelopes, and (iii) temporal envelopes. For the first type, $\mathcal{M}(\mu, C)$ is restricted by constraining the values a landmark vector can take, e.g. by prescribing intervals $J_{i,j} \subset \mathbb{R}$ with $\lambda_{i,j} \in J_{i,j}$. For the second type, we define two functions $\underline{f}, \bar{f} : X \to \mathbb{R}^n$ and restrict $\mathcal{M}(\mu, C)$ to the models f such that $\forall x \in X : \underline{f}(x) \leq f(x) \leq \bar{f}(x)$ (inequalities defined component wise). By introducing temporal envelopes we restrict the space of admissible trajectories $\mathcal{E}$ with two functions $\underline{x}, \bar{x} : \mathbb{R}_+ \to X$ by requiring that $\forall x(\cdot) \in \mathcal{E}, t \geq 0 : \underline{x}(t) \leq x(t) \leq \bar{x}(t)$.

Combinations of these restrictions can be used to infer estimates for the sets of quantitative states which are consistent with a qualitative state v, or for the length of time intervals where a solution of the monotonic landmark ensemble has v as state abstraction. Various partially successful procedures have been proposed, but an efficient solution for systems with more than three variables is still lacking. The basic problem are tight approximations of solutions of interval-valued differential inclusions with state constraints (see section 2.3, p. 42),

e.g. of the form

$$\begin{aligned} \dot{x} &\in [\underline{f}(x), \bar{f}(x)] \\ \text{subject to} & \\ x(0) &\in J, \\ \forall t \in \mathbb{R}_+ &: [\dot{x}(t)] = \mathrm{qdir}(v) \\ \text{and } x(t) &\in [\underline{x}(t), \bar{x}(t)], \end{aligned}$$

where $J \in \mathbb{R}^n$ describes an interval of possible initial values and $v \in S$ a qualitative state under consideration. The new method I develop in section 3.4 (p. 77) is a step in this direction.

Analysis Techniques

When a state-transition graph cannot be restricted or abstracted further, it may still be too large to be displayed or analysed by hand. In this case, we can interpret the graph as a data base and can employ search algorithms on this data to find paths which satisfy prescribed properties (Brajnik and Clancy 1996; Clancy 1997; Shults and Kuipers 1997). Temporal logic together with modal and standard predicate logic is a rich language to formulate expressions like "a path where $[\dot{x}_i] = [\dot{x}_j] = [-]$ after $[\dot{x}_k] = [+]$ is impossible". An introduction to temporal logic is beyond the scope of this text. However, it should become clear that effective algorithms to test whether a state-transition graph fulfils given expressions of this kind can contribute to the understanding and interpretation of large QDEs.

In Chapter 3, I develop one novel abstraction and three new restriction methods which all fit into the scheme proposed in this section. The abstraction technique draws from basic concepts of viability theory, which will be introduced in section 2.4 (p. 45). The other techniques (i) exclude trajectories which represent unlikely developments in some specified sense, (ii) restricts monotonic landmark ensembles to models which satisfy a prescribed order on the coefficients of the Jacobian, and (iii) restricts them to models with quantitative interval bounds on the Jacobian. Before going on to this, I first introduce differential inclusions.

2.3 Differential Inclusions

In this section, basic concepts of differential inclusions are reviewed from the literature. This is a prerequisite for viability theory, which will be indispensable in Chapter 3, and provides a further elaborate example for a model ensemble (cf. section 2.1, p. 17). This contributes to the ongoing discussion about the relation of differential inclusions and QDEs. Differential inclusions are a generalisation of ordinary differential equations. An ODE assigns a single velocity to points in the state space, and is thus a special case of a differential inclusion, where *multiple* velocities can be assigned. We map a state x to a set of possible velocities $F(x)$, and admit a trajectory $x(\cdot)$ as a solution, if $\dot{x}(t)$ is always an element of $F(x(t))$. As in the case of QDEs we cannot generally expect to obtain unique solutions in such a setting, yielding a set-valued solution operator. The first ideas to this approach arose in the 30s of the last century (Zaremba 1936; Marchaud 1934), were the existence of continuous solutions was investigated under the terms of "contingent" or "paratingent" equations. Later, absolutely continuous solutions were considered (Ważewski 1961a; Ważewski 1961b). Filippov (1959) put differential inclusions into the context of optimal control. A broad overview to the fundamentals and subsequent development of the theory is provided by Aubin and Cellina (1984). One basic motivation – similar to QDEs – is to consider uncertainties which cannot be expressed in a probabilistic way. We may have an ODE $\dot{x} = f(x, t; u)$, depending on a parameter or a control u. If we do not know u exactly but can restrict the value, say, to an interval J such that $u \in J$, we obtain a set of possible values $F(x, t) := \{f(x, t; u) \mid u \in J\}$. Since the analysis of differential inclusions is more complicated than for ODEs, open research problems include accurate and efficient numerical schemes for state spaces of higher dimension and for differential inclusions with weak regularity properties. Differential inclusions are applied to problems from, e.g. population dynamics (Křivan and Colombo 1998; Guo et al. 2003), physics (Maisse and Pousin 1997), climate change (Chahma 2003), nonsmooth analysis (Clarke 1983), control theory (Leonov 2000; Lorenz 2005) and differential games (Chodun 1989; Ivanov and Polovinkin 1995).

We now present some basic definitions. Let X and Y be sets. A **set-valued map** $F : X \to \mathcal{P}(Y)$ is a map assigning to any $x \in X$ an element $F(x)$ of the power set of Y; we also write $F : X \rightsquigarrow Y, x \rightsquigarrow F(x)$. We denote the **domain** of F by $\mathrm{Dom}(F) := \{x \in X \mid F(x) \neq \varnothing\}$. A set-valued map F is called nontrivial if $\mathrm{Dom}(F) \neq \varnothing$. The **graph** of F is $\mathrm{Graph}(F) := \{(x, y) \in X \times Y \mid y \in F(x)\}$. A standard example for set-valued maps are parameterised maps. We consider the sets X, Y and U and a (single-valued) map $f : X \times U \to Y$. Then $F(x) := \{f(x, u) \mid u \in U\}$ is called a set-valued map parameterised by U. There are several regularity concepts for set-valued maps which are helpful in practice, e.g. the following three. For a metric space X and $K \subseteq X$, the distance from $x \in X$ to K is defined by $d(x, K) := \inf_{y \in K} d(x, y)$, $d(x, \varnothing) = +\infty$, and $B(K, r) := \{x \in X \mid d(x, K) \leq r\}$ denotes the ball around K with radius $r > 0$.

DEFINITION 12: *Let $F : X \rightsquigarrow Y$ be a nontrivial set-valued map from a metric space X to a metric space Y. It is called*

1. ***Lipschitz** with constant $L \geq 0$ if*

$$\forall x, x' \in X : F(x) \subseteq B\big(F(x'), Ld(x, x')\big).$$

2. ***Upper semicontinuous*** *at $x \in X$ if for any open neighbourhood $N \supseteq F(x)$ there exists a neighbourhood $M \supseteq x$ such that $F(M) \subseteq N$. It is called upper semicontinuous, if it is upper semicontinuous at any $x \in X$.*
3. ***Marchaud*** *if it is upper semicontinuous, has compact convex images and has linear growth, i.e. there exists a positive constant c such that*

$$\forall x \in \mathrm{Dom}(F) : \|F(x)\| \leq c(\|x\| + 1),$$

where $\|F(x)\| := \sup_{y \in F(x)} \|y\|$.

Set-valued maps which are parameterised by a closed set U and a (single-valued) continuous map f with linear growth such that $F(x)$ has convex values are always Marchaud (Aubin 1991, p. 203). It can be shown that Marchaud maps are characterised by nonempty and closed $\mathrm{Dom}(F)$, nonempty and closed $\mathrm{Graph}(F)$, convex values and linear growth (Aubin 2001).

For a given set-valued map $F : X \rightsquigarrow Y$ an "equation" of the form

$$\begin{aligned} \dot{x} &\in F(x), \\ x(0) &= x_0, \end{aligned} \tag{2.14}$$

is called a **differential inclusion**. What we define as solution of a differential inclusion depends on the chosen space of admissible trajectories. The situation is more complicated than for ODEs which have differentiable solutions if the right-hand side is continuous. As a standard case, we call an absolutely continuous function $x(\cdot) : J \to X$ on an interval $J = [0, T]$, possibly $T = \infty$ a **solution** of Eq. (2.14) if $x(0) = x_0$ and $\dot{x}(t) \in F(x(t))$ almost everywhere on J. For sake of simplicity we assume in the following that $X, Y \subseteq \mathbb{R}^n$, although results also hold for more general cases. There are various theorems on the existence of solutions to a differential inclusion (see e.g. Aubin 1991, p. 172).

PROPOSITION 6: *Let X be an open set, and $F : X \rightsquigarrow \mathbb{R}^n$ a set-valued map which is Lipschitz on X. Then, for any $x_0 \in X$ and $v_0 \in F(x_0)$, there exists a $T > 0$ and a solution $x(\cdot)$ to the differential inclusion Eq. (2.14) on $[0, T]$ such that $x(0) = x_0$ and $\dot{x}(0) = v_0$.*

How are differential inclusions related to model ensembles and causal loop diagrams? A Lipschitz map $F : X \rightsquigarrow \mathbb{R}^n$ and the resulting differential inclusion defines a model ensemble (cf. section 2.1, p. 17) by

$$\begin{aligned} \mathcal{M} := \{ f : X \times \mathbb{R}_+ \to \mathbb{R}^n \mid & f(x,t) \text{ measurable with respect to } t \\ & \text{and } \forall t \in \mathbb{R}_+ : f(x,t) \in F(x) \}. \end{aligned}$$

Taking the set of absolutely continuous functions on intervals $J = [0, T]$ as space of admissible trajectories $\mathcal{E}$, we obtain a set-valued solution operator $\mathcal{S}_F(\cdot) : X \rightsquigarrow \mathcal{E}$,

$$\begin{aligned} \mathcal{S}_F(x_0) := \{ x(\cdot) \in \mathcal{E} \mid & x(0) = x_0, \exists f \in \mathcal{M} : \\ & \dot{x}(t) = f(x(t), t) \text{ almost everywhere } \}, \end{aligned}$$

which assigns to an initial value $x_0 \in X$ the set of solutions of Eq. (2.14).

To find all possible trajectories which can be brought about by a causal loop diagram, one could start with a model ensemble $\mathcal{M}$ which contains all (autonomous) models consistent with the causal loop diagram (according to a concise interpretation of the diagram). We could then define a set-valued map by $F(x) := \{f(x) \mid f \in \mathcal{M}\}$ such that the solutions of the differential inclusion Eq. (2.14) describe all trajectories. However, if a causal loop diagram is specified by a sign matrix $\Sigma = (\sigma_{i,j}) \in \mathcal{A}_*^{n\times n}$ and a monotonic ensemble $\mathcal{M}(\Sigma)$, we run into trouble, as the following shows:

Suppose that $f \in \mathcal{M}(\Sigma)$. Since it follows from $\dot{x} = f(x)$ that $\ddot{x} = \mathcal{J}(f)(x) \cdot \dot{x}$, we obtain a second order differential inclusion in the joint state and velocity space:

$$\begin{aligned} \ddot{x} &\in F(\dot{x}, x), \\ F : (\dot{x}, x) &\rightsquigarrow \{\mathcal{J}(f)(x) \cdot \dot{x} \mid f \in \mathcal{M}(\Sigma)\}. \end{aligned}$$

This can be simplified to

$$\ddot{x} \in \hat{F}(\dot{x}) := \{A\dot{x} \mid [A] \approx \Sigma\},$$

where A denotes $n \times n$ matrices over the real numbers. We observe that the components $i = 1, \ldots, n$ of $\hat{F}(\dot{x})$ evaluate to

$$\hat{F}_i(\dot{x}) = \begin{cases} 0 & \text{if } \forall j = 1, \ldots, n : \dot{x}_j \cdot \sigma_{i,j} = 0, \\ \mathbb{R}_+ \setminus \{0\} & \text{else if } \forall j = 1, \ldots, n : [\dot{x}_j] = \sigma_{i,j} \neq 0 \text{ or } \dot{x}_j \cdot \sigma_{i,j} = 0, \\ \mathbb{R}_- \setminus \{0\} & \text{else if } \forall j = 1, \ldots, n : -[\dot{x}_j] = \sigma_{i,j} \neq 0 \text{ or } \dot{x}_j \cdot \sigma_{i,j} = 0, \\ \mathbb{R} & \text{otherwise .} \end{cases}$$

Except the trivial case, this unbounded set-valued map is not Lipschitz and not Marchaud. Also other well-known regularity concepts do not apply to $\hat{F}$. Thus, it cannot be expected that this simple approach provides valuable results. In section 3.4 (p. 77) I develop an alternative with better smoothness properties.

2.4 Viability Theory

Viability theory provides an elegant mathematical framework to consider normative settings in a model-driven analysis in a transparent way. Here, the basic definitions and results are recalled from the literature. They are the foundation for a new abstraction technique in section 3.1 (p. 52). I also outline the viability kernel algorithm, which is needed in section 3.4 (p. 77).

Viability theory considers whether prescribed state constraints can be satisfied by a set of trajectories. A viability constraint is a closed subset K of the state space X. It is assumed that in a state outside K the system is no longer viable. If a trajectory $x(\cdot)$ is given as a solution of an ODE with initial value x_0, and it remains in K forever, we call it viable. If we deal with a model ensemble, admitting multiple solutions starting from an initial state, the situation becomes more interesting: it may be that all solutions are viable, or at least one solution, or none. We can also formulate the control problem of how to steer a system such that it produces a viable solution. It turns out that these problems are also closely related to the reachability of a target set $C \subseteq X$, i.e. whether a viable trajectory reaches C in finite time. For ODEs, the first characterisation of viability by Nagumo (1942) has been forgotten and rediscovered several times. Later it was extended to differential inclusions (Yorke 1969; Gautier 1976; Aubin et al. 1977; Haddad 1981). A very comprehensive introduction is provided by Aubin (1991). The strength of viability theory is that it provides a consistent framework to conceptualise and analyse non-deterministic or uncertain dynamics under state constraints. By concentrating on initial states which admit viable solutions, questions about dynamics are reduced to geometrical considerations. This is in particular valuable to investigate complex control problems. However, as for differential inclusions, numerical schemes for viability theory are still difficult in state spaces of higher dimensions. Fields of applications include economics (e.g. Aubin et al. 2001; Aubin et al. 2005), engineering (e.g. Seube et al. 2002), population dynamics (e.g. Bonneuil and Müllers 1997; Bonneuil and Saint-Pierre 2005), and sustainability science (Petschel-Held et al. 1999; Bene et al. 2001; Eisenack et al. 2006; Cury et al. 2005).

Normative settings, which are important for sustainability issues, can be formalised by defining viability constraints as sets which subsume preferable or problematic states of the system under investigation. For example, it may be relevant to know whether there are trajectories reaching a problematic set from an initial value in finite time. In the “hopeless” case, all trajectories will become non-viable (and thus the system has to be altered structurally to make it sustainable). In the “foolproof” case, all trajectories will remain viable forever. Then, the system is on the safe side although we have to face the uncertainty or the generality expressed by the model. A third possibility is the “critical” case with both viable and non-viable solutions – the “fate” of the system depends on decisions (for a control problem) or on unknown system properties. This approach differs from usual criteria like optimality, which force the modeller to specify unique solutions. Traditionally, viability theory is based on differential inclusions (and difference inclusions for numerical schemes). In Chapter 3 we will see how qualitative reasoning profits from this approach, since we can apply similar concepts to the state-transition graph of a QDE.

2.4.1 The Viability Theorem

To define the basic objects of this section, the viability and invariance kernel, let $K \subseteq X$ be a subset of the the state space called a **constrained set**. For simplicity we assume that $X \subseteq \mathbb{R}^n$, although most results hold for more general cases. A trajectory $x(\cdot) : J \to X$, $J = [0, T]$ (possibly with $T = \infty$) which remains in K, i.e. $\forall t \in J : x(t) \in K$, is called **viable** in K on J. If multiple trajectories start from a given initial value $x_0 \in K$, we ask whether *all* trajectories with $x(0) = x_0$, or whether *at least one* trajectory has this property. These questions are further elaborated by considering a target $C \subseteq K$, which we prefer to reach with a viable trajectory in finite time.

DEFINITION 13: *Let F be a set-valued map on the state space X, defining a differential inclusion and the set-valued solution operator $\mathcal{S}_F(\cdot)$. For $C \subseteq K \subseteq X$ the*

1. ***Viability kernel** of K with target C, denoted by $\mathrm{Viab}_F(K, C)$, is the set of all $x_0 \in K$ such that*

$$\begin{aligned} &\exists\, x(\cdot) \in \mathcal{S}_F(x_0) \\ &\qquad \forall\, t \geq 0 : x(t) \in K \\ &\qquad \text{or } \exists\, T > 0 : x(T) \in C \text{ and } \forall\, t \in [0, T] : x(t) \in K. \end{aligned}$$

2. ***Invariance kernel** of K with target C, denoted by $\mathrm{Inv}_F(K, C)$, is the set of all $x_0 \in K$ such that*

$$\begin{aligned} &\forall\, x(\cdot) \in \mathcal{S}_F(x_0) \\ &\qquad \forall\, t \geq 0 : x(t) \in K \\ &\qquad \text{or } \exists\, T > 0 : x(T) \in C \text{ and } \forall\, t \in [0, T] : x(t) \in K. \end{aligned}$$

3. ***Capture basin** $\mathrm{Capt}_F(K, C)$ is the set of all $x_0 \in K$ such that $\exists\, x(\cdot) \in \mathcal{S}_F(x_0)$ for which $\exists\, T > 0 : x(T) \in C$ and $\forall\, t \in [0, T] : x(t) \in K$.*
4. ***Absorption basin** $\mathrm{Abs}_F(K, C)$ is the set of all $x_0 \in K$ such that $\forall\, x(\cdot) \in \mathcal{S}_F(x_0)$ there $\exists\, T > 0 : x(T) \in C$ and $\forall\, t \in [0, T] : x(t) \in K$.*

If the set-valued map F is clear from the context, it is usually omitted as a subscript. As important examples $\mathrm{Viab}_F(K, \varnothing)$ contains the initial states from which at least one viable solution starts, whereas $\mathrm{Inv}_F(K, \varnothing)$ contains the initial states for which all trajectories are viable. These important cases are simply called the viability kernel of K, $\mathrm{Viab}_F(K)$, or invariance kernel of K, $\mathrm{Inv}_F(K)$. We can thus define the following types of sets which we aim to characterise geometrically by the viability theorem.

DEFINITION 14: *As set K is **viable** if $K = \mathrm{Viab}_F(K)$, i.e.*

$$\forall\, x_0 \in K\, \exists\, x(\cdot) \in \mathcal{S}_F(x_0)\, \forall\, t \geq 0 : x(t) \in K.$$

*It is **invariant** if $K = \mathrm{Inv}_F(K)$, i.e.*

$$\forall\, x_0 \in K, x(\cdot) \in \mathcal{S}_F(x_0), t \geq 0 : x(t) \in K.$$

*It is **locally invariant** if*

$$\forall\, x_0 \in K, x(\cdot) \in \mathcal{S}_F(x_0)\, \exists\, T \geq 0\, \forall\, t \in [0, T] : x(t) \in K.$$

These concepts can also be translated to difference inclusions $x_{i+1} \in F(x_i), F : X \rightsquigarrow X$, where the solution operator $\mathcal{S}_F(x) := \{(x_i) \mid x_0 = x, \forall i \in \mathbb{N} : x_{i+1} \in F(x_i)\}$ contains all sequences of states which start from x and evolve according to the set-valued map F. This discrete version is needed for the viability kernel algorithm which is presented in the next subsection.

DEFINITION 15: *A sequence $(x_i) \in \mathcal{S}_F(x_0)$ is said to be **viable** in $K \subseteq X$ if $\forall i \in \mathbb{N} : x_i \in K$. The **discrete viability kernel** of K with target $C \subseteq K$, denoted by $\mathrm{Viab}_F(K, C)$, is the set of all x_0 such that there exists at least one sequence $(x_i) \in \mathcal{S}_F(x_0)$ which is viable in K or for which there is an $N \in \mathbb{N}$ with $x_N \in C$ and $\forall i \leq N : x_i \in K$. Again, we set $\mathrm{Viab}_F(K) := \mathrm{Viab}_F(K, \varnothing)$.*

We are interested in the structure of viable sets, and how it depends on F and the constrained set K. There are various such characterisations, and a simple one for discrete viability kernels:

PROPOSITION 7: *The discrete viability kernel $\mathrm{Viab}_F(K, C)$ is the largest closed set $D \subseteq K$ such that*

$$\forall x \in D \setminus C : F(x) \cap D \neq \varnothing.$$

In the viability and invariance theorem for differential inclusions, viable sets are characterised by tangential conditions. To state them, we need another definition:

DEFINITION 16: *Let $K \subseteq X$ be nonempty and $x \in K$. The **contingent cone** to K at x is the set*

$$T_K(x) := \left\{ v \in \mathbb{R}^n \mid \liminf_{h \to 0+} \frac{d(x + hv, K)}{h} = 0 \right\}.$$

The contingent cone contains the directions which point into K from a given point x in some sense. If x is on the boundary of a smooth manifold K, it is identical to the tangent space to K at x. For $x \in \mathrm{Int}(K)$ we have $T_K(x) = \mathbb{R}^n$, and if K is a singleton, $T_K(x) = 0$. If K is an open set, the contingent cone is always the whole space. The above definition also works for nonsmooth K. Sets with the following properties are viable or invariant, under certain regularity assumptions for F.

DEFINITION 17: *Let $F : X \rightsquigarrow X$ be a nontrivial set-valued map. We call a set $K \subseteq \mathrm{Dom}(F)$ a **viability domain** of F if*

$$\forall x \in K : F(x) \cap T_K(x) \neq \varnothing,$$

*and an **invariance domain** of F if*

$$\forall x \in K : F(x) \subseteq T_K(x).$$

Consequently, every open set is a viability and an invariance domain. The viability theorem states the following existence result (cf. Aubin 1991, p. 91):

PROPOSITION 8: *Let F be a Marchaud map. A closed set $K \subseteq \mathrm{Dom}(F)$ is viable iff K is a viability domain.*

The invariance theorem holds for Lipschitz maps (cf. Aubin 1991, p. 173).

PROPOSITION 9: *Let F be Lipschitz on $\mathrm{Int}(\mathrm{Dom}(K))$ with compact values. A closed set $K \subseteq \mathrm{Dom}(F)$ is locally invariant iff K is an invariance domain.*

Interestingly, viability and invariance kernels are maximal closed viability domains and invariance domains contained in the constrained set (Aubin 2001).

PROPOSITION 10: *If F is Marchaud, K and $C \subseteq K$ are closed, the viability kernel $\mathrm{Viab}_F(K, C)$ of K with target C is the largest closed subset D with $C \subseteq D \subseteq K$ satisfying*

$$\forall x \in D \setminus C : F(x) \cap T_D(x) \neq \varnothing.$$

PROPOSITION 11: *If F is Lipschitz, K and $C \subseteq K$ are closed, the invariance kernel $\mathrm{Inv}_F(K, C)$ of K with target C is the largest closed subset D with $C \subseteq D \subseteq K$ satisfying*

$$\forall x \in D \setminus C : F(x) \subseteq T_D(x).$$

2.4.2 The Viability Kernel Algorithm

Viability and invariance kernels as well as capture and absorption basins of a broad class of differential inclusions can be numerically approximated by the viability kernel algorithm (Frankowska and Quincampoix 1991; Cardaliaguet et al. 1994; Saint-Pierre 1994; Quincampoix and Saint-Pierre 1995; Cardaliaguet et al. 1999). It can be described in two steps. At first, time is discretised by replacing the differential inclusion by an appropriate difference inclusion (the semi-discrete scheme). Then, the fully discrete scheme introduces a grid for the state space and makes a further modification of the difference inclusion necessary. It can be shown that by refining the grid, the associated discrete viability kernel converges towards the viability kernel of the underlying differential inclusion. There are specialised numerical schemes for different types of differential inclusions and some methods to improve computation efficiency, e.g. the refinement principle which avoids re-computation of grid points in subsequent iterations. Some basic results from Cardaliaguet et al. (1999) are laid out in the following paragraphs. For a difference inclusion $x_{i+1} \in G(x_i)$ on the state space X, the discrete viability kernel (cf. DEF. 15) can be approximated as follows:

PROPOSITION 12: *Let $G : X \rightsquigarrow X$ be an upper semicontinuous map with compact, nonempty values, and $K \subseteq X$ be closed. Define a decreasing (with respect to inclusion) sequence of closed sets (K_i) by*

$$\begin{aligned} K_0 &:= K \\ K_{i+1} &:= \{x \in K_i \mid G(x) \cap K_i \neq \varnothing\}. \end{aligned}$$

Then, the discrete viability kernel is

$$\mathrm{Viab}_G(K) = \bigcap_{i=0}^{\infty} K_i.$$

This is not sufficient for practical computation, because determining K_i requires quantifying over an infinite set. Moreover, to compute the viability kernel of a *differential* inclusion, we have to relate it to a discretisation by an appropriate *difference* inclusion.

Let a differential inclusion be given by a Marchaud map $F : X \rightsquigarrow Y$ which is bounded by M, i.e. $\forall x \in X, y \in F(x) : \|y\| \leq M$. Choose a family of set-valued maps $F_\epsilon : X \rightsquigarrow Y$ such that the following regularity assumptions are satisfied:

(R1) Each F_ϵ is upper semicontinuous with compact, convex and non-empty values,

(R2) $\mathrm{Graph}(F_\epsilon) \subseteq B(\mathrm{Graph}(F), \phi(\epsilon))$ with $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{\epsilon \to 0} \phi(\epsilon) = 0$,

(R3) $\forall x \in X : \bigcup_{\|y-x\| \leq M\epsilon} F(y) \subseteq F_\epsilon(x)$.

The discrete viability kernel with respect to a certain family of maps G_ϵ converges to the viability kernel with respect to F (see Cardaliaguet et al. 1999):

PROPOSITION 13: *Let $F : X \rightsquigarrow Y$ be a bounded Marchaud map, the family $F_\epsilon : X \rightsquigarrow Y$ satisfy (R0) – (R2), and define $G_\epsilon(x_i) := x_i + \epsilon F_\epsilon(x_i)$. Then,*

$$\operatorname*{Lim}_{\epsilon \to 0} \mathrm{Viab}_{G_\epsilon}(K) = \mathrm{Viab}_F(K).$$

Here, Lim denotes the Painléve-Kuratowski limit for sets which is defined as follows. If $A : \mathbb{R}_+ \rightsquigarrow X$ is a map assigning to each $\epsilon > 0$ a subset $A(\epsilon)$ of X, the upper limit

$$\operatorname*{Limsup}_{\epsilon \to 0} A(\epsilon) := \{x \in X \mid \liminf_{\epsilon \to 0} d(x, A(\epsilon)) = 0\},$$

and the lower limit

$$\operatorname*{Liminf}_{\epsilon \to 0} A(\epsilon) := \{x \in X \mid \lim_{\epsilon \to 0} d(x, A(\epsilon)) = 0\}.$$

A subset $K \subseteq X$ is the Painléve-Kuratowski limit $\operatorname*{Lim}_{\epsilon \to 0} A(\epsilon)$ if

$$K = \operatorname*{Liminf}_{\epsilon \to 0} A(\epsilon) = \operatorname*{Limsup}_{\epsilon \to 0} A(\epsilon).$$

EXAMPLE 7: If F is a Marchaud map bounded by M and Lipschitz with constant L,

$$F_\epsilon(x) := B(F(x), ML\epsilon)$$

satisfies the regularity assumptions (R1)–(R3) (Cardaliaguet et al. 1999). □

The next step is the discretisation of the state space. For $h \in \mathbb{R}_+$ we introduce a grid $X^h \subseteq X$ such that for any compact $K \subseteq X$, the intersection $K \cap X^h$ is finite and $\forall\, x \in X\, \exists\, x^h \in X^h : \|x - x^h\| \leq h$. For convenience, we define $K^h := B(K, h) \cap X^h$. On this grid, the discrete viability kernel of a difference inclusion $x_{i+1} \in H(x_i)$ can be computed in a *finite* number of steps (in contrast to PROP. 12, p. 48):

PROPOSITION 14: *Let $H : X^h \rightsquigarrow X^h$ be a set-valued map with finite, nonempty values, and K^h be finite. Define a decreasing sequence of sets (K_i^h) by*

$$\begin{aligned} K_0^h &:= K^h \\ K_{i+1}^h &:= \{x \in K_i^h \mid H(x) \cap K_i^h \neq \varnothing\}. \end{aligned}$$

Then there exists an $N \in \mathbb{N}$ such that the discrete viability kernel is

$$\mathrm{Viab}_H(K^h) = K_N^h.$$

Again, an appropriate family of maps $H_{\epsilon,h}$ can be constructed which approximates $\mathrm{Viab}_F(K)$ when the grid width h and the time step ϵ is refined. Let F be a bounded Marchaud map as before, F_ϵ a set-valued map satisfying (R1)–(R3), and $G_\epsilon(x_i) := x_i + \epsilon F_\epsilon(x_i)$. Then choose $H_{\epsilon,h}$ such that the following additional regularity assumptions hold:

(R4) $\mathrm{Graph}(H_{\epsilon,h}) \subseteq B(\mathrm{Graph}(G_\epsilon), \psi(\epsilon,h))$ with $\psi : \mathbb{R}_+^2 \to \mathbb{R}_+$ such that $\frac{\psi(\epsilon,h)}{h} \to 0$ for $\epsilon \to 0$ and $\frac{h}{\epsilon} \to 0$.

(R5) $\forall x^h \in X^h : \bigcup_{\|y-x^h\| \leq h} B(G_\epsilon(y), h) \cap X^h \subseteq H_{\epsilon,h}(x^h)$.

EXAMPLE 8: When F is a Marchaud map, bounded by M and Lipschitz with constant L,

$$H_{\epsilon,h}(x^h) := B(x^h + \epsilon F_\epsilon(x^h), 2h + Lh\epsilon + ML\epsilon^2) \cap X_h$$

satisfies the regularity assumptions (R4) and (R5) (Cardaliaguet et al. 1999). □

We now end up with the fully discrete scheme, implicitly including an estimate for the approximation error (see Cardaliaguet et al. 1999).

PROPOSITION 15: *Let $F : X \rightsquigarrow Y$ be a bounded Marchaud map. If $H_{\epsilon,h}$ satisfy (R1) – (R5), then*

$$\mathrm{Viab}_F(K) \subseteq B(\mathrm{Viab}_{H_{\epsilon,h}}(K^h), h),$$

and

$$\lim_{\substack{\epsilon \to 0, \\ \frac{h}{\epsilon} \to 0}} \mathrm{Viab}_{H_{\epsilon,h}}(K^h) = \mathrm{Viab}_F(K).$$