Implications of eigenvector localization for dynamics on complex networks
Implications of eigenvector localization for dynamics on complex networks

Dissertation
zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der
Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von
Helge Eberhard Aufderheide,
geboren am 5. Oktober 1985
in Osterode am Harz.

Eingereicht am 14.4.2014,
verteidigt am 8.9.2014.
THE THESIS WAS CREATED AT:

- Biophysics group of Prof. Dr. Frank Jülicher,
  Max-Planck-Institute for Physics of Complex Systems,
  Nöthnitzer Str. 38, 01187 Dresden, Germany.

- Biological network dynamics group of Dr. Thilo Gross,
  University of Bristol, Dept. of Engineering Mathematics,
  Woodland Road, Bristol BS81UB, United Kingdom.

REFEREES (GUTACHTER):

- Prof. Dr. Frank Jülicher
  MPI für Physik komplexer Systeme, 01187 Dresden.

- Prof. Dr. Günter Radons
  TU Chemnitz, Institut für Physik, 09107 Chemnitz.
Abstract

In large and complex systems, failures can have dramatic consequences, such as black-outs, pandemics or the loss of entire classes of an ecosystem. Nevertheless, it is a centuries-old intuition that by using networks to capture the core of the complexity of such systems, one might understand in which part of a system a phenomenon originates. I investigate this intuition using spectral methods to decouple the dynamics of complex systems near stationary states into independent dynamical modes. In this description, phenomena are tied to a specific part of a system through localized eigenvectors which have large amplitudes only on a few nodes of the system’s network. Studying the occurrence of localized eigenvectors, I find that such localization occurs exactly for a few small network structures, and approximately for the dynamical modes associated with the most prominent failures in complex systems. My findings confirm that understanding the functioning of complex systems generally requires to treat them as complex entities, rather than collections of interwoven small parts. Exceptions to this are only few structures carrying exact localization, whose functioning is tied to the meso-scale, between the size of individual elements and the size of the global network. However, while understanding the functioning of a complex system is hampered by the necessary global analysis, the prominent failures, due to their localization, allow an understanding on a manageable local scale. Intriguingly, food webs might exploit this localization of failures to stabilize by causing the break-off of small problematic parts, whereas typical attempts to optimize technological systems for stability lead to delocalization and large-scale failures. Thus, this thesis provides insights into the interplay of complexity and localization, which is paramount to ascertain the functioning of the ever-growing networks on which we humans depend.
Contents

1 Introduction 1

2 Concepts and Tools 7
  2.1 Networks ............................................. 7
  2.2 Food webs ........................................... 9
  2.3 Dynamics on networks ................................ 11
  2.4 Steady state operating modes ........................ 13
  2.5 Bifurcations affecting operating modes ............... 16
  2.6 Dynamical modes ..................................... 17
  2.7 Generalized models for food webs ..................... 18

3 Perturbation Impact 23
  3.1 Impact of perturbations on food webs ................. 24
  3.2 Examples ............................................. 27
  3.3 Impact formulation with dynamical modes ............. 30
  3.4 Influence and sensitivity of species ................. 32
  3.5 Localized dynamical modes .......................... 35
  3.6 Iterative parameter estimation ....................... 37
  3.7 Most important parameters and species ............... 40
  3.8 Discussion ........................................... 42

4 Exact Localization 44
  4.1 Graph symmetries .................................... 47
  4.2 Localized dynamics on symmetries .................... 49
  4.3 Exactly localized dynamics ........................... 53
  4.4 Symmetry reduction in networks ...................... 54
  4.5 Application to food webs ............................ 56
  4.6 Localization on asymmetric structures ............... 61
  4.7 Nearly-exact localization ............................ 66
  4.8 Other systems ........................................ 67
  4.9 Discussion ........................................... 69
5 Approximate Localization

5.1 Spread of a dynamical mode ............................................. 74
5.2 Examples for localized instabilities ................................. 75
5.3 Localization of extreme eigenvalues ................................. 88
5.4 Dependence on the system size ....................................... 91
5.5 Localization in the model of R. May ................................. 94
5.6 Finding motifs that carry localization ............................... 97
5.7 (Self-)stabilization of food webs ................................. 101
5.8 Repairing localized instabilities ................................. 105
5.9 Discussion ................................................................. 110

6 Conclusions ........................................................................ 112

Acknowledgments ................................................................... 117

Appendix .............................................................................. 119

A Parametrization of the Gatun Lake food web .................. 119
B The Master Stability Function approach ....................... 123
C Approximate localization on larger structures ................. 128

Bibliography ........................................................................ 131
1 Introduction

On several levels, humans depend on the functioning of complex networks. Such networks are not only technological networks of communication or electrical power distribution, but also complex food chains in ecosystems, biological protein interactions in cells, and networks of social interactions with friends and colleagues. All of these networks consist of large numbers of individual elements, such as computers or machines, species, proteins or persons, which are connected according to a complicated pattern of interactions. Already the individual elements in such a network can be extremely diverse. But, it is through the complicated network of their interdependencies, that the system becomes different from these parts [1–4]. On the one hand, networks of interactions allow the functioning of the Internet, the formation of biological organisms and cells, or of joint efforts in social groups [5–7]. On the other hand they also exhibit failures, such as blackouts, the spreading of cancerous diseases, or they propagate cascades of extinctions through ecosystems [8–10].

To understand the interdependent structure of a network system, it is often instructive to study the graph of its interactions [1]. In this graph, each element is represented as a node, and edges connect those nodes which depend on each other. The oldest example for such a representation is probably the problem of the Seven Bridges of Königsberg, which was solved by Leonhard Euler in the early 18th century [11]. The problem is to find a path that allows to walk over all seven bridges connecting the four different city parts of Königsberg, without crossing any of these bridges twice. As Euler realized, this problem depends only on the information, which parts of the city each bridge connects, but not on the properties of the city parts or bridges. Thus, representing it as a graph of nodes (city parts) and edges (bridges) reduces the problem to its core.

The problem of the seven bridges can be solved by an intuitive trial-and-error approach. For instance, choosing a node as the beginning of a path, one can try different paths to see if one of them is a suitable solution. After trying all possible paths beginning in all possible nodes, this approach leads to the solution, that none of them crosses each bridge exactly once. However, it is easy to imagine that this approach becomes prohibitively tedious if the number of bridges or city parts is large. Each time a new path is tried out, one has
to keep track of the whole system to decide whether the path is a solution to the problem or not. In other words, the problem with this approach is its dependency on the knowledge of the global path.

The solution to this problem is much simpler when counting the bridges that end in each individual part of the city. When a node is connected by an odd number of edges (bridges) to other nodes, then this node can only be the starting or endpoint of the required path over all bridges because entering and leaving the node always requires two edges. For a node with an odd number of edges and after passing through this node often enough with the path, only one bridge remains unused, such that the node can only be entered once more, but cannot be left afterwards. Thus, a path that uses all bridges is only possible, if either no node has an odd number of edges, or if there are exactly two of them (one is the starting point and one is the end point of the walk). However, at the time the problem was formulated, all four city parts of Königsberg had odd numbers of bridges, such that no suitable path existed.

The approach of counting bridges is much simpler than the trial-and-error approach. Instead of the global knowledge required for testing each possible path, it requires only the independent local knowledge of the number of bridges connecting to the different city parts. Because of this requirement of local information, the approach remains feasible also when increasing the number of bridges or city parts.

The great appeal of Euler’s solution to the Seven Bridges of Königsberg originates not only in the reduction to a graph representation, but also in the ability to identify those nodes and properties that preclude the existence of a suitable path. This ability touches on the fundamental intuition of complex network research to detect the origin of an observed emergent phenomenon in a specific part of a network. Ultimately, if all phenomena in a network were explained by the properties of such small parts, these could then be used as building blocks to understand the network as a whole. Much research aims at this, for instance, by relating phenomena to small structures, called motifs, to closely interlinked parts, called communities or even to the properties of single nodes [1, 12–15]. One successful example is the competitive exclusion principle which precludes the simultaneous existence of similar predators with the same prey in an ecosystem [16]. Already, a few such examples linking network structure and phenomena were identified and studied heuristically [17–19]. However, rigorously linking emergent phenomena to local properties or structures remains a largely unsolved problem [20].

While a network representing simplifies a system and makes new tools available for its analysis, it does not resolve the complexity of the original problem. For instance, for the Seven Bridges, the graph representation merely facilitates finding those quantities that make the problem simple to solve, but the com-
plexity of the initial problem is conserved. Thus, understanding a phenomenon that truly emerges from a complex interaction pattern requires to study a pattern of equal complexity also in the network representation.

When considering systems in nature, it is intuitive to expect global phenomena which affect a large proportion of a system, to have their origin on the meso-scale, in parts of the system consisting of only a few nodes. For instance, the phrase “spreading of a disease” suggests that an epidemic has a specific origin, such as a particular individual with the ability to rapidly transmit the disease to many others. Or, the extinction of populations of large fish in an ecosystem might be the consequence of phenomena that originally affected a few species further down the food chain [21]. Intriguingly, meso-scale phenomena are observed in many network systems. Examples include spider-webs, in which excessive loads lead to local brittleness and destruction of only a small part of the web, localized isolating electronic states in disordered crystals, persistence of diseases on small structures of a network, localization of failures in decentralized power grids, or extinctions in ecological food webs [8, 22–25]. To understand these effects requires methods that are not global, but aimed at the meso-scale of their origin.

In different fields of network research, progress has been made by mapping the phenomenon under consideration to an eigenvalue problem $Mv = \lambda v$, where the matrix $M$ describes the system under consideration, and which one solves for the eigenvectors $v$ and their corresponding eigenvalues $\lambda$. Examples of this general approach include linear stability analysis of dynamical systems, the fundamental Schrödinger equation in quantum mechanics, the master-stability function approach for coupled oscillators, spectral centrality measures, such as the PageRank, or spectral analysis of stochastic processes [1, 26–32].

The essence of an eigenvalue description is that the phenomena in question are decoupled into independent dynamical modes, each associated to one eigenvalue. Each dynamical mode describes one particular elementary phenomenon and the corresponding eigenvalue characterizes if this phenomenon appears or not. For instance, the positive eigenvalues of the Jacobian matrix characterize specific outbreaks of a contagious disease in a network of social interactions, or specific instabilities of the coexistence of species in an ecosystem [26, 27, 33]. Through the eigenvalue description, the phenomena associated to different modes can therefore be studied and understood independently.

Strictly speaking, the eigenvalue computation does not solve the complicated underlying problem. Instead, the eigenvalue computation merely hides the complexity of the problem in a well-established routine, which transforms the phenomena into a base where they are local (the base of the eigenvectors). However, the eigenvalue computation is itself an extremely complex global routine, which implicitly requires knowledge of the entire system (for instance, to
establish the characteristic polynomial). Thereby it severely hampers the understanding of the underlying mechanisms giving rise to each of the resulting modes.

The goal of this thesis is to develop methods linking the eigenvalue description of network structure with meso-scale phenomena. This is possible through localization of eigenvectors, eigenvectors that have non-zero entries only on a small part of a system. While the eigenvalue of dynamical mode characterizes if the associated phenomenon occurs, the corresponding eigenvector indicates where it occurs in the system [26]. Therefore, if a phenomenon originates in a small part of a system, this is reflected by a localized eigenvector. Thus, localization of an eigenvector bridges the gap between the global eigenvalue description and the local phenomenon it describes.

Through an understanding of localization on networks, applying the successful eigenvalue approach to meso-scale phenomena becomes possible without the drawbacks of its global character. For instance, it allows to identify phenomena that admit a local description and which can therefore be studied even in prohibitively large embedding system. Further, one can distinguish local phenomena from those that are truly global in nature and which therefore require an understanding of the complex structure of the entire system. Finally, by allowing the analysis of phenomena on their inherent meso-scale, an understanding of localization provides insights into the interplay of the complexity of a network and the observable emergent phenomena.

Localization of eigenvectors has been observed in different fields of research. Most prominent is Anderson-Localization of electron eigenfunctions around impurities in disordered media [34], or the Mott-Localization of electrons in metals [35] (see e.g. Ref. 23 for an overview containing no less than 19 criteria for localization for electron eigenfunctions). Other examples are random matrix theory [36, 37], real-world data-sets [38], and matrices describing network structures [24, 39–41].

In this thesis, I focus on localization in food webs, the networks of who-eats-whom of species in an ecosystem, which are a paradigmatic example of a complex system. Based on these networks, I develop methods to study localization and build up an understanding for its emergence. Afterward, I generalize these findings to different contexts, such as social or technological networks.

The thesis starts with a short introduction of the central concepts and methods (Chap. 2). This includes, in particular, a short review of networks and their characteristic quantities, of ecological food webs and their dynamics, and an introduction of the eigenvalue description of such dynamics in terms of decoupled dynamical modes.
In Chap. 3, I present a method for the study of impact of perturbations on the dynamics of complex systems. I apply it to assess the effects of a new species invading an ecosystem. I show that the observed phenomena depend primarily on the properties of a few particularly sensitive or influential species. Focusing on these species in a data-gathering process strongly improves the quality of predictions when only incomplete knowledge is available. Further, I link the unusually high influence and sensitivity of these species to localized dynamical modes. Thus, perturbation analysis is an example where the localization of eigenvectors has direct practical consequences.

In Chap. 4, I study small structures in networks that imply the exact localization of dynamical modes, dynamical modes associated with eigenvectors that contain nonzero entries only for the nodes of these small structures. One example for such structures are small symmetric network motifs consisting of a few nodes [42]. The associated symmetry implies localized dynamical modes on the associated motif regardless of the embedding network. Whenever such a network motif appears in a network, it is the origin of specific emergent dynamics. Studying such motifs leads to an exact reduction algorithm which allows to separate the local dynamics due to specific motifs from the global dynamics of the remaining network. For food webs, this separation explains why food webs with different interaction patterns can show identical apparent dynamics if the food webs differ only by motifs associated to localized modes [43]. Finally, investigating which motifs typically lead to exact localization in networks, I conclude that, while a network commonly contains these motifs, it is generally not possible to decompose a network’s dynamics entirely into the dynamics of such motifs. Exact localization therefore rigorously ties dynamics to motifs in a network, but the dynamics of a system is more than the superposition of such localized dynamics.

In Chap. 5, I focus on the localization of dynamical modes in networks that describe the most likely failures or most prominent dynamical phenomena. For instance, I study the most likely outbreak centers of diseases in a network of social interactions, and the loss of synchrony in a network of coupled oscillators. Using the fundamental random matrix Model proposed by R. May (see Ref. 44) for illustration, I show that the particularly large eigenvalues describing these phenomena lead to the approximate localization of the associated eigenvectors. On the one hand, this means that failures in natural systems associated to such eigenvalues generally have a local origin. On the other hand, I show that attempting to repair localized problems in networks generally reduces their localization and may therefore turn small failures into large-scale ones.

In summary, the thesis focuses on localized dynamics in complex systems. This is dynamics which originates in only a small part of a system, before possibly spreading throughout the entire system, such as a disease that is introduced
by one individual possibly becoming a pandemic, or the extinction of a species in an ecosystem possibly causing the loss of entire classes of species. I describe dynamics of complex systems through eigenvalue methods. In these methods, localized dynamics is associated with localized eigenvectors of the matrices describing the system under consideration. Pinpointing such localized eigenvectors in a system, one can trace the occurrence of the corresponding localized dynamics to a manageable small-scale problem. Thus, understanding localized dynamics allows to study and counteract the occurrence of possible dramatic failures, even in the ever-growing systems on which humans depend.


2 Concepts and Tools

In this chapter, I review the concepts and notions that I use throughout this thesis. Roughly, these concepts address three different tasks: First, methods to model networked systems, such as biological food webs, social interaction networks, or technological power grids (see Refs. 1 and 45 for detailed introductions). Second, methods to describe the network properties of these models, which include methods from graph theory, group theory, random matrix theory, and linear algebra (see Refs. 1, 46 and 47). Finally, methods from dynamical systems theory to analyze the dynamics (see Ref. 26 for an introduction).

I start with a short overview about networks (Sec. 2.1) and, especially, food webs (Sec. 2.2). Then, I present general notions to describe network dynamics (Sec. 2.3), steady states which play a central role in my thesis as operating modes of biological systems (Sec. 2.4), and bifurcations in which such states lose stability (Sec. 2.5). Third, I introduce the eigenvalue description which allows the decomposition of the dynamics near steady states into a superposition of independent dynamical modes (Sec. 2.6). Finally, I briefly review the generalized model for food webs [27] I use in examples throughout this thesis (Sec. 2.7).

2.1 Networks

Systems that contain many elements with a complicated structure of interactions or relations, such as individuals in a social group, computers on the internet, or species in an ecosystem, can be represented as networks (see Ref. 1 for a detailed introduction). By representing each element as a node and interactions or relations as edges between nodes, such a network representation achieves a reduction of the inherent complexity of the system’s constituents, but retains the complexity of the interaction pattern [1, 2, 48–50].

In recent years, networks have received attention in a diversity of research areas [1, 4]. Examples include technological networks, such as power grids, computer networks or traffic movements [30, 51], biological networks, such as food webs in ecosystems, metabolic pathways, and interactions underlying swarming behavior [45, 52, 53], social networks to model opinion spreading or dis-
ease transmission [33, 54, 55], and networks to model the interdependencies between businesses in economy [56], and networks to model climate dynamics [57].

The pattern of interactions or relations in a network is called its topology. If these interactions are reciprocal, one calls the network undirected, otherwise the network is directed. For instance, imagine a network in which nodes represent scientists and the edges mark co-authorship on papers. Then this network is undirected because the co-authorship is reciprocal. But, if the edges mark email traffic between these scientists on a specific day, then this network is directed because sending an email to another author does not mean receiving one from in return. Further, the topology in a network can be characterized only by the presence or absence of edges (unweighted network), or it can also contain information about the properties or weight of each edge (weighted network). For instance, in the scientist network, if the relation of emails sent from one author to another is characterized by a binary “yes” or “no”, then the network is unweighted. But, if one retains information about the number of emails sent, then this network is weighted (for a more detailed introduction, see Ref. 1).

The properties of a network can be characterized at different levels of detail. On a system level, the fundamental properties are the size and the connectance. The size is simply the number of nodes in a network and the connectance is the density of edges, given by the number of edges relative to the total number of possible edges in the network. For instance, the food web network in Fig. 2.1 (see the following section for its interpretation) has size $N = 12$ and connectance $C = 21/N^2 \approx 14.6\%$ because 21 of $N^2$ possible directed edges are present.

On an intermediate level, a network topology can be characterized by its tendency to form closely connected communities or clusters of nodes [12, 58–61], while more detail is contained in the distributions of centrality measures, which result from assigning a value to each node based on its position in the network [62–64]. One important example for such a centrality measure is the number of incoming or outgoing edges of a node, called the node’s degree in undirected networks or out-degree and in-degree in directed networks. See Ref. 1 or Ref. 7 for a detailed overview of these intermediate characterizations.

In this thesis, I describe networks on a level that retains all the details of the network topology. Such a description is provided by the network’s adjacency matrix, a square matrix with the dimension of the network size [1, 50]. For each edge of the network, say from node $i$ to node $j$, it contains a nonzero entry $A_{ji}$, while all other entries (without an edge) are zero. If the network is unweighted, these entries are either 1 or 0, but in weighted networks, the entry $A_{ji}$ represents the weight of the corresponding edge. Further, if the network is undirected, then the adjacency is symmetric ($A_{ji} = A_{ij}$), because the edge
from node $i$ to node $j$ implies a reciprocal edge of same weight. In the following, I consider networks to be weighted and directed, such that entries $A_{ij}$ are real numbers, and such that the adjacency matrix is generally not symmetric.

Closely related to the adjacency matrix is the Laplacian matrix $[65]$. This matrix is defined as $L = D - A$, where $A$ is the adjacency matrix, and where $D$ is a diagonal matrix containing the sums over each row of $A$ on its diagonal, i.e. $D_{ii} = \sum_j A_{ij}$. In this simple form, the Laplacian matrix is the network analog to the classical Laplace operator $\Delta$ in physics. It appears, for instance, in the description of diffusion processes, such as electrical flow in a network of wires, liquid flow through a network of pipes, or random walks $[46, 66]$.

Below, I relate the dynamics of a network system to the eigenvalues and eigenvectors of the adjacency matrix, called the graph spectrum, or to those of the Laplacian matrix (see Ref. 67 for details and further applications) $[1–4]$. It is important to note that the type of edges (directed or undirected) already determines the basic properties of the graph spectrum. For example, for an undirected network, the adjacency matrix is symmetric and the eigenvalues and eigenvectors are all real. For a directed network, the adjacency is not symmetric and the eigenvalues can also form pairs of complex conjugate eigenvalues.

### 2.2 Food webs

A paradigmatic class of examples considered in network science are ecosystem networks $[3, 44, 68–72]$. These contain hundreds of species, ranging from small bacteria, via plants, to large mammals which interact in a variety of ways. Most notably, different individuals feed or prey on each other, compete for prey or resources, or evolve and shape their environment. The complexity of ecosystems becomes apparent already in the diversity of the coexisting species, acting on time scales from hours to decades. However, what makes ecosystems of interest in network science, is the large and complex topology of the species’ interactions.

A food web is the network of who-eats-whom between the species of an ecosystem. Instead of accounting for each individual in a system, the food web of an ecosystem describes the interactions between species or groups of similar species. In other words, a food web is a coarse-grained description of the ecosystem $[21, 59, 72–76]$. Further, the food web representation of an ecosystem contains only the predator-prey interactions explicitly (as opposed to indirect effects, such as competition). Despite the omission of an explicit description of indirect effects, a food web retains most of the complex interaction pattern. They are, therefore, generally considered the backbone of most natural ecosystems $[3, 45, 59, 71, 77–86]$. 
2 Concepts and Tools

1 Nanophytoplankton
2 Filamentous green algae
3 Zooplankton
4 Insects
5 Silverside
6 Tetras
7 Sailfin Molly & Mosquito Fish
8 Blackbelt Cichlid
9 Bigmouth Sleeper
10 Tarpon
11 Black Tern
12 Herons and Kingfishers

Fig. 2.1: Food Web description of the Gatun Lake ecosystem [10]. Each circle represents one species or aggregated taxa of similar species. Arrows (edges) indicate the feeding interactions between the species; they originate in the node of a prey species and end in the node of the predator. Edges from a node to itself mark cannibalistic interactions, commonly encountered for fish feeding on younger individuals of their own species. Numbered species and black (solid) edges indicate the established ecosystem. The width of arrows indicates the importance of different prey species to a predator’s diet (see App. A for details).

One example for a food web is shown in Fig. 2.1. Different species or groups of similar species in Gatun Lake in Panama are shown as nodes [10]. Feeding relations are represented as arrows (directed edges) from each prey species’ node to the predator’s node, the width of each arrow indicates the estimated relative biomass flows.

The aggregation performed in the coarse-graining leading to a food web representation depends on the intended application. For instance, in the description of Gatun Lake, most of the fish species are retained explicitly because they are studied in detail in Ref. 10, while whole layers of the remaining ecosystem are aggregated, such as phytoplankton and zooplankton (Fig. 2.1). On an intermediate level, some biological information can be preserved by aggregating functional groups or guilds of species which are similar in morphology, life strategy, or niche [61, 70, 87, 88]. Examples are the four groups of phytoplankton distinguished in the ERSEM North Sea model (three of them defined by body size and one comprising all species with a silicate shell) [89], and the fish and bird species aggregated for the Gatun Lake food web (e.g. node 7 in Fig. 2.1 representing Mosquito Fish and Sailfin Molly) [10].
Based on the position of species in a food web, one can distinguish specific roles play in a food web [45, 86]. For instance, species that do not have an explicit prey (in the food web), such as algae or phytoplankton in Fig. 2.1, are called primary producers. Further, species that have no direct predator, such as the birds and tarpon in Fig. 2.1, are called top predators.

In the last decades, databases with hundreds of real-world food web topologies have been compiled [90]. However, a meaningful numerical analysis requires large numbers of realistic food webs with specific properties, such as a given size and connectance. For this, one must therefore rely on food web models (see Ref. 45 for an overview about food web models and their comparison to real-world data).

In this thesis, I establish the model food web topologies required for numerical analysis using the niche model for food webs. This model, which was presented in Ref. 91, assumes that each species in an ecosystem occupies a different ecological niche, such as a specific specialization on. Explicitly, the niche model assigns to each species a niche value, which is typically interpreted as a measure for the species’ fitness or body size [92]. Further, it assumes that different niches are opened for species through their specialization on morphologically different prey, such as large predatory fish which may specialized on prey with different body masses about one to three orders of magnitude smaller [93]. This specialization is captured by assigning a feeding range to each species, which is a range of niche values on which the species preys (see Ref. 91 for details). Ultimately, the topology of a niche model food web then results from drawing links from each species towards species to whose feeding range they belong.

The niche model reproduces much of the topological properties of real-world food webs [91]. While other models have been presented recently which take into account spatial effects, or probabilistic feeding relations [18, 94, 95], I use the niche model because of its simplicity and performance.

### 2.3 Dynamics on networks

The topology of interactions gives some indications about the time evolution, or dynamics, of a system, prescribing, for instance, which species in a food web become extinct following the death of all of their prey [71, 96]. More generally, one distinguishes two kinds of networks, depending on which type of dynamics they show: dynamics of the network, or dynamics on the network [1]. The dynamics of a network describes changes of a network topology over time, such as friendship relations between individuals in a social network that may be created or broken. These networks are called dynamic networks. The dynam-
ics on a network describes the evolution of the different elements of a network system in which the underlying interaction topology does not change. In this thesis, I focus on this second type of dynamics, which is also called dynamics on a static network (see Ref. 1 for a more detailed introduction). To study the dynamics on a static network, one studies dynamics in terms of the time-evolution of a system's state [26]. In static networks, this state is given by the states of all individual nodes in a network, which are captured by one or multiple variables describing the momentary properties of each node [1]. The dynamics on the network thus describes the time-evolution of its node’s states.

Food web dynamics is an example for dynamics on a static network. Because the feeding topology is due to morphological or spatial constraints, it remains the same over long periods of time. The dynamics manifest itself through the time-development of the abundances of the different species, measured, for instance, as the absolute number of individuals of each species in the ecosystem, or as a density indicating the amount of organically bound carbon attributed to each species relative to the total amount of carbon [27, 71]. Typically, the state of the food web contains one variable \( X_i \) for each node which indicates the abundance of the corresponding species [27, 71, 96]. Studying dynamics in food webs therefore means to investigate the time-evolution of these abundances, and of the resulting variations of the interaction strengths [45].

The topology of a static network already gives some indication about the dynamics of the system, but is generally not sufficient for a reliable description [13, 62, 82, 97, 98]. Instead, studying dynamics on a network requires more precise information on how the state of one node in a network affects the states of adjacent nodes. This information is provided by a model of the effects of the node’s interactions.

Throughout this thesis, I consider models of the form \( \frac{dX}{dt} = F(X) \), where the (vector-valued) function \( F \) describes the dependencies of all node’s states (the entries of \( X \)) on the states of the others. For instance, below, I consider a generalized model for food webs that describes the dynamics of each species’ abundance by an equation of the form [27]

\[
\frac{dX_i}{dt} = G_i(X) + S_i(X_i) - L_i(X) - M_i(X_i),
\]

(2.1)

where \( G_i, L_i, M_i, \) and \( S_i \) describe the biomass gain \( (G_i) \) and loss \( (L_i) \) of this species due to predator-prey interactions, the loss \( (M_i) \) due to mortality, and the gain \( (S_i) \) by reproduction. Note that, in this model, the topology of the food web is implicitly included in the dependencies of the predation gain \( G_i(X) \) and loss \( L_i(X) \) on the abundances of prey or predators of species \( i \).
2.4 Steady state operating modes

Nonlinear interactions, as modeled by a function $F(X)$, can lead to extremely complicated dynamics [26]. Even in systems described by only a few variables, one can simultaneously observe trajectories $X(t)$ with different character, depending on their initial state $X(0)$, such as steady states (states where $X(t) = X^*$ with each entry of $X^*$ being constant over time), periodic orbits ($X(t) = X(t - T)$ for some period $T$), and chaotic trajectories [99]. Furthermore, even small changes to the model function $F(X)$ can dramatically change the character of a system’s dynamics (see Ref. 26 or Ref. 100 for detailed introductions of trajectories and their characterization).

In summary, I study the dynamics on networks through the time-evolution of the states of all of their nodes. Explicitly, I consider dependencies of the form $\frac{dX}{dt} = F(X)$, where $X$ is a vector containing the state variables, and where $F$ is a function that models the interactions between the different nodes.

2.4 Steady state operating modes

In the following, I consider steady states ($X(t) = X^*$) of network systems and study if the dynamics allows the system to remain in such a state over long times. Formally, a steady state for a model of the form $\frac{dX}{dt} = F(X)$ has to satisfy $F_i(X^*) = 0$ for each variable $i$. It is called stable, if the dynamics restores this state after a small deviation by driving $X(t)$ back towards $X^*$, such that the state can persist even under small perturbations. It is called unstable, if the dynamics drives away from this state, such that the state is left after a small perturbation (see Ref. 26 for details).

To facilitate the analysis of the stability of a given steady state, one can use the variables $Y_i = X_i - X^*_i$, which indicate the distance of each variable from its steady state value. To study stability of a steady state against small perturbations, one can then study a linearization of $F_i(Y)$ around $Y_i = 0$. In this linearization, a stable steady state is characterized by dynamics where $Y$ tends towards zero after any small perturbation [26].

As an example, consider the (toy) model for the abundance $X$ of a single species $\frac{dX}{dt} = rX - X^2$, where $r$ is a constant parameter indicating the reproduction rate of the species, and where the quadratic term characterizes a mortality that grows strongly with $X$ (e.g. due to a limited availability of nutrition). The steady states of this system are $X^* = 0$ and $X^* = r$.

Now, I linearize the system for each of these states in terms of the distances $Y = X - X^*$. First, for $X^* = 0$, the linearization results in $\frac{dY}{dt} = rY$, which is solved by $Y(t) = Y(0)e^{rt}$. Thus, for $r > 0$, the distance to the steady state increases with time following a small deviation $Y(0) \neq 0$ from the steady state; $X^* = 0$ is unstable, and the system tends to leave this state. Second, for
the steady state $X^* = r$, the linearization results in $dY/dt = -rY$, which is solved by $Y(t) = Y(0)e^{-rt}$. Thus, $Y(t)$ tends towards zero after any small perturbation, and the state $X^* = r$ is stable against small perturbations. In other words, the species can exist with constant abundance $X^* = r$ in the model system over long times.

Similarly to the one-species example, one can also study the stability of steady states in systems containing multiple variables through a linearization. More explicitly, consider a system $dX/dt = F(X)$, which admits a steady state $X^*$. To study the dynamics near this state, I linearize in terms of the distances of each variable from the steady state $Y_i = X_i - X_i^*$. This leads to

$$\frac{dY}{dt} = JY,$$  \hspace{1cm} (2.2)

where $J$ is the Jacobian matrix of the system near the steady state [26, 100]. This Jacobian matrix contains the entries $J_{ij} = \partial F_i/\partial X_j |^*$, where $|^*$ indicates that the derivatives are taken at the steady state $X^*$.

The dynamics of $Y(t)$ solving (2.2) can be found explicitly in terms of the eigenvalues and eigenvectors of the Jacobian matrix. More precisely, the dynamics results from the superposition of independent elementary solutions, each associated with one eigenvalue and its corresponding eigenvector$^1$. Explicitly, the elementary solutions are $e^{\lambda_k t} v^{(k)} C^{(k)}$, where $\lambda_k$ denotes one (possibly complex) eigenvalue of the Jacobian matrix, $v^{(k)}$ denotes the corresponding (right) eigenvector, and $C^{(k)}$ denotes a constant factor depending on the initial state $Y(0)$.

Because the entries of the Jacobian matrix are real, the Jacobian eigenvalues are either real and belong to eigenvectors with real entries, or they form pairs of complex conjugate eigenvalues and belong to pairs of complex conjugate eigenvectors [101, 102]. For a real eigenvalue, the dynamics characterized by an elementary solutions of (2.2) is given by $e^{\lambda_k t} v^{(k)}$, while for a pair of complex eigenvalues, $\lambda_k \pm i\tilde{\lambda}_k$, a pair of complex conjugate elementary solutions exists, which sum to $e^{\lambda_k t}(\cos(\tilde{\lambda}_k t)\Re\{v^{(k)}\} + \sin(\tilde{\lambda}_k t)\Im\{v^{(k)}\})$, where $\Re$ denotes the real part and $\Im$ the imaginary part of the entries of the eigenvector. Thus, for both types of solutions, the real part of an eigenvalue $\lambda_k$ indicates if the

$^1$For completeness, I remark, that the decomposition of (2.2) is based on the assumption that the Jacobian matrix admits a full basis of independent eigenvectors. This condition is met, if the Jacobian matrix contains only distinct eigenvalues [101, 102], but can otherwise be resolved through the use of generalized eigenvectors defined by the equation $(J - \lambda)^r v = 0$, for $r = 1, 2, \ldots$. For instance for $r = 2$, the solutions are of the form $Y(t) = (\Sigma t + v)e^{\lambda t}v$, where $\Sigma$ is a linear combination of ordinary eigenvectors corresponding to $\lambda$. See Ref. 102 for a detailed introduction.
elementary solution tends towards or away from the steady state ($\lambda_k < 0$ or $\lambda_k > 0$), while for complex pair, this dynamics is superimposed with a rotation around the steady state with a frequency $\tilde{\lambda}_k/2\pi$ (see Ref. 26 for details).

The full solution $Y(t)$ of (2.2) tends to zero, if all its elementary solutions do. Thus, if the real parts $\lambda_k$ of all the eigenvalues of the Jacobian matrix are negative for a given steady state, then this steady state is stable.

In the following, I call stable steady states the operating modes of the complex systems I consider. This is motivated by the assumption that these states describe the typical functioning of these system. For instance, for food webs, interpreting stable steady states as their operating modes follows the standard literature [10, 84, 103–106]. Motivating this assumption are numerical and real-world observations, that suggest that stable steady states allow all species of an ecosystem to coexist, while oscillations or other dynamics is often temporary or precedes extinctions [27, 107, 108]. A similar reasoning also applies to other biological systems, such as networks describing metabolic pathways governing the functioning of cells [52, 109].

Apart from allowing the long-term persistence of biological systems, stable steady states have a distinct interpretation when studying infection spreading, such as the spreading of a computer virus on the internet, species migration between habitats, spreading of a rumor through a social network, or the spreading of forest fires between scattered patches of woodland [9, 24, 33, 54, 110]. In such applications, they describe the states in which the spreading phenomenon is stationary, i.e. situations in which no spreading takes place. Most prominently, when the infection-free state of a system is a stable steady state, then an accidental appearance of the infection typically disappears in the system. Thus, this state has a direct interpretation as a functioning operating mode of a system in which the infection hampers functioning, such as a virus in a computer network [24, 33]. Similar states describe situations where species cannot permanently invade habitats where they are not native [111], of networks that are fragmented into homogeneous clusters of nodes with different opinions [55], or of forest fires that do not spread to patches far from their outbreak [9].

In summary, I study in the following the stable steady states of complex systems. This is motivated by the observation that such states allow the persistence of a presumably functioning configuration of a complex system over long times, whereas oscillations or chaotic dynamics might lead to failures and therefore the destruction of the system.
2.5 Bifurcations affecting operating modes

When studying steady states as the candidates for operating modes of a given system, one searches for the phenomena, that may cause the instability of the steady states underlying these operating modes. In terms of parameters characterizing the network interactions functions $F(X)$, one searches for those combinations of parameters where a given steady state becomes unstable or disappears. Such situations where dynamics changes qualitatively under a small change of parameters, are called bifurcations (see Ref. 26 for a rigorous introduction). Thus, I search for bifurcations that separate parameter combinations for which a steady state is an operating mode, from those where it is not.

Normally, a full analysis of all bifurcations in complex nonlinear systems is difficult because many different types of bifurcations, and many different types of dynamics can occur [26, 100]. However, when studying dynamics near operating modes, only those specific types of bifurcations are relevant which separate stable from unstable steady states. Because the trajectory of a steady state is only a point in the space of possible states, these bifurcations generally depend only on a small neighborhood of the steady state, thereby allowing for the linearized analysis in (2.2). More formally, concentrating on steady states as operating modes allows to omit global bifurcations for which a local description does not suffice, but to concentrate on local bifurcations (for a more detailed review of local and global bifurcations, see Refs. 26 and 100).

Steady states are unstable, if one or more of the eigenvalues of the Jacobian matrix in (2.2) have a positive real part [26]. Thus, bifurcations in which operating modes of a system lose their stability take place when one or multiple eigenvalues of the Jacobian matrix cross the imaginary axis under a small change of the system parameters. More precisely, they take place, when the largest eigenvalue of the Jacobian matrix crosses the imaginary axis.

The type and number of eigenvalues crossing the imaginary axis in a bifurcation provides insight into the physical nature of the associated instability appearing in the bifurcation [26, 112]. For instance, when a stable operating mode undergoes a transcritical bifurcation (crossing of a single real eigenvalue through the imaginary axis), then the instability manifests initially by a dynamics in the direction of the corresponding eigenvector driving away from the steady state. However, when it undergoes a Hopf bifurcation (crossing of a pair of complex conjugate eigenvalues through the imaginary axis), then the instability manifest as oscillations with increasing amplitude and with frequency $\tilde{\lambda}_k/2\pi$, either towards an emerging limit cycle or far away from the steady state (for a more detailed analysis and more examples for local bifurca-
2.6 Dynamical modes

The superposition of the different elementary solutions of (2.2) is central to the study of the instabilities causing the break-down of an operating mode. I recall that each of these elementary solutions describes an independent part of the (linearized) dynamics of the system near the steady state. Intuitively, they can therefore be interpreted as specific independent instabilities that may cause the break-down of the operating mode if the corresponding eigenvalues is positive. In other words, the elementary solutions decouple the dynamics into independent dynamical modes, that are either stable, or cause an instability.

The decoupling of dynamical modes is based on the eigenvalues and eigenvectors of the Jacobian matrix $J$. More explicitly, each eigenvalue $\lambda_k$ of $J$ corresponds to one right eigenvector $v^{(k)}$, defined by $Jv^{(k)} = \lambda_k v^{(k)}$, and to one left (row) eigenvector $w^{(k)}$, defined by $w^{(k)} J = \lambda_k w^{(k)} [101,102]$. These are normalized, such that $v^{(k)} \cdot v^{(k)} = 1 = w^{(k)} \cdot w^{(k)}$, where $\cdot$ denotes the complex scalar product ($w^{(k)} \cdot v^{(m)} = \sum_i \overline{w}_i^{(k)} v_i^{(m)}$; $\overline{w}$ is the complex conjugate of $w$). Two right or two left eigenvectors are not necessarily orthogonal, but right and left eigenvectors are $(w^{(k)} \cdot v^{(m)} = \delta_{km} C^2_k$, where $C^2_k = |w^{(k)} \cdot v^{(k)}|$ is a normalization constant, and $\delta_{km}$ denotes the Kronecker delta, i.e. $\delta_{km} = 0$ if $k \neq m$ and 1 if $k = m$).

The orthogonality relations of the Jacobian matrix eigenvectors allow to rewrite the linearized dynamics near a steady state in (2.2) as $[101,102]$

$$\frac{dY}{dt} = JY = \sum_k \frac{\lambda_k}{C^2_k} v^{(k)} \left( w^{(k)} \cdot Y \right). \quad (2.3)$$

Thus, the effect of the Jacobian matrix becomes the sum over the contributions from different dynamical modes, each associated to one eigenvalue $\lambda_k$, and of the corresponding right and left eigenvectors.
In summary, the dynamics described by the Jacobian matrix can be decomposed into independent dynamical modes. Each of the dynamical modes is characterized by one eigenvalue and the left and right corresponding eigenvectors. Superimposing all dynamical modes results in the full (linearized) dynamics.

### 2.7 Generalized models for food webs

In the following, I study the dynamics of food webs based on a generalized model. Here, I therefore give a brief overview over the underlying method and the food web model. Generalized models allow to study the stability of steady states of systems of ordinary differential equations in which not all processes are restricted to specific functional forms [27]. Recently, generalized models were applied successfully to study problems in various contexts, from food web stability, over the robustness of metabolic cycles to the remodeling of bones [92, 112–119]. In this thesis, they play a central role because they allow to study the stability of operating modes in a system with a known network structure, but in which the functional forms of some or all interactions are unknown.

The model I use was introduced in Ref. 27, and was previously applied to investigate stabilizing factors in model food webs [92]. Here, I only give a short introduction to the model; more details can be found in Ref. 27, and mathematical proofs are given in Ref. 120.

The food web model presented in Ref. 27 describes the dynamics of \(N\) populations \(X_1, \ldots, X_N\) by \(N\) differential equation of the form of (2.1). These equations explicitly contain the topology of the food web, but the species’ interactions are unknown functions. Explicitly these unknown functions describe the species’ predation gain \((G_i(X))\), predation loss \((L_i(X))\), natural mortality \((M_i(X))\), and primary production \((S_i(X))\).

Normally, unspecified (and in practical applications often unknown) functions hinder finding the steady states of a model and therefore prohibit the application of local stability analysis. Generalized modeling circumvents this problem through the assumption that an operating mode must exists to allow coexistence of all \(N\) species. Thus, one assumes that an (unknown) steady state \(X^* = (X^*_1, \ldots, X^*_N)^T\) exists, in which each of the species’ densities is positive (see Ref. 120 for a discussion of the special case of extinctions, \(X^*_i = 0\)). Through a change to normalized variables \(x_i = X/X_i^*\), one can then formally compute the Jacobian matrix near the steady state \(x_i = 1\) without requiring explicit information about the absolute abundances of species in the steady state (see Ref. 27).
2.7 Generalized models for food webs

<table>
<thead>
<tr>
<th>Name Interpretation</th>
<th>Value(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Turnover rates</strong> - each species’ biomass turnover in the steady state.</td>
<td></td>
</tr>
<tr>
<td>$\alpha_i$ Rate of biomass turnover of $i$</td>
<td>$\alpha_i = \tilde{\alpha}^n_i$</td>
</tr>
<tr>
<td><strong>Scale parameters</strong> - defining the relative composition of biomass flows.</td>
<td></td>
</tr>
<tr>
<td>$\beta_{i,j}$ Contribution of predation by $i$ to loss rate of $j$</td>
<td>$[0,1]$</td>
</tr>
<tr>
<td>$\chi_{i,j}$ Contribution of $i$ to the prey of $j$</td>
<td>$[0,1]$</td>
</tr>
<tr>
<td>$\rho_i$ Fraction of growth in $i$ gained by predation</td>
<td>0 or 1</td>
</tr>
<tr>
<td>$\tilde{\rho}_i$ Fraction of growth in $i$ through primary production</td>
<td>$1 - \rho_i$</td>
</tr>
<tr>
<td>$\sigma_i$ Fraction of mortality of $i$ resulting from predation</td>
<td>$(0,1]$</td>
</tr>
<tr>
<td>$\tilde{\sigma}_i$ Fraction of mortality of $i$ not from predation</td>
<td>$1 - \sigma_i$</td>
</tr>
<tr>
<td><strong>Elasticities</strong> - sensitivities of interactions to state variables.</td>
<td></td>
</tr>
<tr>
<td>$\gamma_i$ Sensitivity of predation of $i$ to $i$’s prey density</td>
<td>$[0.5,1.5]$</td>
</tr>
<tr>
<td>$\mu_i$ Sensitivity of mortality of $i$ to $i$ (exponent of closure)</td>
<td>$[1,2]$</td>
</tr>
<tr>
<td>$\phi_i$ Sensitivity of primary production of $i$ to $i$</td>
<td>$[0,1]$</td>
</tr>
<tr>
<td>$\psi_i$ Sensitivity of predation of $i$ to the dens. of predators</td>
<td>$[0.5,1.5]$</td>
</tr>
</tbody>
</table>

Tab. 2.1: Generalized model parameters as defined in Ref. 27, where $i$ and $j$ refer to species. Following Ref. 92, the turnover rates $\alpha_i$ are allometrically scaled with each species’ niche value $n_i$, with the scaling constant $\tilde{\alpha} = 0.08$. For random sampling of model food webs, I draw the parameters from the indicated ranges of biologically sensible parameters [92, 122].

Explicitly, the diagonal and offdiagonal entries of the Jacobian matrix are [27]

\[
J_{ii} = \alpha_i \left( \tilde{\rho}_i \phi_i + \rho_i (\gamma_i \chi_{ii} + \psi_i) - \tilde{\sigma}_i \mu_i - \sigma_i \sum_k \beta_{ki} ((\gamma_k - 1)\chi_{ki} + 1) \right)
\]

\[
J_{ij} = \alpha_i \left( \rho_i \gamma_i \chi_{ij} - \sigma_i \beta_{ji} + \sigma_i \sum_k \beta_{ki} (\gamma_k - 1)\chi_{kj} \right),
\]

(2.4)

where the different parameters are given in Tab. 2.1 (see Ref. 27 for the derivation of (2.4), or Ref. 121 for an implementation of this formula). What makes this formulation interesting is that the Jacobian, and thus the stability of the normalized steady state, are expressed as function of simple parameters, called generalized parameters, which have an intuitive biological interpretation and can be measured directly in applications [27]. Explicitly, there are three types of parameters, which I now study in more detail.

First, each entry of the Jacobian matrix scales with a turnover rate ($\alpha_i$) which describes a species’ biomass intake or output relative to its own abundance. Such rates have the dimension of inverse time and can be directly in-
interpreted as characteristic per-capita growth or mortality rates of an individual of the species. For instance, for a top predator $i$, this rate is the biomass loss rate due to natural mortality in the steady state relative to its abundance, i.e. $M^*_i/X^*_i$. In applications, they can be estimated well from allometric scaling laws [72, 123, 124].

Second, scale parameters weight different contributions to the total biomass gain and loss of a given species. In (2.4), $\tilde{\rho}_i$ indicates the biomass gain of species $i$ from feeding on other species, as opposed to biomass gained by production ($\rho_i$). The dietary parameters $\chi_{ij}$ capture the fraction of biomass gained by a species $i$ through predation on different prey species $j$. For instance, for a predator ($P$) gaining biomass from feeding equally on two prey species ($A, B$), $\rho_P = 1$, $\tilde{\rho}_P = 0$, while $\chi_{PA} = \chi_{PB} = 1/2$. Further, $\sigma_i$ indicates the fraction of biomass loss of species $i$ through predation, as opposed to natural mortality ($\tilde{\sigma}_i$), and the dietary parameters $\beta_{ji}$ capture the fraction of this loss to different predators $j$. For instance, if the biomass loss of species $A$ is caused by 30% through predation by $P$, by 50% through predation by another predator $S$, and by 20% through natural mortality, then $\sigma_A = 0.8$, $\tilde{\sigma}_A = 0.2$, while $\beta_{PA} = 0.3$, and $\beta_{SA} = 0.5$. The scale parameters are closely linked to the food web topology, and they can often be estimated from allometric scaling laws for predator-prey body mass ratios [97, 125].

The dietary contributions in the sums over $k$ in (2.4) identify the terms in these sums as the result of indirect apparent competition between species with a common predator $k$. These effects describe that when species have a common predator $k$, then an abundance increase in one prey species alleviates the predations pressure on all prey species of this predator.

The third type of parameters, called elasticities, are logarithmic derivatives of the functions $G_i(X)$, $L_i(X)$, $M_i(X)$, $S_i(X)$ by the state variables in the steady state, such as $\partial \log G_i/\partial \log X_i|*$. They have been proposed in economic theory, and are used widely in metabolic control theory, and ecology [52, 114, 118, 126].

The elasticities can be interpreted straightforwardly as a measure for a function’s nonlinearity at the steady state. For example, the elasticity for a power-law, $f(x) = Ax^p$, is $\partial \log f/\partial \log x = p$, independently of $A$ or $x$. Thus, any linear function has an elasticity of $p = 1$ regardless of the slope. Also for functions that are not power-laws, the elasticity still provides an intuitive measure of the sensitivity in the steady state. For instance, for food webs, one specific model function describing predation as a function of prey abundance is $f(x) = x/(x + a)$, called the Holling type-II functional response, where $a$ is a parameter corresponding to the time a predator spends inactive after killing a prey [127]. For this response, the logarithmic derivative approaches 1 for small prey abundances $x$ (predation is proportional to $x$ because the predator consumes all prey individuals it encounters), while the derivative approaches 0 for large $x$. [53x574]
2.7 Generalized models for food webs

for large $x$ (prey is so abundant that the inactive time $a$ limits the amount of prey the predator kills). In applications, the elasticities can thus be estimated well from observational data of a system in a steady state, or from biological intuition (see App. A for examples) [126].

In summary, for a given food web topology, the generalized model presented in Ref. 27 allows to write the Jacobian matrix describing the (linearized) dynamics near any steady state in terms of biologically meaningful parameters. If the Jacobian matrix has only negative eigenvalues for a given combination of these parameters, then it describes a stable steady state in which species coexistence is possible. In other words, these parameters then characterize a possible operating mode of the food web.

Through the generalized model, it is straightforward to analyze the dynamics of different specific models on a given food web. For this purpose, one fixes those generalized parameters to specific values which are known, e.g. because they were measured in a system, and samples the remaining ones from biologically sensible ranges to find those combinations that lead to a stable operating mode [27].

In conventional models, it is often prohibitively complex to numerically compute a steady state explicitly. However, while in conventional models, one must recompute the steady states and the Jacobian matrix when adjusting the modeling functions, generalized models have the advantage that they avoid this computation entirely, and that testing new modeling functions is possible by simple variation of the generalized parameters. Therefore, one achieves a numerical efficiency even for large food webs, that enables a detailed and statistically sound exploration [92].

For applications, generalized models offer further advantages. In contrast to half-maximum concentrations and maximal growth rates used in conventional models, the generalized parameters are defined in the state observed in nature and in terms of parameters with sensible biological interpretation [27, 45]. Therefore they can be measured directly (as opposed to a fitting procedure) and their estimation from noisy data converges much faster than in conventional models [119, 126].

In this thesis, I apply the generalized model for food webs to investigate the dynamics of both a real-world food web, and to numerically investigate the dynamics of model food webs. While I base the choice of generalized parameters values for real-world examples on data or biological intuition (see App. A), I sample these parameters for numerical investigations following the approach presented in Ref. 92, for which I provide a numerical implementation [121]. More precisely, for a given food web topology, typically established using the niche model discussed above [91], I first establish the generalized model Jacobian matrix according to (2.4). Then, I sample the generalized parameters
independently and uniformly from the ranges given in Tab. 2.1. If a combination of parameters leads to a Jacobian matrix with only negative eigenvalues, then the chosen combination of parameters describes a stable operating mode. Thus, one can investigate dynamics close to this mode using the Jacobian matrix [92].
3 Perturbation Impact

Predicting the results of perturbations on a complex network is challenging because perturbation of one element can percolate across the network of interactions. Even perturbations acting only on a small part of a system may propagate through the entire network and lead to systemic changes. This is particularly evident in ecosystems, where the arrival of new species or habitat change, can lead to extinctions or systemic changes abundances [84, 103–106], such as the destruction of native fish populations following the introduction of a new species [10], or the loss of biodiversity with climate change [128].

Extensive research aims to deduce the impact of perturbations on a food web from its topology. For instance, one can establish notions of robustness, such as the likelihood of a species to become extinct following the extinction of another species (secondary extinctions) [63, 75, 92, 129–134]. In general, however, the topology alone is insufficient for reliable predictions of perturbation effects, and approaches taking into account biomass flows between the different species correlate much better with the results of numerical simulations than topological indices alone [13, 62, 82, 97, 98].

An accurate prediction of the impact of a perturbation requires precise information about underlying biomass flows and the control coefficients characterizing the nonlinearity of the underlying processes. Such parameters require extensive measurements, and errors in their estimation quickly reduce the accuracy of predictions [69, 82]. Predictions for systems of more than 25 species were found practically impossible with current methods, unless detailed information is available [135].

In this part of my thesis, I use a generalized model for food webs (see Sec. 2.7) to investigate the predictability of responses to perturbations. First, I develop a framework to estimate the effect of perturbations using these models (Sec. 3.1). After illustrating the process with two examples (Sec. 3.2), I reformulate the response of a food web to perturbation as a superposition of decoupled dynamical modes (Sec. 3.3). This allows to identify species in food webs that are particularly important when assessing perturbation responses (Sec. 3.4). Studying localization, I show that this importance is due to the localization of particularly excitable dynamical modes on these species (Sec. 3.5). Using numerical
3 Perturbation Impact

experiments, I illustrate a feasible iterative strategy for field studies that identifies such species (Sec. 3.6), and study statistical association of this importance of species with their properties or positions in the food web topology (Sec. 3.7). These experiments show that prioritizing the important species when gathering data highly improves the efficiency of impact assessments.

3.1 Impact of perturbations on food webs

In typical descriptions of food webs, the species in the food web are characterized by a set of state variables $X_1, \ldots, X_N$, denoting the abundances of the established species. For instance, in the generalized model presented in Sec. 2.7, the variables $X_1, \ldots, X_N$ are governed by the set of ordinary differential equations (2.1), which have the form $dX_i/dt = A_i(X_1, \ldots, X_N)$, where $A_i$ is a function representing the right-hand side of the differential equations. In particular, I consider the case where the system initially resides in the stable equilibrium $X^*$, in which each species exists at a constant positive abundance $X^*_i$.

The system is then subject to a perturbation characterized by another set of variables $Y_1, \ldots, Y_M$. Those can denote, for instance, the abundances of newly arriving species, or the severity of environmental changes. Following Ref. 135, I apply a perturbation that is characterized by small and constant values of $Y^*$. For instance, for a newly arriving species, the value $Y^*_j$ might indicate the low, but constant, abundance at which this species persists in the ecosystem due to initially positive growth or constant influx. This new species affects the behavior of the established species, such as feeding or breeding of the established species. Mathematically, these effects appear in the right-hand side $A_i(X, Y^*)$ of the differential equation.

The dependence of $A_i(X, Y^*)$ on $Y^*$ means that the steady state values $X^*_i$, which solve the equations, also depend on $Y^*$. In other words, the abundance of the established species changes in response to the abundance of the new species, i.e. $X^*_i = X^*_i(Y^*)$. In the absence of the new species, $X^*(Y^* = 0)$ is the solution of the original system.

The impact $I_{i,j}$ of a perturbation variable $Y^*_j$ on a resident species’ abundance $X^*_i$ is the change of $X^*_i$ with $Y^*_j$,

$$
I_{i,j} = \frac{\partial X^*_i}{\partial Y^*_j} \bigg|_0 ,
$$

where $|_0$ indicates that the derivative is evaluated in the limit of vanishing densities of the arriving species $Y^*_j \to 0$. 

24
3.1 Impact of perturbations on food webs

For two, or sometimes three, species, the impact can be computed explicitly by first defining the explicit model functions, then solving \(A(X^*, Y^*) = 0\) to find the stationary solution \(X_i^*(Y^*)\), and finally taking the derivative by \(Y_j^*\) as defined in (3.1). However, for more than three species, the analytical computation of the stationary solution becomes prohibitively difficult. Furthermore, the whole process has to be repeated each time the functional forms in the model are adjusted, for example, to reflect new data. For these reasons, a more general approach than the explicit computation of the stationary solution is required.

Computing the stationary solution can be avoided by recognizing that the stationary density of a resident species \(X_i^*\) is an implicit function defined as the solution of the stationarity condition \(A_i(X^*(Y^*), Y^*) = 0\). Using a corollary to the implicit function theorem in multiple dimensions \(^1\), one obtains the fundamental equation

\[
I = -J^{-1}K, \quad (3.2)
\]

where the superscript \(-1\) indicates the matrix inverse [136], and where \(J = \partial A/\partial X\) is the Jacobian matrix, defined as the derivatives of \(A_i\) with respect to the abundances of established species. The matrix \(K\) is defined by \(K_{ij} = \partial A_i/\partial Y_j\) and thus captures the direct impact of each perturbation variable \(Y_j\) on the functions describing the properties of the established species \(i\) at the steady state.

For example, the direct impact of a newly arriving species \(j\) on the properties of an established species \(i\) is captured by the entry \(K_{ij}\). Typical examples for such a direct impact are a reduction in \(i\)'s primary production rate, an increase in \(i\)'s mortality because of increased predation pressure, or a decrease in species \(i\)'s predation gain if species \(j\) competes for the same prey.

Beyond the example of perturbations caused by an arriving species, (3.2) applies to press perturbations on an established community in general (press perturbation are perturbations acting over longer periods of time, see Ref. 137 for details). For instance, to estimate the ultimate impact of a slight drought, the entry \(K_i\) of the perturbation vector indicates the direct impact of the drought on the behavior of species \(i\). Explicitly, \(K_i\) indicates the derivative of \(A_i\) by the severity of the drought \(Y^*\) which can be approximated by the finite value \(K_i = \Delta A_i(X^*)\), where \(\Delta A_i(X^*)\) denotes the expected absolute change of \(A_i\) due to the drought.

---

\(^1\) \[
\frac{dA(X^*(Y^*), Y^*)}{dY^*} = 0 = \frac{\partial A}{\partial Y^*} + \frac{\partial A}{\partial X} \frac{\partial X^*}{\partial Y^*}, \text{ see Ref. 136 for details.}
\]
3 Perturbation Impact

In summary, (3.2) establishes a relationship between the direct proximal impact of a press perturbation, and the indirect ultimate impact $I$ when the internal interactions in the unperturbed system $J$ are known. As a note of caution, I remark that (3.2) holds up to linear order. The impact-approximation therefore applies only to reasonably small perturbations.

The equations above refer to the steady state of the system, which seems to imply that information about this state is required. However, this is resolved by the use of a generalized model, which provides the Jacobian matrix $J$ in terms of species turnover rates and interaction elasticities, independently of an explicit steady state (Sec. 2.7). Therefore, using this Jacobian matrix in (3.2), one can study the effect of a perturbation independently of the specific steady state under consideration.

Since the linearization captured by the Jacobian matrix in a generalized model applies to normalized variables $x_i = X_i/X_i^*$, also the direct impact ($K$) and the ultimate impact ($I$) have to be formulated in terms of these variables. The impact $I$ therefore represent the change of the abundance of a species relative to its abundance in the unperturbed system. Further, the direct impact $K$ contains the derivative of products of the form $\alpha G(X)/G^*$, consisting of the affected species’ per-capita turnover rate $\alpha$, and a biomass flow $G(X)$ normalized by its steady state value $G^*$ (see Ref. 27 for details on these quantities). One can assume that the per-capita turnover rate remains unchanged through the perturbation (it is a species-inherent property), such that the direct impact is measured as the change in biomass flow relative to the unperturbed steady state value, but weighted by the turnover rate $\alpha$. Thus, the use of generalized models reduces the impact analysis to a study of relative changes in flows, which are often easier to interpret than their absolute counterparts [27].

Note that the approach to assess impact is closely related to the method presented in Ref. 135. The main novelty is to apply this approach to generalized models, which yields a series of advantages. Most prominently, the numerical efficiency of generalized models allows a more detailed, and statistically sound, numerical exploration, and the use of generalized parameters permits to study impact in terms of intuitive parameters that facilitate a comparison with observed biological data [27, 92].
3.2 Examples

For illustration, I consider two examples: a simple predator-prey system and the real-world food web of Gatun Lake. While the predator-prey system is analytically tractable, the second example requires numerical computations (e.g., with the simple program in Ref. 138).

The first example is a class of predator-prey models in which a predator of abundance \( X_1 \) consumes a producer of abundance \( X_2 \) (Fig. 3.1, see Ref. 139 for a detailed treatment and discussion of the stability of this system). The Jacobian matrix of this system near the steady state is

\[
J = \begin{pmatrix}
\alpha_1(\psi - \mu) & \alpha_1 \gamma \\
-\alpha_2 \sigma \psi & \alpha_2(\phi - \sigma \gamma - (1 - \sigma) \mu)
\end{pmatrix},
\]

where \( \alpha_i \) represents each species’ turnover rate, \( \sigma \) the relative loss of the producer due to predation (instead of natural mortality), \( \phi \) the elasticity of primary production with regard to the producer abundance, \( \gamma \) the elasticity of predation with primary producer abundance, \( \psi \) the elasticity of predation gain to predator abundance, and \( \mu \) the elasticity of natural mortality to a species’ own abundance (c.f. Tab. 2.1).

The inverse of the Jacobian, required for the impact computation, is

\[
J^{-1} = \frac{1}{\det J} \begin{pmatrix}
\alpha_2(\phi - \sigma \gamma - \tilde{\sigma} \mu) & -\alpha_1 \gamma \\
\alpha_2 \sigma \psi & \alpha_1(\psi - \mu)
\end{pmatrix},
\]

where \( \det J = \alpha_1 \alpha_2 [(\phi - \sigma \gamma - \tilde{\sigma} \mu)(\psi - \mu) + \sigma \gamma] \) denotes the determinant of \( J \), which is positive in a stable operating mode.

Now, consider the impact of the arrival of a competing predator (node “P” in Fig. 3.1) in the established producer-predator system. It can be assumed that this new predator has a direct negative effect on the primary producer, but no direct effect on the established predator, such that the perturbation matrix \( K \) contains the entries \( K_1 = 0 \) and \( K_2 < 0 \) (I drop the second index for simplicity). Then, (3.2) results in an impact on the predator, \( I_1 \), and an impact on the producer, \( I_2 \), given by

\[
I_1 = \frac{\alpha_1}{\det J} \gamma K_2 \quad \text{and} \quad I_2 = \frac{\alpha_1}{\det J}(\mu - \psi)K_2.
\]

The impact on the established predator is negative (Fig. 3.1) because the established predator competes with the arriving predator for the established prey species. Of particular interest is the case where the established predator suffers from linear loss (\( \mu = 1 \)) and has an effect on the producer that scales

\[
^2\text{det } J \text{ is the product of the two eigenvalues of } J, \text{ which are negative or complex conjugate if the system is in a stable steady state.}
\]
Fig. 3.1: Estimated response to the perturbation caused by the introduction of a new species into a simple predator-prey model (left), and into the Gatun Lake food web (right). Shown are species (circles) and predator-prey relationships (arrows, in the direction of the biomass flow). The new species and its feeding predation relations are marked in green. The blue-and-red color-code indicates the sign and strength of the response estimated for this perturbation. In the predator-prey system (left) the new consumer competes with the established consumer for the common prey, such that the response of the established consumer is negative (color-coded red). For the Gatun Lake food web example, the introduced peacock bass (node “P”) directly affects five of the established fish species (nodes 5–9). Refer to Ref. 10 for details on the Gatun Lake food web, Fig. 2.1 for a list of the species corresponding to each node, and to App. A for the parametrization and perturbation estimation.

linearly with the predator abundance ($\psi = 1$). In this limit, $\det J = \alpha_1 \alpha_2 \sigma \gamma$. The impact on the producer, $I_2$, is zero in this case. However, the impact on the established predator, $I_1 = 1/(\alpha_2 \sigma)$, is of same order as the perturbation. Intuitively speaking, the new predator takes over part of the predation, but because the established predator decreases, the total pressure on the prey remains unchanged. This is a manifestation of the well-known competitive-exclusion principle, which precludes the coexistence of two such predators in a food web, because the growth of one predator generally leads to the extinction of the other [16, 68].

The assessment of impact in larger food webs can be carried out analogously, but requires numerical computations in which the generalized parameters are set to specific values. I illustrate this assessment with the example of the Gatun Lake food web [10]. The Gatun Lake food web was invaded in 1967 by a new
3.2 Examples

species, the peacock bass, *Cichla ocellaris*, which caused the disappearance of more than half of the established fish species [10]. I assess the impact one would expect for this invasion, based on a particular parametrization of the generalized model, and compare it to the observed changes in species numbers in Gatun Lake [10].

To parametrize the Gatun Lake food web, I set the generalized parameters to biologically meaningful values: I set the turnover rates of different species $\alpha$ and the dietary preferences $\chi$ according to allometric scaling rules for turnover rates and typical predator-prey body size ratios with the species body mass. Further, I set the loss contributions of a species to different predators, $\beta \sigma$, such that biomass is conserved. Finally, I set elasticities based on qualitative arguments, such as assuming mortality to be linear with species’ abundance (see App. A for details on the parametrization process).

To assess the impact of the peacock bass requires information about its direct effect on the established species. Here, I assume that the main direct effect of the bass is due to its predation on the established species because it increases the predation pressure on these species. Further, I assume that the perturbation is characterized by the relative contributions of each species to the bass’ diet, established using the same allometric scaling laws as before (see App. A for a step-by-step breakdown of this perturbation assessment). The estimated impact of this perturbation is shown by the color-code in Fig. 3.1.

The impact assessment predicts that peacock bass have a strong and generally negative impact on the secondary consumers on which they feed (Fig. 3.1). Furthermore, they have a generally positive impact on the consumer’s prey, and a negative impact on other top predators. These predictions are consistent with basic ecological reasoning. The results are counter-intuitive for species 11 (black tern) because the species benefits from the strong decrease of its competing predators 9 (bignose sleeper) and 10 (tarpon), and for species 7 (sailfin molly, mosquito fish), which benefits from the reduction of 6 (tetras) with whom they are both in competition for prey and in apparent competition.

Real world observations of the Gatun Lake showed [10], that the introduction of the peacock bass strongly decreased the secondary consumers (5–7) and their predators (9–12), but increased the consumers prey (3–4) and, counter-intuitively, species 8 (blackbelt cichlid). Comparing the predictions of the model with the real data, I note that the model correctly predicted the change in the producers 1–4, the decrease in the consumers 5 and 6, and the decrease in predators 9, 10, and 12.

However, the model predictions do not agree with the observed decrease in species 11 (black tern), and 7 (sailfin molly, mosquito fish) and with the increase in species 8 (blackbelt cichlid). A likely reason for the discrepancy between predicted and observed results is the inaccuracy of the parametrization used here,
that is based on simple allometric scaling relationships instead of direct observations. Reviewing the results, one can guess that the incorrect prediction for an increase in species 11 is due to an overestimation of the sensitivity of its competing top predators (9–10) to the decrease in their common prey 5 (silverside). Further, the incorrect prediction for an increase in 7 and a decrease in 8 might result from incorrectly estimating the benefit to the prey of 12 (herons and kingfishers) due to the strong decrease in 6 (tetras). This exercise shows how food webs often follow the basic logic of direct and indirect effects, but that the details can be sensitive to incorrect parametrization.

3.3 Impact formulation with dynamical modes

While the results in the previous section allow to compute the impact of a perturbation, an understanding of the underlying principles is hampered by the formulation as a matrix product in (3.2). To allow for such an understanding, I now formulate (3.2) in terms of eigenvalues and eigenvectors of the Jacobian matrix to decouple the effects into those of different dynamical modes (c.f. Sec. 2.6).

To formulate (3.2) in terms of dynamical modes, I write $J^{-1}$ in terms of the eigenvalues $\lambda_k$ of $J$, and the corresponding (right and left) eigenvectors, $v^{(k)}$ and $w^{(k)}$. Because of $v^{(k)} = J^{-1}Jv^{(k)} = J^{-1}\lambda_k v^{(k)}$, each eigenvalue $\lambda_k$ of $J$ corresponds to an eigenvalue $1/\lambda_k$ of $J^{-1}$. Further, the eigenvalues $1/\lambda_k$ of $J^{-1}$ and $\lambda_k$ of $J$ correspond to the same eigenvectors. Writing (3.2) in terms of dynamical modes (Sec. 2.6) then results in

$$I = J^{-1}K = \sum_k \frac{v^{(k)}(w^{(k)} \cdot K)}{-\lambda_k C_k^2},$$

(3.4)

where $C_k^2 = |w^{(k)} \cdot v^{(k)}|$ is a normalization constant. In other words, the perturbation is given as a superposition of the responses of each dynamical mode $k$.

This superposition of the dynamical modes $k$ in (3.4) can best be understood in analogy with vibrations traveling through a drum. In this analogy, different dynamical modes correspond to different notes that simultaneously sound when striking a drum. Their superimposition is the perceived sound of the drum.

In (3.4), the right eigenvector $v^{(k)}$ indicates, which variables are affected by a given dynamical mode $k$. For the drum analogy, $v^{(k)}$ corresponds to the specific pattern of vibration appearing when playing a specific note. If a short (pulse)
perturbation hits a food web, then the right eigenvector governs how the system returns to the steady state (if the food web is stable) or points in the direction in which the system leaves the steady state (if it is unstable).

In (3.4), the product $w^{(k)} \cdot K$ of the perturbation and the left eigenvector indicates, if a perturbation excites the specific dynamical mode $k$. For the drum analog, it indicates the different volume of a given note depending on where the drum is struck. If a perturbation affecting only one species directly hits a food web ($K$ has only one nonzero entry), then the entry of the left eigenvector on the affected species governs the resulting excitation of the associated dynamical mode.

The excitability $1/\lambda_k$ indicates the overall strength of the response caused when exciting a dynamical mode. In the drum analog, it indicates how strongly the drum resonates at a given frequency relative to other frequencies. For instance, in a food web Jacobian matrix, a dynamical mode with a negative eigenvalue $\lambda_k$ resists a perturbation of this mode and quickly returns to the steady state. The closer the eigenvalue of a dynamical mode is to the imaginary axis, and thus to a bifurcation, the larger the response of the perturbation due to an excitation of this dynamical mode.

As an example, consider a system with the two following dynamical modes: First, $\lambda = -1$ with $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $w = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, and second, $\tilde{\lambda} = -1/2$ with $\tilde{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\tilde{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. A perturbation that affects only the first species, $K = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then only affects the first dynamical mode because $\tilde{w} \cdot K = 0$. The system’s response is therefore $I = v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; only the first species responds to the perturbation. However, a perturbation that affects only the second species, $K = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, excites both dynamical modes. The response is therefore their superposition $I = 1/2(2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$; the decrease of the first species is twice as strong as decrease of the second.

Separating dynamical modes corresponding to real eigenvalues of the Jacobian, and those that form complex conjugate pairs, one obtains two different types of dynamical modes. For a real eigenvalue $\lambda$, the contribution of this dynamical mode to the summed impact in (3.4) is a real number. A pair of dynamical modes corresponding to complex conjugate eigenvectors is slightly more subtle, because their contributions to (3.4) are possibly complex. Only when taking their sum, the complex contributions cancel. Such pairs of complex modes are closely linked and can therefore be best studied when interpreting them as the constituents of one single dynamical mode.

In summary, one can think of each dynamical mode as a possible response of the system to a specific perturbation. The right eigenvector of the dynamical mode denotes the direction of response (which species “feel the vibrations”), while the corresponding left eigenvector denotes the type of perturbation, that
can trigger a particular response (which specie needs to be perturbed to “play a given note”). The strength of this response is given by the inverse eigenvalue corresponding to the dynamical mode.

3.4 Influence and sensitivity of species

The evaluation of impact presented above allows to understand the effect of perturbation for which the direct effect on the ecosystem is known exactly. Prior information is therefore required about the way an invading species will interact with the established species. However, if one wants to assess which species are at risk due to invading species without this explicit knowledge on an invading species, there are two types of species that are of particular importance. Influential species that are likely to cause great changes to other species when perturbed, and sensitive species that are likely to respond strongly to perturbations of other species. To quantify these notions, I now develop a method to assign a sensitivity and influence value to every species in a food web.

For the sensitivity of a species, I estimate the expected strength of a response of this species, when every dynamical mode of a system has an equal probability to be excited. For a given mode $k$, the response is determined by the product of the excitability $-1/\lambda_k$ of the mode with the absolute value of the right eigenvector $|v_i^{(k)}|/C_k$ on the species $i$ ($C_k$ is the eigenvector normalization to make different modes comparable). The sensitivity can thus be measured as the sum over these effects from all dynamical modes. Explicitly, the sensitivity of node $i$ is then

$$S_{ei} = \log \left( \sum_k \frac{|v_i^{(k)}|}{-\lambda_k C_k} \right),$$

where the logarithm is used to bring the numerical values into a more manageable range.

For the influence of a species, I estimate the expected strength of the excitation of all dynamical modes caused by a perturbation of this species. For a given mode $k$ and a species $i$, this strength is given by the product of the excitability $-1/\lambda_k$, with the (absolute) entry of the left eigenvector $|w_i^{(k)}|/C_k$ on the species. The influence $In_i$ can thus be measured as the sum over this strength for all the dynamical modes. Explicitly, this leads to the influence of node $i$,

$$In_i = \log \left( \sum_k \frac{|w_i^{(k)}|}{-\lambda_k C_k} \right).$$
3.4 Influence and sensitivity of species

Fig. 3.2: The importance of each species in a food web compared to the response to perturbations of single species. Panels a–l) show the food web response to the arrival of specialist predators (P, yellow) affecting each a single prey species. Panels n) and m) show the sensitivity and influence assigned to each species. Comparing these panels with panels a–l) shows that influential species cause a strong response, when perturbed, while sensitive species respond strongly to perturbations of other species.

For example, for the Gatun Lake food web, the sensitivity and influence of the different species correspond well to the response of the system to (hypothetical) specialist predators (Fig. 3.2). If an influential species is perturbed (e.g. Fig. 3.2k), many species show a strong response but when a species with low influence is perturbed, only few species show a strong response (e.g. Fig. 3.2a). Further, sensitive species respond more strongly to perturbations elsewhere than species with low sensitivity.

To confirm the relationship between the actual impact of a perturbation, and the sensitivity and influence of species, I numerically study an ensemble of $10^6$ model food webs with 50 species and average connectance 0.04. The topology of these food webs is generated using the niche model (see Sec. 2.2) [91]. For the
Fig. 3.3: The average absolute impact on a species with given sensitivity if a species of given influence is perturbed. Generally, the more influential the species that is perturbed, and the more sensitive the species for which the response is tracked, the greater the observed response (impact). Parameters $N = 50, C = 0.04$.

Numerical computations, the generalized parameters of the species are drawn uniformly and independently from the ranges given in Tab. 2.1. The color-code in Fig. 3.3 indicates the average impact that a focal species of given sensitivity in these food webs experiences when a species of given influence is perturbed by a specialist predator. This reveals a strong correlation of the impact with both the sensitivity of the focal species and the influence of the perturbed species.

Further, I also study the typical distribution of influence and sensitivity values found in model food webs. For this purpose, I establish the distribution of influence and sensitivity values of species in each food web, relative to the maximal and minimal value of the respective quantity encountered in that web. Typically, in a given web, the number of species with a high sensitivity or influence (relative to the remaining species in that web) is low (Fig. 3.4).

Summarizing the above, one can use dynamical modes to identify which species are potentially sensitive to perturbations or influential when perturbed. In the food webs under consideration, typically a small number of species have unusually high influence or sensitivity. Therefore, when assessing the effects of a perturbation or invasion, one can focus the assessment process on perturbations that affect the influential species and on the responses of the sensitive ones.
3.5 Localized dynamical modes

The results presented so far show that generally only few species have a high influence or sensitivity, making them key players for an impact assessment. From (3.4), one can therefore conclude that the dynamical modes that cause the strongest responses, those modes with large eigenvalues $\lambda_k$, do not depend uniformly on the whole network, but are tied to these key species. In other words, one might expect, that these most excitable dynamical modes, are localized on the influential and sensitive species of a food web.

I recall that the left and right Jacobian eigenvectors corresponding to a dynamical mode indicate, how strongly the mode is excited by a perturbation of different species in a food web, and how strongly it affects different nodes when triggered. Thus, a dynamical mode is localized, if its associated eigenvectors have large amplitudes (absolute squared entries) on only few nodes. In other words, the localization of a dynamical mode is determined by the localization of its corresponding eigenvectors.
Fig. 3.5: The dynamical modes of the Gatun Lake Food Web. Each panel shows the dynamical mode corresponding to one real eigenvalue (c–e,g) or to a pair of complex conjugate eigenvalues (a,b,f,h) of the Jacobian matrix. The panels are sorted by decreasing real parts of the associated eigenvalues (the rounded values are given in each panel); panel a) depicts the most excitable dynamical mode and panel h) the least excitable one. The color-code represents the contributions of the left eigenvector of the mode to a node (top row, blue) and of the right eigenvector (bottom row, red). When the blue nodes are perturbed the mode is triggered and the red ones respond to this excitation.

For illustration, I consider the localization of the eigenvectors of the Jacobian Matrix of the Gatun Lake food web (see App. A for the underlying parametrization of the web). In this food web, the eigenvectors associated with the most excitable dynamical modes (Fig. 3.5a) are localized on nodes with high positions in the food chain, while the least excitable ones (Fig. 3.5g,h) localize on nodes with low positions in the food chain. Localization is most pronounced for the most excitable and least excitable modes (Fig. 3.5a,g,h). Also the slightly
3.6 Iterative parameter estimation

Knowledge of the Jacobian matrix of a food web allows to predict the impact of specific perturbations. The main challenge for impact assessment is thus to collect the data necessary for constructing this Jacobian. It is intuitive to assume that accurate predictions hinge on precise measurements of the parameters of the most influential and sensitive species. However, the approach to identify these species builds itself on an analysis of the Jacobian matrix, such that it is not possibly to determine a-priori which species are sensitive or influential.

To address this dilemma, I now propose an iterative strategy in which existing preliminary information is used to estimate the impact and sensitivity of species. This assessment is then used to obtain improved parameter estimates on seemingly important species. Once additional data on these species
becomes available, it can be used to further improve the estimates of the influence and sensitivity of species. Thus, a cycle is formed in which the necessary information for precise impact predictions is iteratively improved.

I now explore the quality of impact prediction in the presence of measurement uncertainty in a series of numerical experiments. In each experiment, the task is to predict the impact of a random perturbation to a food web. The food web is generated as above based on a generalized model describing an underlying niche model topology such that I obtain a set of generalized parameters (Tab. 2.1) describing the true Jacobian of the food web.

Imagine that the true food web is observed. Based on the true Jacobian, I therefore generate an estimated Jacobian with slightly different generalized parameter values to simulate measurement errors. More precisely, I draw each generalized parameter value for the estimated Jacobian matrix from a log-normal distribution centered on the corresponding parameter value used for the true Jacobian matrix; the log-normal distribution is chosen to allow large errors while keeping the sign of parameters consistent. Then, I compute the true impacts of random perturbations, $\mathbf{I}$, based on the true Jacobian, and the estimated impacts of these perturbations, $\tilde{\mathbf{I}}$, based on the estimated Jacobian. The quality $Q$ of the estimated impacts is then evaluated as the cosine of the angle between the true and the estimated impact vectors $Q = \cos(\mathbf{I} \angle \tilde{\mathbf{I}})$. For instance, if $\mathbf{I} = \tilde{\mathbf{I}}$, i.e. if the estimation is exactly correct, then $Q = \cos(0) = 1$. Further, if $\mathbf{I} = -\tilde{\mathbf{I}}$, i.e. if the estimation is exactly the opposite of the true impact, then $Q = \cos(\pi) = -1$.

Now, I study the quality of impact assessments in a numerical implementation of the iterative strategy to improve these assessments proposed above. Specifically, I consider numerical experiments in which the knowledge of the Jacobian is initially poor, such that the estimated generalized parameters are drawn with a standard deviation of 10% around their true values. Further, I assume that additional empirical work can be carried out on one specific species at a time that brings the error in all parameters of this species down to 2%. The aim is to carry out such precise measurements on different species in the order that leads to the most rapid increase in the quality of overall impact prediction.

For demonstration, I consider four different protocols to choose the next species for precise measurement: a) precise measurements are carried out in random order, b) species are measured in the order of decreasing influence, c) species are measured in the order of decreasing sensitivity, d) species are measured in the order of the decreasing sum of sensitivity and influence. The choice of species to measure next, is always based on the estimated Jacobian that is available at the time. Thus, only information is used that would also be available in the real world at the respective time.
3.6 Iterative parameter estimation

Fig. 3.6: Focusing on sensitive and important species can reduce the measurement effort when species are measured successively with higher precision. Starting on the left, only low-precision information is available for all food web species. Advancing to the right, the measurement error is reduced for one species at a time until all nodes have been measured with high precision. The different curves refer to different strategies for selecting which species to measure precisely. For the dashed line (empty symbols), the species are selected randomly. For the solid curves, I first evaluate the sensitivity or influence of each species based on current knowledge and then select the species with the highest value of either sensitivity or influence. For the dotted curves, I select the species with the highest product of influence and sensitivity \( (S_i I_i) \). Error-carrying parameters of each node: \( \alpha, \mu, \psi, \phi, \gamma, \sigma, \beta, \chi \). Other Parameters: Initial error of each parameter 10%, final error 2%, connectance \( C = 0.04 \). Higher connectance results in similar graphs, but the overall prediction quality decreases.

Estimating influence and sensitivity of the species prior to a refined measurement strongly increases the accuracy of predictions (Fig. 3.6). This is particularly pronounced if measurements focus on the species with the highest sum of sensitivity and influence. For instance, after measuring 20% of all species according to this protocol, I attain a quality of prediction comparable to measuring 60% – 80% of all species when species are chosen randomly. The estimation of influence and sensitivity to focus observational efforts thus significantly reduces the amount of empirical work that is needed to achieve a given prediction quality.
When studying localization, the refinement algorithm iteratively identifies the nodes on which the most excitable dynamical modes localize. Thus, from this perspective, the increase in prediction efficiency is an example for the direct practical implications of the localization of these modes.

### 3.7 Most important parameters and species

The iterative refinement procedure proposed above, needs some initial information on the system as a starting point. In this section, I therefore explore what types of parameters and what types of species should be prioritized when measuring a new food web, when no prior knowledge about the Jacobian matrix is available.

To get an intuition for the importance of different parameters for an impact prediction, consider a situation where the estimated Jacobian is identical to the true Jacobian, except for a single parameter that carries an error (of 10%). The quality of the estimated impact decreases with increasing influence and sensitivity of the species described by the estimated parameter (Fig. 3.7). Furthermore, the decrease is more pronounced for some parameters than for others: it is most pronounced for an error in the elasticities of mortality, predation, and primary production ($\mu$, $\Psi$ and $\phi$); it is less pronounced for the dietary contributions of different prey to a predator ($\chi$), and the elasticity of predation to prey abundance ($\gamma$); and it is weak for errors in species turnover rates ($\alpha$) and for the contributions of different predators to a species mortality ($\beta$). Without further knowledge, preliminary measurements should therefore focus on the elasticities $\mu$ and $\psi$, and on the dietary contributions $\chi$.

To get an intuition for the importance of different species, I explore the correlations between sensitivity or influence and species properties in a set of $10^6$ model food webs. In the analysis, I consider the correlations of sensitivity and influence with the following potential biological indicators:

- Generality, the number of prey species of a species (in-degree).
- Vulnerability, the number of predators of a species (out-degree).
- A binary value that is 1, if the focal species is a primary producer, and 0 otherwise.
- The trophic level $TL$ indicating how high a species position is in the food chain. The trophic level is 1 for primary producers. For other species it is $\text{mean}(TL_{\text{prey}}) + 1$, where $\text{mean}(TL_{\text{prey}})$ denotes the mean trophic level of the species’ prey.
3.7 Most important parameters and species

Fig. 3.7: Errors in different parameters have different effects on the average quality of an impact prediction. Each graph shows the average quality of impact prediction if one parameter of one node with the specified sensitivity (bottom) or influence (top) is subject to a measurement error. The different data-sets refer to errors in different parameters (c.f. Tab. 2.1). Other parameters are: system size $N = 50$, connectance $C = 0.04$, and the standard deviation of the relative measurement error is 10%. For higher $C$, the effect of $\chi$ and $\beta$ increases, but the qualitative behavior is identical. Noise at high importance and sensitivity values is due to the relatively rare occurrence of these values in the numerical experiments.

Fig. 3.8: Correlations of species’ properties with their sensitivity and influence. The pattern of correlations is consistent with large predators playing an important role for a system’s response to perturbations. Parameters $N = 50$, $C = 0.04$, results are similar for higher connectance.
• The biomass turnover rate (generalized model parameter $\alpha$), indicating the amount of biomass an individual consumes in comparison to its own mass.
• The network degree of a species, i.e. the total number of direct feeding interactions with other species.
• The weighted topological importance of a species $WI^s$ as introduced in detail in Ref. 140. For each species, the value of $WI^s$ indicates the indirect interactions from other species, based on the topology and biomass flows of a food web. The step parameter $s$ indicates the maximum number of direct interactions, through which indirect effects are perceived.

The correlation analysis (Fig. 3.8) shows that high trophic levels and low biomass turnover rates (e.g. long life span) correlate strongly with sensitivity and influence. This result suggests that top predators and other large and long-lived species, despite their typically small total biomass (compared to the total biomass of very abundant small species), play a disproportionate role in the systems’ response to perturbations. This corresponds well with observations in real-world systems [10, 141, 142]. Without specific information on biomass flows, these species should therefore be targeted for initial parametrization.

The sensitivity of a species is highly correlated with its generality, while its influence appears to be independent of its generality. Intuitively, this can be interpreted as generalist species being sensitive to changes in all of its prey species, while itself having relatively little impact on those species.

Applying the findings to improve the impact assessment of the Gatun Lake ecosystem in Sec. 3.2 (see also App. A), field work should prioritize better measurements of tetras (node 6 in Fig. 3.1), sailfin molly and mosquito fish (node 7), and, to a lesser degree, the large predators because these species have the highest combined sensitivity and influence (Fig. 3.2). Already without the explicit results on sensitivity and influence, the correlation would suggest to prioritize these species because of their low turnover rates, relatively high generality and their high trophic levels. Furthermore, the most significant improvements should be obtained by focusing in the measurement process on the elasticities of mortality and of predation of these species.

### 3.8 Discussion

Previous work has suggested that it may be intractable to predict the effects of perturbations to large, complex systems without near-perfect information on the system properties [135]. In this chapter, I presented a method to predict the
impact of perturbations on complex systems more efficiently by making use of
the localized character of the underlying dynamics which depends mainly on
a few key species.

The method I proposed is based on the linear stability of steady states in
food webs. Strictly speaking, it therefore describes the consequences of small
perturbations to a state of stable species coexistence. However, linear stability
is often found to agree well with other stability criteria for food webs, such
as robustness against noise [96], or permanence, measuring the boundedness
of a trajectory in a plausible part of the state space [143]. One can therefore
expect that the methods give at least some indications about perturbation con-
sequences and key species in systems that are not in a steady state.

The results in this chapter suggest that the potential impact of environmen-
tal perturbations on food webs can be predicted with reasonable accuracy if the
most relevant parameters for only a small number of important species in the
web are measured well. This is explained by the localization of the dynamical
modes which capture the strongest responses of a food web to a perturbation.
More precisely, the Jacobian matrix eigenvectors associated with these dynam-
ical modes have large amplitudes on only a few species. Therefore, precise in-
formation on these key species is sufficient to predict the responses of the most
relevant dynamical modes, making reasonable predictions possible despite a
lack of precise data on the remaining system.

Finding important species in a food web is the key to a good and efficient
impact assessment. I propose two methods to find important species by a)
pre-selecting species based on their biological properties, and b) applying an
iterative refinement procedure once some initial information is available. The
correlation analysis, used for pre-selection, suggests that it is most important
to obtain precise parameter estimates for long-lived, generalist consumers at
high trophic levels. Furthermore, the analysis suggests that for these species, it
is most important to precisely estimate the dependence of their mortality and
predation on their abundance.

While I have focused exclusively on food webs, the same approach can be ap-
plied to other networks of nonlinear interactions that are found in metabolism
control [109], gene regulation [115], and cellular population dynamics [113].

Understanding the origin and the implications of localized dynamical modes
is the goal of the remaining chapters of this thesis. Intriguingly, high sensitiv-
ity and influence values, which are closely related to localization of dynamical
modes, correlate not only with properties inherent to species (such as turnover
rates), but also with topological measures (such as the network degree). This
suggest that the localization of dynamical modes is not a feature inherent to
the properties of different elements of a network alone, but emerges, at least
partially, from the complex interaction topology.
4 Exact Localization

In the previous chapter, I showed that the inhomogeneity or localization of dynamical modes plays a crucial role for the response of a system to perturbations. In this chapter, I analyze a specific type of localized dynamical modes, modes that are exactly localized. These are dynamical modes which correspond to eigenvectors with only a few nonzero entries, and for which the remaining entries are exactly zero. Such an exactly localized dynamical mode therefore describes dynamics which is tied exactly to those nodes corresponding to the nonzero entries.

Exactly localized dynamical modes have been observed occasionally in the past. One example is the fundamental competitive-exclusion principle in ecology [16]. This principle describes the situation where two top predator species in a food web compete for the same prey (c.f Fig. 3.1). Normally, one of the two species out-competes the other, leading to that species’ extinction and disappearance from the ecosystem. As detailed below, this principle is explained by a special localized dynamical mode, which affects only the two competing species. This mode describes unstable dynamics, which increases one species (the “fitter” one), but decreases the other one, such that coexistence of both of them in a steady state is not possible [16].

The competitive-exclusion example links the occurrence of a small motif in a network to the presence of specific dynamics. Thus, one might hope to find more such motifs, and then use them like building-blocks to construct the entire system dynamics from their local dynamics. This idea has lead, for instance, to intense research of the occurrence of small structures in food webs and other networks [12, 14, 17, 19]. However, while it was possible to identify a certain three-node motif that has a destabilizing effect on species coexistence [18], implications of other motifs remain unclear, and also experimental findings suggest that the occurrence and properties of small motifs alone cannot explain the dynamics of a whole network [20]. Exact localization can address this discussion because it allows to rigorously distinguish which motifs are associated to localized dynamics and do therefore form building blocks, and which motifs require an understanding of the embedding network.
Fig. 4.1: Equivalence of local bifurcation diagrams of food webs with different topologies (top row from Ref. 43, bottom row adapted). If plotted as a function of suitable generalized parameters, the three-parameter bifurcation diagrams (top row) of different food webs (shown schematically in the top-right corners) are identical. The parameter space is spanned by the nonlinearity of mortality $m$ (equivalent to $\mu$ in the model presented in Sec. 2.7), the sensitivity of predators to prey abundance $g$ (equivalent to $\gamma, \Psi$), and the ratio between the timescales of predator and prey dynamics $r$ (equivalent to $\alpha_{pred}/\alpha_{prey}$). In the diagrams, every point in the parameter space corresponds to a specific steady state. Steady states are stable if the corresponding set of parameters $(g, m, r)$ lies in the front volume of the bifurcation diagram. The surfaces in the diagrams mark the location of bifurcation points of Hopf (red, green) and saddle-node (blue) bifurcations, where stability is lost. The lower row shows more examples for food webs for which equivalent dynamics (noted “=” was conjectured (see Ref. 43 for details).

A second example for an observation of (the effects of) exactly localized modes is dynamical equivalence found for certain food webs Ref. 43. Food webs with different topologies can give rise to identical (local) bifurcation diagrams if the food webs differ only by specific (symmetric) motifs (see Fig. 4.1 for an example). In particular, the authors of Ref. 43 conjectured that two species, that a) interact with the same set of topological neighbors, and b) are described by the same generalized model parameters, can be aggregated into a single vari-
able without changing the local bifurcation points\(^1\). Below, I show that the dynamical equivalence is explained by the presence of exactly localized dynamical modes resulting from the small network symmetries described by this conjecture.

The dynamical equivalence of food webs suggests that some part of the dynamics can be removed, for instance, through an aggregation of some species in a food web, without losing information on the apparent dynamics [43]. In fact, the assumption that such an aggregation is possible is implicitly made when aggregating species that hold topological positions in a network [61, 70, 144] and of strongly connected groups, known as communities in network terminology (in contrast to the ecological usage of the term) [12, 87, 145]. For both approaches, one builds on the intuition that the dynamics between the aggregated elements can be removed because it is redundant, or because it equilibrates more rapidly than the studied dynamics of the remaining system. However, rigorously, an aggregation without changing the observed dynamics is only possible if the removed dynamics is exactly localized and does therefore not depend on the remaining network.

In this chapter I study when dynamical modes can be attributed to specific network motifs, and how their localized dynamics can be removed through suitable aggregation of the species in these motifs. For this, I build on recent advances studying the implications of small symmetries in networks, called graph orbits. The presence of these symmetric motifs in an undirected unweighted network implies exactly localized eigenvectors in the corresponding graph spectrum [42, 146]. By extending these results to directed and weighted networks, they become applicable to study food web dynamics.

I start with a brief introduction of the notion and implications of graph symmetries (Sec. 4.1). By interpreting the Jacobian matrix as the adjacency of a hypothetical weighted network, I derive that in a food web, such symmetries lead to an independent subspace of the Jacobian matrix (Sec. 4.2). This subspace is the base for localized dynamical modes (Sec. 4.3) which can be removed through a coarse-graining algorithm that leaves the dynamics of the remaining network unchanged (Sec. 4.4). Application of these results to food web examples explains the observation of dynamical equivalence in different food webs (Sec. 4.5). To generalize the results, I study exact localization for structures that are not symmetric (Sec. 4.6) and the consequences of motifs that are only similar to a motif allowing exact localization (Sec. 4.7). Finally, I point out important aspects when considering systems different from food webs (Sec. 4.8), and summarize the implications of exact localization (Sec. 4.9).

\(^1\)The conjectured identity of bifurcation diagrams is not trivially true because of the nonlinearity of the underlying equations. Indeed, the identity is much harder to observe in conventional models and was hence not noted before the publication of Ref. 43.
4.1 Graph symmetries

Below, I study the localization of dynamics observed for local symmetries in a network. Here, I therefore introduce the terminology to characterize those symmetries. My presentation follows in large parts Ref. 42, but extends the theory presented therein to directed networks.

Graph symmetries are defined through node permutations. For instance, to permute two nodes \( n_1 \) and \( n_2 \) in a network, all links previously connecting to \( n_1 \) are rewired to \( n_2 \) and all links previously connecting to \( n_2 \) are rewired to \( n_1 \). In the same way, three or more nodes may be permuted by cyclically rewiring all links from each node to another node. However, permutations may involve more than a single cycle of nodes. For instance, a more complex permutation is formed by exchanging the nodes inside two pairs of nodes at the same time, thus forming two cycles of nodes that are simultaneously permuted.

A permutation of nodes corresponds to a symmetry, if the network remains unchanged during the permutation. Or, in terms of the adjacency matrix \( A \) of a network, node permutations correspond to simultaneous row and column permutations of this matrix. A network therefore contains a symmetry, represented by a permutation matrix \( P \), if application of this permutation leaves the adjacency relations unchanged, i.e. if \( PA = A \).

If multiple symmetries exist in a network which permute different nodes, then they can be combined into larger symmetries containing more nodes by element-wise addition of their permutation matrices. In the following, I consider only symmetries that are minimal, meaning that they cannot be decomposed into symmetries permuting fewer nodes. I denote all nodes that participate in such a minimal symmetry as a symmetric structure. Further, a set of nodes that are mapped onto each other under a symmetry permutation are said to form an orbit.

So far, I have reviewed the concept of orbits for unweighted networks (more details in Ref. 42), but it intuitively extends also to weighted and directed nets. However, in a weighted and directed web, symmetry between nodes in a weighted network requires, that the permutation of nodes preserves not only the adjacency relations, but also the weights and direction of these links.

For illustration of the concept of orbits, I now present some small networks containing symmetric structures. For each of them, nodes corresponding to the same orbit are colored identically. In the simplest case, a symmetric structure contains only one orbit (Fig. 4.2a). Some symmetric structures comprise multiple orbits, such that multiple nodes have to be exchanged in parallel to preserve the adjacency. For instance, in Fig. 4.2b), exchanging only the two
nodes at the top means that the nodes in the lower orbit become connected to a different node of the upper orbit, but exchanging the nodes in both orbits simultaneously retains edges between the same nodes.

I refer to the number of nodes belonging to an orbit as the size of the orbit. The size is the number of identical graphs generated by the corresponding symmetry operation, unless an orbit is nested inside another orbit (Fig. 4.2c), in which case it is a multiple of this number. For example, for Fig. 4.2 the number of identical graphs is always 2, and all orbits are of size 2, except the nested upper orbit in Fig. 4.2c), which is of size 4.

Symmetries have a distinct impact on the graph spectrum of an undirected and unweighted graph [42]. In fact, if such a graph contains a symmetry, then the spectrum of the graph contains two types of eigenvectors. The first type of eigenvectors is exactly localized on the symmetric structure. Furthermore, adding the entries of this eigenvector \( \mathbf{v} \) for all nodes of an orbit one obtains zero \( \left( \sum_p v_p = 0 \right) \), where \( p \) runs over all nodes of a given orbit of the symmetry). Following Ref. 42 I call these eigenvectors and their eigenvalues redundant. The second type of eigenvectors are all remaining eigenvectors. These eigenvectors have arbitrary entries outside the symmetric structure, but identical entries for all nodes of the same orbit. These results suggest that reducing the network by collapsing all nodes in each orbit onto each other leaves the spectrum of the adjacency invariant, except for removing the redundant eigenvalues.

In the following section, I show that similar results hold for the spectrum of the Jacobian matrix, when a food web contains a symmetry. Ultimately, this leads to a reduction algorithm for food webs, that allows to collapse symmetric species.
4.2 Localized dynamics on symmetries

The goal of this chapter is to find localized dynamics, dynamics that is tied to a small part of a network. In other words, I want to find dynamical modes corresponding to localized eigenvectors of the Jacobian matrix $J$ describing a systems dynamics near a steady state. Thus, I want to relate symmetries in a food web network with the occurrence of localized eigenvectors of the matrix $J$.

More formally, localized dynamics is guaranteed by the existence of a localized independent subspace of $J$. Localized means, that this subspace contains only localized vectors which have nonzero entries for the same nodes. Further, independent means that for every vector $v$ in the subspace, the vector $Jv$ is also in the subspace. Together these conditions imply, that if a state vector $Y$ is inside such a subspace, then it remains in the subspace under the application of $J$ prescribing the time-development of the system.

In the following, I show that if a food web contains a symmetry, then a localized independent subspace exists for the Jacobian matrix. This is facilitated by interpreting the Jacobian matrix as the weighted adjacency matrix of a (hypothetical) Jacobian matrix network. Intuitively speaking, the Jacobian matrix entries represent dependencies between the different variables in a linearized system. Thus, $J$ can be seen as the adjacency matrix of the weighted and directed network in which each node corresponds to one variable of the system, and in which each edge corresponds to a dependency between two variables.

The Jacobian matrix network and the underlying food web are closely related, but not identical. For instance, in the model used in this thesis (Sec. 2.7), each node in the food web is characterized by one variable, and therefore also corresponds to one node in the Jacobian matrix network [27]. However, the edges in the original and in the Jacobian matrix network are different. For example, a predator-prey relationship between a prey $i$ and a predator $j$ leads to only one edge in the food web. But in the Jacobian matrix network, the predator-prey interaction leads to multiple edges. For instance, it leads to one edge with positive weight (from $i$ to $j$) representing the gain of the predator, one edge with negative weight (from $j$ to $i$) representing the loss of the prey, and to additional edges between $i$ and the other prey species of $j$ because of apparent-competition effects. Generally, the Jacobian matrix network therefore has a similar structure as the original system, but contains more edges.

A food web contains a symmetry if the adjacency matrix of the food web contains an orbit, and if the symmetric species corresponding to this orbit are described by symmetric generalized parameters. Therefore, such a symmetry means that also the nodes corresponding to these species in the Jacobian matrix
network form an orbit. Thus, a symmetry in a food web is equivalent to a symmetry in the Jacobian matrix network describing the food web near a given steady state.

Note that species characterized by identical generalized parameters near a specific steady state need not be identical. For instance, consider two species that predate on the same prey species. For the first species, the predation gain increases linearly with its own abundance, such that the sensitivity of predation is \( \Psi = 1 \) regardless of the species abundance. The second species has an advantage when hunting in small groups, such that this species has a large sensitivity of predation for small abundances, but a low sensitivity of predation if the abundance (and therefore group size) becomes too large. For some intermediate abundance of the second species, the elasticities of both species are 1 and the two different species can be symmetric in that specific state (see Ref. 43 for a more detailed discussion).

Now, I extend the results for localization in undirected networks (see Ref. 42) to orbits in the Jacobian matrix network. Explicitly, I show that symmetries in a food web imply the existence of an independent localized subspace \( V \) of the Jacobian matrix. This subspace then acts as the foundation for the set of eigenvectors that are localized on the symmetric structure.

To prove the existence of the localized independent subspace \( V \), I show that in the presence of a symmetry, there is a non-empty set of localized vectors \( v \) that remain inside the subspace \( V \) under application of the Jacobian matrix. In particular, I show in the following that for a given symmetry this is the case for the following vectors \( v \): The vectors \( v \) are zero in elements corresponding to nodes outside the symmetric structure as required for localization. Further, the elements of \( v \) that correspond to nodes in the same orbit sum to zero (zero-sum condition).

Explicitly, I have to show that vectors \( z = Jv \) remain in \( V \) under the application of \( J \). For this, I proceed in two steps. First, I show that \( z \) remains localized on the symmetric structure, and second, I show that for any \( v \) also \( z \) fulfills the zero-sum condition.

First, to show that \( z = Jv \) remains localized, I confirm that the vectors \( v \) lead to \( z_k = \sum_i J_{ki}v_i = 0 \) on all nodes \( k \) outside the symmetric structure, where \( i \) runs over all nodes. The terms summing over nodes outside the symmetric structure are zero because \( v_i = 0 \). Further, the sum over all remaining nodes can be split up into sums \( S_n \) over the different orbits \( n \) of the symmetric structure, such that \( z_k = \sum_n S_n \), where \( S_n = \sum_j J_{kj}v_j \), with \( j \) running only over the nodes in a given orbit \( n \). Now consider the sum \( S_n \) over the nodes of a given orbit of the symmetric structure. The symmetry implies that node \( k \) must be connected to all nodes \( j \) in the orbit \( n \) with the same connection weight \( J_{kj} = b \). Therefore, \( S_n = \sum_j J_{kj}v_j = b \sum_j v_j = 0 \), where I used the zero-sum
4.2 Localized dynamics on symmetries

condition in the final step. Thus the zero-sum conditions on the different orbits guarantees that all terms $S_n = 0$ individually and hence $z_k = 0$. Therefore, the vector $z = Jv$ remains localized.

Second, to show that any vector $v$ in $V$ remains localized also under repeated application of $J$, I must show that also $z = Jv$ satisfies the zero-sum condition and therefore belongs to $V$. Thus, I want to confirm that for each orbit individually $\sum_k z_k = \sum_k \sum_i J_{ki} v_i = 0$, where $k$ runs over the nodes of the orbit. To show this, I group all edges between two given orbits together. Explicitly, consider the edges from the nodes inside an orbit $m$ towards the nodes of an orbit $n$ in the network of the Jacobian matrix. By reordering the variables in the system such that the variables describing nodes in $m$ and nodes $n$ have consecutive indices, all terms representing these edges appear in a single block in the matrix $J$. In the following, I call such blocks orbit connections and denote them by $C^{nm}$, because they contain all edges from nodes in orbit $m$ to nodes in orbit $n$.

To show, that $z$ satisfies the zero-sum condition, I write the sum $\sum_k z_k$ over the nodes $k$ of a given orbit as a sum over the contributing orbit connections. Explicitly, the contributions of the links from nodes inside an orbit $m$ to the entries of $z$ on nodes in an orbit $n$ are given by $C^{nm}v$ where in slight abuse of notation $v$ represents only those entries in $v$ referring to nodes in orbit $m$. Thus, I want to show that $\sum_k z_k = \sum_k (C^{nm}v)_k = 0$, where $k$ runs over all nodes in orbit $n$. However, one can change the order of summation over the rows of $C^{nm}$ with the summation over its columns performed in the product with $v$. In other words, I now first sum over the entries in each column of $C^{nm}$ (sum over $k$) and then multiply the resulting column sums by the corresponding entries of $v$.

The column sums of $C^{nm}$ respectively correspond to the sum over all links originating from one node in $m$ towards nodes in the orbit $n$. However, by symmetry these columns are permutations of each other such that their column sums are identical. Let this sum be $c$. Then, $\sum_k (C^{nm}v)_k = c \sum_j v_j = 0$, where $k$ runs over nodes in $n$ and $j$ runs over all nodes in $m$. The last step follows again from the zero-sum condition for the entries of $v$ on orbit $m$. Thus, under application of the Jacobian to a vector in $V$, the edges towards a given orbit originating in nodes of any other orbit sum to 0. Therefore, also summing over the edges from multiple orbits leads to zero, such that $z = Jv$ satisfies the zero-sum condition on each orbit.

Because the vectors $Jv$ meet both the localization and zero-sum conditions, $V$ is a localized independent subspace of $J$. In analog with the results on undirected graphs, I call this subspace the redundant subspace of the corresponding symmetry [42]. For a symmetry consisting of $o$ orbits with sizes $s_n$, the dimension of this subspace is $\sum_n s_n - o$. To see this, I observe that the sum
4 Exact Localization

corresponds to the number of nodes in the symmetry and thus to the number of nonzero entries of a vector \( \mathbf{v} \), which are subject to the zero-sum conditions. For example, if all orbits in the symmetry have the same size \( s \), the dimension of the subspace \( V \) is \( o(s-1) \).

To the right redundant subspace \( V \), spanned by column vectors \( \mathbf{v} \), corresponds a left redundant subspace \( W \), spanned by row vectors \( \mathbf{w} \) that remain localized under application of \( J \) from the right. To see this, consider the network of the transposed Jacobian \( J^T \) (because \( \mathbf{w}J = J^T \mathbf{w}^T \)). This network is identical to the network of \( J \) with inverted direction for all edges. These two networks therefore have the same orbits. Thus, for a given orbit, there is a left subspace \( W \) spanned by vectors \( \mathbf{w} \) that are both localized and fulfill the zero-sum conditions.

Taking the complement of a redundant subspace \( U = V^\perp \), one obtains another independent subspace. This subspace \( U \) contains all vectors that are orthogonal to the left redundant subspace \( W \). Since the vectors in \( W \) have arbitrary entries on the symmetric structure, but satisfy the zero-sum conditions, vectors in \( U \) are vectors with identical entries for nodes of the same orbit, and arbitrary entries outside the symmetric structure.

In practice, the goal is to find localized dynamical modes, which are associated to eigenvectors that belong to the redundant subspace \( V \). In other words, one has to diagonalize the Jacobian matrix inside the subspace \( V \) to obtain these eigenvectors, which I call redundant eigenvectors in the following. Since the eigenvectors of the Jacobian matrix generally form a complete base, any vector \( \mathbf{v} \) in \( V \) is then a linear combination of redundant eigenvectors. Remarkably, the difficulty of the required diagonalization to find them depends only on the complexity and size of the symmetric structure, but not on the complexity of the embedding system.

For example, consider a symmetric structure, that contains only a single orbit \( m \) of size \( s \). Considering only localized vectors \( \mathbf{v} \) in \( V \), \( J \) in the eigenvalue equation \( J \mathbf{v} = \lambda \mathbf{v} \) can be replaced by orbit connection \( C^{mm} \), because \( \mathbf{v} \) is zero elsewhere. Further, the rows and columns of \( C^{mm} \) are symmetric under cyclic permutation because the nodes of the orbit \( m \) are. Mathematically speaking, \( C^{mm} \) is a circulant matrix, for which the entries of the \( k \)‘th eigenvector are \( v_j^{(k)} = e^{2\pi ikj/s} \) [147]. The eigenvector for \( k = 0 \) violates the zero-sum condition \( (v_i^{(0)} = 1) \). The redundant subspace \( V \) is therefore spanned by the remaining \( s-1 \) eigenvectors corresponding to \( k = 1 \ldots s-1 \).

For completeness, I remark that the symmetry also facilitates solving the eigenvalue problem for symmetric structures containing multiple orbits. In this case, one has to diagonalize a so-called block-circulant matrix, containing the
4.3 Exactly localized dynamics

The presence of a symmetry implies the existence of localized eigenvectors of the Jacobian matrix satisfying the zero-sum conditions for each orbit, and of eigenvectors that are constant for the nodes of each orbit. This implies the existence of two corresponding types of dynamical modes: *Localized dynamical modes* that govern the dynamics acting anti-symmetrically on the nodes of the symmetric structure, and which are independent from the embedding network. *Constant dynamical modes* that govern the symmetric dynamics and which capture the dependence between the symmetric nodes and the remaining network.

The localization and structure of redundant eigenvectors has a distinctive effect on the corresponding dynamical modes. As an example, consider a perturbation to a system containing a symmetry. Only if the perturbation affects the nodes in the symmetric structure, it can trigger a response of the associated localized dynamical modes. Further, if such a mode is triggered, then also its response affects only the symmetric nodes.

The localization of the redundant eigenvectors furthermore means that also the stability of the corresponding dynamical modes depends only on the symmetric structure. The stability is characterized by the eigenvalues corresponding to the redundant eigenvectors. Because of the localization of their eigenvectors, these eigenvalues depend only on the Jacobian matrix entries referring to the symmetric structure. Thus, not only the dynamics is localized, but also its stability is characterized locally by the properties of the symmetric structure.

If a localized dynamical mode is stable, then the corresponding localized dynamics tends to disappear. For instance, for a symmetry, the (anti-symmetric) localized dynamics disappears, such that the system tends towards a symmetric state. However, if a localized dynamical mode is unstable, then the localized dynamics grows. For instance, for a symmetry case, the asymmetry between
the states of the symmetric nodes grows. Because of this growth, the dynamics leaves the neighborhood around the steady state, such that the linearization provided by the Jacobian matrix is no longer valid. In a nonlinear system, this generally means that the situation leading to localization (such as symmetry) is destroyed, such that the originally local growth can spread through the embedding network, possibly leading to large-scale effects. Thus, localized dynamics is an example for a possible system-wide dynamical effect, that can be traced to a local instability in a small network motif.

### 4.4 Symmetry reduction in networks

One can now imagine a mapping which collapses all nodes in a given orbit in the Jacobian network onto each other, such that the orbit is replaced with a single substitute node. It is intuitive that this mapping removes the anti-symmetric localized dynamical modes corresponding to redundant eigenvalues of \( \mathbf{J} \), but retains the symmetric ones associated to the remaining eigenvalues with eigenvectors inside the complement subspace \( \mathbf{U} \). For confirmation, I rewrite \( \mathbf{J}\mathbf{u} \) in terms of new variables that translate that nodes of a given orbit have identical entries, i.e. I replace all entries of \( \mathbf{u} \) referring to an orbit \( n \) by the value \( u'_n \), corresponding to collapsing all the nodes corresponding to these entries onto a single new one.

If a vector \( \mathbf{u} \) belongs to the independent subspace \( \mathbf{U} \), then also the vector \( \mathbf{z} = \mathbf{J}\mathbf{u} \) does\(^2\). Therefore, also \( \mathbf{z} \) is constant on nodes of the same orbit and can be written in terms of the new variables \( z'_n \). Thus, for vectors in \( \mathbf{U} \), the mapping that replaces all nodes of an orbit with a single one is consistent with the application of \( \mathbf{J} \).

Now, I establish the **reduced Jacobian matrix** \( \mathbf{J}' \), which corresponds to the Jacobian matrix \( \mathbf{J} \), but acts on the vector \( \mathbf{u}' \) of new variables. In other words, \( \mathbf{J}' \) is equivalent to \( \mathbf{J} \) in a system, in which orbits are collapsed into single nodes. First, I formally derive how the entries of \( \mathbf{J}' \) result from those in \( \mathbf{J} \), and second, I summarize the reduction of \( \mathbf{J} \) to \( \mathbf{J}' \) in a simple algorithm that directly reduces the network of \( \mathbf{J} \) to a network corresponding to \( \mathbf{J}' \).

\(^2\)Explicitly, the entries of \( \mathbf{z} \) on an orbit \( n \) are given by \( \sum_m \mathbf{C}^{nm} \mathbf{1}_m u'_m \), where \( m \) runs over all orbits, where \( \mathbf{C}^{nm} \) denotes the orbit connection from \( m \) to \( n \), and where \( \mathbf{1} \) is a vector that contains the entry 1 for each node in \( m \). Because each of the row-sums of \( \mathbf{C}^{nm} \) represents the sum over the incoming links of one node in \( n \) from nodes in \( m \), the entries of \( \mathbf{C}^{nm} \mathbf{1} \) representing these row-sums are identical. Therefore the entries in \( \sum_m \mathbf{C}^{nm} \mathbf{1}_m u'_m \), where \( m \) runs over all orbits, are also identical and can be replaced identical value \( z'_m \).
First, I write \( z = Ju \), in terms of the collapsed variables \( z' \) and \( u' \). This leads to
\[
z'_n = \sum_m C^{nm} 1_m u'_m,
\]
where \( C^{nm} \) denotes the orbit connection from an orbit \( m \) to an orbit \( n \), and where \( 1_m u'_m \) is a vector that is \( u'_m \) for each node in the orbit \( m \) (formally, \( 1_m \) is a vector with entry 1 on all of the nodes in \( m \) and entry 0 elsewhere). In analogy to the equation for the Jacobian matrix, \( z_j = \sum_i J_{ji} u_i \), this defines the entry of the reduced Jacobian matrix \( J'_{nm} \) as \( C^{nm} 1_m \).

Thus, replacing each orbit connection inside \( J \) by its row-sum, and removing the rows referring to nodes in the same orbit results in the reduction from \( J \) to \( J' \).

The reduction of \( J \) to \( J' \) can be summarized in a simple algorithm that reduces the network of \( J \) to a network corresponding to \( J' \). Starting with the network representation of the Jacobian matrix \( J \), remove all symmetric nodes except one from each orbit, such that this one node replaces the orbit in the reduced graph. Links that pointed towards a removed node are removed and links that originating from a removed node are rewired, such that they originate from the remaining node in the same orbit. Finally, the resulting network is simplified by replacing all links originating and ending in the same nodes by a single new link. The weight of this new link is the sum over the replaced ones. This network corresponds to the network of \( J' \), and I call it the *quotient* network of the network of \( J \) in analogy with the terminology introduced for undirected networks [146].

Because the application of \( J \) is identical to the application of \( J' \) for vectors in the subspace \( U \), \( J' \) contains the same eigenvalues as \( J \), except for the redundant ones. If the localized dynamical modes are stable (the redundant eigenvalues of \( J \) have a negative real part), then their dynamics tends to disappear. Therefore, on a meso-scale level, the dynamics described by \( J \) and \( J' \) are then exactly equivalent.

Summarizing, dynamical equivalence exists between two systems, if their original Jacobian matrices reduce to the same reduced Jacobian matrix \( J' \), and if the localized dynamical modes removed in their respective reductions describe stable dynamics. This equivalence is of particular interest because it describes not only the equivalence of simple transcritical and Hopf bifurcations, but extends also to local bifurcations of higher co-dimension, such as a Bogdanov-Takens bifurcation found at the intersections of local bifurcation surfaces (see Ref. 26 for details). However, the existence of such bifurcations has direct implications not only for the local dynamics near steady states, but also for the global dynamics of the system [26, 139]. Therefore, the proposed reduction should also conserve at least some features of the global dynamics.
4.5 Application to food webs

So far, the focus of this chapter was to develop general results for localization of dynamical modes due to symmetries. Now, I illustrate these results by studying the dynamics of example food webs. For simplicity, I use the model presented in Ref. 43 for which dynamical equivalence was observed empirically (Fig. 4.1). In this model, only top predators have significant biomass loss through natural mortality, while for all other species biomass loss is due to predation only. Further, all feeding interactions carry the same weight, i.e. the diet of a predator is composed equally from all of its prey species, and that the loss of a prey species is due equally to all of its predators. In terms of the generalized parameters characterizing the model presented in Sec. 2.7, this means $\sigma = 1$ for all species except top predators, $\chi_{ij} = 1/p$, where $p$ is the number of prey species $j$ of $i$, and $\beta_{ij} = 1/r$, where $r$ is the number of predators $i$ of a species $j$ (c.f. Tab. 2.1).

Each of the following examples is studied in four steps. First, I establish the Jacobian matrix of the original food web. Second, I collapse the symmetries in the Jacobian matrix network to obtain the reduced Jacobian matrix. Third, I identify the dynamical system corresponding to the reduced Jacobian matrix. Fourth, I compare the original food web and the system corresponding to the reduced Jacobian matrix. This shows, that the reduction of the Jacobian matrix is generally equivalent to a reduction that collapses species in the food web.

As the first example, I consider the competitive-exclusion structure shown in Fig. 4.3a). In this structure, two top-predators (populations 1, 2) feed on the same prey (population 3). The Jacobian matrix governing the stability of any steady state in this structure has the form

$$
J = \begin{pmatrix}
J_{11} & 0 & J_{13} \\
0 & J_{22} & J_{23} \\
J_{31} & J_{32} & J_{33}
\end{pmatrix}.
$$

The two top-predators are symmetric, if they are described by identical generalized parameters, leading to $J_{13} = J_{23}$, $J_{11} = J_{22}$, and $J_{31} = J_{32}$. Simple diagonalization then yields the eigenvalues $\lambda_{\pm} = 1/2(J_{11} + J_{33} \pm \sqrt{\Delta})$, where $\Delta = (J_{11} - J_{33})^2 + 4J_{31}J_{31}$, and $\lambda_{3} = J_{11}$. As expected, the eigenvector corresponding to $\lambda_{3}$, $v^{(3)} = (-1, 1, 0)$, is redundant since it is localized and satisfies the zero-sum condition $v_1 + v_2 = 0$.

Following the procedure outlined in the previous section, I collapse populations 1 and 2 in the Jacobian matrix network to obtain the reduced Jacobian matrix

$$
J' = \begin{pmatrix}
J_{11} & J_{13} \\
2J_{31} & J_{33}
\end{pmatrix}.
$$
Fig. 4.3: Examples for food webs that differ only by dynamical modes localized inside their meso-scale symmetries. Different colors denote different orbits. Two systems are marked as dynamically equivalent ("=") if the spectra of their Jacobian matrices are identical up to the localized dynamical mode characterizing the mesoscale effects inside the symmetries. If nodes of a symmetry have a common prey or predator with a node outside the symmetry, as e.g. in i), then the generalized parameters of the common prey or predator in the reduced network are weighted, which is shown in j) as a double-headed link. See text for more details.

Straightforward diagonalization confirms that this reduced matrix $J'$ retains the eigenvalues $\lambda_{\pm}$, while the eigenvalue $\lambda_3$, associated with the symmetry, has disappeared. Thus, $J'$ describes similar dynamics as $J$, but the localized dynamical mode is removed.

To find the reduced system corresponding to the reduced Jacobian matrix $J'$, I compare $J'$ to the Jacobian of a food web in which the symmetric species are collapsed. For the competitive-exclusion motif, this system is a simple predator-prey system (Fig. 4.3b). In terms of the generalized parameters (c.f. Tab. 2.1), the entries of $J'$ are given by $J_{11} = J_{22} = \alpha_1(\psi_1 - \mu_1)$, $J_{13} = J_{23} = \alpha_1\gamma_1$, $J_{33} = \alpha_3(\phi_3 - \gamma_3)$ and $J_{31} = J_{32} = -\alpha_3\psi_1/2$, where the factor of $1/2 (= \beta_{13} = \beta_{23})$ in the last expression arises because each of the species 1 and 2 is responsible for half of the biomass loss of species 3 in the original system. Plugging these expressions into (4.2), $J'$ is exactly identical to the Jacobian matrix (3.3) of the predator prey-system in Fig. 4.3b), in which the predator is described by the generalized parameters $\alpha_1, \psi_1, \gamma_1, \mu_1$, and the prey species by $\alpha_3$ and $\phi_3$. Remarkably, that the factor 2 in $J'$, that arises from the reduction, is indeed required to make this direct correspondence possible.
4 Exact Localization

The dynamics of the reduced (predator-prey) system is equivalent to the original (competitive-exclusion) system if the dynamics in the direction of the redundant eigenvector is stable, i.e. if the eigenvalue \( \lambda_3 \) that disappears in the reduction is negative. For the competitive-exclusion motif, \( \lambda_3 = \alpha_1(\psi_1 - \mu_1) \). Assuming that the predation gain increases linearly with predator abundance (\( \psi_1 = 1 \)), the original and reduced systems are therefore dynamically equivalent when predators suffers from super-linear mortality (\( \mu_1 > 1 \)), while \( \mu_1 \leq 1 \) leads to a localized instability.

While the competitive-exclusion motif was treated in isolation for illustration, the same reduction holds when it occurs inside food webs of arbitrary size. In particular, using the results from the previous section, it is not necessary to compute the complete spectrum of the Jacobian to apply the reduction. For instance, consider the competitive-exclusion motif inside the complicated food web in Fig. 4.3d) for which writing all the Jacobian matrix eigenvalues is complicated. Nevertheless, the symmetry guarantees that the competitive-exclusion motif in the web carries the redundant eigenvector associated with the eigenvalue \( \lambda_3 \). Removing the associated dynamical mode leads to the Jacobian matrix of the reduced food web in Fig. 4.3e), which is dynamically equivalent if \( \lambda_3 < 0 \).

As the second example, I consider the apparent-competition motif in Fig. 4.3c), consisting of one predator (population 1) feeding on two prey species (populations 2, 3). If the prey species are symmetric, then the Jacobian matrix of this structure is

\[
J = \begin{pmatrix}
J_{11} & J_{12} & J_{12} \\
J_{21} & J_{22} & J_{23} \\
J_{21} & J_{23} & J_{22}
\end{pmatrix}.
\]

In contrast to the Jacobian of the competitive-exclusion motif, this matrix contains the elements denoted by \( J_{23} \) that couple the symmetric species due to mutualistic apparent-competition effects.

The dynamical mode localizing on the symmetric prey has the eigenvector \((0, 1, -1)^T\) and the eigenvalue \( \lambda_3 = J_{22} - J_{23} \). Therefore, if \( J_{22} < J_{23} \), then the mode is stable and removing it leads to a dynamically equivalent system.

Following the procedure outlined in Sec. 4.4, the reduced Jacobian matrix becomes

\[
J' = \begin{pmatrix}
J_{11} & 2J_{12} \\
J_{21} & 2J_{22} + J_{23}
\end{pmatrix},
\]

which retains the eigenvalues of \( J \) except for \( \lambda_3 \). Explicitly, the entries of the Jacobian \( J \) of the apparent-competition motif in Fig. 4.3c) are \( J_{11} = \alpha_1(\psi_1 - \mu_1) \), \( J_{12} = \alpha_1 \frac{1}{2} \gamma \), \( J_{22} = \alpha_2(\phi_2 - (\gamma + 1)/2) \), \( J_{21} = -\alpha_3 \psi_1 \), and \( J_{23} = -\alpha_2(\gamma - 1)/2 \), where the factors \( 1/2(= \chi_{12} = \chi_{13}) \) arise because each
of the two prey species is responsible for half of the biomass gain of species 1. Plugging these terms into (4.4), the apparent-competition terms (containing $\gamma$) cancel, and the resulting Jacobian matrix is identical to the Jacobian matrix of a simple predator-prey system (Fig. 4.3b) in which the predator is described by the generalized parameters $\alpha_1, \psi_1, \gamma, \mu_1$ and the prey by $\alpha_2, \phi_2$.

In the two simple examples presented so far, the reduced Jacobian matrix was always identical to the Jacobian matrix one would have written for a simple predator-prey system in which species are described by the same generalized parameters as the symmetric species in the original system. In other words, the reduced Jacobian matrix, obtained by collapsing nodes in the Jacobian matrix network, was identical to the Jacobian matrix one would have written for a food web in which orbits were collapsed. This simple relation between the original and reduced system explains, why the dynamical equivalence between systems with such a simple symmetry could be observed heuristically (Fig. 4.1).

The reduction presented here allows to find equivalent systems, even if the relationship between the original food web and the system described by the reduced Jacobian matrix is less simple. For instance, consider a network where symmetric species are in competition or apparent-competition with species outside the symmetric structure, such as in Fig. 4.3i). There, a top predator (population 1) feeds on two symmetric prey species (populations 2, 3), but also on an additional species (population 4), which does not feed on the prey species (population 5) of the symmetric nodes. Collapsing the symmetric nodes in the Jacobian matrix network means summing up the contributions of the symmetric nodes to their common prey, and their contributions to their common predator. For the primary producer 5, collapsing the symmetry is identical to the competitive-exclusion situation above, such that in the reduced Jacobian, the entry describing the edge from the node replacing the orbit to the prey 5 is $J_{52} + J_{53} = 2(-\alpha_2 1/2\psi_2) = -\alpha_2 \psi_2$. The resulting entry is therefore identical to the entry one would have written for the food web in Fig. 4.3k). However, for the top-predator 1, each of the symmetric nodes is responsible for only $1/3$ of its total diet. Therefore, in the reduced Jacobian matrix, the entry describing the edge from the node replacing the symmetric nodes to the predator is $J_{12} + J_{13} = 2/3\alpha_1 \gamma_1$. This entry differs from the one in the food web in Fig. 4.3k) obtained by collapsing the original food web. This the case also for other entries of the reduced Jacobian matrix, because the weight of $1/3$ also appears in the mutualistic interactions between the symmetric species and with species 4. Thus, simply collapsing the nodes in the food web in Fig. 4.3i), but retaining the species properties as before, does not lead to a system described by the reduced Jacobian matrix.
While retaining species properties when collapsing symmetries in the food web does not always lead to an equivalent web, the food web corresponding to the reduced Jacobian matrix $J'$ can still be obtained by collapsing symmetric nodes. However, the reduced food web must retain the relative strength of the effects due to the collapsed orbit compared to the effects from other species. For instance, I use a double-headed arrow in Fig. 4.3j) to mark that the effect on the top predator from the node replacing the two symmetric species is twice as strong as the effect of the additional prey species. In terms of generalized parameters, the dietary parameters of the reduced food web must reflect the contributions in the original system. Generally speaking, collapsing the orbits of the original food web leads to a dynamically equivalent reduced system with different generalized parameters.

Collapsing all orbits in a food web into single nodes, reduces also symmetries containing multiple orbits. The main difference is that the dynamics inside such a symmetry may itself be complicated. For instance, in the food web shown in Fig. 4.3h), the symmetric structure consists of two interconnected orbits of top predators (1, 2) and their respective prey (3, 4), which feed on the common primary producer (5). When collapsing the orbits, this web reduces to a simple chain (Fig. 4.3g). To satisfy the zero-sum conditions on the two orbits, the redundant eigenvectors of this food web must have the structure $v = (a, -a, b, -b, 0)^T$. Solving $Jv = \lambda v$ for such vectors leads to two redundant eigenvalues $\lambda_\pm = 1/2(J_{11} + J_{33} \pm \sqrt{\Delta})$ with $\Delta = (J_{11} - J_{33})^2 + 4J_{31}J_{13}$. The redundant eigenvalues can thus form a pair of complex-conjugate eigenvalues that may lose stability in a Hopf-bifurcation and thus lead to oscillatory instabilities.

In summary, in networks, every mesoscale symmetry is associated with localized dynamical modes which can be studied by considering only the symmetric structure. If the dynamical modes are found to be stable, then they can be removed from the system without changing the remaining dynamics according to the rules described in Sec. 4.4. For food webs, this reduced Jacobian generally describes a physically sensible, yet smaller, food web, with possibly altered parameter values.

The reduction of food webs to dynamically equivalent systems explains the empirical observations in Ref. 43 and proves the conjecture made there. Moreover, it extends this conjecture to cases that are much harder to observe empirically because the parameters describing species of the original food web and of the reduced food web corresponding to the reduced Jacobian are different. This is the case, for instance, for symmetric species competing with species outside a symmetry (Fig. 4.3i), or when considering more elaborate models. Neverthe-
less, even in these cases, the reduced Jacobian matrix can be computed straightforwardly using the procedure proposed here and therefore allows to establish rules that relate the original and the reduced system.

## 4.6 Localization on asymmetric structures

Inspired by the observations of localization and aggregation for symmetries, one might ask, if similar localization occurs also on structures that are not symmetric. One might even hope that a sufficient number of such structures exists to reconstruct the complete dynamics of a network from localized dynamics. However, using food webs, I show here that only specific classes of network motifs lead to exactly localized dynamical modes. Therefore, a reconstruction of the whole network dynamics is generally not possible.

The symmetries considered so far are symmetric in two ways. First, they are topological symmetric structures, and second, they are described by symmetric parameters. Thus, when attempting to generalize to different structures, one has to consider a) more general topologies, and b) parameters that are not identical.

For illustration, I first study localization for the competitive-exclusion motif (Fig. 4.3a), when the topologically symmetric predators have different elasticities and turnover rates. The simplest way to retain localization in this case are parameter combinations that are not symmetric, but which still lead to a symmetric Jacobian matrix. In other words, I consider parameter combinations for which the local eigenvalue problem $J_{loc}v_{loc} = \lambda v_{loc}$ remains symmetric, where $J_{loc}$ contains only those entries of the Jacobian matrix that refer to the two predators which carry the localization. Explicitly, using the formulation (4.1), this symmetry requires that the diagonal entries of the topologically symmetric species are identical ($J_{11} = J_{22}$), which yields $\alpha_1(\psi_1 - \mu_1) = \alpha_2(\psi_2 - \mu_2)$.

Assuming that the competing species have similar elasticities ($\psi, \mu$), there are two fundamental solutions to the equation $\alpha_1(\psi_1 - \mu_1) = \alpha_2(\psi_2 - \mu_2)$. First, the species are symmetric and therefore have identical turnover rates ($\alpha_1 = \alpha_2$). Second, the elasticities satisfy $\psi = \mu$, such that arbitrary turnover rates $\alpha_{1,2}$ are permitted. This case is of particular interest because it includes the biologically reasonable case of linear predation and mortality ($\psi = \mu = 1$) [16], and because it allows localized dynamics, even when the species carrying the localization are inherently different. Remarkably, the corresponding eigenvalue $\lambda = 0$ marks a bifurcation in the system, suggesting that this type of localization is only likely to occur close to changes in the system dynamics.
In addition to the asymmetry of parameters describing the properties of the predators in the competitive-exclusion case, I now consider that also their interactions with their common prey are asymmetric. Such an asymmetry leads to an asymmetry of the entries describing these interactions in the Jacobian matrix, such as $J_{31}$ and $J_{32}$ in (4.1). Localization requires that the entry of a localized eigenvector on the common prey vanishes, $v_3 = 0 = J_{31}v_1 + J_{32}v_2$; a condition that generalizes the zero-sum condition for symmetries ($0 = v_1 + v_2$). In the following I call this condition an attachment condition because it describes the prey node through which the predators, and the localized dynamics, attach to the remaining network.

A localized eigenvector $v$ for the competitive-exclusion motif with different generalized parameters for the predators must simultaneously satisfy the attachment condition ($v_1/v_2 = -J_{32}/J_{31}$), and the local eigenvalue equations $J_{11}v_1 = \lambda v_1$ and $J_{22}v_2 = \lambda v_2$. This is possible exactly when $J_{11} = J_{22}$, independently from $J_{32}$, and $J_{31}$. Therefore, for the competitive-exclusion motif, localization of $v$ is possible for asymmetric interactions of the predators, but the parameters describing the inherent properties of the predators ($\alpha$, $\psi$, $\mu$) must lead to the symmetric entries in the Jacobian matrix I studied above.

In the competitive-exclusion structure, the biologically sensible case of linear mortality and predation leads to localization despite possibly different turnover rates of the competing species. While this explains, why the competitive-exclusion principle became such an important ecological principle, other structures generally require species to be similar to allow localization. For instance, consider the apparent-competition structure containing two producers (Fig. 4.3c) described by the Jacobian matrix (4.3). For this structure, the attachment condition (of the common predator of the prey species 2 and 3) requires that $v_2/v_3 = -J_{13}/J_{12}$, and the local eigenvalue problem leads to

$$\alpha_2(\phi_2 - \mu_2) = \alpha_3(\phi_3 - \mu_3).$$

While one can assume mortality to be linear ($\mu = 1$), the sensitivity of production $\phi_3$ is typically sub-linear, because the production saturates with increasing abundance of a species [92, 122]. Therefore, localization on the apparent-competition structure requires producer species described by similar parameters, such as symmetric turnover rates and elasticities.

Now, I extend the findings for species that are topologically symmetric, but which are described by different parameters, to structures that are also topologically asymmetric. For this, I select arbitrary nodes in a given Jacobian matrix network and study the conditions an eigenvector $v$ must satisfy to localize on them. In the following, I call these nodes localization nodes.

---

$^3$Explicitly, $J_{22}v_2 + J_{23}v_3 = \lambda v_2$ and $J_{33}v_3 + J_{32}v_2 = \lambda v_3$. Inserting the attachment condition then cancels all terms except $\alpha_i(\phi_i - \mu_i)v_i = \lambda v_i$ with $i = 2, 3$. 

---

62
As for the topologically symmetric cases above, localization on a set of localization nodes requires that the vector $v$ simultaneously solves the local eigenvalue problem on these nodes, and satisfies the attachment conditions on the nodes separating the localization nodes from the remaining network. The local eigenvalue problem for $v$ is $\lambda v_i = \sum_j J_{ij} v_j$, where $j$ runs over the localization nodes. If all the eigenvalues $\lambda$ are different (which is typically the case in an inhomogeneous structure), then the solutions for a given eigenvalue is a single eigenvector. In other words the eigenspace of each eigenvalue is of dimension 1. Only if some of the eigenvalues are degenerate, then the corresponding eigenspace can have higher dimension (up to the degeneracy of the eigenvalue, see Ref. 102 for details).

The attachment conditions on an eigenvector $v$ must be satisfied for localization of this vector on the localization nodes. Explicitly, each node $i$ that has an incoming edge from a localization node in the Jacobian matrix network leads to a condition $v_i = 0 = \sum_j J_{ij} v_j$ which must be simultaneously satisfied by $v$. Note, that this may be impossible, for instance, if a node has only one relevant incoming edge.

For a given set of localization nodes, I use dimensional analysis of the eigenvalue problem and attachment conditions to study if they allow the existence of a localized eigenvector. More precisely, I consider the dimension of each eigenspace and consider the attachment conditions as restrictions of this space. Explicitly, if an eigenvalue is $P$-fold degenerate, then its eigenspace is (at most) of dimension $P$. Further, if there are $R$ attachment conditions (which must be independent; see below), these conditions reduce the dimension of the eigenspace in which an eigenvector also satisfies the attachment conditions to at most $P - R$. Localization therefore requires $P - R > 0$, i.e. that the degeneracy of an eigenvalue in the local eigenvalue problem is higher than the number of independent attachment conditions.

For illustration, consider again the competitive-exclusion motif with asymmetric parameters. This motif has one attachment condition. Localization therefore requires a two-fold degenerate eigenvalue, which exists only if $J_{11} = J_{22}$. Thus, the dimensional approach leads to the same condition as in the explicit calculations above.

The number of independent attachment conditions, $R$, needs some discussion. In particular, for food webs, there are two important cases in which this number is lower than the number of nodes that are the destination of edges originating in localization nodes.

First, if the local eigenvalue problem is governed by the same interactions as an attachment condition, then the resulting eigenvectors may generically match this condition. The matched attachment condition then puts no restriction on the eigenspace. For instance, this is the case for symmetries, where the
Fig. 4.4: Example for a structure that is topologically asymmetric and contains species described by asymmetric parameters, but leads to exact localization. The structure (schemas in the first and last column) contains three localization nodes (solid gray): two form a topologically symmetric kite motif, and one is a top predator. $\Delta$ indicates the asymmetry of the predator feeding on the nodes in the kite motif ($\chi_{\text{pred.}, \pm} = \chi(1 \pm \Delta)$) and of the feeding of the kite motif species on their common prey ($\beta_{\pm, \text{prey}} = \beta(1 \mp \Delta)$). For $\Delta = 0$, the kite structure is symmetric (schema in the first column). For $\Delta > 0$ the attachment conditions are different (black edges in the schema in the last column mark higher biomass flows than gray ones). The color-code of the nodes shows the proportion of the localized eigenvector to different nodes of the network ($|v_i|^2$, solid blue is $1/2$, white 0).

Symmetry applies to the attachment conditions and to the eigenvector equations, explaining, why symmetries always lead to localized dynamics. Another example are the mutualistic interactions between the prey species in the apparent-competition motif, which define the eigenvectors and correspond to the interactions of the prey species with the common predator. Because of this, the eigenvalue problem is always solved by an eigenvector which also satisfies the attachment condition. This explains why localization occurs in this structure despite the non-degenerate eigenvalues.

Second, attachment conditions may be identical, such that they are not independent restrictions. Again, this is the case for symmetries for which all attachment conditions are multiples of each other. Furthermore, it is the case for attachment conditions resulting from apparent-competition of species with a localization node due to a common predator. Intuitively, apparent-competition describes that the predator’s behavior towards all of its prey changes when the abundance of one of them changes (for instance, saturation due to the increase of one prey’s abundance). Therefore, when the effect of the localization nodes on the predator cancel, i.e. when the predator’s attachment condition is satisfied, then no mutualistic effects occur and also the attachment conditions arising from the mutualistic effects are satisfied.
In summary, determining if a specific structure in a network allows for localization of an eigenvector requires two steps. First, one has to solve the local eigenvalue problem on the nodes of the structure to obtain the degeneracy \( P \) of each eigenvalue. Alternatively, to simplify, one might in a preliminary step use the number of nodes of a structure as the maximally possible value for \( P \). Second, one has to find the number of (independent) attachment conditions \( R \) by counting the number of nodes through which the structure is connected to the remaining network, but excluding identical ones. Localization is then possible if \( P - R > 0 \). For food webs, \( R \) generally corresponds to the number of species outside the structure, that predate on nodes in the structure or are prey to these nodes. Exceptions are only a few special cases, which lead to identical attachment conditions (such as symmetries), but which are contingent in specific parameter combinations, and are therefore easy to exclude in the counting process.

For an inhomogeneous system, such as a food web, eigenvalues normally have low degeneracy, such that localization is rare. However, there is one class of food web structures and parameters for which such degeneracy, and therefore localization, is plausible. Topologically, these are structures, in which nodes only interact through competition or apparent-competition, but do not predate on each other directly. The simplest examples for such structures are the competitive-exclusion and apparent-competition examples studied above. More generally, if there are \( P \) such nodes for which their respective parameters satisfy \( \alpha_i (\psi_i - \mu_i) = \lambda \), then the degeneracy of the eigenvalue \( \lambda \) is \( P \). Localization therefore occurs on these nodes if there are less than \( P \) species preying on these nodes, or being prey of these nodes.

An example for a topologically and parametrically asymmetric structure that carries localization is shown in Fig. 4.4. For this example, all species have linear mortality and predation sensitivities (\( \mu = \psi = 1 \)) such that localization can occur on species with different turnover rates. For \( \Delta = 0 \), the two blue nodes are symmetric. The local eigenvalue problem on them leads to a two-fold degenerate eigenvalue \( \lambda = 0 \), which allows localization, because the attachment conditions from their common predator and common prey are identical. Furthermore, localization remains possible also if the attachment conditions are different (\( \Delta > 0 \)) because the additional top predator leads to a three-fold degenerate eigenvalue in the local eigenvalue problem, such that the two attachment conditions can be simultaneously satisfied.
4.7 Nearly-exact localization

Above, I show that exact localization requires specific combinations of generalized parameters. However, because the systems under consideration generally change smoothly with these parameters, one may expect that localization persists to some degree also for situations, where the conditions for exact localization are only approximately satisfied. To test this hypothesis, I numerically study two specific structures, the apparent-competition motif with asymmetric turnover rates, and the kite motif (two nodes with one common predator and one common prey), with asymmetric attachment conditions.

For the apparent-competition structure, localization is lost when the turnover rates of the competing species are asymmetric (Sec. 4.6). Here, I quantify their asymmetry, $\alpha_\pm = \alpha(1 \pm \delta)$, through a parameter $\delta > 0$. When $\delta$ increases, the sum of the contributions of the eigenvector on the two competing nodes quickly decreases (Fig. 4.5).

For the kite motif, localization is lost when its attachment condition on the common prey and on the common predator become incompatible. I characterize this incompatibility through the parameter $\Delta$, which quantifies the difference in the interactions with the common predator and prey as $\chi_\pm = \beta_\pm = 1/2(1 \pm \Delta)$ (see also Fig. 4.4). When $\Delta$ increases, the decrease in localization differs between systems, but generally it occurs slower than for a change of turnover rates (Fig. 4.5).

In summary, localized dynamics requires specific parameter combinations, but persists approximately when parameters are close to such a combination. This justifies the aggregation of similar species and communities in food webs even if their interactions with the remaining web differ slightly, which is commonly applied to food webs [61, 70, 144]. At the same time, however, the examples show that if the species in a food web are different (i.e. if they have different turnover rates), then highly localized eigenvectors are only possible on the particular structures identified in Sec. 4.6.

---

4 A measure for this is the vector entry $v_{pred}$ of the common predator. Linear matrix perturbation suggest that it grows as $|v_{pred}|^2 \propto \delta^2 \alpha_{pred}/\alpha$, where the proportionality constant for the typical values $\gamma = 0.95$ and $\sigma \approx 1$ is close to 2000.

5 Linear matrix perturbation predicts, that the contribution of the eigenvector to the common predator and prey (pr.) grow as $|v_{pr.}|^2 \propto \Delta^2 \alpha_{pr.}/\alpha$, where the proportionality constant is approximately 1.
4.8 Other systems

Localization hinges on the degeneracy of the eigenvalues solving the local eigenvalue problems of common motifs in a network. While this degeneracy, and thus localization, is rare in food webs, I now identify systems for which exact localization is more prominent, and discuss aspects to consider when generalizing the results in the last chapters to other systems.

Searching for network structures that carry exactly localized eigenvectors is particularly promising for unweighted networks with identical nodes. For instance, consider localization of the adjacency matrix eigenvectors for the graph in Fig. 4.6. Solving the local eigenvalue problem on any clique of $M$ identical nodes (two or more nodes connected all-to-all) results in the redundant eigenvalue $\lambda = -1$ with degeneracy $M - 1$. If such a clique is subject to less than $M$ independent attachment conditions, then the clique carries a localized eigenvector (yellow nodes in Fig. 4.6).

Fig. 4.5: Loss of localization when parameters only approximately satisfy the conditions for localization. The left-hand panel shows the loss of localization on the two nodes of the apparent-competition motif when their turnover rates become asymmetric ($\alpha_\pm = \alpha (1 \pm \delta)$). The right-hand panel shows the loss of localization on the nodes of the kite motif when the attachment conditions become incompatible (see Fig. 4.4 for details). Localization is measured as the norm of $\mathbf{v}$ on the symmetric nodes. Red lines mark curves for different example systems; the black curve marks the average over 100 realizations. Parameters: $N = 12, C = 0.12$. 

4.8 Other systems

Localization hinges on the degeneracy of the eigenvalues solving the local eigenvalue problems of common motifs in a network. While this degeneracy, and thus localization, is rare in food webs, I now identify systems for which exact localization is more prominent, and discuss aspects to consider when generalizing the results in the last chapters to other systems.

Searching for network structures that carry exactly localized eigenvectors is particularly promising for unweighted networks with identical nodes. For instance, consider localization of the adjacency matrix eigenvectors for the graph in Fig. 4.6. Solving the local eigenvalue problem on any clique of $M$ identical nodes (two or more nodes connected all-to-all) results in the redundant eigenvalue $\lambda = -1$ with degeneracy $M - 1$. If such a clique is subject to less than $M$ independent attachment conditions, then the clique carries a localized eigenvector (yellow nodes in Fig. 4.6).
Fig. 4.6: Localization in an undirected unweighted network. Nodes with the same color belong to a structures on which eigenvectors of the adjacency matrix localize. Colored edges correspond to attachment conditions; different colors mean that attachment conditions are independent (for instance, the six yellow edges belong to three identical conditions). Localization takes place on cliques (nodes connecting all-to-all, yellow and turquoise structures) or on structures combining cliques (blue, red, green structures). The corresponding eigenvalues are all $-1$.

Since all cliques in an unweighted graph give rise to the same eigenvalue in the local eigenvalue problem, eigenvectors can localize on structures combining unconnected cliques (green, red and blue structures in Fig. 4.6). This requires that all attachment conditions that refer only to one clique are satisfied, and that the number of cliques in the whole structure is larger than the number of attachment conditions involving multiple cliques. For instance, in the blue structure in Fig. 4.6, the attachment conditions corresponding to the cyan and magenta edges are satisfied because of the symmetry of the cliques they connect to. Localization occurs on the blue structure because the 3 individual blue cliques connect to the remaining network via a single attachment condition (blue edges).

The localization of eigenvectors on the graph in Fig. 4.6 suggest that localized dynamical modes are frequent in systems in which the dynamics is captured by the eigenvectors of an unweighted matrix. Examples for such systems are networks of coupled identical oscillators, or networks of social interactions.
in which the individuals are replaced by identical nodes\textsuperscript{6}. In these systems, the symmetry reduction algorithm (Sec. 4.4) leads to great simplifications (see also Ref. 146). Remarkably, the reduction process might result in new symmetries in such systems, such that a recursive reduction becomes possible.

The Jacobian matrix network of a food web contains reciprocal edges because, if a predator depends on a prey, then the prey also depends on the predator. However, not all systems have this property, such as networks describing movement patterns between different locations\textsuperscript{7}. In these networks, a localized eigenvector \(v\) no longer implies the localization of the corresponding left eigenvector \(w\), or vice-versa. In other words, one might observe effects with a localized origin but delocalized consequences, and vice-versa.

Exact localization is particularly important in the field of control theory, which asks how many variables of a network system need to be regulated, i.e. set to a specific value, to control the values of all others \[149, 150\]. Structures leading to exact localization are important because the localized eigenvectors introduce additional degrees of freedom that can only be controlled by explicitly regulating one of the localization nodes.

In summary, the results on exact localization presented in this chapter generalize well to systems other than food webs. In fact, while exactly localized modes are rare in food webs, they occur prominently in more homogeneous systems, such as networks of identical nodes.

### 4.9 Discussion

In this chapter, I discussed the origin and implications of exactly localized dynamical modes. For this, I first extended a result for symmetries on undirected unweighted graphs to show that certain dynamical modes originate on symmetric structures in food webs. Based on these findings, I developed a reduction algorithm, that separates the localized from the delocalized dynamics. Afterward, I generalized the results for symmetries to structures that are not symmetric and discussed implications of exact localizations when studying different systems.

\textsuperscript{6}I study such systems in the next chapter. For instance, I use a model for infection spreading in which infection outbreaks are characterized by the eigenvectors of the underlying undirected unweighted network topology (Sec. 5.2). There, exactly localized dynamical modes lead to the distinct peak at \(-p\) in the histogram in Fig. 5.1, corresponding to eigenvectors localizing on cliques in the network with eigenvalue \(-1\). Similarly, the peak at 0 in the left-most histogram in Fig. 5.6 is due to eigenvectors localizing on stars (multiple nodes connected only to one common neighbor) favored by the underlying preferential-attachment model.

\textsuperscript{7}For example, the model for the transport of an invasive species through the network of global cargo-ship movements in Sec. 5.2.
4 Exact Localization

When a small network motif is identified to carry a localized dynamical mode, the associated possible instability can be studied regardless of the embedding system. If the dynamical mode is stable, then the mode drives the system back towards its normal operating state and the dynamics has can be omitted from further investigations. However, if the mode is unstable, then it leads to an instability, that may affect the entire system. For instance, for food webs, an unstable localized mode generally leads to dynamics that destroys the symmetry of the linearized system, such that the instability can have system-wide consequences. A motif allowing localization is therefore an example for a meso-scale structure with distinct dynamical implications for the dynamics of a network as a whole.

Because localized dynamical modes are independent from the remaining network, one can locally remove them to obtain a coarse-grained system. In particular, if the localized dynamics is stable, then this removal has no effect on the observable (non-local) dynamics; the dynamics of the coarse-grained system is exactly equivalent to the original system. Performing this coarse-graining on food webs generally leads to a smaller, but dynamically equivalent, food web in which symmetric species are aggregated. This reduction explains the dynamical equivalence of food webs with different topologies observed heuristically [43], and it justifies the aggregation of similar species and communities commonly applied in food webs [61, 70].

Taking the opposite direction from the coarse-graining process, exact localization also allows to engineer specific dynamics in a network by adding suitable small motifs. For instance, my results on symmetries were used in Ref. 151 to engineer specific instabilities in a system describing species migrations between multiple habitats through a network of delay-coupled patches.

Exact localization is possible only on specific topological motifs, which give rise to degenerate eigenvalues when studying their local eigenvalue problem, and which have few connections with the remaining system. These motifs, such as symmetric stars or cliques, but also parallel paths from the bottom to the top of the underlying network, are common in unweighted networks of identical elements. Therefore, in systems characterized by such networks, insights into the dynamics are possible by studying typical motifs that carry localization, and application of the reduction algorithm leads to a strong reduction.

If a network contains nodes with different properties, such as species in a food web which have different intrinsic timescales, then exact localization is possible only for specific combinations of the parameters describing these properties. For instance, for food webs, localization is only possible close to bifurcation points, such as in the competitive-exclusion case, or on species with similar properties, such as species in symmetries.
The requirement for species with similar properties does not preclude localization in food webs because similar species, and therefore localization, may arise naturally. For instance, they might arise for meta-populations where the same species is found in more than one geographical location [152], or right after speciation events when an ancestral species splits into two newly formed species that are still similar. Moreover, even if the specific parameter combinations for localization are only approximately met, localization is retained approximately. Therefore evolutionary or environmental effects might lead to situations with localized dynamics. For instance, species holding similar topological positions in a food web often have similar properties, as reflected in allometric scaling laws relating body mass and metabolic rates with the trophic level of a species [72, 123, 124, 153].

While motifs carrying localized dynamics are common in unweighted networks, the small number of topological motifs allowing localization shows that not all dynamical modes of a network can be exactly localized. Thus, while counting the occurrence of small motifs gives a heuristic indication about design principles underlying a specific network, only those few motifs carrying exact localization allow to draw conclusions on the functionality of the network. Therefore the results on exact localization challenge typical attempts to describe network dynamics as the sum of their part’s dynamics [17, 154] and opens the way to rigorously link specific motifs to their function in a network.
Approximate Localization

In the previous chapter, I studied dynamics associated with exactly localized eigenvectors, which are exactly zero on most nodes of a network. Here, I study a less extreme form of localization, *approximate localization*, associated with localized eigenvectors which are localized predominantly on a few nodes, but which can have small amplitudes also on the remaining nodes of a network.

Examples for approximate localization appear in the previous chapters. For instance, structures for which the conditions for exact localization are approximately satisfied carry an eigenvector which has small nonzero entries outside the structure (Sec. 4.7). A second example are the dynamical modes of Gatun Lake, which I study in the context of perturbation impact (Fig. 3.5). There, the dynamical modes corresponding to the largest eigenvalues are associated to eigenvectors with large entries on three species, while all other entries are smaller by at least two orders of magnitude.

For perturbation assessments, approximately localized dynamical modes play a central role because they explain why few key species are most important for assessing perturbation responses. More explicitly, the difference in importance results from the approximate localization of the most excitable modes, those dynamical modes corresponding to eigenvalues closest to the imaginary axis. Because these modes characterize the most prominent responses of a food web to a perturbation, the species on which they localize are the key species.

The approximate localization of excitable dynamical modes is important beyond the perturbation assessments. For a system in a steady state operating mode, the most excitable dynamical modes are closest to bifurcation situations in which this operating mode loses its stability. This makes the excitable modes particularly important in the context of *slow-moving risks* describing systems that slowly drift towards less stable conditions. In this context, the modes closest to bifurcations are generally the first to show unstable behavior [119, 155]. The most excitable dynamical modes are thus directly tied to the most prominent instabilities such complex systems face.

A fundamental debate in complexity science, called the complexity-stability debate, was started through a seminal paper by R. May in 1972 [44]. This debate revolves around the apparent discrepancy between the stability of the complex
networks observed in nature, and the theoretical observation that the probability of randomly drawing a stable network topology decreases exponentially with the complexity of this topology [44, 92, 107, 122]. However, in the light of the localization of prominent instabilities, one might imagine that the decrease in stability of complex systems observed theoretically is due to the occurrence of localized instabilities. Thus, one might ask if networks in nature achieve stability through the removal or modification of only those few nodes on which these instabilities localize, thereby achieving stability with only a minimal loss in complexity.

Recent simulation results suggest that unstable food webs typically find a stable configuration after the extinction of only a few species [25, 56, 95, 156]. One might thus imagine that this stabilization is the manifestation of a process in which an unstable food web stabilizes its prominent localized instabilities, while retaining most of its structure and complexity. Localization of prominent instabilities might therefore explain the complexity of networks observed in nature as the result of an evolutionary process in which networks stabilize through the local break-off of problematic parts to the benefit of the remaining network.

For illustration, I use spider webs to visualize a process that locally repairs problems, but preserves the remaining system’s complexity. In these webs, the elasticity of the silk normally enables the web to endure high loads by diffusing tension [22, 157]. However, when the load exceeds a certain threshold in a given part of the web, a molecular rearrangement makes the silk stiff and brittle. Arguably, this mechanism has evolved as a strategy to withstand extreme forces; sacrificing the affected strands locally rather than transmitting the forces and endangering the whole web. While this specific mechanism in spider silk may well be an extreme example that evolved due to strong selective pressure, it nevertheless shows a mechanism where the stability of a larger (global) network results from sacrificing problematic parts.

Also many artificial networks, such as social or technical systems with an underlying network structure, show features that suggest a localization of their prominent dynamics. For instance, recent results suggest that outbreaks of infections on networks, such as a computer virus in a network of computers, or a disease in a network of socially interacting humans, localizes on a small number of nodes of the network that are re-infected again and again [24]. Other examples are the localization observed for eigenfunctions of the Schrödinger operator around impurities in condensed matter which reduce conductivity [34, 35], or failures in power-grids that appear to have their origin on a specific small parts of such grids [158].
Recent results for power grids suggest a trade-off between the stability of a grid, i.e. the risk of a black-out occurring, and the probable size of such a black-out [8, 158]. This means that repairing a system by increasing its stability comes at the cost of increasing the risk of large-scale failures. In other words, these results suggest that a trade-off exists between the robustness of a network, i.e. its ability to withstand perturbations, and its resilience, i.e. its ability to recover from large failures with minimal losses. To understand this trade-off and develop methods to repair systems effectively, it is therefore instructive to investigate localization with regard to attempts to stabilize network systems.

To reflect the diversity of the systems in which approximate localization is observed, I study in this chapter not only food webs, but complex systems in general. First, I develop a quantitative measure for the number of nodes affected by an approximately localized dynamical mode (Sec. 5.1). Then, I present the localized instabilities occurring in four different contexts: infection spreading on social networks, bioinvasion of port areas by new species due to cargo-ship movements, instabilities in ecological food webs, and instabilities causing asynchrony in power grids (Sec. 5.2). I show that localization, as in these examples, is linked to the typical spectral properties of network matrices describing natural systems (Sec. 5.3), and that this localization is independent from the size of the system under consideration (Sec. 5.4). Using the paradigmatic model introduced by R. May as a prototype model for complex systems [44], I investigate localization of instabilities analytically (Sec. 5.5), and identify properties of nodes on which instabilities localize (Sec. 5.6). Finally, I show that food webs in nature might stabilize through the extinction of species that carry their prominent instabilities (Sec. 5.7), and contrast this stabilization process with attempts to stabilize artificial networks through human intervention (Sec. 5.8).

### 5.1 Spread of a dynamical mode

Below, I study approximate localization through the localization of dynamical modes which characterize the prominent instabilities that may drive a system away from its operating mode. For this purpose, I require a measure for the number of nodes that are associated with a given instability. Here, I introduce a measure for this number.

For a given normalized vector \( \mathbf{v} \), I define its spread as

\[
\text{Spread}(\mathbf{v}) = \frac{1}{\text{IPR}(\mathbf{v})} = \left( \sum_i |v_i|^4 \right)^{-1}.
\]  

(5.1)
In this formula, IPR stands for inverse participation ratio which is a standard measure for the localization of a vector (see e.g. Ref. 23 for an overview over possible measures). Intuitively speaking, \( \text{IPR}(v) = \sum_i |v_i|^4 \) measures localization as the statistical variance of the amplitudes \( |v_i|^2 \) of a vector on different nodes, and spread is therefore the inverse of this variance.

In a network, the spread of a vector corresponds the number of nodes on which it localizes. To see this, consider a vector \( v \) which is exactly localized on \( L \) nodes and has equal amplitude on all of them, i.e. \( |v_i|^2 = 1/L \) on these nodes. Then, the spread is \( (L/L^2)^{-1} = L \). Thus, for an exactly localized vector with identical nonzero entries, the spread corresponds exactly to the number of nodes on which the vector localizes. Or, consider an approximately localized vector with approximate amplitudes \( 1/2, 1/4, 1/4 \) on three specific nodes, and negligible amplitude elsewhere. This vector is slightly more localized than a vector contributing equally to all three nodes, and the spread is 2.67. Thus, also for approximately localized vectors, the spread provides an effective number of nodes on which a vector localizes.

To improve readability, I refer in the following to the spread of the eigenvectors associated to a given dynamical mode simply as the spread of the dynamical mode. If this mode characterizes an instability, then I call it the spread of this instability.

5.2 Examples for localized instabilities

In this chapter, I discuss typical localized instabilities in four example systems: first, the SIS model, which is a standard model for infection spreading on networks [9, 24, 33, 54, 110]. Second, a model for bioinvasion events, which I apply to a real-world data-set of cargo-ship movements around the globe [159, 160]. Third, the generalized model for food webs used in the previous chapters [27]. Finally, a model for machines coupled in a power grid, for which synchrony is described by the Master Stability approach for coupled oscillators [29, 30]. For each model, I briefly review the underlying spectral approach that identifies instabilities and present the observed localization. Afterward, I summarize and discuss their common features (Fig. 5.6).

**SIS-model for infection spreading**

Infection spreading phenomena in complex systems are a typical application of network models [1, 54, 161]. Examples are the spreading of a computer virus on the internet [110], the spreading of an infectious disease or rumor through a network of socially interacting humans [33], or the spreading of a forest fire through scattered patches of woodland [162]. To study these systems, one represents each computer, individual, or patch as one node of a network, and
represents interactions through which the infection can spread as edges. For instance, edges represent the social interactions between humans that might transmit the studied disease, the exchange of data that allows a computer virus to infect a new computer, or patches of woodland that are geographical proximity close enough to allow for mutual ignition. In these systems, one asks if an infection can spread throughout the network after occurring on a few nodes.

A standard model to capture spreading phenomena on a network is the SIS model (SIS stands for Susceptible-Infected-Susceptible) [1, 33, 54]. In this model, each node of a network is either infected (I), or susceptible to infection (S). A node can become infected, if it has an edge towards another node that is infected. Infection due to such an edge takes place with an infection rate \( p \), which indicates the probability that the susceptible individual becomes infected within a given period of time. Further, an infected node can recover, thereby becoming again susceptible to infection later. This recovery is governed by the recovery rate \( r \), which indicates the probability that an infected individual spontaneously recovers during a given period of time. Phenomenologically, the infection spreads with a rate \( p \) through the network of interactions, but is hampered by each individuals potential to recover \( r \).

I study spreading on two types of undirected and unweighted static networks. For the introduction of the dynamics and the basic results, I focus on random geometric graphs in which edges indicate spatial proximity of nodes. They are generated by randomly distributing the nodes in a plane and then drawing an undirected edge between those nodes that are closer than a given threshold value. These networks model, for instance, scattered patches of woodland that are close enough to ignite each other [161]. Further, I study Barabasi-Albert graphs which have the small-world properties that are commonly observed in networks of social interactions, computer networks and other networks growing through preferential attachment [2, 163, 164].

Typically, spreading phenomena in the SIS-model are studied through a probabilistic approach [9, 33]. One studies the dynamics of the infection probabilities \( I_i(t) \) which indicate the probability that a node \( i \) is infected at time \( t \). In an infinite system, and in the time-continuous limit, the dynamics of the infection probability of each node is approximately governed by the differential equation
\[
\frac{dI_i}{dt} = -rI_i + (1 - I_i)(1 - \prod_j (1 - A_{ij} p I_j)),
\]
where \( A_{ij} \) is the adjacency matrix of the network [33].

Whether a spontaneous occurrence of an infection in some nodes dies out or spreads through a given network, is determined by the stability of the infection-free state \( I_i \approx 0 \). If this state is (linearly) stable, then a small nonzero probability of infection in some nodes generally disappears; if it is unstable, then the infection spreads [33]. To study this stability, one linearizes the dynamics around the infection-free state, leading to
\[
\frac{dI_i}{dt} = -rI_i + p \sum_j A_{ij} I_j.
\]
Thus, the Jacobian
matrix describing the dynamics near the disease-free state is $J = -r \mathbf{1} + p \mathbf{A}$, where $\mathbf{1}$ is the identity matrix. If all of the eigenvalues of $J$ are negative, then the infection-free state is stable, while positive eigenvalues are associated to dynamical modes that promote spreading of the infection.

The stability of the infection-free state depends only on the eigenvalues of the adjacency matrix $\mathbf{A}$ and on the parameters $p$ and $r$. If $\lambda_k$ denotes the eigenvalues of the adjacency matrix $\mathbf{A}$, then the eigenvalues of the Jacobian matrix are $p\lambda_k - r$. A mode is therefore stable if $\lambda_k p / r < 1$, where this quotient is called its reproductive number [165]. Of particular interest is the reproductive number $R_0$ corresponding to the largest eigenvalue $\lambda_0$. If $R_0 < 1$, then the infection-free state is stable, but if $R_0 > 1$, then the infection spreads.

If the reproductive number $R_0$ is slightly larger than 1, then the infection spreads and persists on only a small number of nodes [24]. This suggest, that the corresponding eigenvector of the adjacency matrix is localized. For instance, starting from an almost infection-free state, the dynamical mode corresponding to $R_0$ initially promotes increasing infection probabilities $I_i$ for nodes with large amplitudes of the corresponding eigenvector $\mathbf{v}$. However, at the same time, second-order terms $(-p \sum_k I_k^2 A_{ik})$ become significant and halt the spreading. Thus, if $R_0$ is only slightly larger than 1, infection spreads mainly on those nodes on which $\mathbf{v}$ localizes.

The connection between localized outbreaks and the localization of the eigenvectors corresponding to the largest eigenvalues of the adjacency matrix can be observed in explicit simulations of the SIS model. For example, for the network in Fig. 5.1, the average infection risk is large for those nodes on which the eigenvectors with large eigenvalues localize. For small $r$, the infection risk localizes mainly on two clusters (Fig. 5.1a), higher values of $r$ result in a more distributed infection risk, but the clusters are concentrated around the same nodes of the network (Fig. 5.1b). Further, the eigenvectors corresponding to the five largest eigenvalues localize around the same nodes as the most prominent clusters of high infection risk (Fig. 5.1e). While in Fig. 5.1a) large infection risk exists mainly for nodes with large entries in the eigenvectors of the three largest eigenvalue, large risk exists in Fig. 5.1b) and in Fig. 5.1c) increasingly for nodes with large contributions of the subsequent eigenvectors.

---

1I study the SIS-dynamics in a network by performing time-steps with simultaneous update of all nodes. In each time step, an infected node recovers with probability $r$. Further, a susceptible node remains susceptible with probability $(1 - p)^n$, where $n$ denotes the number its neighbors that are infected at the beginning of the time-step (c.f. Ref. 166). To study infection risk, I randomly infect 10% of the nodes of a network and then perform 250 time-steps to allow the system to reach a steady state. Repeating this process 1000 times, I assess the infection risk of a node as the number of repetitions in which it is found infected after the 250th update.
Fig. 5.1: Distribution of infection risk and centers of outbreak. The panels in the top row represent infection risk for each node in a network for different recovery rates $r$, while the infection rate, $p = 0.0391$, is constant. The color code marks nodes with zero risk in white, and those with the maximum observed risk as red. Panel d), shows the eigenvalue density of the graph (scaled by $p$). The five largest eigenvalues are marked triangles marked in different colors (see the color code for their values). In panel e), the intensity of each of these colors indicates the amplitude of the corresponding eigenvectors on each node of the network. Notably, the eigenvectors localize on the same nodes as the most prominent infection risk in panels a–c).
5.2 Examples for localized instabilities

The situation where only a few dynamical modes promote infection spreading corresponds to a situation in which first outbreaks are observed in a context of slow-moving risks. For instance, such a process might start with a high recovery rate which slowly decreases over time. Initially, infected nodes then quickly recover after infection and therefore the infection disappears. However, with increasing risk, infected individuals can infect enough others to allow the infection to spread. From a spectral perspective, the eigenvalues of the Jacobian matrix, \( p\lambda_k - r \), are initially negative, but become positive one by one. Spreading of the infection is therefore first possible on those nodes that have a large entry in the eigenvector of the largest eigenvalue (marked in green in Fig. 5.1e). As \( r \) decreases further, the infection also spreads onto those nodes of the next smaller eigenvalue (marked in blue), and the subsequent ones.

As a note of caution, I remark that a positive eigenvalue \( p\lambda_k - r \) does not guarantee the spreading of an infection if the system is finite. Since the spreading process is stochastic, it may accidentally reach the absorbing infection-free state, even if a dynamical mode, in theory, promotes spreading [165, 167]. For instance, for the network in Fig. 5.1, the largest Jacobian eigenvalue is \( p\lambda_0 - r = 0.3 - r \). Nevertheless, for \( r = 0.2 \) (Fig. 5.1a), only few simulation runs result in a persistence of the infection for more than 250 time-steps.

The five eigenvectors corresponding to the largest eigenvalues of the graph in Fig. 5.1 are localized (their spread is approximately 13). Because these modes are the first to promote infection spreading for decreasing \( r \), their localization explains the observations of the distinct localized outbreak centers in the explicit simulations.

The localization of the eigenvectors corresponding to the largest eigenvalues of networks occurs not only for the random geometric graphs studied so far, but also for graphs generated by the Barabasi-Albert preferential attachment algorithm (Fig. 5.2). In a context of slow-moving risk, the first infection outbreaks in these networks, i.e. their most prominent problems, are localized.

In summary, when studying infection spreading on a given network using the SIS-model, then largest eigenvalues of the network adjacency matrix correspond to the most prominent outbreaks of the infection. The larger the eigenvalue, the earlier the associated outbreak occurs in a situation of slow-moving risk. In the networks studied so far, the eigenvectors corresponding to these eigenvalues are localized, such that the corresponding outbreaks affect only a few nodes.

Bioinvasion due to cargo-ship movements

The movements of ships between distant parts of the world has allowed species, such as rats [168] or zebra mussels [169], to populate habitats far beyond their natural ranges [160]. Here, I formulate a simple model for the population of port areas by a species that is not native in these areas, but which arrives on a cargo-
ship originating in a populated port. I apply the model to study the localization of such population events based on a data set of cargo-ship movements around the globe [159].

Imagine a port that is populated by a given species. Each time a ship leaves from this port, it can carry some individuals of this species to the port of its destination. Under favorable conditions, the species then populates this new port. When each port is represented as one node in a network, then the movement of cargo-ships from each port to the next can be represented as edges between these nodes. Since the likelihood of transporting a species is proportional to the frequency at which ships travel on a specific route, the network of species transport is directed and weighted. It is thus captured by a weight-matrix $W$, in which entry $W_{ij}$ indicates the frequency at which ships travel from port $j$ to port $i$.

To model the dynamics of species populations in different ports, I denote the abundance of the species in a given port $i$ as $X_i$, relative to the maximum possible population $X_i = 1$. Further, I denote the probability that individuals of a species survive a transport from a populated port to another port as $p$. Finally, I assume that the mortality of the individuals of the new species in a

---

Fig. 5.2: Spread and eigenvalues of the dynamical modes in the SIS-model. Shown is the spread of dynamical modes as a function of the associated eigenvalue. The left panel corresponds to a random geometric graph (a, graph shown in Fig. 5.1), and the right panel corresponds to a graph generated using the Barabasi-Albert preferential attachment (b). Regions, where eigenvalues lead to infection outbreaks (positive eigenvalues for a choice of $p$ and $r$) are shaded in red. Parameters: Number of nodes: $N = 250$ (a) and $N = 5000$ (b). Average number of neighbours: 9 (a), 3 (b). Infection rate: $p = 0.0384$ (a) and $p = 0.0169$ (b). Recovery rate: $r = 0.27$ (a) and $r = 0.18$ (b).
port where they are not native is \( r_i(1 - X_i) \), where \( r_i \) indicates how adverse the port environment is for the invading species, while \( (1 - X_i) \) captures that the chance for survival of each individual increase with the population density. If \( p \) is reasonably small, then dynamics of the species abundances in a port where the species is not native is given by \( \frac{dX_i}{dt} = -r_iX_i(1 - X_i) + (1 - X_i) \sum_j W_{ij}pX_j. \)

In the following, I assume that the matrix \( W \) describes the network of cargo-ship movements between ports in which the species is not native, but that movements from ports in which the species is native towards a port of the network of ports where the species is not native are captured by the vector \( d \). Explicitly, the entry \( d_i \) of this vector indicates the frequency at which cargo-ships from ports where the species is native arrive at the unpopulated port \( i \). The linearized dynamics near a state in which the ports \( i \) are generally not populated \( (X_i = 0) \) is then \( \frac{dX_i}{dt} = JX + pd \), where \( J \) is the Jacobian matrix with entries \( J_{ii} = -r_i \) and \( J_{ij} = pW_{ij} \).

Phenomenologically, the bioinvasion model allows for two different types of population dynamics in an initially unpopulated port. First, the continuous arrival of cargo ships from populated ports can sustain a small abundance of species. This abundance is \( X^* = -pJ^{-1}d \), where the exponent \(-1\) denotes the matrix inverse. Second, if the density of the species becomes high enough, for instance, due to temporarily favorable environmental conditions, then the port may become permanently populated \( (X_i = 1) \) [159, 160]. In line with the previous arguments, one can expect that such an event resulting in permanent population of a port is most likely to occur in ports with large values of the temporary population \( X_i^* \).

The temporary population density \( X^* \), which is sustained by the arrival of cargo-ships from ports where the species is native, is analogous to a perturbation response to this sustained introduction \( pd \) of the species (c.f. Sec. 3.1, with \( K = pd \), and \( I = X^* \)). In this picture, the most prominent population phenomena are determined by the most excitable dynamical modes of \( J \). Translating the results in Sec. 3.3 to the bioinvasion system, the amplitudes of a right eigenvector corresponding to a large eigenvalues of \( J \) indicate, which ports in the network are likely to be populated, if a port with a large amplitude of the corresponding left eigenvector is perturbed (by a large number of incoming ships \( pd_i \)). Thus, if the right eigenvectors are localized, permanent population is likely to occur only for few ports, while localization of the left eigenvectors implies that prior arrival of the species to a few specific ports is required.

Now, I apply the bioinvasion model to study the population of new ports by an invasive species based on a data set of worldwide cargo-ship movements [159]. Since, normally ports are differently suitable for a temporary survival of the invasive species [160], I draw the values of \( r_i \), from a log-normal
Fig. 5.3: Spread of the dynamical modes in a model for bioinvasion. In each plot, the blue data points indicate the spread of the right eigenvector (blue) and left eigenvector (yellow) associated to a dynamical mode with an eigenvalue that has the real part indicated on the horizontal axes. The different panels correspond to different parameter combinations of $p$ and $\sigma$; $p$ indicates the invasion probability, and $\sigma$ indicates the width of the (lognormal) distribution that characterizes the suitability of different ports for survival of the invading species. The region shaded in red contains the largest eigenvalues, for which the associated dynamical modes are most likely to cause the permanent population of ports where the species is not native.

distribution with variance $\sigma^2$ around 1. A low value of $\sigma$ means that all unpopulated ports are similarly suitable for the survival of the species, while a large value means that suitability varies.

The distribution of the spread values for the different dynamical modes in the real-world cargo-ship network is qualitatively similar when making different choices for the parameters $p$ and $\sigma$ (Fig. 5.3). Quantitatively, the spread decreases with decreasing transportation survival rate $p$, and with increasing $\sigma$. Intuitively speaking, the spread of each eigenvector decreases when the system becomes more inhomogeneous.
For all parameter combinations, the eigenvectors associated to the dynamical modes with the largest eigenvalues have low spread (Fig. 5.3). Therefore, the most prominent population events generally have well-defined local origins, and should lead to the population of only a few new ports.

**Food Webs**

The results in Chap. 3 suggested that in a food web the dynamical modes corresponding to the largest eigenvalues are localized. Here, I study this localization in more detail.

I recall that in the generalized model description for food webs I use in this thesis (Sec. 2.7), instabilities to the state of species coexistence arise from dynamical modes associated to eigenvalues of the Jacobian matrix that have large positive real parts. If such a dynamical mode exists, then the steady state is unstable against perturbations in the direction of the corresponding left eigenvector, and such a perturbation will cause the system to leave the steady state in the direction of the corresponding right eigenvector.

Because food webs typically contain tens or hundreds of species [90, 91], making general statements about the typical spread of their eigenvectors requires combining data from many sample webs. Here, I therefore average over the spread found for dynamical modes with similar eigenvalues in a set of $10^7$ model webs.

For model food webs, the average spread is small for dynamical modes corresponding to very small or very large eigenvalues (Fig. 5.4). The most prominent instabilities, associated to very large eigenvalues, affect less than 5 species. Further, the overall spread increases when increasing the connectance of the food web, but the qualitative properties remain the same. These results support the observation that generally the most prominent instabilities in a food web have a localized origin and an initially localized effect.

**Synchrony of coupled oscillators**

Networks of coupled oscillators are a standard model to study the synchrony of elements with periodic dynamics [30]. They are applied, for instance, to study the synchrony of biological and technical networks, such neuronal networks [170], metabolic cycles [30], or power grids [158].

I study the synchrony in networks of coupled oscillators using the Master Stability Function approach, which was introduced in Ref. 29. Here, I summarize this approach briefly, a more detailed introduction is given in App. B; for a detailed review, see Ref. 30.

The Master Stability Function approach describes synchrony in networks of identical oscillators. Each of the oscillators is a small subsystem, described by its *internal state* $\mathbf{x}$, consisting of one or more variables. In the absence of
coupling between the oscillators, the evolution of the internal states of each oscillator is described by $\frac{dx}{dt} = F(x)$, where $F$ is the (vector-valued) function characterizing the oscillators’ internal dynamics. Further, if two oscillators are coupled, then their states affect each other. The effect of an oscillator $j$ on a (coupled) oscillator $i$ is given by the difference of the (vector-valued) coupling function $H$, i.e. the effect of oscillator $j$ on the internal state of oscillator $i$ is $H(x^{(j)}) - H(x^{(i)})$, where the superscript indicates to which oscillator the state $x$ refers.

From the internal dynamics $F$ and the coupling function $H$, one can establish the Master Stability Function $\text{MSF}(\mu)$ (see App. B or Ref. 30 for details). When considering a specific coupling topology, this function characterizes if the topology allows synchronous operation of the oscillators based on the eigenvalues $\mu$ of the Laplacian matrix of the coupling network. More precisely, if the Master Stability Function $\text{MSF}(\mu) < 0$ for all the eigenvalues $\mu$ of
5.2 Examples for localized instabilities

Fig. 5.5: Spread of the dynamical modes describing possible instabilities of the synchronous operation of a network of randomly coupled oscillators. Each data-point shows the spread of the eigenvectors associated to a given eigenvalue $\sigma \mu_k$, where $\mu_k$ are the eigenvalues of the Laplacian matrix of the coupling network, and where $\sigma$ is a scaling factor that depends on the properties of the oscillators. Eigenvalues lead to a breakdown of synchrony if they are very large or very small (shaded regions for a given choice of oscillators and coupling functions). Shown are results for random networks with 5000 oscillators and independent probability for the presence of each link of $C = 0.002$ (a) and $C = 0.01$ (b). Scaling parameters: $\sigma = 0.045$ (a) and $\sigma = 0.0074$ (b).

the coupling network, then the synchronous state is (linearly) stable. However, if $\text{MSF}(\mu) > 0$ for an eigenvalue $\mu$, then this eigenvalue corresponds to an instability of the synchronous state.

Because the Master Stability Function depends only on the internal dynamics $\mathbf{F}$ and on the coupling function $\mathbf{H}$, one must solve $\text{MSF}(\mu) > 0$ only once to identify those values of $\mu$ which lead to instabilities for any network topology. Typically, the Laplacian Matrix eigenvalues $\mu$ lead to instabilities if they are outside a well-defined interval. An example is given in App. B, where a simple toy model for coupled power generators or machines leads to a function $\text{MSF}(\mu)$ that is negative only for intermediate values of $\mu$. This structure is also observed for other oscillators, such as Rössler systems, and is often considered the standard case [29, 30, 171]. In practice, one can therefore expect synchronous operation in coupling networks with eigenvalues $\mu$ inside a specific interval, while large and small eigenvalues of the coupling network outside this interval lead to instabilities.
The amplitude of an eigenvector corresponding to a specific large or small eigenvalue $\mu$ of the network Laplacian indicates in which oscillators the associated instability manifests itself (see App. B). Large entries indicate which oscillators initially lose synchrony, and thereby cause the (possibly global) breakdown of the synchronous state. Low spread of such an eigenvector therefore indicates an initially localized break-down of synchrony.

As an example, I consider identical oscillators coupled in an Erdös-Renyi random network [48, 172]. These networks can be obtained by taking $N$ unconnected nodes, and assigning an undirected and unweighted edge with an independent probability $C$ between any pair of them [172]. The spread of the Laplacian matrix eigenvectors for a large connected network of this type is shown in Fig. 5.5. The spread is low for eigenvectors corresponding to the dynamical modes with large and small eigenvalues, which are associated to the most prominent instabilities.

In summary, for a system of oscillators with identical internal dynamics $F$, and coupled through a function $H$, one can establish the Master Stability function $MSF(\mu)$. This function identifies which eigenvalues $\mu$ of the Laplacian matrix of a coupling network allow synchronous operation of the oscillators, and which eigenvalues lead to instabilities. Typically, these instabilities appear for very large eigenvalues, or very small eigenvalues, or both. For systems of randomly coupled oscillators, the spread of these instabilities is low.

**Summary & Discussion**

The four examples presented here are network systems in which instabilities are associated to particularly large or small eigenvalues of their different characteristic matrices. All systems have in common that these extreme eigenvalues belong to regions of low eigenvalue density, and that the associated problems are generally localized (Fig. 5.6).

Notably, a connection between low density of eigenvalues, and localization of the corresponding eigenvectors, has been observed for various types of random matrices [36–39, 173–175]. Below, I show that such a connection exists generally between the low density of extreme eigenvalues in natural systems and the localization of the dynamical modes describing their most prominent dynamics.

Arguably, low density of extreme eigenvalues observed in the examples might be typical also for the characteristic matrices of other real-world network systems. First, real-world complex systems are generally inhomogeneous because they contain elements operating on different timescales (e.g. different species in a food web), or because the individual elements have different topological positions (e.g. hubs in a network). This inhomogeneity translates to a wide distribution of eigenvalues, which has low-density tails for eigenvalue
### Model Examples for Localized Instabilities

<table>
<thead>
<tr>
<th>Model</th>
<th>Disease spreading</th>
<th>Bioinvasion</th>
<th>Food Web</th>
<th>Oscillators</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Graph</strong></td>
<td>Barabasi-Albert graph</td>
<td>cargo-ship movements</td>
<td>niche model food webs</td>
<td>random graph</td>
</tr>
<tr>
<td><strong>Matrix</strong></td>
<td>adjacency matrix</td>
<td>weight matrix</td>
<td>Jacobian matrix</td>
<td>Laplace matrix</td>
</tr>
<tr>
<td><strong>Size</strong></td>
<td>5000</td>
<td>951</td>
<td>50</td>
<td>5000</td>
</tr>
</tbody>
</table>

**Fig. 5.6:** Summary of the spread and spectra of dynamical modes in four examples. Each column corresponds to one example, showing the type of the characteristic matrix, the type of network, and the network size. The graphs represent the spread of the eigenvectors of the respective matrix and a histogram of their eigenvalue density. For each example, the ranges of problematic eigenvalues are shaded in red. In all examples, these extreme eigenvalues have low spread and belong to ranges of low eigenvalue density.
region that separate the high-density region of common eigenvalues from regions without eigenvalues [38, 39, 41, 173, 175–177]. Second, low density of eigenvalues at extreme eigenvalues is also linked to the sparsity of such networks [36, 37, 178]. For instance, the eigenvalue density for random matrices with entries drawn from a normal distribution forms a semi-circle (called Wigner’s law) if all entries of the matrix are set, but becomes a triangle with low-density tails on both sides if the matrix is sparse [179].

I remark that special isolated eigenvalues appear in some characteristic matrices. However, the corresponding dynamical modes generally have a specific interpretation, such as a globally conserved quantity, which makes them obsolete in the study of localized problems. One example is the eigenvalue 0 in the spectrum of the Laplacian matrix, which cannot lead to asynchrony in an oscillator network (App. B) [30]. Other examples are eigenvalues that correspond to the row or column sums of a matrix, such as the adjacency matrix of a network in which all nodes have the same degree [151, 180].

Summing up, the extreme eigenvalues of suitable network matrices describe the prominent instabilities in different systems. In real-world networks, one can expect that such eigenvalues belong to regions of low eigenvalue density. In the following, I show that such low eigenvalue density explains the observed localization of the eigenvectors corresponding to these eigenvalues, which suggest the localization of prominent instabilities also in other networks.

### 5.3 Localization of extreme eigenvalues

In the previous section, I suggest that a connection exists between the localization of prominent instabilities of a characteristic matrix, and the low density of the associated extreme eigenvalues. I now study this connection between low eigenvalue density and eigenvector localization for general characteristic matrices describing network dynamics.

In the four examples studied in the previous section, phenomena in a network were captured by the eigenvalues of different types of characteristic matrices. Here, I imagine a network phenomenon described by a matrix $J$. Regardless of the physical interpretation of this matrix (e.g. as Jacobian, adjacency matrix or Laplacian), this matrix $J$ can be interpreted as a weight matrix describing directed edges in an abstract network, such that each element $J_{ij}$ indicates the weight of one edge from a node $j$ to a node $i$ in this abstract network. Further, when considering an eigenvalue $z$ of this matrix $J$, then the eigenvalue equation $zv_i = \sum_j J_{ij}v_j$ relates the entry of its corresponding (right) eigenvector $v$ on a specific node $i$ to those of its neighbors $j$ in the abstract network. Thus, the abstract network gives an intuitive understanding of the interactions al-
5.3 Localization of extreme eigenvalues

allowing an eigenvector entry to contribute to the entries of neighboring nodes. In other words, it allows to study to which degree the eigenvector is spread out over the network and therefore describes a delocalized phenomenon in the original system.

In the following, I use the network of the matrix $\mathbf{J}$ to investigate the localization of eigenvectors corresponding to its extreme eigenvalues. Explicitly, I denote the largest eigenvalue of the characteristic matrix $\mathbf{J}$ as $z$ and I call the corresponding eigenvector $\mathbf{v}$. Further, I denote the node in the network for which $\mathbf{v}$ has the largest absolute value $q$, such that this largest entry is $v_q$. Informally speaking, I interpret the node $q$ as the origin of the eigenvector to study how this eigenvector spreads throughout the network of $\mathbf{J}$.

Following roughly the developments introduced in Ref. 39 for Laplacian matrices of a special type of undirected networks, I separate the properties of node $q$ from those of the remaining network. Explicitly, I remove the row $q$ and column $q$ from $\mathbf{J}$, and denote the resulting matrix $\tilde{\mathbf{J}}$. $\tilde{\mathbf{J}}$ then describes the network obtained by removing node $q$ from the network of $\mathbf{J}$ with all incoming and outgoing edges. Analogously, I define $\tilde{\mathbf{v}}$ as the vector obtained when removing $v_q$ from $\mathbf{v}$. Finally, to capture the properties of node $q$, I define $a$ as the $q$th column of $\mathbf{J}$ without the diagonal element $J_{qq}$, i.e. $a$ is the vector containing the weights of all edges originating in node $q$ in the network of $\mathbf{J}$. Thus, $a$ and $v_q$ describe the quantities related to node $q$, while $\tilde{\mathbf{J}}$ and $\tilde{\mathbf{v}}$ describe the remaining network.

To study the localization of the eigenvector $\mathbf{v}$ corresponding to an extreme eigenvalue $z$, I now establish a relation between its largest absolute entry ($v_q$) and its remaining entries ($\tilde{\mathbf{v}}$). For this purpose, I start from the eigenvalue equation $\mathbf{J}\mathbf{v} = z\mathbf{v}$, and substitute $\mathbf{J}$ by $a$ and $\tilde{\mathbf{J}}$. Dropping the $q$th line of these equations, this leads to $\tilde{\mathbf{J}}\tilde{\mathbf{v}} + a v_q = z \tilde{\mathbf{v}}$. Rewriting this equation, I obtain the required relation

$$\tilde{\mathbf{v}} = X a v_q \text{ with } X = (z\mathbf{1} - \tilde{\mathbf{J}})^{-1},$$

(5.2)

where $\mathbf{1}$ is the identity matrix, and where the superscript $-1$ denotes the matrix inverse.

Intuitively speaking, (5.2) describes how the amplitude of the eigenvector $\mathbf{v}$ on node $q$ spreads through the abstract network of $\mathbf{J}$. The entries of $a$ indicate which nodes of the network have incoming edges from node $q$, such that $v_q$ directly contributes to their entries, while the matrix $X = (z\mathbf{1} - \tilde{\mathbf{J}})^{-1}$ characterizes the further spreading through the network of $\tilde{\mathbf{J}}$. In more mathematical terms, $X$ is the Green’s Function of $\tilde{\mathbf{J}}$. This function is commonly used to describe propagation phenomena (see for example Ref. 181 for its application to Anderson localization).
Now, I ask what proportion of the eigenvector $v$ is localized on node $q$ under the assumption that the network, and the edges from $q$ towards this network are known. Because $v$ is normalized, i.e. $v \cdot v = 1$, this proportion is the amplitude $|v_q|^2 = 1 - \tilde{v} \cdot \tilde{v}$. Intuitively, the stronger $v$ spreads into the network, the smaller the amplitude $|v_q|^2$ becomes. Substituting (5.2) for $\tilde{v}$ and solving for $|v_q|^2$, I obtain the localized proportion

$$|v_q|^2 = \frac{1}{1 + (Xa) \cdot (Xa)}.$$  \hspace{1cm} (5.3)

For completeness, I remark that analogous derivations apply also to the left eigenvector corresponding to an extreme eigenvalue $z$. For the localized proportion of this left eigenvector, $|w_q|^2$, the developments above lead to an equation with the same structure as (5.3). In this equation $Xa$ is replaced by $X^T b$, where $X^T$ is the transposed matrix of $X$, and where $b$ is the vector containing the weights of all edges ending in the node $q$ in the network of $J$.

In summary, for a given extreme eigenvalue $z$, the relation (5.3) expresses the localized proportion $|v_q|^2$ of the corresponding eigenvector on a node $q$ to the interactions of this node with its neighbors (given by $a$), and to the properties of the remaining system (given by $X$). The mathematical challenge that needs to be addressed is thus to compute the system-dependent expression $S(z) = (Xa) \cdot (Xa)$ which can be interpreted as the non-localized proportion of the eigenvector. I note that a generalization of this description to study the localization of eigenvectors on structures consisting of multiple nodes is straightforward (see App. C).

For illustration, I write $S(z)$ in terms of dynamical modes of $\tilde{J}$. If the eigenvalues of $\tilde{J}$ are denoted $\lambda_k$, and if $\tilde{v}^{(k)}$ and $\tilde{w}^{(k)}$ are the corresponding right and left eigenvectors, then the eigenvalues of $X$ are $\mu_k = (z - \lambda_k)^{-1}$ and correspond to the same eigenvectors as $\lambda_k$. Thus, in terms of dynamical modes, $Xa = \sum_k C^{(k)} \cdot \frac{\tilde{w}^{(k)} \cdot a}{|\tilde{w}^{(k)} \cdot \tilde{v}^{(k)}|}$, where $C^{(k)} = (\tilde{w}^{(k)} \cdot a) \tilde{v}^{(k)}$.

$$Xa = \sum_k C^{(k)} \cdot \frac{\tilde{w}^{(k)} \cdot a}{|\tilde{w}^{(k)} \cdot \tilde{v}^{(k)}|} \hspace{1cm} (5.4)$$

is a system-dependent constant vector that is independent of $z$.

In the formulation using dynamical modes, the contribution of a dynamical mode $k$ to the sum resulting in $Xa$ increases the closer the eigenvalue $z$ is to the $\lambda_k$. One can therefore bound the non-localized proportion of the eigenvector.
vector of $z$ by $S(z) < |\sum_k C^{(k)}|^2/\Delta^2 = C \sigma_a^2/\Delta^2$, where $\Delta = |z - \lambda|$ is the minimal distance between $z$ and any eigenvalue $\lambda_k$, where $\sigma_a$ denotes the norm of $a$, and where $C$ is a system-dependent constant (typically $C$ is of order one\(^2\)).

The spectral formulation of $Xa$ for extreme eigenvalues shows that the localized contribution of the corresponding eigenvector depends mainly on two properties. First, how extreme the eigenvalue $z$ is in comparison to the eigenvalues encountered in the remaining system (its distance $\Delta$ from its closest neighbor in the spectrum of $\tilde{J}$), and second, the strength of the connection from the node $q$ towards the remaining network (the norm $\sigma_a$ of the vector of the outgoing edge weights). The more extreme the eigenvalue, and the weaker the connection of the associated node to the remaining network, the more localized is the associated eigenvector. Thus, these results support the hypothesis that, in general, problems in networks are localized if the corresponding extreme eigenvalues belong to regions of low eigenvalue density.

In principle, if the network of $\tilde{J}$ is known, one can establish an explicit result for the localized proportion by developing $S(z)$ as a series in $\tilde{J}$. Stopping this series after order zero in this series means to consider only neighbors of node $q$, order 1 means to consider also these neighbor’s neighbors, etc., until the entire network is included in the computation. This series is analogous to expansions in terms of Greens functions used in quantum mechanics, and one might therefore attempt to tackle this expansions with methods from this field, such as path diagram formulations. However, in practice, these methods become difficult when considering network matrices because these matrices are non-hermitian and because the complex structure makes it difficult to find universal formulations for terms in the expansions [182–184].

As a tractable approximation for the localized proportion of the vector $v$ on a node $q$, I use in the following a formulation that corresponds to an expansion of order 1 in terms of the Green’s function, while capturing all higher orders in a mean-field description. More precisely, I use $a$ to describe the spreading of $v$ from node $q$ to $q$’s neighbors, and describe the spreading through the remaining network using results on the typical spectrum of the matrix $\tilde{J}$.

### 5.4 Dependence on the system size

With increasing size of a system the number of eigenvalues increases. Therefore, if the overall density of eigenvalues remains unchanged through such an increase, then the average difference between eigenvalues decreases. Since, in

---

\(^2\) $C$ depends on the non-orthogonality of the different eigenvectors $v^{(k)}$. If the eigenvectors are orthogonal (as for a normal matrix) or oriented randomly, then $C \approx 1$, while in the case where all eigenvectors point in a similar direction, $C$ can be as large as the system size. In the inhomogeneous network systems considered here, the vectors typically point in different directions, such that $C$ is of the order of 1.

---

91
the simple approximation of (5.4) above, the spread of an eigenvector corresponding to an extreme eigenvalue \( z \) is inversely proportional to a distance between eigenvalues (\( \Delta \)), one might expect that the spread of such an eigenvector increases with system size. Here, I show that this is true for delocalized eigenvectors, but that it is not the case for localized ones.

To illustrate the effect of increasing the system size on the spread of eigenvectors, I first study this dependence for the examples introduced in Sec. 5.2. In particular, I study the size dependence for all examples except the bioinvasion model, for which the network size is fixed by the real-world data. For the three remaining examples I compare the spread of the eigenvectors of their characteristic matrices with the spread of those found for networks that have half their original size. To approximately retain the overall eigenvalue density, each smaller network is generated with the same average node degree as the larger one.

A change in system size reflects differently in the spread of localized and delocalized eigenvectors. The spread of the delocalized eigenvectors, corresponding to eigenvalues in regions of high eigenvalue density in the center of the spectrum, increases with the system size (top row of Fig. 5.7). However, the spread of localized eigenvectors corresponding to the extreme eigenvalues in regions of low eigenvalue density at the borders of the spectrum remains practically unchanged (bottom row of Fig. 5.7).

The observation of localized and delocalized eigenvectors can be explained intuitively in a process that merges two similar networks into a large one by adding new edges. In this process, the additional edges connecting the two systems allows the percolation of delocalized eigenvectors through both systems, explaining why their spread doubles. At the same time, a localized eigenvector on one initial system will normally remain localized also in the final system. Thus, when increasing system size, delocalized eigenvectors percolate through the entire system, but localized vectors remain unchanged.

In the mathematical formulation of the delocalized proportion of a localized eigenvector, (5.4), the persistence of localization when increasing system size is a consequence of the product \( \tilde{w}^{(k)} \cdot a \) that weights the contribution of a dynamical mode \( k \) to the spreading of the eigenvector of \( z \). If an eigenvalue \( \lambda_k \) is close to an extreme eigenvalue \( z \), then it belongs to a region of low eigenvalue density such that \( \tilde{w}^{(k)} \) is localized. For instance, consider that \( \tilde{w}^{(k)} \) has large entries on \( s_w \) nodes, while vector \( a \) has \( s_a \) large entries. In a large system of size \( N \), the expected number of nodes on which both vectors are large scales then as \( s_a s_w / N \), reflecting that the two vectors are increasingly likely to belong to different parts of a system. This scaling in the nominator in (5.4) cancels the scaling of \( \Delta \propto N^{-1} \) in the denominator, such that increasing the system size does not lead to a delocalization of the eigenvector of \( z \).
5.4 Dependence on the system size

Fig. 5.7: Dependence of the spread on the network size. Each column shows the spread of eigenvectors in one example system (c.f. Sec. 5.1): the SIS-model (a), food webs (b), and coupled oscillators (c). The lower and upper panels show the same curves for different ranges of spread. The solid curves show the spread of right (blue) and left (yellow, only food web) eigenvectors of the large example networks in Fig. 5.6. The dashed curves (red, green) show the spread for networks with identical degree, but half the system size. For eigenvalues in the center of each spectrum, the spread scales with the system size (upper panel), but for extreme eigenvalues, it remains practically unchanged (lower panel). For visibility, the curves in a) and c) show spread averaged over 5 consecutive eigenvalues.

In summary, the spread of localized eigenvectors with extreme eigenvalues is independent from the size of the embedding system. When combining two similar systems into a larger one, then nodes that carried a localized eigenvector in one of the original systems will generally do so also in the combined one. Thus, increasing the system size increases the number of localized eigenvectors, but does not increase their spread.
5.5 Localization in the model of R. May

Above, I use a basic approximation to study the overall behavior of the non-localized proportion $S(z)$ of an eigenvector corresponding to an extreme eigenvalue $z$. To illustrate the properties of the non-local proportion in a more specific system, I now establish $S(z)$ for the paradigmatic model for complex systems introduced by R. May, which started the complexity-stability debate [122, 130].

In his seminal paper [44], May proposed to interpret sparse random matrices as prototypes for Jacobian matrices describing the dynamics of a complex system of the form $\frac{dx_i}{dt} = F_i(x) - x_i$. Explicitly, he proposed to establish these Jacobian matrices $J$ as follows: the diagonal elements are set to $J_{ii} = -\frac{\partial x_i}{\partial x_i} = -1$. Second, each off-diagonal entry is nonzero with probability $C$ (the probability that two variables affect each other). The nonzero entries are drawn from a normal distribution of width $\sigma$ indicating the typical strength of interactions. Thus, the Jacobian matrix $J$ in May’s model is $J = -1 + G$, where $G$ is a sparse Gaussian random matrix with diagonal entries set to zero. The defining properties of $G$ are its dimension (or size of the system) $N$, the connectance (or interaction probability) $C$, and the typical strength of interactions $\sigma$.

The probability of randomly finding a matrix $J$ that describes a stable system quickly decreases with the system complexity (measured as $NC$) [129]. For stability, all eigenvalues of $J$ must be smaller than zero, i.e. those of $G$ must be smaller than 1 (by real part). However, the eigenvalues of a Gaussian random matrix, such as $G$, are uniformly distributed inside a disc of radius $\sqrt{CN}\sigma$ around the origin of the complex plane [184, 185]. For large systems, the disc is densely filled with eigenvalues, such that one obtains almost surely an eigenvalue of $G$ larger than one if $\sqrt{CN}\sigma > 1$ [129].

Now, I compute the non-localized proportion of the eigenvector corresponding to an extreme eigenvalue $z$ outside the disc of typical eigenvalues. For this purpose, I use $\sigma^2 = 1/(NC)$, such that the eigenvalues of $G$ are distributed inside a unit disc around the origin of the complex plane. Further, for simplicity, I consider that the extreme eigenvalue $z$ is real\(^3\).

The non-localized proportion is defined as $S(z) = \mathbf{X}a \cdot \mathbf{X}a = \mathbf{a}^T \mathbf{X}^T \mathbf{X} \mathbf{a}$, where $\mathbf{X}^T$ denotes the transposed matrix of $\mathbf{X}$. This is identical to summing the terms $(\mathbf{X}^T \mathbf{X})_{ij} a_i a_j$, with $i$ and $j$ independently running over all $N$ nodes of the network. In the May Model, each interaction (entry of $J$) is drawn independently, such that the matrix $\mathbf{X}^T \mathbf{X}$, and each of the entries of $\mathbf{a}$

\(^3\)The system is invariant against rotation around $-1 + 0i$, such that results for complex $z$ only require a rotation before proceeding.
are independent. Except when \( i = j \), the products \( a_i a_j \) are zero on average, such that for a large system, the terms with \( i \neq j \) cancel, resulting in \( S(z) \approx \sum_i (X^T X)_{ii} a_i^2 \). Further, because \( a \) and \( X \) are independent, in a large system \( S(z) \approx \sigma_a^2 \text{Tr} \{ X^T X \} / N \), where \( \text{Tr} \) denotes the trace (the sum over all diagonal entries of a matrix), and \( \sigma_a \) the norm of \( a \). Thus, \( S(z) \) is the average over the eigenvalues of \( X^T X \), also called the squared singular values of \( X \), multiplied by the typical weight of edges in \( a \).

To obtain an explicit expression for \( S(z) \), I require an expression for the average over the singular values of \( X \) based on knowledge of the spectrum of \( J \). However, for an asymmetric network matrix \( J \), finding this average is difficult (see Ref. 186 for a review of current methods). The difficulty resides in the non-orthogonal eigenvectors of \( J \), because products of such eigenvectors appear when writing the product of \( X^T X \) which defines the singular values.

To obtain an approximation for \( S(z) \), I proceed as if the eigenvectors of \( J \) were orthogonal. Under this assumption, the singular values of \( X \) are the absolute values of the eigenvalues \( 1/(z - \lambda_k) \) of \( X \), where \( \lambda_k \) are the eigenvalues of \( J \). Therefore, \( S(z) \approx \sigma_a^2 / N \sum_k 1/|z - \lambda_k|^2 \). To evaluate the sum, I integrate over a constant eigenvalue density inside the disc \( \bigcirc \) of radius 1 around \(-1 + 0i\). This leads to the approximation

\[
S_o(z) = \sigma_a^2 \int_0^1 \int_0^{2\pi} \frac{r \, d\phi \, dr}{\pi |z - \lambda|^2} = \sigma_a^2 \int_0^1 \int_0^{2\pi} \frac{r \, d\phi \, dr}{\pi |z + 1 - re^{i\phi}|^2}
= \sigma_a^2 \int_0^1 \frac{2r \, dr}{(z + 1)^2 - r^2} = \sigma_a^2 \ln \left( \frac{(z + 1)^2}{z^2 + 2z} \right).
\] (5.5)

\( S_o(z) \) tends to infinity for \( z = 0 \) (where \( z \) enters the disc of the densely distributed eigenvalues \( \lambda \)), and to \( 1/z^2 + O(1/z^3) \) for large \( z \).

The localized proportion \( |v_q|^2 \) obtained by plugging \( S_o(z) \) into (5.3) agrees well with numerical evaluations of sample systems with finite size (Fig. 5.8a). However, it overestimates the localized proportion if \( z \) is close to 0. This may be partially due to the slight concentration of eigenvalues of \( J \) on the real axis in the finite systems studied numerically (see Fig. 5.8c), an effect known as Girkos law [47, 187]. Approximating \( S(z) \) based on a constant density in the interval \([-2, 0]\) on the real axis, I find \( S(z)_- \approx (z^2 + 2z)^{-1} \), which leads to a better approximation for small \( z \) (dashed line in Fig. 5.8).

For comparison with the results in terms of the spread of eigenvectors, I use an approximation for spread based on the localized proportion \( |v_q|^2 \). For this purpose, I consider that the system is tree-like near the node \( q \), i.e. I omit small network cycles. Using the eigenvalue equation, the average amplitude of \( v \) on a neighbor \( i \) of \( q \) is \( |v_i|^2 = |v_q|^2 p \), where \( p = \sigma^2 / (z + 1)^2 \), the average amplitude on neighbors \( j \) of nodes \( i \) is \( |v_j|^2 = |v_i|^2 p \), etc. Thus, if \( N_d \) denotes the
Approximate Localization

Fig. 5.8: Localization of eigenvectors corresponding to extreme eigenvalues $z$ in the model proposed by May [44]. Panel a) shows the localized proportion and panel b) the spread of an eigenvector of the Jacobian matrix, depending on the value of the corresponding eigenvalue $z$. The data-points represent numerical computation performed on matrices of size $N = 1000$ (blue □) and 250 (red △) and with an average interaction probability $C = 0.1$. The analytical approximations (lines) are based on the typical eigenvalue density of the Jacobian matrices. For the solid line, I consider a uniform density of the eigenvalues $\lambda$ inside the complex unit disc around $-1$ (circular law) [47]. For the dashed line, I consider a constant eigenvalue density on the real axis between 0 and $-2$ (Girko’s law) [187]. The empirical distribution of the eigenvalue density ($10^6$ samples with $N = 1000$) is color-coded in panel c).

number of nodes that can be reached by passing exactly $d$ edges after starting from node $q$, the spread is $\left(\sum_i v_i^4\right)^{-1} = |v_q|^{-4} (\sum_d N_d p^{2d})^{-1}$. In a tree-like graph, $NC$ indicates the average number of neighbors of each node, such that $N_d \approx (NC)^d$. Solving the geometric series over $d$ leads to the required approximation of spread. The results of this approximation agree well with the numerical computations (Fig. 5.8b).
In conclusion, in this paradigmatic model, the spread of an eigenvector quickly decreases for extreme eigenvalues when the distance of the eigenvalue to the disc of typical eigenvalues increases. The fast transition from the delocalized eigenvector regime to the localized regime is explained by the sharp boundary between the high density of eigenvalues on the disc associated to delocalized eigenvectors, and the practically zero density outside it (Fig. 5.8c). While I consider here the localization of an eigenvector on a single node, the same approach extends to the localization of extreme eigenvalues on larger structures (App. C).

5.6 Finding motifs that carry localization

In the previous chapter, I showed how to identify network motifs, such as symmetries, that carry exactly localized eigenvectors. Analogously, I present here how to identify nodes or structures which carry approximate localization.

As in the previous section, I start by considering localization of an eigenvector centered around one node (see App. C for the generalization to larger structures). Generalizing the methods presented in Ref. 39 for localization of eigenvectors of specific network Laplacian matrices, I identify which properties allow a specific node \( q \) to be such a center of localization.

Because I want to identify specific properties of a particular node \( q \) in the abstract network of the characteristic matrix \( J \), I describe the properties of this node separately from those of the remaining system. As above, I denote by \( \tilde{J} \) the matrix containing the weights of the network obtained by removing node \( q \) from the network of \( J \), by \( a \) the vector containing all the edges originating in node \( q \), and denote by \( \tilde{v} \) the vector containing all entries of \( v \) except \( v_q \). In addition to the notation used previously, I now denote by \( b \) the vector containing all the edges ending in node \( q \), and by \( e = J_{qq} \) the diagonal entry of node \( q \) in \( J \).

Writing the eigenvalue equation \( Jv = zv \) in terms of \( \tilde{J}, a, b, e \) leads to two equations. First, I obtain the (vector-)equation for the entries of \( \tilde{v} \) which I used above to describe how node \( q \) contributes to the entries of \( v \) on the remaining network. Second, I obtain the equation \( b \cdot \tilde{v} = (z - e)v_q \), which describes how the remaining network contributes to the entry \( v_q \). Intuitively speaking, an eigenvector with eigenvalue \( z \) can only localize on \( q \) if the contributions from \( v_q \) to the network and the effect from the network to \( v_q \) are consistent. To study this consistency, I combine the two equation (I replace \( \tilde{v} \) by (5.2)). This results in \( (b \cdot Xa)v_q = (z - e)v_q \), where \( X = (1z - \tilde{J})^{-1} \). In the following, I call \( b \cdot Xa = (z - e) \) the compatibility condition because its left-hand-side
must be compatible with its right-hand-side to allow the eigenvalue \( z \) to exist in the spectrum of the matrix \( \tilde{J} \) and to allow localization of the corresponding eigenvector on node \( q \).

From the converse perspective, choosing a value for \( z \) for a known system, that is described by a matrix \( \tilde{J} \), fully defines \( X \). Thus, I can use the compatibility condition to find those values of \( a, b \) and \( e \) that allow localization of an eigenvector \( v \) on node \( q \). In other words, I can use the compatibility condition to identify the properties of a node \( q \) that make it compatible with the localization of the eigenvector of a given eigenvalue \( z \).

To illustrate the application of the compatibility condition to find the properties of nodes that carry localization, I study the model proposed by R. May [44]. I start by approximating the left-hand side, \( b \cdot Xa \). For large systems, I obtain

\[
b \cdot Xa = \sum_{ij} b_j a_i X_{ji} \approx \mu(z) (b \cdot a),
\]

where \( \mu(z) = \operatorname{Tr}\{X\}/N \) is the average diagonal entry of \( X \). For this approximation, I used that in the model \( X \) is independent from the entries of \( a \) and \( b \), and that the entries \( a_i \) and \( b_j \) referring to different nodes \( i \) and \( j \) are independent. Thus, this formulation separates the left-hand-side into the product \( b \cdot a \), which depends only on the edges originating and ending in the node \( q \), and into \( \mu(z) \) which depends on \( z \) and on the remaining system.

In principle, \( b \cdot a \) tends to zero for May’s model because the entries \( a_i \) and \( b_i \) are independent random variables, also when referring to the same node. However, I retain the product to study structures with correlated edges ending in node \( q \), and originating in \( q \). Explicitly, I write \( b \cdot a = \sigma_a \sigma_b \mathcal{R}, \) where \( \sigma_a \) and \( \sigma_b \) are the norms of the respective vectors, and where \( \mathcal{R} \) indicates the degree of their correlation. Explicitly, \( \mathcal{R} = 1 \) if edges originating in \( q \) are reciprocated by edges of the same weight, \( \mathcal{R} = -1 \) if edges originating in \( q \) are reciprocated by edges with the same weight but opposite sign, and \( \mathcal{R} = 0 \) if edges originating and ending in \( q \) are not correlated. Rewriting the compatibility condition in terms of \( \sigma_a \sigma_b \mathcal{R} \), I obtain

\[
\mu(z) \sigma_a \sigma_b \mathcal{R} = (z - e).
\]

It remains to find \( \mu(z) \), the average over the eigenvalues of \( X \). Explicitly, the individual eigenvalues of \( X \) are \( 1/(z - \lambda_k) \), where \( \lambda_k \) are the eigenvalues of the matrix \( \tilde{J} \) (see Sec. 5.5). Furthermore, for a model with a typical interaction strength \( \sigma^2 = (NC)^{-1} \), the eigenvalues \( \lambda_k \) of \( \tilde{J} \) are distributed uniformly inside a disc \( \mathbb{D} \) of radius 1 centered around \( -1 + 0i \) in the complex plane (see Fig. 5.8) [44, 47]. To approximate \( \mu(z) \) for large systems, I replace the average over the eigenvalues of \( X \) by an integral over \( \mathbb{D} \), leading to

\[
\mu(z) \approx \int_{\mathbb{D}} \frac{d\lambda}{\pi} \frac{1}{z - \lambda} = \int_0^1 dr \int_0^{2\pi} \frac{r d\phi}{\pi} \frac{1}{z + 1 - re^{i\phi}} = e^{-\theta i} \frac{e^{-\theta i}}{z + 1}, \tag{5.6}
\]
5.6 Finding motifs that carry localization

Fig. 5.9: Node properties and compatible extreme eigenvalues $z$ for May’s model. If a node is described by a diagonal entry $e$, and its interaction with the remaining network by $\sigma_a\sigma_b R$, then an eigenvector with eigenvalue $z$ localizes on it if the compatibility condition, $\mu(z) = (z - e)/(R\sigma_a\sigma_b)$, is satisfied. The solid black curve corresponds to the left-hand side of this equation $\mu(z)$. The dashed curves correspond to the right-hand with different values for $e$ (0: blue, 1.5: green), and $R\sigma_a\sigma_b$ (0: dashed, 1: dash-dotted, −1: dash-dot-dotted). Crossings between the solid black curve and a dashed curve mark situations where the eigenvector corresponding to the given eigenvalue $z$ localizes on the node characterized by $e$ and $\sigma_a\sigma_b R$. Data points represent numerical computations of $\mu(z)$ in matrices of size $N = 1000$, connectance $C = 0.1$, where $\circ$: $R\sigma_a\sigma_b = −1$, $\square$: $R\sigma_a\sigma_b = 0$, $\triangle$: $R\sigma_a\sigma_b = 1$, blue: $e = 0$, green: $e = 1.5$.

where $\theta$ is the angle between a line from $z$ towards $-1$ and the real axis. For real values of $z$, $\theta = 0$, such that $\mu(z) = 1/(1 + z)$ is simply the inverse distance of $z$ from the center of the disc $\circ$. This completes the consideration of the left-hand side of the compatibility condition.

Now, I use the compatibility condition to study the properties of a node $q$ that make this node compatible with a specific extreme eigenvalue $z$. In other words, I study for which values of $\sigma_a\sigma_b R$ and $e$ the two sides of $\mu(z)\sigma_a\sigma_b R \approx (z - e)$ are consistent.

Explicitly, the compatibility condition is satisfied for the eigenvalues $z_{\pm} = 1/2(e - 1 \pm \sqrt{(e + 1)^2 + 4R\sigma_a\sigma_b})$, if these values satisfy $z_{\pm} > 0$. To get a more intuitive understanding of these solutions, I study the compatibility condition graphically. For this purpose, I divide the equation by $\sigma_a\sigma_b R$, and study the
crossing of the function $\mu(z)$ (solid line in Fig. 5.9) with those of the linear function $(z - e)/(\sigma_a \sigma_b R)$ (dashed and dash-dotted lines in Fig. 5.9) for different values of $R$ and $e$.

For $R = 1$, i.e. for a positive correlation between the edges ending in $q$ and those originating in $q$, there is a unique solution $z_+ > 0$ if $e > -\sigma_a \sigma_b$ (data points marked $\triangle$ in Fig. 5.9). If $q$ is a typical node of the model ($e = -1$, $\sigma_a \sigma_b = 1$), then the crossing occurs on the border of the disc of typical eigenvalues at $z_+ = 0$. An eigenvalue $z_+ > 0$ with a localized eigenvector therefore requires a node with an unusual diagonal entry $e > -1$, or with unusually strong correlated interactions with other nodes $\sigma_a \sigma_b > 1$.

For $R = -1$, i.e. for a negative correlation between the edges ending in $q$ and those originating in $q$, there is one solution if $e > \sigma_a \sigma_b > 1$, or two solutions if $e > \sigma_a \sigma_b \leq 1$ (data points marked $\circ$ in Fig. 5.9). A solution is therefore not possible for typical nodes of the model proposed by May, but requires a node with an unusually large diagonal entry.

Finally, for $R = 0$, i.e. for uncorrelated edges ending and originating in $q$, there are two identical solutions $z_+ = z_- = e$ if $e > 0$ (data points marked $\square$ in Fig. 5.9). Also in this case, extreme eigenvalues are not possible for typical nodes in the original model proposed by May. Intuitively speaking, the missing correlation decouples node $q$ from the remaining system.

In summary, the method presented here allows to identify the properties of a node that lead to the localization of an extreme eigenvalue on this node. The difference between the diagonal entry $e$ of this node in the characteristic matrix and a compatible eigenvalue $z$ is limited by the correlated proportion of the node’s interactions with other nodes $\sigma_a \sigma_b R$. For the model proposed by May, $\sigma_a \sigma_b R \approx 0$ because of the absence of correlations. Thus, this model is an extreme example in which localization is generally possible only if $z \approx e$. Since typically $e = -1$, extreme eigenvalues (outside the disc of radius 1 around $-1$) are rare. This explains the sharp transition from localized to delocalized eigenvectors when $z$ approaches the imaginary axis (Fig. 5.8).

While I focus in this section on the properties of single nodes that allow them to carry a localized eigenvector, it is straight-forward to generalize the approach to structures consisting of multiple nodes (see App. C). This allows, for instance, to identify structures leading to oscillatory instabilities due to pairs of localized eigenvectors with complex conjugate eigenvalues, and it allows to relate this approach to the description of exact localization in symmetries (see App. C).

When studying large networks with a given dynamics, then the procedure applied here allows to establish small structures that carry the localized eigenvectors of extreme eigenvalues $z$. Thus it allows to identify motifs that are
centers of localized instabilities. Analogous to the motifs leading to exact localization, one can study them independently from the embedding network, allowing to understand the associated instabilities regardless of network size.

5.7 (Self-)stabilization of food webs

The results in Sec. 5.3 suggest that the most prominent instabilities of food webs are localized. In this light, it is instructive to revisit the discrepancy between results suggesting that the probability of randomly drawing a stable food web configuration decreases exponentially with its complexity, and the apparent stability of complex food webs observed in nature [44, 92, 130].

The localization of the prominent instabilities suggests that a stabilization of an unstable system may be possible through the extinction of only a few species. One can therefore imagine that the stability of observed food webs results from an evolutionary process that removes small problematic parts from an existing large system, thereby driving towards a stable configuration with a minimal loss of complexity. Here, I study this hypothesis by explicitly exploring how an unstable food web stabilizes.

Instabilities in a food web can lead to oscillations or strong changes in species abundances, followed by extinctions [85, 96, 107]. Therefore, the dynamics following an instability cannot be analyzed through a generalized model which describes the neighborhood of a steady state of coexistence. Instead, numerical integration of this dynamics requires to chose explicit functional forms that prescribe the species’ dependencies throughout the space of possible states.

To study the dynamics of model food webs explicitly, I use the following specific model functions: the dependence of primary production and mortality of a species are modeled as power laws of the abundance of the species. Explicitly, in terms of the functions and parameters used in the generalized model (Sec. 2.7), I use 

\[ S_i(X_i) \propto X_i^\phi \]

for primary production \((S_i = 0\) for predators), and

\[ M_i(X_i) \propto X_i^\mu \]

to model mortality, where the exponents \(\phi\) and \(\mu\) are constant parameters that correspond directly to the sensitivity parameters used in the generalized model (c.f Tab. 2.1).

For the feeding interactions, I use the well-established Holling type-II functional response [127]. Explicitly, if species \(i\) predates on species \(j\), then the predation gain \(G_{ij}\) of species \(i\), and the predation loss \(L_{ij}\) of species \(j\) are

\[ G_{ij}(X) \propto L_{ij}(X) \propto \frac{X_j}{T_i(X) + A_i} X_i^\psi, \]
where \( T_i(X) \) is the sum over all prey available to predator \( i \), where the parameter \( A_i \) indicates the initial saturation in prey of the predator, and where \( \psi \) corresponds to the sensitivity of predation in the generalized model. Intuitively, these functions can be understood as the probability of a predator meeting a prey individual (the numerator) multiplied with the probability that the predator kills it (the denominator) [127]. For low prey abundance \( T_i \), the predation gain and loss is linear in the prey abundance (the predator kills every prey individual it encounters), but for high abundances \( T_i \gtrsim A_i \), they saturate (the predator kills as many individuals as it can process). The complete dynamics of a species is \( \frac{dX_i}{dt} = S_i - M_i + \sum_j G_{ij} - \sum_k L_{ki} \), where \( k \) runs over the species predating on \( i \), and \( j \) runs over the species that are prey to \( i \).

I choose the parameters and the proportionality constants in the specific model in accordance with the generalized model description used throughout my thesis. This is possible through the following procedure [27, 122]: for a given food web topology, I draw generalized parameters. These parameters describe a (generally unstable) steady state of the food web. Using this state as the initial configuration for numerical simulations, it prescribes the parameters and proportionality constants of the interaction functions in the specific model. Explicitly, the generalized parameters \( \phi, \mu, \psi \) directly translate into the exponent parameters of the specific model, and the value of \( A_i \) results from solving \( \frac{\partial G_i}{\partial T_i} = \gamma \) at the initial state. The proportionality constants denoting the initial biomass flows are given by the turnover rates (\( \alpha \)) and the dietary contributions (\( \beta, \chi \)).

To study a process in which a food web stabilizes, I apply the specific model to a niche model topology in a slightly unstable initial configuration (I discard stable food webs and food webs whose largest Jacobian eigenvalue is larger than 0.05), and numerically integrate its dynamics.

The abundances of species throughout a process stabilizing one example food web are shown in Fig. 5.10a). Initially, all species coexist at a (normalized) abundance \( X_i = 1 \). Because the initial state is unstable, the dynamics drives the system away from this state. After the extinction of one species (at \( \times \), where the red curve drops below \( 10^{-6} \)), the system settles into a new steady state.

To study how the food web dynamics stabilizes, I trace the eigenvalues of the Jacobian matrix over time together with their localization on the species becoming extinct (Fig. 5.10b,c). It is important to note that outside a steady state, these eigenvalues do not denote the stability of the system, but the stability of the current dynamics against perturbations. Intuitively, a large eigenvalue indicates that the system is brittle because small changes of the system may lead

\[ \text{The integration was performed with the Dormand-Prince algorithm (DOPRI45) [188], and with Modified Extended Backward Differentiation Formulae (MEBDF) [189], using an implementation by M. Morvai. Both methods lead to the same results.} \]
Fig. 5.10: Stabilization of food webs through local brittleness. Panel a) represents the normalized abundances of species in an example food web throughout a stabilization process, in which one species goes extinct (the species is removed at × where its abundance drops below $10^{-6}$). Panels b) and c) represent the eigenvalues of the Jacobian matrix over time, the color-code indicating the localization of the corresponding right (b) and left (c) eigenvectors on the species becoming extinct. The point where the largest eigenvalue is maximal, is marked by △. After reaching this point, the eigenvalue drops and disappears in the extinction (at □). Panel d) summarizes the localization for the maxima of the largest eigenvalue preceding extinctions in $10^5$ model food webs. Red data points mark the combined localization of the eigenvectors (of the brittleness preceding the extinction) corresponding to this eigenvalue on species becoming extinct, and blue data points mark localization on individual other species. The data points show the localization for 50 webs, the histograms combine the data from all $10^5$. Parameters: $N = 15$, $C = 0.1$. 

\[\text{Abundances} \quad \begin{array}{c}
0 & 1 & 2 \\
0 & 1 & 2 \\
0.0 & 0.5 & 1.0
\end{array} \]

\[\text{Eigenvalues} \quad \begin{array}{c}
0.0 & 0.5 & 1.0 \\
0.0 & 0.5 & 1.0 \\
0.0 & 0.5 & 1.0
\end{array} \]

\[\text{Time} \quad 0 \quad 2000 \quad 4000 \]

\[\text{Loc. on species becoming extinct} \quad \begin{array}{c}
\triangle \text{on surviving species} \\
\Diamond \text{on species becoming extinct}
\end{array} \]
to large changes in the dynamics. The amplitude of the right eigenvector of a large eigenvalue therefore indicates in which part of the system this eigenvalue leads to brittleness (c.f. sensitivity in Sec. 3.4). Further, the amplitude of the left eigenvector indicates which part of a system may cause the dynamical changes that break the brittle ones (c.f. influence in Sec. 3.4).

The maximum of the largest eigenvalue, i.e. the point where the dynamics of the system is most brittle, is marked △ in Fig. 5.10b,c), while the disappearance of an eigenvalue due to the extinction is marked □. These points divide the stabilization process into three phases: Initially, the example food web has one positive eigenvalue and the corresponding unstable modes drives the food web away from the initial state. During this process, the brittleness of the system increases and localizes on the species that goes extinct later. After reaching △, the largest eigenvalue decreases quickly. In the process, the left eigenvector of the eigenvalue becomes increasingly localized on the species going extinct. Finally, after the eigenvalue disappears at □ due to the extinction, all eigenvalues are negative, and the system settles in the new steady state.

One can interpreted the three phases leading to an extinction as phases of a stabilization process: in the first, the instability of the initial state drives the food web towards a state of localized brittleness near the most problematic species. In the second phase, this brittleness leads to strong dynamics that ends in the decoupling and extinction of the species at the center of the instability. Finally, after the extinction, the remaining system finds a stable configuration. Thus, the process shows an example for dynamics that leads to the self-stabilization of the food web through the localized removal of a small problematic part.

To study the (self-)stabilization of food webs through localized brittleness more generally, I investigate the process in \(10^5\) unstable model food webs. In each food web, I determine the points where the largest eigenvalue is maximal and study the localization of the associated brittleness on species that become extinct in the process (Fig. 5.10d). The histograms show that in most networks, a large part of the dynamical mode with the largest eigenvalue, i.e. of the associated brittleness, localizes on the species that becomes extinct in the process, while other species generally carry only small contributions to these eigenvectors. This supports the hypothesis that food webs stabilize by becoming locally brittle prior to an extinction, thereby allowing the break-off of species contributing to the instability, without endangering the remaining system.

In summary, I observe a process through which food webs can stabilize. In this process, an initial instability leads to the increasing localization of a dynamical mode with a large eigenvalue, i.e. to a localization of its associated brittleness. In the subsequent extinction, the process eliminates the instability, while preserving most of the complexity of the remaining web. This mecha-
nism, which stabilizes while retaining most of a food webs complexity, might therefore explain how complex systems in nature achieve their stability despite their complexity.

5.8 Repairing localized instabilities

Above, I show that localized eigenvectors occur for the extreme eigenvalues inside regions of low eigenvalue density at the borders of the high-density regions of typical eigenvalues. However, if the regions of high and zero density are separated by a sharp edge in the eigenvalue density function, then these regions of low density are absent, such that even the extreme eigenvalues occur in regions of high eigenvalue density. From the reasoning above, one can expect for this situation that even the most prominent dynamical modes are relatively delocalized. In the following, I show that such sharp boundaries for spectra may emerge due to attempts to stabilize a system by local interventions, and that these attempts thereby lead to the occurrence of delocalized problems.

As a first example, I study a system of coupled oscillators as a toy model for a power grid coupling machines and generators. In particular, I study the effect of adding or removing couplings between oscillators that are the centers of problems. This mimics attempts to stabilize the operation of a power grid by adding or removing power lines to or from parts in which the origin of a break-down of the global synchrony was detected.

In an oscillator network, instabilities leading to break-down of synchrony are associated with very large or very small eigenvalues of the Laplacian matrix of the underlying coupling network (Sec. 5.2). Attempts to stabilize a specific instability must therefore move the corresponding eigenvalue from a problematic region of extreme eigenvalues towards the stable region at the center of the spectrum by adding or removing couplings. One can implement this, for instance, through the following simple protocol: first, to increase a small problematic eigenvalue, select the node on which the associated instability originated (I choose the one with the largest entry of the corresponding eigenvector), and add a link (build a power line) to a randomly chosen other oscillator. Second, to reduce a very large eigenvalue, select the node on which the associated instability originated, and delete a link (cut a power line) ending in this oscillator. Since adding a link increases the diagonal entry of the Laplacian matrix of the most problematic node, while deleting a link decreases it, an eigenvalue with an eigenvector localizing on this node generally changes accordingly. Thus, this process achieves a stabilization of the associated instability.
Now, I consider a network of oscillators that is subject to a slow-moving risk, which slowly increases the risk of failures, such as a constant reduction of the resources allocated for maintenance of the generators in a power grid. In this context, one will observe a succession of break-ups of synchrony related to the most prominent instabilities of the system. After each break-up, one would therefore repeatedly attempt to stabilize the corresponding instability.

To investigate the consequences of repeated attempts to stabilize instabilities in a network of coupled oscillators, I start with oscillators coupled in an Erdös-Renyi network [172], and iterate the stabilization procedure. In each iteration, I simultaneously attempt to increase the smallest eigenvalue and to decrease the largest eigenvalue of the spectrum, such that total number of couplings between oscillators remains constant.

Iterating the stabilization process leads to a compression of the spectrum of the Laplacian matrix of the initially random coupling network (main panel of Fig. 5.11). Thus, the process leads to a stabilization of the prominent instabilities. However, at the same time, this compression leads to an increase of the overall spread of eigenvalues, in particular, for eigenvalues on the borders of the spectrum. Thus, stabilizing the localized instabilities associated to extreme eigenvalues that are the first to occur in a process of slow-moving risk, leads to a delocalization of instabilities that occur slightly later.

The graphs in the upper row of Fig. 5.11 quantify the trade-off between stability and localization. The gray envelope in these graphs results from the superposition of gray curves corresponding to repeated stabilization attempts in different example networks. Each curve represents the spread of the eigenvector corresponding to the largest or smallest eigenvalue of the spectrum of the network throughout the stabilization process. The solid red line shows the average spread observed when the largest or smallest eigenvalues have the value represented on the horizontal axis. In other words, the red line indicates the expected spread of the most prominent instability as a function of the stability of the system. The graphs show, that already for an eigenvalue of 0.5 or 1.75, leading to a problem with spread 100 in the original system, the spread is typically higher than 200 in a system in which previously occurring instabilities were repaired. Thus, stabilizing the instabilities associated to the localized extreme eigenvalues, leads to a strong delocalization of the remaining ones.

To study in more detail the consequences of attempts to stabilize a system, I return to the SIS-model studied in Sec. 5.2. I recall that in this model, an infection, such as a disease, a computer virus or forest fire, spreads from one infected node in a network to its neighbors with an infection rate $p$, while an infected node recovers spontaneously with a recovery rate $r$. 
5.8 Repairing localized instabilities

Fig. 5.11: Stabilizing a network of oscillators by successive rewiring. The main graph represents the spread of the instabilities leading to the break-up of synchrony in a network of randomly coupled oscillators. Each data set shows the spread of the Laplacian matrix eigenvectors after the given number of rewiring steps were undertaken to increase the stability of the synchronous state by compressing the problematic extreme eigenvalues towards more stable values at the center of the spectrum. For visibility, the spread was averaged over the eigenvectors inside small intervals of eigenvalues. No data point (and no line) for a value on the horizontal axis indicates, that no dynamical mode with this eigenvalue exist, such that extent of the line directly indicates the width of the compressed spectrum. Otherwise, the size of the symbols denotes the number of eigenvalues that fell into each given interval. The graphs in the top row show the spread of the dynamical modes corresponding to the largest and smallest eigenvalues throughout such a repairing process. 25 such trajectories, obtained for different random networks, are shown in grey, the red line indicates the average spread of the dynamical mode with a given eigenvalue in 100 sample networks. Parameters: $N = 1000$ oscillators, on average 10 links per oscillator.
As a first method to stabilize a system against outbreaks of infections, I consider a process in which nodes showing instabilities gain faster access to treatment than others. This models a *reallocaion* of treatment resources following the observation of an infection outbreak allowing to treat a disease, to maintain a computer, or to prevent forest fires. In the SIS-model, this difference in access to treatment resources means that recovery rates are different for different nodes, but that their average remains constant throughout the reallocation.

For a given network, described by an adjacency matrix \( A \), I perform the stabilization through repeated application of the following reallocation procedure: first, I identify the prominent outbreak centers of the infection. These are the nodes with the large entries in the eigenvectors corresponding to the largest eigenvalues of the matrix \( Ap - r \), where \( r \) is a diagonal matrix containing the recovery rates of the different nodes. Then, I slightly increase the recovery rates for these nodes forming the outbreak centers, balanced by a small overall decrease of recovery rates to preserve their average. This reallocation generally achieves a reduction of the largest eigenvalues of \( Ap - r \), and thus increases the stability of the infection-free state.

As an alternative method to stabilize a system against an outbreak, I study a process in which nodes at the center of outbreaks is removed from the network. This *removal* is inspired by the self-stabilization process in food webs observed in the previous section, and can model, for instance, a vaccination of an individual, a firewall installed on a computer. For a given network, I perform this stabilization by removing the node which is the center of the most prominent outbreak (node with largest entry in the eigenvector corresponding to the largest eigenvalue of \( Ap - r \)). This removal generally achieves a stabilization by the removal or reduction of the eigenvalue associated to the outbreak.

In the following, I repeatedly apply the reallocation and the removal processes repeatedly to an identical initial network. This models the application of each process to counter observed infection outbreaks on this network in a process of slow-moving risk caused, for instance, by the slow decrease of overall treatment resources. More precisely, I iterate the two stabilization procedures independently, until the largest eigenvalue of their characteristic matrix drops below \( \frac{3}{4} \) of its initial value.

The consequences of stabilizing a network on the typical size of outbreaks are shown in Fig. 5.12 as a function of the average recovery rate \( r \). In the original system, outbreaks occur for \( r < 0.15 \). As \( r \) decreases further, i.e. as risk increases, the sizes of these outbreak, increases towards 1. Repairing the system through repeated reallocation of resources reduces the threshold for the appearance of an outbreak to an average recovery rate of \( r = 0.11 \) (\( \times \) in Fig. 5.12). However, as \( r \) decreases further, the size of the observed outbreaks increases much faster than in the original system. Repairing the system through removal
5.8 Repairing localized instabilities

![Graph showing infection outbreaks](image)

Fig. 5.12: The size of infection outbreaks in the SIS-model, before and after stabilizing the network. Each curve represents the typical size of infection outbreaks on a random geometric graph (c.f. Sec. 5.2) as a function of the (average) recovery rate. Considering a slow-moving risk (decreasing recovery rates from left to right), the curves show the increasing size of such outbreaks over time for the initial network (blue ○), after the system was stabilized by modification of the recovery rates (resource reallocation; yellow ×), and after the same initial system was stabilized by removing nodes at the center of outbreaks (black +). The networks have \( N = 1000 \) nodes and average degree 10, in the removal process, 55 nodes were removed. The size of outbreaks is computed based on the total number of nodes found infected after performing 250 updates in 100 runs of the same system, each started by randomly infecting 10\% of all nodes.

of nodes at the center of outbreaks (marked + in Fig. 5.12) also reduces the threshold for the appearance of an outbreak to \( r = 0.11 \). As the average recovery rate \( r \) decreases, the size of the observed outbreaks increases slower than for the system repaired through reallocation of resources, and the relative size of outbreaks stays below the size of those observed in the original system.

In summary, the methods for repairing networks presented in this chapter lead to a stabilization of the most prominent instabilities in a network of coupled oscillators, and in the SIS-model for infection spreading. If these systems are subject to a slow-moving risk, they therefore lead to a delayed occurrence of these instabilities compared to the original system. At the same time, however, stabilizing while preserving all elements of a network, such as in the rewiring process for oscillators, or in the reallocation process for the infection spreading model, leads to a delocalization of the instabilities that occur slightly later. This delocalization might therefore explain the occurrence of large failures in
networks, such as large blackouts in power grids, or financial break-downs in networks of interdependent companies, as a result of their optimization for stability [8, 190].

## 5.9 Discussion

In this chapter, I show that prominent instabilities in real-world network systems tend to be approximately localized. In the model proposed by R. May [44], which is a standard model for complex systems, I study such localization quantitatively and identified characteristics of nodes that carry such approximately localized instabilities. Finally, I observe that attempting to stabilize such localized instabilities increases the risk of delocalized failures.

In the network systems I study, the localization of prominent instabilities is due to the low eigenvalue density for the associated extreme eigenvalues in the spectra of their characteristic matrices. The common observation of such low density for the inhomogeneous and sparse networks describing real-world systems [36–39, 41, 173–178] implies that such localization of prominent instabilities is a general feature beyond the examples I study explicitly.

Understanding the consequences of a specific localized instability requires considering the system-specific dynamics. Already in the four examples presented explicitly, instabilities can lead to different phenomena. In the SIS-model, and in the bioinvasion model, they manifest as the persistent occurrence of an infection on a few individuals, and in the population of a few ports by an introduced species. In food webs, an instability often leads to species extinctions, and for power grids, an instability means breakdown of synchrony or black-outs. Despite the possibly large-scale consequences of a localized instability, for instance due to cascades of overloaded power lines, or cascades of species extinctions, the instability arises initially from only a small part of a network. Therefore, the results in this chapter allow to understand their origin and to develop methods to contain them, even if the system is too large for global understanding of all of their possible consequences.

Localization of prominent instabilities might explain the discrepancy between the instability of randomly created complex system, and the observation of complex systems in nature, discussed in the complexity-stability debate [107]. The results on food web stabilization suggest that, while it is hard to establish a stable system randomly, a stable configuration may result from a stabilization of prominent localized instabilities through a break-off of the responsible parts. Such a processes which decouples and then sacrifices small
problematic parts to stabilize an unstable system, and which thereby retain most of a system’s complexity, might explain the formation of the complex stable networks observed in nature.

The localization of dynamical modes with extreme eigenvalues is particularly important in situations where a system faces a slow-moving risk, because in this context, these modes describe the likely first instabilities. Due to their localization, one can expect these instabilities to be localized in natural systems. However, at the same time, attempts to optimized networks for stability may lead to large-scale failures because they lead to a delocalization of the dynamical modes associated with their slightly less prominent instabilities. In other words, a trade-off exists between the robustness of a network, i.e. its ability to withstand perturbations, and its resilience, i.e. its ability to recover from large failures with minimal losses. In particular, for power grids, this view is supported by recent results suggesting that decentralizing power grids instead of adding wires makes them more vulnerable to perturbations (they become less robust), but also makes them more resilient against large-scale failures [158].

While even localized instabilities can lead to delocalized consequences, the inherently delocalized modes arising from stabilization attempts are particularly problematic. They describe instabilities that manifest on large parts of a network simultaneously and emerge from many nodes, thereby hampering understanding their consequences and origin. In the light of these results, one might therefore ask if a better understanding of localization might allow to design more efficient ways to stabilize instabilities in networks which avoid the large-scale failures commonly observed in optimized network systems, such as pandemic infections in networks of humans or computers, or financial and economical crises in networks of financially dependent businesses.
Conclusions

In the thesis, I used spectral methods to study dynamics on complex systems. In particular, I studied localized dynamics, dynamics that possibly affects the entire system, but which originates in a small part, such as the pandemic spreading of a forest fire following an initial local occurrence, black-out cascades in power grids due to a failed power line, or the loss of entire classes of species in an ecosystem following the loss of a resource.

I studied complex systems through their representation as networks, which achieves a reduction of the inherent complexity of a system’s constituents, but retains the complexity of its interaction pattern. Using this representation, I investigated the dynamics of complex systems near stable steady states, which can be interpreted as their typical functioning operating modes.

I studied systems in which the dynamics near an operating modes is determined by a suitable characteristic matrix, such as a Jacobian matrix describing the coexistence of species in an ecosystem, or an adjacency matrix capturing the spreading of a virus through a network of computers. Investigating the spectrum of these matrices decouples their dynamics into dynamical modes, each describing an independent part of the dynamics, such as a distinct possible instability, that can drive the system away from its operation mode. More precisely, each dynamical mode is associated to one eigenvalue of the characteristic matrix and its eigenvectors. The eigenvalue determines if the mode leads to an instability or not, while the amplitudes of the corresponding eigenvectors determine where in a system this instability originates and manifests itself. To study localized dynamics, I therefore focused on the origin and implications of dynamical modes with localized eigenvectors, eigenvectors which have large amplitudes on only a few nodes of a network.

I distinguished two different types of localized dynamical modes, exactly and approximately localized modes. First, exactly localized modes are tied exactly to specific nodes of a network because their corresponding eigenvectors have exactly zero amplitude on all other nodes. I showed that such exact localization is tied to the occurrence of specific motifs in a network, small structures consisting of only a few nodes, such as small symmetric graph orbits.
Second, *approximately localized modes* localize on a few predominant nodes of a network, but depend to a lesser degree also on the remaining network. I showed that approximate localization occurs typically for dynamical modes with eigenvalues that belong to regions of low eigenvalue density, such as unusually large or small, *extreme*, eigenvalues of the characteristic matrix.

Studying exact localization allowed to establish analytical results on the implications of localized dynamics. Most prominently, I showed that if localized dynamics occurs for a small motif in a network, then the localized dynamics can be separated exactly from the remaining dynamics of the network. Thus, on the one hand, motifs that carry exactly localized dynamics causally link the topology and dynamics of a network. This result was recently used to engineer network motifs with specific dynamical properties in delay-coupled networks [151].

On the other hand, omitting localized modes and considering only the remaining dynamics, I established an exact coarse-graining process that reduces the size of networks, while retaining their non-localized dynamics. For instance, applied to food webs containing small symmetries, this coarse-graining explained the heuristic observation that such webs can have exactly identical local bifurcation diagrams despite having different topologies [43].

While exact localization allowed to understand localization analytically, approximately localized modes generally had greater implications for the observable dynamics of network systems. For example, when assessing the response of a complex system to perturbations, those few nodes on which dynamical modes with extreme eigenvalues localized are the *key nodes* that determine most of the system’s response. Further, in network systems that are undergoing a slow destabilization due to a slow-moving risk, approximate localization surfaced in the manifestation of their most prominent instabilities on only a few nodes, independently from the network size.

Since Euler solved the Seven Bridges of Königsberg [11], network representations of complex systems have drawn their appeal from the intuition that through them, one might attribute the functioning or failure of a system to specific smaller parts of a system [1, 14]. The connection drawn in Euler’s solution, between the local properties of city parts and the failure of finding a suitable path that crosses all bridges, may be much simpler than the connection modern methods draw between network properties and their consequences [15, 88, 191]. Nevertheless, it is still the original intuition to pinpoint phenomena in parts of a systems that underlies these modern methods, and which gave also rise to the idea pursued in this thesis, of tracing phenomena described by spectral methods to their localized origin.
The methods I presented allow to identify the key players and structures carrying localization from which dynamical phenomena emerge in a large network. However, taking a step back, they also provide a solid foundation to reflect on the underlying intuition of explaining a system’s functioning from its parts. For instance, one might ask, if the functioning of complex systems depends on their complexity as a whole, or if it is an expression of the functioning of closely interwoven, yet possibly simple, individual parts. Simply put, one might ask if the functioning of a complex system is the sum of the functioning of its parts.

The observation of delocalized dynamical modes in the thesis showed that a large proportion of a network’s dynamics is typically dependent on the entire system. This manifests, for instance, when studying perturbations because these perturbations can percolate throughout large networks [106]. Intuitively, this observation of global interdependency inside a network is not surprising, as it explains why complex systems are typically perceived as large entities rather than as collections of interdependent individual parts. Ultimately, one might therefore conclude that at least some of the functioning of complex systems emerges from a functioning of the system on a global level. A detailed understanding of this functioning will therefore require an understanding of the complexity of the entire system.

While delocalized modes show that much of a complex system’s dynamics is dependent on the complexity of a system on the global level, the occurrence of exact localization ties dynamics to the level of small structures. More precisely, exact localization implies the existence of dynamics on a meso-scale level, dynamics that exists on a level between single nodes and the entire network.

On the one hand, the observation of meso-scale dynamics means that it is possible to understand part of a system’s functioning from methods investigating a system on the meso-scale. By finding motifs that carry localization, one can turn observation of their occurrence into rigorous statements about their role in a network, thereby allowing to advance an understanding for the interplay between complexity and dynamics. Notably, studying the possibility for exact localization on a motif allows to distinguish rigorously between motifs whose dynamics admits an understanding on the meso-scale, and motifs that require considering the complexity of the embedding network as a whole. For instance, localization explains, why heuristic studies correlating the occurrence of small motifs in networks with their assumed functioning [14, 15, 17, 19] allowed to identify the destabilizing effect of the competitive-exclusion motif in food webs (which carries localized dynamics), while they remain inconclusive for other motifs, such as small chains [17, 20].
On the other hand, the scarcity of motifs carrying meso-scale dynamics leads to the somewhat sobering conclusion that meso-scale approaches can only capture a small fraction of the dynamics of a complex system. Therefore, achieving an understanding of all details of the functioning of complex systems may remain an unsolved problem, if the complexity of the system cannot be reduced.

However, not being able to understand the entire functioning of a system without understanding the consequences of its global complexity does not mean that all network phenomena are too complex to understand. In particular, the approximate localization of typical failures shows that often the most relevant phenomena in complex systems can be understood even in otherwise prohibitively large systems; a view that is supported by the observation of key players and self-stabilization for food webs, and by the observation of localized failures and dynamics in various network systems [24, 35, 119, 158, 181]. One might even argue that the persistence of complex systems in nature would be impossible without the localization of their prominent failures. In this reasoning, the localization might be nature’s way to locally contain failures, instead of allowing them to spread and possibly affect an entire system.

The observation, that normally the prominent failures of complex systems are localized, gives hope for understanding and mitigating the failures in these networks. For example, if a small structure is identified to cause a localized failure, one can study the interactions of this structure with a possibly large remaining system by modeling the structure and its direct neighborhood in a network explicitly, but using a mean field background to capture the complex remaining system. Thereby, one can understand the consequences of such a failure, such as the probability of it spreading into the remaining network, even in networks that are prohibitively complex to understand their functioning. Thus one might take appropriate action to mitigate the associated failure when the structure is observed as part of a large complex system.

In conclusion, the functioning of a complex system depends on its global complexity, such that its understanding requires to take into account the entire system. However, at the same time, the most prominent failures typically have a localized origin, such that their understanding is possible by tracing them to manageable small-scale problems.

One might even speculate, if an intrinsic link exists between the global properties of the functioning of a complex system, and the localization of its failures. Starting with the discovery of Anderson-Localization [34] for materials with impurities, localization was repeatedly linked with the disorder or inhomogeneity of complex systems, such as different potentials for sites in crystal lattices, or the different time-scales characterizing the interacting elements [35, 175, 177, 182]. Thus, one might wonder, if the complexity that makes
the functioning of a system dependent on the global level, also implies that the interaction pattern is inhomogeneous, thereby causing the localization of the system’s failures.

The thesis provides a solid starting point to tackle the challenge of understanding localization and its implications in many real-world networks. In particular, the methods presented straight-forwardly generalize to other generalized models, thereby making networks from metabolic control and cell signaling to delay effects of species migrations promising systems to study in future research [109, 113, 115, 118, 119, 180]. Further important insights and methodological progress may also result from understanding the connection of localization in networks with localization in other fields such as random matrix theory [36, 40, 182, 192, 193], or condensed matter physics [23, 35, 36, 40, 41, 194].

We live in a world, in which networks grow constantly in size and complexity. Therefore it will be paramount for ensuring their functionality to understand the trade-off between their robustness and their resilience against large-scale failures. This need for understanding is underlined by the contrast between the localization of instabilities that can isolate unstable parts in natural systems, and the delocalization in human-optimized systems that can lead to large-scale failures, from black-outs in power grids, over pandemics of infectious diseases or computer viruses, to economic or financial crises in networks of interdependent companies. The insights of this thesis into the interplay between complexity and localization therefore present a starting point towards an understanding of the functioning or failure of the complex systems on which we humans depend.
Acknowledgments

My thanks go to

- Thilo Gross, for proposing the interesting problems I studied in this thesis, for always being available for discussions, and for his patience in teaching me “science”.

- Thilo Gross and Lars Rudolf for great adventures in physics, Endeara and beyond.

- Kevin D. Lafferty (UCSB), for his help with the biological applications of my results.

- my fellow “Buncarites” in Bristol, and my fellow PhD students in Dresden, for inspiring discussion and challenging problems from physics, engineering, biology, and (pub)-lication and -location.

- the Studienstiftung des deutschen Volkes for financial support during my stay in Bristol, and the Bristol University Alumni Foundation for financial travel support.

- my family and friends for their support, particularly Janina Hesse for proofreading the thesis.
Appendix

A Parametrization of the Gatun Lake food web

Here, I apply the generalized model for food webs (Sec. 2.7) to parametrize the real-world food web of Gatun Lake in Panama (Fig. 2.1). I recall that the generalized model was introduced in Ref. 27, and that the Gatun Lake Food Web was proposed and studied in Ref. 10. This example demonstrates the method of generalized modeling and gives a hands-on intuition for understanding such a parametrization.

I start with the parametrization in which the parameters defining the generalized model are set to specific values. To keep the parametrization process clear, I choose a basic parametrization in which parameters follow from allometric scaling laws with the species’ body mass. Afterward, I use the generalized model description of perturbation impact (Chap. 3) to analyze the perturbation consequences of an invasion of this food web by the predatory peacock bass *Cichla ocellaris* recorded in Ref. 10. For the numerical analysis, I wrote two small programs, which are available online, one to parametrize food webs [121], and one to evaluate Impact [138].

**Parametrization**

For a given food web, the generalized model presented in Sec. 2.7 describes a broad class of different food web models. Testing a specific model in this class requires parametrizing the model by setting the generalized parameters to specific values (see Tab. 2.1 for a list). Following Ref. 92, I chose the elasticity parameters \((\gamma, \mu, \phi, \psi)\) based on biological intuition, and the turnover rates \((\alpha)\) based on allometric scaling with each species’ body mass. Further, I establish the dietary parameters \((\chi, \beta, \sigma, \rho)\) based on estimated strengths of feeding relations between different species. These strengths were computed by Kevin D. Lafferty (University of California at Santa Barbara, and US Geological Survey) based on allometric scaling laws of predator-prey body mass ratios. Kevin D. Lafferty also provided me with the data for the estimated body mass of each species and with typical biomass turnover rates for some of them.
The parameters used throughout the thesis to describe the Gatun Lake food web are set using the following rules (see Tab. A.1 for the resulting values):

- \( \alpha_i = (0.00485 M_i)^{-1/4} \). The species turnover rates are scaled allometrically with their body mass \( M_i \), unless more precise explicit empirical knowledge is available. The pre-factor 0.00485 follows from fitting the measured turnover data with this rule.
- \( \gamma_i = 0.95 \). The biomass gain through predation of a predator species increases slightly slower than the abundance of its prey. Intuitively, this is due to a saturation in prey of the predator, as in the well-known Holling type-II functional response.
- \( \mu_i = 1 \). The loss of biomass through natural mortality of each species increases linearly with the species’ abundance.
- \( \phi_i = 0.5 \). The biomass gain of primary producer species’ in the food web increases sublinearly with the primary producer’s abundance, because of space or nutrient constraints.
- \( \psi_i = 1 \) for fish and 1.02 for birds. The predation gain increases linearly with the abundance of fish. For birds, predation gain increases slightly faster because increasing abundance decreases fish’s refugium from predation.
- \( \lambda_{ij} = 1 \). A predator has no dietary preferences and switches passively between different prey species (see Ref. 27 for details).
- \( \rho_i = 0 \) for primary producers and \( \rho_i = 1 \) for predators. Each species gains biomass either by primary production or by predation.
- \( \chi_{ji} \) reflects the relative contribution of species \( i \) to species \( j \’s \) diet. Values are set based allometric relationships between consumer and prey body sizes. The following method was proposed by Kevin D. Lafferty: for each predator, one estimates a preferred prey body mass and a typical prey body mass range from Table 1 in Ref. 125, individually considering allometric scaling for birds, fishes, and invertebrate predators. Then, one considers that for each predator with a preferred prey body mass \( M_{\text{pref}} \) the feeding on prey with a body mass \( M_{\text{prey}} \) is distributed normally with \( \log(M_{\text{prey}}/M_{\text{pref}}) \), where the width of the normal distribution is estimated from the typical prey body mass range of each consumer. Thus one obtains for each predator-prey relation in the food web the preference \( e_{ki} \) of the predator \( k \) to consume the prey \( i \). To obtain \( \chi_{ki} \), I normalize the contributions, i.e. \( \chi_{ki} = e_{ki} \sum_j e_{kj} \).
• $\sigma_i$ indicates the biomass loss due to predation, relative to the total loss, which increases with predation pressure on a species. Using the dietary preferences $e_{ki}$, I set $\sigma_i = \frac{\sum_k e_{ki}}{(\sum_k e_{ki} + m_i)}$, where $m_i$ is a measure for the importance of the loss through mortality relative to the feeding pressure measured by $\sum_k e_{ki}$.

• $\beta_{ji}$ represents the relative loss of a species $i$ to its different predators $j$. In general, this predation loss of a species to different predators depends on the remaining diet of the predators, i.e. the parameters $\beta$ are dependent on $\chi$. To resolve this dependency consistently, I assume that when a predator kills a prey it consumes all its biomass, such that biomass is conserved throughout the process. I denote the total biomass loss rate of a species $i$ as $B_i$. For top predators, this loss is due only to natural mortality and $B_i = 1$. For all other species, the conservation of biomass implies that the loss $B_i$ equals the intake of its predators, i.e. $B_i = \sum \chi_{ji}B_j/\sigma_j$, where $j$ runs over all of $i$’s predators. This system of linear equations can be solved for all the species biomass out-flows $\mathbf{B}$. Then the parameters $\beta$ are the contributions of each predator $j$ to the total loss of $i$, i.e. $\beta_{ji} = \chi_{ji}/B_iB_j/\sigma_j$.

I note that these rules are implemented in the code I provide online [138].

After the parametrization, one can write the Jacobian matrix $\mathbf{J}$ using the expressions given in Sec. 2.7. Stability of any particular steady state (corresponding to a particular set of generalized parameters) is ensured by checking that all eigenvalues of the Jacobian have negative real parts [26]. For the Gatun Lake food web and the parametrization above, the eigenvalues with the largest real parts are the pair of complex conjugate eigenvalues $\lambda_{\pm} \approx -10^{-6} \pm 0.02i$. Therefore, the chosen parametrization corresponds to a steady state of coexisting species.

**Estimation of the impact caused by the introduction of peacock bass**

Now, I compute the impact of a perturbation to the Gatun Lake food web. More specifically, I estimate the initial effect of the invasion of the lake by a predatory fish, such as the peacock bass, as described in Ref. 10.

I consider the increase of predation pressure on species that are prey to the peacock bass as the main direct perturbation of the established ecosystem. To quantify this perturbation for the different established species $i$, I now estimate the effect of the additional population of peacock bass $Y_B$. The perturbation of species $i$ due to increased predation by $B$ is

$$K_{iB} = - \left. \frac{\partial l_i}{\partial Y_B} \right|_* = -\alpha_B \chi_{Bi} \kappa_i,$$  

(1)
### Appendix

#### Tab. A.1: Modeling Parameters estimated for the Gatun Lake food web.

The parameters $\alpha_i$, $\psi_i$, $\rho_i$, $\sigma_i$, $\chi_{ij}$ and $\beta_{ij}$ are generalized parameters as given in Tab. 2.1. All other generalized parameters are assumed identical for all species, i.e. $\gamma = 0.95$, $\phi = 0.5$, $\lambda = 1$, $\mu = 1$, $\psi_{fish} = 1$, $\psi_{bird} = 1.02$, $m_i = 0.12$. $M_i$ is the used typical body mass of individuals from each species, and $\chi_{Bi}$ is the relative diet composition of the peacock bass. All values for species body masses, $M_i$, and biomass turnover rates $\alpha$ marked by a star were collected from literature by Kevin D. Lafferty, University of California.

where $l_i = \frac{L_i}{L_i^*}$ is the total biomass loss normalized by its steady state value, and where I split the derivative into the turnover rate $\alpha_B \chi_{Bi}$ ($\alpha_B$ is the turnover rate of the peacock bass and $\chi_{Bi}$ the relative contribution of $i$ to $B$’s diet), and an elasticity $\kappa_i = \partial \log l_i / \partial \log Y_B$. For simplicity, I assume in the following that $\kappa$ is identical for all species $i$ and therefore I drop the corresponding index.

The peacock bass feeds on the secondary consumer fishes (tetras, silverside, and sailfin molly/mosquito fish, blackbelt cichlid), and on the bigmouth sleeper [10]. To obtain the relative contributions of each of these species to its diet, $\chi_{ai}$, I again employ the allometric scaling rules introduced to estimate the diets of species in the food web. This leads to the values indicated in Tab. A.1.

Plugging $J$ and the perturbation for the peacock bass $K$ into $I = -J^{-1}K$ leads to the ultimate impact of the perturbation shown in Fig. 3.1. The resulting response (Fig. 3.1) is given in Tab. A.2 alongside the observed changes [10].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i Name</td>
<td>$\alpha_i$ $\rho_i$ $\sigma_i$ $\chi_{ii}$ $\beta_{ii}$ $M_i$ [g] $\chi_{Bi}$</td>
<td></td>
<td>$i$</td>
<td>$j$</td>
</tr>
<tr>
<td>1 Nanophytoplankton</td>
<td>203.3$^*$ 0 0.62 - - $10^{-6}$ -</td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2 Filam. green algae</td>
<td>30$^*$ 0 0.50 - - 0.0003 -</td>
<td></td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3 Zooplankton</td>
<td>30$^*$ 1 0.9 0.25 0.32 0.002 -</td>
<td></td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4 Insects (mosquitoes)</td>
<td>$5^*$ 0 0.82 - - 0.003 -</td>
<td></td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>5 Silverside</td>
<td>2.509 1 0.80 0.09 0.11 5.2 0.30</td>
<td></td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6 Tetras</td>
<td>1.432 1 0.59 0.08 0.13 49 0.17</td>
<td></td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>7 Sailfin m.&amp;Mosqu. f.</td>
<td>3.89 1 0.66 0.20 0.30 0.89 0.39</td>
<td></td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>8 Blackbelt Cichlid</td>
<td>0.847 1 0.26 0.14 0.56 1716 0.05</td>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>9 Bignmouth Sleeper</td>
<td>0.38$^*$ 1 0.26 0.14 0.56 1716 0.05</td>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>10 Tarpon</td>
<td>0.33$^*$ 1 0.13 0.08 1 24000 -</td>
<td></td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>11 Black Tern</td>
<td>1.39 1 0 - - 55 -</td>
<td></td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>12 Herons and Kingf.</td>
<td>0.71 1 0 - - 800 -</td>
<td></td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>B Peacock bass</td>
<td>0.64 - - - - 1213 -</td>
<td></td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

Tab. A.1: Modeling Parameters estimated for the Gatun Lake food web. $\alpha_i$, $\psi_i$, $\rho_i$, $\sigma_i$, $\chi_{ij}$ and $\beta_{ij}$ are generalized parameters as given in Tab. 2.1. All other generalized parameters are assumed identical for all species, i.e. $\gamma = 0.95$, $\phi = 0.5$, $\lambda = 1$, $\mu = 1$, $\psi_{fish} = 1$, $\psi_{bird} = 1.02$, $m_i = 0.12$. $M_i$ is the used typical body mass of individuals from each species, and $\chi_{Bi}$ is the relative diet composition of the peacock bass. All values for species body masses, $M_i$, and biomass turnover rates $\alpha$ marked by a star were collected from literature by Kevin D. Lafferty, University of California.
<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
<th>Observed change [10]</th>
<th>Impact $I_B/\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nanophytoplankton</td>
<td>decrease</td>
<td>-0.06</td>
</tr>
<tr>
<td>2</td>
<td>Filamentous green algae</td>
<td>?</td>
<td>+0.06</td>
</tr>
<tr>
<td>3</td>
<td>Zooplankton</td>
<td>increase</td>
<td>+0.05</td>
</tr>
<tr>
<td>4</td>
<td>Insects (Mosquitoes)</td>
<td>increase</td>
<td>+0.04</td>
</tr>
<tr>
<td>5</td>
<td>Silverside</td>
<td>-50%</td>
<td>-0.6</td>
</tr>
<tr>
<td>6</td>
<td>Tetrass</td>
<td>-95%</td>
<td>-0.3</td>
</tr>
<tr>
<td>7</td>
<td>Sailfin Molly &amp; Mosquito Fish</td>
<td>-100%</td>
<td>+0.24</td>
</tr>
<tr>
<td>8</td>
<td>Blackbelt Cichlid</td>
<td>+50%</td>
<td>-0.06</td>
</tr>
<tr>
<td>9</td>
<td>Bigmouth Sleeper</td>
<td>-90%</td>
<td>-0.16</td>
</tr>
<tr>
<td>10</td>
<td>Tarpon</td>
<td>decrease</td>
<td>-1.25</td>
</tr>
<tr>
<td>11</td>
<td>Black Tern</td>
<td>decrease</td>
<td>+3</td>
</tr>
<tr>
<td>12</td>
<td>Herons and Kingfishers</td>
<td>decrease</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

Tab. A.2: Comparison of the observed long-term changes in the species abundances after the introduction of the peacock bass [10] with the assessed impact of the perturbation corresponding to this introduction. The impact, $I_B$, representing the response to the population of peacock bass, is normalized by the unknown elasticity $\kappa$, which I assume as constant.

B The Master Stability Function approach

Here, I give a more detailed summary of the Master Stability Approach used in Sec. 5.2. The approach was introduced in Ref. 29, and a review with applications is given in Ref. 30. After a short summary of the main concept, I apply it to a simple model of power generators.

The Master Stability Function approach considers systems of identical oscillators. Each of the oscillators is a small subsystem, described by its internal state $x$ consisting of one or more variables. In the absence of coupling between the different oscillators, the evolution of this internal state of each of the oscillators is independently described by the identical (yet possibly complicated) equations $\frac{dx}{dt} = F(x)$, where $F$ is the (vector-valued) function describing each oscillators internal dynamics. Denoting the internal state of each oscillator $i$ as $x^{(i)}$, and starting all of them in the same synchronous state, all of the internal states evolve identically as $x^{(i)} = s(t)$. 
Appendix

When introducing coupling between the oscillators, the state of one oscillators can affect the state of other oscillators. More explicitly, the effect on the internal state of oscillator $i$ caused by a coupled oscillator $j$ is the difference of the (vector-valued) coupling function $H$ on these two oscillators, i.e. the effect on an oscillator $i$ by another oscillator $j$ is $H(x^{(j)}) - H(x^{(i)})$.

Following Ref. 29, I consider an undirected coupling network in which the pairwise coupling effects on an oscillator are summed up from all of its neighbors. More precisely, if $A$ is the adjacency matrix of the coupling network ($A_{ij} = A_{ji} = 1$ if oscillators $i$ and $j$ interact), then the dynamics of an oscillator $i$ is governed by the equation

$$\frac{dx^{(i)}}{dt} = F(x^{(i)}) + \sigma \sum_j A_{ij} \left( H(x^{(j)}) - H(x^{(i)}) \right), \quad (2)$$

where $\sigma$ represents the weight of the coupling relative to the internal dynamics of the oscillators.

As long as all oscillators are synchronized, i.e. $x^{(i)} = s(t)$, the coupling terms cancel out. However, when this synchrony is slightly perturbed, the coupling can either drive the system back towards the synchronous state, or the asynchrony tends to grow. Thus, for the synchrony to persist, the synchronous state has to be linearly stable against small perturbations at each point in time. To analyze this, I write the deviation of the state of each oscillator from the synchronous state as $\delta x^{(i)} = x^{(i)} - s(t)$ and then linearize (2) around $\delta x^{(i)} = 0$. In other words, I place myself in the reference frame that co-evolves with the synchronous state, and study the stability of the steady state 0 in this frame. If this state is stable against perturbations, all oscillators tend towards the synchronous state, and the original system will remain synchronous.

Linearizing (2) around $\delta x^{(i)} = 0$ leads to

$$\frac{d\delta x^{(i)}}{dt} = J_F \delta x^{(i)} + \sigma J_H \sum_j A_{ij} (\delta x^{(j)} - \delta x^{(i)}), \quad (3)$$

where $J_F$ is the Jacobian matrix of $F$ describing the linearized internal dynamics of each oscillator near the synchronized state, and where $J_H$ is the Jacobian matrix of $H$ describing the linearized properties of the coupling function around the synchronous state $s(t)$. I note, that both matrices $J_H$ and $J_F$ depend on time\(^1\).

\(^1\)They are linearizations around the synchronous state $x^{(i)} = s(t)$, which changes over time, such that the derivatives taken to obtain the Jacobian matrices may change if the functions $F$ and $H$ are nonlinear.
The stability of the synchronous state, $\delta x^{(i)} = 0$, is determined by the Jacobian matrix describing the combined dynamics of all oscillators as given by (3). When ordering the variables suitably, this Jacobian matrix consists of square blocks of identical size; the block in the $i^{th}$ row of blocks and in the $j^{th}$ column of blocks contains the derivatives of the dynamics of oscillator $i$, as given in (3), by the state variables of oscillator $j$. For instance, the $i^{th}$ block on the diagonal is $J_F - J_H \sigma \sum_j A_{ij}$. Further, the off-diagonal blocks are either zero (if two oscillators $i$ and $j$ are not coupled) or $\sigma J_H$ if they are coupled. Thus, the structure of the blocks exactly corresponds to the structure of the entries in the adjacency matrix.

Because of the block-structure of the Jacobian matrix, one can block-diagonalize the Jacobian matrix by treating $J_F$ and $J_H$ as if they were simple numbers. The blocks on the diagonal of the resulting matrix are then given by $J_F - \sigma \mu_l J_H$, where each $\mu_l$ is one eigenvalue of the Laplacian matrix $L$ of the coupling network\(^2\) [30].

Each block $J_F - \sigma \mu_l J_H$ characterizes the stability of the synchronous state against a specific perturbation. If one eigenvalue of the block $J_F - \sigma \mu_1 J_H$, is positive, then the synchronous state $s(t)$ is unstable against a perturbation of the oscillators in the direction of the eigenvector corresponding to the eigenvalue $\mu_l$. For instance, consider the eigenvalue $\mu_0 = 0$, which corresponds to a special eigenvector of $L$ that is constant on all oscillators. If one eigenvalue of $J_F - \mu_0 J_H$ is positive, then a perturbation affecting all oscillators identically grows, such that the states of all oscillators deviate from the dynamics $s(t)$ without coupling.

The eigenvalue $\mu_0$ is a special case. Because the corresponding eigenvector has identical entries for all oscillators, an instability described by this eigenvalue affects all oscillators identically, and synchrony is maintained although the trajectory becomes different of the trajectory of uncoupled oscillators $s(t)$. For all other eigenvalues $\mu_l$, however, the corresponding eigenvectors of $L$ have different entries on different oscillators, such that a positive eigenvalue of $J_F - \mu_1 J_H$ for them means that the deviations from $s(t)$ grow differently for different oscillators, leading to the break-up of synchrony.

In summary, finding instabilities of a synchronous state for a given topology requires diagonalizing $J_F - \sigma \mu_l J_H$ for each of the eigenvalues $\mu_l$ corresponding to the Laplacian matrix $L$ of the coupling network, where $J_F$ denotes the

\(^2\) Explicitly, $J = -L \otimes J_H + 1 \otimes J_F$, where $\otimes$ denotes the Kronecker product (see Ref. 195), and $1$ the identity matrix with the same dimensions as $L$. Writing the first factor of the products in the eigenvector base of $L$ leads to $-D_L \otimes J_H + 1 \otimes J_F$, where $D_L$ is the diagonalized Laplacian matrix, containing the eigenvalues $\mu_l$ as entries.
Jacobian matrix of the internal oscillator dynamics, and $\mathbf{J}_H$ the Jacobian matrix of the coupling. If an eigenvalue of $\mathbf{J}_F - \sigma \mu_l \mathbf{J}_H$ is positive, then the synchrony is lost in the direction indicated by the eigenvector of $\mathbf{L}$ corresponding to $\mu_l$.

When studying the same oscillators for different topologies, establishing the Master Stability Function, which I note $\text{MSF}(\mu)$ in the following, avoids the repeated diagonalization of the blocks $\mathbf{J}_F - \sigma \mu_l \mathbf{J}_H$ [29]. Explicitly, one first establishes the largest eigenvalue of $\mathbf{J}_F - \sigma \mu \mathbf{J}_H$ observed along the synchronous trajectory $\mathbf{x}_i = s(t)$ as a function of $\mu$. Then, studying the synchrony of coupling these oscillators in a given network only requires to verify that $\text{MSF}(\mu_l) < 0$ for all eigenvalues $\mu_l$ of the Laplacian matrix.

Using the Master Stability function facilitates studying implications of the network topology because once $\text{MSF}(\mu)$ is established, the regions are known where eigenvalues $\mu_l$ of a Laplacian matrix can lead to a instabilities. In particular, when the internal state of each oscillator is described by only few variables, $\text{MSF}(\mu)$ is often simple enough to establish these regions explicitly. For instance, the Master Stability Functions for a simple model of power generators given below and for the commonly studied Rössler oscillators are negative only for a well-defined range of values $\mu$ (Fig. B.1) [171]. Thus, for these systems, one can quickly verify whether synchrony is stable in a coupled system by determining if the largest and smallest eigenvalues of the Laplacian matrix are inside the stable region.

To give a more intuitive understanding of the Master Stability Approach, I now consider a simple toy model of coupled machines or generators. These may be coupled, for instance, as elements of a power grid. For optimal efficiency of such a network, one would therefore aim at a synchronous operation of all these oscillators at a given system frequency $\Omega$, which I set to 1, for simplicity, after the model description.

Here, I consider machines with an internal dynamics comprising two variables. First, each machine has a momentary operation frequency $\omega$. Second, each machine runs at an amplitude $r$ denoting, for instance, the amount of power it generates, or its current power consumption. I assume that $\omega$ is regulated internally for each machine to tend towards the system frequency, while the amplitude increases at a constant rate $\alpha$ until its increase is stopped by friction or other losses. Explicitly, I consider

$$\frac{d\omega}{dt} = \Omega - \omega \quad \text{and} \quad \frac{dr}{dt} = \alpha r - r^2,$$

where $\alpha$ is the rate of amplitude increase of the friction-less machine.
$B$ **The Master Stability Function approach**

Fig. B.1: Master stability function for a simple model of coupled machines and for Rössler Oscillators. In the main panel, each curve represents the eigenvalues of $J_F - \sigma \mu J_H$ for simple machines (see text for details), where $J_F$ and $J_H$ are the Jacobian matrices of the internal dynamics and of the coupling function, and where $\sigma$ (set to 1) weights the coupling against the internal dynamics. At each value $\mu$, the larger eigenvalue defines the Master Stability Function $\text{MSF}(\mu) = \max(\alpha - \mu, -1 + \mu)$, shown as solid line for $\alpha = 0.1$. Where $\text{MSF}(\mu)$ is positive (shaded areas), eigenvalues of the coupling Laplacian $\mu$ cause an instability of the synchronized state. The inset shows the Master Stability Functions for coupled Rössler-systems, one coupled via the first state variable (yellow), and one via the second (green), see Ref. 171 for a detailed study of the Rössler systems. Parameters $a$, $b$, $c$, $\sigma$ for the curves were chosen, such that both configurations have approximately the same problematic regions as the simple machines. Explicitly, $a = 0.2$, $b = 2$, and $c = 4$, $\sigma = 3.43$ for the system coupled via the first variable, or $c = 5.16$, $\sigma = 9.26$ for the system coupled in the second variable.

Further, the oscillators are coupled in both their amplitudes and frequencies. More precisely, I consider that the amplitudes couple directly with the amplitudes of neighboring machines, while the coupling of the frequencies decreases with frequency as $1/\omega$, for example because of losses in the connection lines. Thus, the coupling function is $H(r, \omega) = (r, 1/\omega)$.

Now, I study the synchrony of the coupled machines in two modes of operation. First, in their natural operating state, $r = \alpha$, $\omega = \Omega$, and second, when they are simultaneously switched on and off again, therefore moving back and
forth between two states, \((r_0, \omega) \approx (0, \Omega) \leftrightarrow (\alpha, \Omega)\). When the machines are in their natural operating state, then \(\text{MSF}(\mu)\) is simply the larger eigenvalue of \(-\alpha - \sigma \mu\) and \(-1\). Thus, \(\text{MSF}(\mu) < 0\) for any value of \(\sigma\), such that the machines remain synchronously in their operation state. However, when the coupled machines are switched on and off, then the maximal values of the two eigenvalues of \(J_F - \sigma \mu J_H\) for this process are \(\alpha - \sigma \mu\) and \(-1 + \sigma \mu \Omega^2\), such that \(\text{MSF}(\mu)\) is negative for \(0.1 < \sigma \mu < 1\), but positive everywhere else (Fig. B.1).

From the MSF in Fig. B.1 one can conclude, that synchronously switching on the machines is possible only when the eigenvalues \(\mu\) of the Laplacian matrix that describes the coupling network fulfill \(0.1 < \sigma \mu < 1\). By design of the considered machines, the unstable region \(\sigma \mu < \alpha\) contains eigenvalues where the amplitudes of the oscillators become asynchronous, while for \(\sigma \mu > 1\) their frequencies diverge. Thus, a direct relation between the extremal eigenvalues and the breakdown of synchrony during a switch-on process of the considered machines exists.

I emphasize, that the problematic (large and small) eigenvalues found for the presented simple model of power grid elements agrees with other models of oscillators. For instance, the qualitative features of Master Stability Function agrees with those of Rössler oscillators which are typically studied with this approach (inset in Fig. B.1) [29, 171]. Thus, also for these systems, which have a more complicated internal dynamics, the most problematic dynamical modes are those associated to eigenvalues of the Laplacian matrix \(L\), that are very small \(\mu\) or very large; an observation that has lead to the definition of the ratio of the smallest (nonzero) eigenvalue and the largest eigenvalue of a network, \(\mu_1/\mu_{\text{max}}\), as its synchronizability [30].

In summary, if a system of oscillators is described by differential equation of the form of (2), then the master stability function identifies regions of eigenvalues \(\mu\) of a coupling graph Laplacian, where these eigenvalues lead to instability of the synchronous state. Typically, these instabilities appear for the largest and smallest eigenvalues of the coupling network.

### C Approximate localization on larger structures

Here, I extend the results for approximate localization on single nodes to structures consisting of more than one node. This is possible by first substituting the scalar variables describing single nodes in Sec. 5.3, Sec. 5.5, and Sec. 5.6 by vectors or matrices, but subsequently applying the same developments. Explicitly, for structure containing the \(s\) nodes \(q_1, \ldots, q_s\), the vector \(a\) becomes a \(N-s \times s\) matrix \(A\) describing all edges in the network of \(J\) originating in these nodes, vector \(b\) becomes a \(s \times N-s\) matrix \(B\) describing all edges ending in these
nodes, \( e \) becomes a \( s \times s \) matrix \( \mathbf{E} \) containing all interactions between these nodes, and \( \mathbf{J} \) becomes a \( N - s \times N - s \) matrix that describes the network of all nodes except the nodes \( q_1 \ldots s \). Further, the eigenvector of an eigenvalue \( z \) question is split into the entries of the nodes \( q_1 \ldots s \) which I call \( \hat{\mathbf{v}} \) (formerly the entry \( v_q \) for the single node \( q \)) and the remaining entries \( \tilde{\mathbf{v}} \). Using these quantities in the eigenvalue equation one can then perform the same operations as for the single-node case to establish the delocalized proportion \( S(z) \) of the vector \( \mathbf{v} \), and a compatibility condition of the eigenvalue \( z \) with the properties of a structure described by \( \mathbf{A}, \mathbf{B}, \) and \( \mathbf{C} \).

First, I study the localized proportion of the eigenvector \( \mathbf{v} \) corresponding to an extreme eigenvalue \( z \). Following the developments in Sec. 5.3, I obtain for the non-localized entries of \( \mathbf{v} \), that \( \tilde{\mathbf{v}} = \mathbf{X} \hat{\mathbf{v}} \). Together with the normalization of \( \mathbf{v} \), this leads to the \((s\text{-dimensional})\) vector equation

\[
1 - S(z) = [1 - \tilde{\mathbf{v}} \cdot \tilde{\mathbf{v}}] \hat{\mathbf{v}} \cdot \hat{\mathbf{v}} = 1 - \hat{\mathbf{v}}^T \mathbf{A}^T \mathbf{X}^T \mathbf{X} \mathbf{A} \hat{\mathbf{v}},
\]

where the superscript \( T \) indicates the transposed matrix. Solving the equation given by the last two term of (5) is required to find the localized proportion explicitly.

**Application to May’s model**

To solve (5) for the expected value of the localized proportion in the May Model, one can split the matrix product into the product of the expected values of the independent terms, \( \text{Tr}\{\mathbf{X}^T \mathbf{X}\}/N \), which I approximate in Sec. 5.5, and \( \sigma_a^2 \hat{\mathbf{v}}^T \mathbf{C} \hat{\mathbf{v}} \), where \( \sigma_a \) denotes the norm of the columns in \( \mathbf{A} \) representing the weights of the links originating in the nodes \( q_i \). The new term in this formula compared to the single-node case is the matrix \( \mathbf{C} = \mathbf{A}^T \mathbf{A}/\sigma_a^2 \). This matrix, however, is simply the covariance matrix between the edges originating in the nodes \( q_i \). For instance, for two nodes with identical edges (symmetrical nodes), \( \mathbf{C} = 1 + (0 1 1 0) \), where \( 1 \) is the \( s \times s \) identity matrix, but for nodes with edges to different nodes only \( \mathbf{C} = 1 \). Using these expressions together with the normalization condition, and setting \( \sigma_a = 1 \) for simplicitly, I obtain the localized proportion as

\[
\hat{\mathbf{v}} \cdot \hat{\mathbf{v}} = \frac{1 + \hat{\mathbf{v}}(\mathbf{C} - 1)\hat{\mathbf{v}} \text{Tr}\{\mathbf{X}^T \mathbf{X}\}/N}{1 + \text{Tr}\{\mathbf{X}^T \mathbf{X}\}/N}.
\]

As an example, consider the localized proportion for a structure of two nodes. If these two nodes have connections to different nodes of the network, then \( (\mathbf{C} - 1) = 0 \); the covariance term vanishes and \( \hat{\mathbf{v}} \cdot \hat{\mathbf{v}} \) simply reduces to the one-node localized proportion found in Sec. 5.3. Intuitively speaking, the two nodes are two independent one-node localization centers. However, if the two nodes

129
are symmetric, i.e. they have edges to the same nodes, then \((C - 1) = (0 \ 1 \ 1 \ 0)\), and their entries of the vector \(\hat{v}\) are \(\hat{v}_1 = -\hat{v}_2\). Plugging these quantities in, I obtain exact localization, \(\hat{v} \cdot \hat{v} = 1\), in agreement with the exact localization one would expect for such a symmetry.

**Finding localization centers**

To find the conditions that allow a structure to carry a localized eigenvector, I study the compatibility condition following along the same lines as in Sec. 5.6. Using the same development as for a single node, but using the matrix quantities defined above, the compatibility condition is given by the requirement of consistency for both sides of the vector equation

\[
B \cdot XA = (z1 - E). \tag{7}
\]

The compatibility condition (7) for localization on \(s\) nodes has itself the form of an \((s\)-dimensional) eigenvalue problem. Only, if \(z\) is an eigenvalue of \(E + B \cdot XA\), then the condition is satisfied.

For the May Model, the compatibility condition simplifies to \(B \cdot XA \approx \mu(z)M\), where \(M = B^T A\) is the covariance matrix of the entries of the columns of \(A\), with the entries in the rows of \(B\), and where \(\mu(z)\) is the average eigenvalue of \(X\), as established in Sec. 5.6. If the columns of \(A\) and of the rows of \(B\) are normalized to 1, i.e. if the average interaction strengths of the nodes \(q_1 \ldots q_s\) with the remaining system are 1, and if the edges originating and ending in different nodes are independent, then \(M = 1R\) is diagonal and \(R\) indicates the correlation between the edges originating and those ending in the same node. Under these conditions, the compatibility conditions becomes \(E = z - \mu(z)R1\), which is solved, if \(z - \mu(z)R\) is an eigenvalue of the matrix \(E\) describing the internal interactions between the nodes \(q_1 \ldots q_s\). Thus, based on this equation, one can conduct similar investigations as in Sec. 5.6 to find the structure of small motifs (i.e. matrices \(E\)), which allow the existence of an extreme eigenvalue \(z\).

Of particular interest is localization on structures of two nodes because they allow to study the occurrence of localized pairs of complex conjugate dynamical modes, and therefore oscillatory instabilities. In particular, I consider a structure described by a matrix \(E = (\epsilon \ \bar{\epsilon} \ \epsilon \ -\bar{\epsilon})\) with identical diagonal entries and opposite off-diagonal entries. The eigenvalues of \(E\) are \(\epsilon_{\pm} = \epsilon \pm i\bar{\epsilon}\). Plugging this into the compatibility equation, the existence of a pair of complex conjugate eigenvalues \(z_{\pm}\) requires that \(\bar{\epsilon} = \Im z - \mu(z)R\), and that \(\epsilon = \Re z - \mu(z)R\). These conditions fully define the properties of a structure on which eigenvectors with such an eigenvalue can localize. It thus allows to find oscillatory instabilities on two-node structures by searching for structures with properties that are similar to those identified for \(E\).
Bibliography

Bibliography


Bibliography


<table>
<thead>
<tr>
<th>Bibliography</th>
</tr>
</thead>
</table>
Hiermit versichere ich, dass

1. ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

2. die Arbeit unter der wissenschaftlicher Betreuung von Dr. Thilo Gross an der University of Bristol, Bristol, UK, und unter der Betreuung von Prof. Dr. Frank Jülicher am Max-Planck-Institut für Physik komplexer Systeme, Dresden, angefertigt wurde.

3. keine frühere Promotionsverfahren stattgefunden haben.

4. ich die derzeit gültige Promotionsordnung der Fakultät Mathematik und Naturwissenschaften der Technischen Universität Dresden anerkenne.