THESIS / THÈSE

DOCTOR OF SCIENCES

Foundations of diffusion and instabilities in nonlinear evolution equations on temporal graphs and graphons

PETIT, Julien

Award date:
2020

Awarding institution:
University of Namur

Royal Military Academy
Polytechnic Faculty
Department of
Mathematics

University of Namur
Faculty of Sciences
Department of
Mathematics

Foundations of diffusion and instabilities in nonlinear evolution equations on temporal graphs and graphons

A thesis submitted by
Julien Petit
in fulfillment of the requirements for the
degree of Doctor of Science of the University of Namur and the
degree of Doctor of Engineering Science of the Royal Military Academy

Composition of the Jury

Timoteo Carletti (Supervisor)
Duccio Fanelli
Johan Gallant
Ben Lauwens (Supervisor)
Alexandre Mauroy (President)
Hiroya Nakao

June 2020

Graphisme de couverture : ©Presses universitaires de Namur

Rue Grandgagnage 19
B-5000 Namur (Belgique)

Imprimé en Belgique

ISBN : 978-2-39029-113-8
Dépôt légal: D/2020/1881/12

Université de Namur
Faculté des Sciences
rue de Bruxelles, 61, B-5000 Namur (Belgique)

Abstract

Evolution equations on graphs describe the development over time of a system on a domain made of nodes and links, starting from an initial state. They apply to phenomena so diverse that they will likely never develop into a single coherent field. However, a central question is to reveal the impact of the structural properties of the graph on the trajectory of the system.

This thesis is an attempt to further uncover this interplay between structure and dynamics, beyond the simple paradigm of static networks of moderate size. Empirical evidence indeed shows that some networks possess inherent temporal properties. On another note, massive graphs have become commonplace in real-life and scientific research, and a range of graph-theoretical methods and algorithms face scalability issues, demanding to consider graphs as if they were continuous objects.

We study two classes of problems: linear diffusion and then nonlinear variants, where local reactions combine with some form of diffusion through the edges of the graph. We revisit well-known stability-instability properties for such systems, and reveal significant effects due to the temporal nature of the graph. As a side note, we examine the inclusion of delay in the diffusive process, as a reasonable way to improve the models and refine the match with observations, in systems where the time for communication, reaction or decision-making cannot be neglected.

We also prove the validity of the continuum-limit approach to random walks on dense weighted graphs, in discrete and in continuous time, relying on the framework provided by graph-limit theory. We answer positively to the question of transferability, showing that the cost of duplicating the analysis for each finite graph can be avoided by considering instead the continuum problem. The dissertation ends with a study of diffusion-driven instabilities in reaction-diffusion systems on graphons which are the continuum-limit version of graphs.

Foundations of diffusion and instabilities in nonlinear evolution equations on temporal graphs and graphons, by Julien Petit
PhD thesis, June 25, 2020
Supervisors: Professor Timoteo Carletti and Military Professor Ben Lauwens

Acknowledgment

When the door came open, it revealed a large window overlooking the university library and a wall covered with books. I vividly remember that day which marked in my eyes the official start of this joint thesis between the University of Namur (UNamur) and the Royal Military Academy (RMA) where I came from. With this partnership in place I was ready to get started.

Back then, now more than five years ago, Anne Lemaître was the head of the Department of Mathematics in Namur and she had just introduced me to my new office, where I would be working half the time. As of today, I still feel privileged by her welcome, and appreciative for her decision to lead me to Timoteo Carletti's office, who became my co-supervisor at UNamur, together with Ben Lauwens at the RMA in Brussels. I whish to sincerely thank them both for agreeing to take on this role, and for their complementary approach to it.

Teo, you never seemed in short supply of availability, suggestions or interest, regardless of the hour of the day or the day of the week. You gave me all the freedom I could wish for, while still keeping in close touch at all times, providing mathematical or personal guidance when needed. Above all, you have shown me over the past years how to be a mentor who listens, genuinely values open discussion and helps build self-confidence. Your kindness and simplicity counted for a lot, and always made me feel fortunate. For many reasons, thank you very much.

I would like to thank the other two members of my steering committee: Renaud Lambiotte (Oxford University and UNamur) and Duccio Fanelli (University of Florence). Each time, you managed to bring an incredible amount of good ideas, insightful suggestions, challenging questions and a lot of friendliness into our discussions. I am thankful for all of that.

To the members of the Jury, Timoteo Carletti, Duccio Fanelli, Johan Gallant (RMA), Ben Lauwens, Alexandre Mauroy (UNamur) and Hiroya Nakao (Tokyo Institute of Technology), whose mention of graph-limit theory in 2016 did not go unnoticed: to you all, I express my sincere gratitude. I saw your agreement to become part of this thesis as an honor and an extra source of motivation throughout the writing of the manuscript. I am in particular grateful to Alexandre Mauroy for presiding over this jury, and for his meticulous list of corrections to this dissertation.

As far as I know, a thesis is anything but a lonely undertaking. I would like to thank my colleagues in Brussels and Namur who have witnessed every step of

the journey first hand. You have been able to add warmth, surprise and fun to the work days. I was lucky that you stopped at my door or let me in for a more or less mathematical debate. I am grateful for the time we spent working on a common topic, the help I received with the paperwork, the happy lunch breaks and all the social and sporting activities at different times of the day and night. This thesis was a part-time job full of good times.

Finally, thank you to my loved ones, friends and family. You mean everything to me.

Contents

Abstract ... iii

Acknowledgment ... v

Notation ... xi
Abbreviations ... xi
Related to graphs and graphons ... xi
Related to matrices and operators ... xii
Relations, operations, and probabilities ... xii

Chapter 1. Introduction ... 1
1.1. Evolution equations on graphs ... 1
1.2. How to read ... 2
1.3. Core concepts ... 2
1.3.1. Graphs and complex networks ... 2
1.3.2. Random walks ... 5
1.3.3. Reaction-diffusion equations ... 7
1.4. Aim ... 11
1.4.1. First question ... 12
1.4.2. Second question ... 12
1.4.3. Third question ... 12
1.5. Structure and contributions ... 13
1.6. Publications ... 15
Journal Publications ... 15
Preprint ... 16

Chapter 2. Mesoscopic foundations of diffusion on temporal networks ... 17
2.1. Introduction ... 17
2.1.1. Temporal networks ... 17
2.1.2. Single-timescale random walks ... 18
2.1.3. Beyond standard models ... 19
2.2. Mesoscopic foundations ... 21
2.2.1. Multiple-timescales random walks ... 21
2.2.2. Master equation ... 23
2.3. Resting time density ... 26
2.4. Mean resting time ... 33
2.4.1. Derivation of the mean resting time ... 34
2.4.2. Discussion ... 35
2.5. Steady state ... 37

2.5.1. Derivation of the steady state 37
2.5.2. Discussion 41
2.6. Conclusion 42

Chapter 3. Emergence of memory with competing timescales 45
3.1. Introduction 45
3.2. The semi-Markovian case 46
3.2.1. Equivalent node- and edge-centric models 47
3.2.2. Time-domain derivation of the propagator 48
3.3. Emergence of memory 51
3.3.1. Neglecting timescales 52
3.3.2. Neglecting the cycles in the graph 52
3.4. Mesoscopic modeling with cycles 52
3.4.1. Propagator with corrections for 2-cycles 53
3.4.2. Transition kernel with correction for 2-cycles 56
3.5. Numerical validation 65
3.6. Conclusion 66

Chapter 4. Random walks on graphons 69
4.1. Introduction 69
4.2. Preliminaries 70
4.2.1. Related to functional analysis 70
4.2.2. Related to semigroups 71
4.2.3. Related to graphons 72
4.3. Continuum limit of random walks 77
4.3.1. Continuum limit of the discrete heat equation 77
4.3.2. Continuum limit of the node-centric walk 78
4.4. Well-posedness 79
4.4.1. Connectedness and integrability of W/k 79
4.4.2. The IVP on $L^2[0,1]$ 80
4.4.3. Positivity 82
4.4.4. The IVP with probability density functions 83
4.5. Convergence on dense graphs 84
4.5.1. Convergence on the quotient graph $W/\mathcal{P}$ 84
4.5.2. Convergence on the sampled graph $W_{[n]}$ 89
4.5.3. Convergence for a sequence of discrete problems 90
4.6. Relaxation 91
4.7. Extension to the discrete-time walk 94
4.8. Conclusion 95

Chapter 5. Diffusion- and delay-driven instabilities on graphs 97
5.1. Introduction 97
5.1.1. Definitions of stability 97
5.1.2. Diffusion-driven instability 98
5.1.3. Pro's and flaws of the Turing approach 105
5.2. Asymptotic steady state 106
5.2.1. Bifurcation from a simple eigenvalue 106
5.2.2. Next-order characterization of the solution 110

5.3. Delay-driven instabilities . . . 113
5.3.1. Local reactions . . . 114
5.3.2. Global feedback . . . 118
5.4. Conclusion . . . 119

Chapter 6. Diffusion-driven instability on temporal networks . . . 123
6.1. Introduction . . . 123
6.2. The continuous case . . . 124
6.2.1. The method of averaging . . . 125
6.2.2. Stability via averaging . . . 127
6.2.3. Critical network timescale . . . 131
6.3. Extension to switched networks . . . 133
6.3.1. Linear stability on switched networks . . . 133
6.3.2. Examples . . . 134
6.4. Conclusion . . . 138

Chapter 7. Stability in reaction-diffusion equations on graphons . . . 141
7.1. Introduction . . . 141
7.1.1. On graphon Laplacian operators . . . 142
7.1.2. On our assumptions . . . 142
7.2. Spectral analysis of the combinatorial Laplacian . . . 144
7.2.1. The adjacency and degree operators . . . 144
7.2.2. Diagonalization of the Laplacian . . . 145
7.2.3. Perturbative approach . . . 152
7.3. Linear stability of reaction-diffusion on graphons . . . 153
7.3.1. Stability of the heat equation . . . 153
7.3.2. Linear stability of reaction-diffusion with one variable . . . 155
7.3.3. Linear stability of reaction-diffusion with two variables . . . 158
7.4. On the stability of the nonlinear system . . . 161
7.4.1. On the Brusselator . . . 161
7.4.2. On the principle of linearized stability . . . 162
7.4.3. The case $p = \infty$ . . . 163
7.4.4. The case $p = 2$ . . . 168
7.5. Conclusion . . . 171

Chapter 8. Conclusions and perspectives . . . 173
8.1. A look back . . . 173
8.1.1. On the first question . . . 173
8.1.2. On the second question . . . 175
8.1.3. On the third question . . . 176
8.2. A look ahead . . . 177
8.2.1. Other directions . . . 177
8.2.2. Final note . . . 178

Appendix A. Spectral measures . . . 179
A.1. Decomposition of a measure . . . 179
A.2. Spectral measures . . . 179
A.3. Towards the spectral theorem . . . 180

Appendix B. Decompositions and bounds on the spectrum 183
B.1. Decomposition based on spectral measures 183
B.2. Decomposition into the discrete and the essential spectrum 184
B.3. Bounds 185

Bibliography 187

Notation

Abbreviations

ACP	abstract Cauchy problem
CTRW	continuous-time random walk
DAG	directed acyclic graph
DDI	diffusion-driven instability
iff	if and only if
IVP	initial value problem
MRT	mean resting (or residence) time
ODE	ordinary differential equation
PDE	partial differential equation
PDF	probability density function
RD	reaction-diffusion
RV	random variable
SS	steady state

Related to graphs and graphons

$G = (V, E)$	graph G with vertex set V and edge set E
$\deg v$	degree of node v
$\operatorname{str} v$	strength of node v
V_v	set of out-neighbors of node v
V'_v	set of in-neighbors of node v
k_v	out-degree of node v
k'_v	in-degre of node v
$\sim$	symmetric adjacency relation
$\to$	directed adjacency relation
$d_G(y, x)$	graph distance between nodes x and y
$\psi(t)$	distribution of the walker's waiting-time
$\mathsf{U}(t)$	distribution of the up-time of an edge
$\mathsf{D}(t)$	distribution of the down-time of an edge
μ	rate of exponential distribution for the walker's waiting-time
η	rate of exponential distribution for the up-time of an edge
λ	rate of exponential distribution for the down-time of an edge
δ_p	metric based on L^p norms
$\delta_\square$	cut metric

$\mathbb{W}$	space of graphons after identification of U and V if $\delta_\square(U,V)=0$
$W/\mathcal{P}$	quotient graph from graphon W on partition $\mathcal{P}$
$W_{[n]}$	sampled sampled graph from graphon W with n nodes

Related to matrices and operators

$\mathbb{M}_n$	space of $n \times n$ matrices
$\mathrm{diag}(\boldsymbol{v})$	diagonal matrix of the entries of v
D	diagonal matrix of the strengths or degrees
L^{rw}	random walk Laplacian matrix of a graph
L	combinatorial Laplacian matrix of a graph
$\boldsymbol{J}$	Jacobian matrix
$\Omega(t)$	Magnus series
ϱ	spectral radius
$L^p[0,1]$	Lebesgue spaces on $[0,1]$ with the Lebesgue measure
$\|\cdot\|$	norm (definition depending on context)
T^*	adjoint of T
$R_\lambda(T)$	resolvent of T
$\rho(T)$	resolvent set of T
σ	spectrum
σ_{pp}	pure point spectrum
σ_{cont}, σ_{ac}	continuous, absolutely continuous spectrum
σ_{sing}	singular spectrum
s	spectral abscissa, spectral bound
ω_0	growth bound
$\mathcal{L}$	combinatorial Laplacian on graphons
$\mathcal{L}^{rw}$	random walk Laplacian on graphons
$\mathcal{L}^{norm}$	normalized (symmetric) Laplacian on graphons
$\mathcal{L}^{cons}$	consensus Laplacian on graphons
Df, df	Fréchet derivative, Gâteaux derivative of f

Relations, operations, and probabilities

$a \wedge b$	minimum between a and b
$a \vee b$	maximum between a and b
$a \bmod b$	a modulo b
$\lesssim$, $\gtrsim$	inegality up to a positive constant
$\|\cdot\|$	absolute value or cardinal
$\propto$	proportional to
Re	real part
Im	imaginary part
χ_S	indicator function of S
$f * g$	convolution between functions f and g
f^{*n}	convolution repeated n times, $f * f * \ldots * f$
$u \circ v$ or $\langle u, v \rangle$	scalar product between vectors of $\mathbb{R}^n$

(f, g)	scalar product between functions
$\otimes$	kronecker product
$\binom{n}{p}$	binomial coefficient
$\cup$	union of sets or events
$\cap$	intersection of sets or events
$\bar{E}$	complement of E
$P\{\cdot\}$	probability
$E\{\cdot\}$	mathematical expectation

CHAPTER 1

Introduction

1.1. Evolution equations on graphs

Linear diffusion problems on graphs and nonlinear variants resembling the semilinear heat equation make up the core of this thesis. They belong to the wider body of evolution equations which model the time evolution of a system, starting from an initial state. The motivation to study those equations on a discrete domain comes from the various fields in the natural sciences, medicine, economy, and many others where networks found their way. The marriage of evolution equations and graphs has produced a dynamic, expanding research community, with very broad interests. Hence, it is important to properly define where we stand, especially when it comes to the type of networks that support of our evolving systems.

In our setting, graphs are inherently discrete objects, made of edges or links representing interaction, or connectedness, or reachability between a set of vertices, but there is no notion of distance or travel-time associated to them. Proximity between nodes is therefore a matter of hop-count. This thesis is not concerned with what is called in the literature metric graphs or quantum graphs. Those graphs are continuous objects, where propagation across the links is ruled by a differential equation. Although largely relevant, metric graphs are only considered in the concluding chapter as a first-choice target for further research.

The evolution equations studied in this work originate in a continuum-domain version. Due to our choice of a discrete support, the differential diffusion operators that appear in the original version of the equations, will be replaced by matrices. However we will also work on equations defined on a continuum, and this will again affect the nature of the diffusion operator. Indeed, the last decade saw the development of graph-limit theory, an elegant mathematical framework for the continuum limit of graphs, when the number of nodes becomes arbitrarily large. The graph-limit version of the two main equations of this work is studied, hence closing the loop of going from a continuum medium to a discrete domain, then back to a continuum. In the process, the nature of the diffusion operator cycles from differential operator, to finite matrix, to integral operator.

To conclude this overview, let make three observations which helped focus our interest and inspired the content of the next chapters. Firstly, empirical evidence shows that real-life networks may vary over time. Secondly, we know that delay may impact diffusion between neighboring nodes. And finally, partially due to the ever increasing availability of data in general, and to some specific

fields of research, in some instances graphs tend to become very large objects. At least one of these considerations made its way into every problem studied in this dissertation.

1.2. How to read

Before we go any further, let us state our reading conventions. Instead of using a different symbol for all different objects, context differentiates between possible meanings. For instance, D denotes a diffusion coefficient, or diagonal matrix of degrees. The index on page xi lays out the symbology and abbreviations, sorted by category.

This document does not start with a preamble chapter dedicated to definitions, theorems, and other preliminaries. We instead tried to preserve the logical flow by introducing the mathematical concepts where they appear naturally in the storyline. This being said, this thesis revolves around graphs, random walks, and reaction-diffusion equations and we believe that as part of the introduction, a global understanding of those core concepts is beneficial for the reader. This will help set the stage and draw a clearer line between contextual information or standard material in this introduction, and contributions in the subsequent chapters. That is the goal of the next section.

This thesis contains many pointers to the literature, but we tried to make it technically speaking, mostly self-contained. To achieve that, we have included two appendices covering material designed to introduce the unfamiliar reader to elements of spectral theory and semigroups. Explicit reference to these appendices will be made along the text.

1.3. Core concepts

There are many excellent works on the matters of networks, random walks and reaction-diffusion. We elected to mainly follow [**134, 96**] for sections 1.3.1 and 1.3.2 about graphs and random walks, and [**96, 100, 101**] for section 1.3.3, which introduces reaction-diffusion processes.

1.3.1. Graphs and complex networks

Graphs, or also networks, model the adjacency relation between elements of a discrete set of objects. There appears to be an agreement in the terminology to refer with graphs to the set of vertices connected by edges, whereas for networks the terms nodes and links are used. For some authors, graphs would be the mathematical objects underlying real-life networks. We use those terms interchangeably.

Beyond terminology, what remains is the indisputable amount of past and recent successes of graph-theoretic tools in a wide array of applications. For instance, there is currently a sustained interest in biological networks, such as protein-protein interaction networks. One goal may be to identify motifs, namely

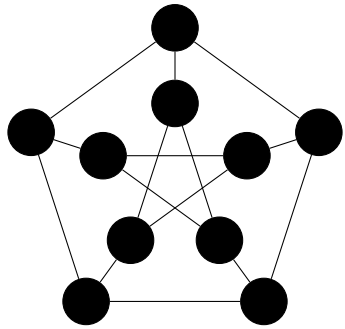


FIGURE 1.1. The Petersen graph is an unweighted connected symmetric graph with 10 vertices and 15 edges. It is regular because all nodes have the same degree equal to two. In spite of its apparent simplicity and modest size, the Petersen graph is well-known in graph-theory for providing many examples and counter-examples, due to its very specific properties.

sub-networks that appear a disproportionately large number of times given the overall structure. Another contemporary examples are climate networks, which allow to monitor and predict the effects of global warming. Networks help law enforcement and intelligence agencies to keep track of terrorist threats and identify emerging leaders. Banks use networks for fraud detection, and many of today's mathematical tools in graph-theory originate in applications in sociology.

In short, every time a situation involves entities interacting through links of some nature, networks offer a conceptually simple modeling approach, although the resulting graph may very well be a complicated object. Complex networks - as they are called - are not necessarily made of independent building blocks, and this does not go without analytical challenges, as reflected by the pages to come. In this section, we first provide some necessary formalism, and then comment on generalizations of networks that have emerged recently.

1.3.1.1. *Formalism*

We settle on the following notation: $G = (V, E)$ is a graph, where V is a finite set of vertices, and $E \subset V \times V$ is a set of edges. The vertices, also called nodes, are labeled by $\{1, 2, \dots, \}$ or by letters such as $\{v_1, v_2, \dots\}$. Two vertices v and w form an edge $[v, w]$ whenever $v \sim w$ where $\sim$ is an adjacency relation. For now we assume this relation to be symmetric,

$$v \sim w \iff w \sim v, \tag{1.1}$$

such that the resulting graph is undirected, see fig. 1.1. Let $|V|$ (resp. $|E|$) denote the number of vertices (resp. edges). The density ρ of the graph is the fraction of edges that are actually present, compared with the maximum possible number of connections:

$$\rho = \frac{|E|}{\binom{|V|}{2}}. \tag{1.2}$$

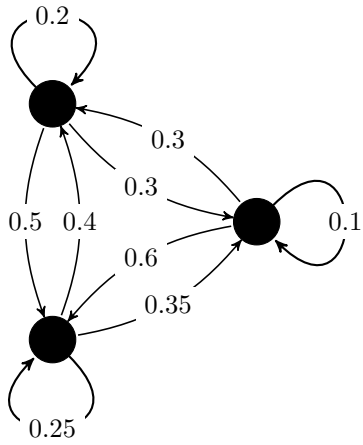


FIGURE 1.2. A weighted directed graph with three nodes. The graph is strongly connected. Each node has out-degree equal to three, and in-degree also equal to three. The weighted out-degree or out-strength is always one. The same holds true for the weighted in-degree.

When it makes sense to take the limit $|V| \to \infty$, one says the graph is dense if $|E| = O(|V|^2)$, and sparse otherwise.

Each edge may be attributed a weight, making the otherwise unweighted graph into a weighted one. The number of neighbors of node v is denoted by $\deg(v)$. In weighted graphs, $\mathrm{str}(v)$ denotes the weighted degree or strength of node v, that is to say the sum of the weights of all edges $[v, w]$ for which $w \sim v$. For the sake of simplicity, the notation $\mathrm{str}(v)$ will henceforth also apply to unweighted graphs and will refer to the degree, thereby identifying unweighted graphs with graphs with binary weights, either 0 or 1.

A path of length n between two nodes v, w is an ordered sequence of n nodes

$$[v_1, \ldots, v_n] \tag{1.3}$$

such that $v = v_1$, $w = v_n$ and $v_i \sim v_{i+1}$, $i = 1, \ldots, n-1$. A graph is connected if every pair of nodes is linked by a path. The graph distance $d_G(v, w)$ between two nodes v, w is the minimal length of all paths from v to w.

Let $\mathbb{M}_n$ be the space of $n \times n$ matrices. The adjacency matrix $A \in \mathbb{M}_n$ of a finite graph G with n vertices is the square matrix where A_{ij} is the weight of the edge between nodes with labels i and j and zero if no such edge exists. Unweighted graphs have binary adjacency matrix, where the ones indicate existing edges. We denote

$$D = \mathrm{diag}(\mathrm{str}(v_1), \ldots, \mathrm{str}(v_n)) \tag{1.4}$$

the diagonal matrix of the strengths (or degrees in unweighted graphs).

If G is actually a directed graph as on fig. 1.2, the adjacency relation is asymmetric and we write as $v \to w$ for an edge from node v to w, and call

$w \to v$ the reciprocal edge. Let $V_v = \{w \in V \mid v \to w \in E\}$ be the neighborhood of v, and its cardinal $k_v := |V_v|$ is called the out-degree of node v. Similarly, $V'_v = \{w \in V \mid w \to v \in E\}$ and $k'_v := |V'_v|$ is called the in-degree of v. We will refer to the weighted out-degree or out-strength (weighted in-degree, in-strength) when the edges have weights. The directed graphs in this work are almost always strongly connected, meaning every two nodes are linked by a directed path.

1.3.1.2. *Networks beyond networks*

The simplicity of networks as presented above has all but tempered the enthusiasm they generate nowadays. For example, in Physical Review Letters, American Physical Society's flagship publication that exists since 1958, the amount of network-science related papers published over the last decade amounts to approximately one third of all publications on the subject until 2009 [**99**]. Geographically speaking, it is interesting to note that besides China and the United States, the UK, Italy and Japan are the three countries to make it into the top 5 of the greatest contributors in terms of number of publications in the field since 2010.

It makes little doubt that networks in their original form will remain a useful tool in data science and other emerging fields. Moreover, generalizations have (re)surfaced in various directions, and keep gaining attention. For instance there is a growing interest in hypergraphs, and simplicial complexes in particular. These are objects where an edge can join any number of vertices, which hence go beyond pairwise interactions. Another current theme is that of higher order models, where one goal is to constraint individual paths in networks in an optimal way, thereby removing the transitivity property of the adjacency relation. Finally, viewing the network as a static entity may not fit to some empirical scenarios, and the temporal dimension needs to be accounted for. This temporal dimension is a main concern in chapters 2 and 3.

1.3.2. Random walks

Diffusion on networks is an extensive topic of research. Its archetype is the study of random walks on graphs, which are random processes that describe a path that consists in a succession of random steps of a walker across the edges. Despite their apparent simplicity, random walks remains an active domain of research [**40, 5, 3, 4, 74**], lying at the crossroads of probability, analysis, graph-theory. They are used as a baseline model for the diffusion of items or ideas on networks [**89**], but also as a tool to characterize certain aspects of their organization such as the centrality of nodes based on the density of walkers [**21, 80, 77**] or the presence of communities where a walker remains trapped for long times [**119, 118, 34, 76**]. Random walks find applications in biology, particle physics, financial markets, and many other fields [**9, 13**]. This diversity contributes to the numerous existing variants of random walks, including Levy flights [**68**], correlated walks [**116**], elephant walks [**123**], random walks in heterogeneous media [**47**], in crowded conditions [**7**], or even quantum walks [**67**] (see [**68, 58**] and the many references therein).

1.3.2.1. *Discrete time*

Consider first discrete time models and let X be a state space together with a transition matrix

$$\mathbf{T} = (T(y,x))_{x,y\in X}$$

where $T(y,x) \geq 0$ is the probability to move from x to y in one step, with normalization

$$\sum_{y\in X} T(y,x) = 1,\ \forall x \in X.$$

The couple $(X,\mathbf{T})$, together with a starting distribution on X, defines a Markov chain. Equivalently, we have a sequence $(Z_n)_{n\in\mathbb{N}}$ of X-valued random variables, where Z_n is the position in X at the n-th time step. The sequence $(Z_n)_{n\in\mathbb{N}}$ is defined on the sample space $\Omega = X^{\mathbb{N}}$, equipped with the product σ-algebra on Ω arising from the σ-algebra on X. For a Markov chain starting in $x \in X$, the discrete probability measure P_x is such that

(1.5)
$$P_x\{Z_n = x_n, Z_{n-1} = x_{n-1}, \ldots, Z_0 = x_0\} = T(x_n, x_{n-1})\ldots T(x_1,x_0)\delta_x(x_0),$$

where δ_x is the dirac distribution in x. As anticipated, a discrete-time Markovian random walk on a symmetric graph $G = (E,V)$ with adjacency relation $\sim$ is a Markov chain where the transition probabilities are adapted to the graph structure, for instance for an unweighted connected graph,

$$T(y,x) = \begin{cases} 1/\deg(x) & \text{if } y \sim x, \\ 0 & \text{otherwise.} \end{cases} \tag{1.6}$$

The scope of this thesis is limited to nearest-neighbor random walks, meaning $T(y,x) > 0$ only if $d_G(y,x) \leq 1$. This means teleportation of the walker, who is then allowed to traverse the whole network to easier reach peripheral nodes like happens in Google's PageRank algorithm, is not permitted. Further, our strong-connectedness assumption on G implies the Markov chain is irreducible, that is,

$$\forall x,y \in X,\ \exists n \in \mathbb{N} : T^n(y,x) > 0,$$

where $T^n(y,x)$ is the (y,x)-entry of the n-th matrix power of $\mathbf{T}$.

Not all discrete-time random walks $(Z_n)_{n\in\mathbb{N}}$ are Markovian in trajectory, or trajectory Markovian. Only those with trajectories satisfying (1.5), or put differently,

(1.7)
$$P_x\{Z_n = x_n \mid Z_{n-1} = x_{n-1}, \ldots, Z_0 = x_0\} = P_x\{Z_n = x_n \mid Z_{n-1} = x_{n-1}\}$$

are labeled as such.

1.3.2.2. *Continuous time*

This document concentrates on the more general continuous time versions, where space is still discrete but where the timings of the moves of the random walk generally follow a renewal process. Therefore, the walk is a stochastic process $\{Z(t)\}_{t\geq 0}$, where $Z(t)$ is X-valued and with $\tau_n := t_n - t_{n-1}$ an associated sequence $(\tau_n)_{n\in\mathbb{N}_0}$ of inter-arrival times.

A continuous-time random walk (CTRW) is semi Markovian, or trajectory Markovian, if for all $n \in \mathbb{N}$ and $s_1 < \ldots < s_n < t$ and all $x_1, \ldots, x_n, y \in X$ such that if $P_x\left\{\cap_{i=1}^n \left(Z(s_i) = x_i\right)\right\} > 0$, we have

$$P_x\left\{Z(t) = y \,|\, \cap_{i=1}^n \left(Z(s_i) = x_i\right)\right\} = P_x\left\{Z(t) = y \,|\, Z(s_n) = x_n\right\}. \tag{1.8}$$

meaning the underlying discrete-time process in the state space X,

$$Y_n = Z(t_n \leq t < t_{n+1}),\ n \geq 0, \tag{1.9}$$

is a Markov chain. Further, the walk is Markovian in time (or time Markovian), if it is a time-homogeneous process,

$$P_x\left\{Z(t+s) = z \,|\, Z(s) = y\right\} = P_x\left\{Z(t) = z \,|\, Z(0) = y\right\} \tag{1.10}$$

for all $x, y, z \in E$ and $s, t \geq 0$, in which case the timings of the moves follow a Poisson process.

Summing up, a Markovian CTRW on a graph with edge set E is a stochastic process where the underlying discrete-time chain on E is Markovian, and the time-homogeneous transition probabilities are adapted to the discrete support G. This thesis mainly contributes to random walks which elude this definition, since we start with semi Markovian CTRW's which are trajectory Markovian but not time Markovian, and eventually study models where even the trajectory Markov property is lost.

1.3.3. Reaction-diffusion equations

Reaction-diffusion (RD) equations proved a successful modeling tool based on the combination of a rate or kinetic equation with the heat equation. They lie at the heart of our understanding of spatial pattern formation and wave propagation in disciplines such as, but not limited to, ecology, chemistry, molecular biology or neuroscience. Indeed, Turing instabilities are one of the signature phenomenon present in RD equations and a central point of our interest. This phenomenon refers to the instability, seeded by diffusion, of a uniform steady state of the system, which may generate stable and stationary spatial patterns of concentration.

We start by introducing rate equations in section 1.3.3.1, and show in section 1.3.3.2 how they combine with diffusion to yield a complete reaction-diffusion model. We then discuss in section 1.3.3.3 the validity of those models including their limitations and possible workarounds.

1.3.3.1. *Rate equations*

Rate equations may describe the evolution of the concentration of particles, species or reactants, depending on the application. They have the form

$$\frac{d\boldsymbol{u}}{dt} = F(\boldsymbol{u}) \tag{1.11}$$

where $\boldsymbol{u}$ is an n-components real vector function of time, F is a rate or kinetic function typically depending on one or more real parameters. The conditions

$$F(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_n) \geq 0 \tag{1.12}$$

ensure the rate function F is positivity preserving, meaning that if $\boldsymbol{u}(0) \geq 0$, then $\boldsymbol{u}(t) \geq 0$ for all $t > 0$ [**96**]. For all our numerical examples in the main text, we have chosen to use two standard models from the literature, one in chemistry dubbed the Brusselator, and one in ecology that is called the Mimura-Murray model of interaction.

Brusselator. The Brusselator is a chemical scheme introduced by Prigogine and Lefever [**113**], will be our main example throughout this document. It is the two-variable, four-step model

$$\begin{aligned} \mathrm{A} + \mathrm{U} &\xrightarrow{k_1} \mathrm{U} \\ \mathrm{B} + \mathrm{U} &\xrightarrow{k_2} \mathrm{V} + \mathrm{D} \\ 2\mathrm{U} + \mathrm{V} &\xrightarrow{k_3} 3\,\mathrm{U} \\ \mathrm{U} &\xrightarrow{k_4} \mathrm{E} \end{aligned}$$

with rate equations

$$\frac{d[U]}{dt} = k_1[A] - (k_2[B] + k_4)[U] + k_3[U]^2[V] \tag{1.13a}$$

$$\frac{d[V]}{dt} = k_2[B][U] - k_3[U]^2[V], \tag{1.13b}$$

where $[X]$ denotes the concentration of X. Scaling time $k_4 t \mapsto t$ and defining new the variables,

$$a = \sqrt{\frac{k_1^2 k_3}{k_4^3}}[A], \quad b = \frac{k_2}{k_4}[B], \quad u_1 = \sqrt{\frac{k_3}{k_4}}[U], \quad u_2 = \sqrt{\frac{k_3}{k_4}}[V], \tag{1.14}$$

the rate equations become

$$\frac{du_1}{dt} = a - (b+1)u_1 + u_1^2 u_2, \tag{1.15a}$$

$$\frac{du_2}{dt} = bu_1 - u_1^2 u_2. \tag{1.15b}$$

The Jacobian matrix $\boldsymbol{J}$ of partial derivatives of the rates evaluated at the steady state $u_e = (u_{e,1}, u_{e,2}) = \left(a, \frac{b}{a}\right)$ is given by

$$\boldsymbol{J} = \begin{pmatrix} b-1 & a^2 \\ -b & -a^2 \end{pmatrix}, \tag{1.16}$$

indicating that if $b > 1$, $\operatorname{sign} \boldsymbol{J} = \left(\begin{smallmatrix} + & + \\ - & - \end{smallmatrix}\right)$ indicating the Brusselator is a cross activator-inhibitor type.

Mimura-Murray. Our secondary scheme is the Mimura-Murray biological model of interactions, selected for its use as a benchmark in the literature. It is given by

$$f(u_1, u_2) = \left(\frac{a + bu_1 - u_1^2}{c} - u_2\right) u_1 \tag{1.17a}$$

$$g(u_1, u_2) = (u_1 - (1 + du_2))\, u_2, \tag{1.17b}$$

where a, b, c, d are positive parameters. Out of the six equilibrium points, two are always unstable, and two more are trivial with one of the variables u_1, u_2 being zero. The remaining two read

(1.18a)

$$(u_{e,1}, u_{e,2}) = \left(1 + \frac{1}{2d}\left(bd - c - 2d + \sqrt{\Delta}\right), \frac{1}{2d^2}\left(bd - c - 2d + \sqrt{\Delta}\right)\right)$$

(1.18b)

$$(u_{e,1}, u_{e,2}) = \left(1 + \frac{1}{2d}\left(bd - c - 2d + \sqrt{\Delta}\right), \frac{1}{2d^2}\left(bd - c - 2d + \sqrt{\Delta}\right)\right)$$

with $\Delta = (bd - c - 2d)^2 + 4d^2(a + b - 1)$. The Jacobian matrix given by

$$\boldsymbol{J} = \begin{pmatrix} 2\frac{b}{c}u_{e,1} - \frac{3}{c}u_{e,1}^2 - u_{e,2} - \frac{a}{c} & -u_{e,1} \\ u_{e,2} & u_{e,1} - 2du_{e,2} - 1 \end{pmatrix}. \tag{1.19}$$

1.3.3.2. *Local kinetics and diffusive coupling combined*

Even if we are mainly interested in two-component reaction-diffusion systems on graphs, we will show in this section how to derive the RD equation for a single variable on a continuum, based on a conservation law and Fickean diffusion. So consider the introductory instance of a single particle with density $u(x,t)$ in a volume $V \subset \mathbb{R}^3$. If J is the particle flux, and F the net rate of production (birth and death combined), we have

$$\frac{\partial}{\partial t}\int_V u(x,t)dV = -\int_S J \cdot ds + \int_V F(u,x,t)dV, \tag{1.20}$$

where S is the boundary of V. The divergence theorem allows to write

$$\int_S J \cdot ds = \int_V \nabla \cdot J dV, \tag{1.21}$$

and (1.20) becomes

$$\int_V \left(\frac{\partial}{\partial t}u + \nabla \cdot J - F\right) dV = 0. \tag{1.22}$$

Since V is arbitrary, a conservation equation for u is obtained:

$$\frac{\partial}{\partial t}u + \nabla \cdot J - F = 0. \tag{1.23}$$

Fick's first law of diffusion asserts that

$$J = -D\nabla u \tag{1.24}$$

where D is the diffusion coefficient, in units of area over time. If this law applies (1.23) becomes

$$\frac{\partial}{\partial t}u = F + \nabla \cdot (D\nabla u), \tag{1.25}$$

where as noted F may be a function of u, x, t and D a function of x and t.

Equation (1.25) allows an extension to a vector (multi-species) setting, where D then is a matrix where off-diagonal terms indicate cross-diffusion. In this

work, D will always remain a diagonal matrix of constant coefficients, such that in the scalar case of (1.25), we can write

$$\frac{\partial}{\partial t}u = F + D\nabla^2 u. \tag{1.26}$$

Building on this example, it appears legitimate to add a Fickean diffusion term to (1.11), resulting in a general equation of the form

$$\boldsymbol{D}_t\boldsymbol{u} = F(\boldsymbol{u}) + D\boldsymbol{\Delta}\boldsymbol{u}, \tag{1.27}$$

where $\boldsymbol{D}_t$ is time-derivative operator, and $\boldsymbol{\Delta}$ is a diffusion operator. If space is continuous, then eq. (1.27) is a partial differential equation (PDE), $\boldsymbol{D}_t = \frac{\partial}{\partial t}$ is a partial time-derivative, and $\boldsymbol{\Delta}$ is typically a second order differential operator. For instance, $\boldsymbol{\Delta}$ is the Laplacian operator

$$\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

in a three-dimensional Euclidean space in the Cartesian coordinates x_1, x_2, x_3 as in our example. Otherwise space is discrete, (1.27) is an ordinary differential equation (ODE) , $\boldsymbol{D}_t = \frac{d}{dt}$, and $\boldsymbol{\Delta}$ is a matrix operator.

1.3.3.3. *Validity of RD models*

We first discuss possible avenues to explain the combination of a rate and a diffusion equation in RD models. Next we question the basic assumptions of these models, and describe the most common extensions.

Macroscopic, mesoscopic and microscopic descriptions. There exist different levels of justifications for combining two terms in the right-hand side of eq. (1.27). Phenomenological explanations are based on the laws of conservation, and eqs. (1.20) to (1.25) form an example of such argument. Equation (1.27) is a macroscopic equation, because it can be obtained based on a long-time, large-scale limit procedure applied to a mesoscopic description. Such mesoscopic description follows from the microscopic rules, that is, the particle-level motion rules, where random fluctuations are averaged out of the resulting mean-field equations. This two-step procedure conveniently assumes that the microscopic jump lengths and times are small comparatively to the macroscopic scales, an assumption which may fail to hold [**96**, Chapter 3]. But when it succeeds, the mesoscopic derivation ensures a sound mathematical model, where the physical constraint of positivity and normalization of densities is guaranteed by construction. We are limited to mesoscopic equations in this work since the combined long-time, large-scale limit toward the macroscopic description does not make sense on graphs which are inherently discrete objects.

Disputed assumptions. In spite of their widespread use, two core assumptions of reaction-diffusion equations may fail to hold in real-life systems. The first one is spatial homogeneity, which requires all reacting units to be identical. The second one is transport following Fick's second law of diffusion, which

is physically unrealistic. Indeed, recall that in a 1-dimensional setting the fundamental solution to

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{1.28}$$

with initial condition $u(x,0) = \delta(x)$ is given by

$$u(x,t) = (4\pi Dt)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{4Dt}\right) \tag{1.29}$$

for all $t > 0$. This solution implies infinitely fast spreading of the initial disturbance, $u(x,t) > 0$ for all x and all $t > 0$. Mathematically, this pathology is due to the parabolic nature of (1.28). From a phenomenological point of view, it is due to Fick's first law (eq. (1.24)) which implies an instantaneous adaptation to the gradient of the concentration. And from a mesoscopic perspective, Brownian particles have no inertia since their directions of motion in however small successive time intervals are uncorrelated. These model-related properties can each be addressed at their level. Firstly, adding a small term to the left-hand side of (1.28) turns it into an hyperbolic PDE,

$$\tau\frac{\partial^2 u}{\partial t^2} + \frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{1.30}$$

the reaction-telegraph equation. Loosely speaking, the solution of this equation approaches that of (1.28) when $\tau \to 0$. Moreover the solution of (1.30) with a point source at $x = 0$ in $t = 0$ is zero on $|x| \geq \sqrt{\frac{D}{\tau}}t$. Secondly, Cattaneo's equation

$$\tau\frac{\partial J}{\partial t} + J = -D\frac{\partial u}{\partial x} \tag{1.31}$$

may replace Fick's first law. Finally, microscopic models of persistent random walks may be designed [**96**, Chapter 2]. This is beyond the scope of this work. However, when it comes to so-called Turing instabilities studied in chapters 5 to 7, the conclusions in the unmodified RD system might carry over [**96**, Chapter 10].

1.4. Aim

A central question when network science was still a novel, rapidly expanding field of research, was to understand the structural properties of complex networks. While a lot of progress has been made on this front, the next step - at least from a dynamical systems point of view - is to understand how the structure of the graph affects various properties of the dynamics. Even in the simplest case of symmetric, unweighted networks, this problem proved both multi-faceted and fascinating. In addition, only a fraction of real-life graphs are symmetric, unweighted, static, and computationally speaking of reasonable size. The need to encompass this reality in the study of dynamical systems on graphs is not open to debate, and is the background motivation for each of the next three research questions.

1.4.1. First question

When networks model the shifting structure of relationships between people, ideas, or even locations, they inherently change in time. These temporal networks have been at the center of attention recently. Researchers have logically started with the study of linear diffusion processes and synchronization phenomena on simple models of such networks. These existing models are mostly single timescale models, which does not necessarily reflect reality. In many outcomes, for instance in some contact networks, it appears that the evolution of the network is the result of more that one independent process, each characterized by its own timescale. On top of that, considering diffusion, the moving entities obey a set of microscopic rules with yet another independent characteristic timescale.

The aim is to first design new random walk models that account for the shortcomings of the single timescale walks, in order to have a tool at our disposal to capture and study, in a second part of the same question, the main differences with respect to the existing walks. The overall objective is to gain a better understanding of diffusion on temporal networks, relying on the study of these evolved walks.

1.4.2. Second question

This work is evenly split between the study of linear and nonlinear problems, but questions to address and ideas to solve them may cross that organizational boundary. It is such a crossing that will produce the second research question, which necessitates two preliminary observations. On the one hand, the practical relevance of Turing instabilities is recurrently challenged because the conditions for the emergence of patterns are too restrictive to be satisfied out of a fully controlled ad-hoc environment. On the other hand, as detailed in the previous theme of research, diffusion is not confined to static graphs. And besides that, diffusion is not immune to time-delay. Indeed, generally speaking delays are key to our understanding of the interactions between neurons in biology, in models of traffic flow, and cannot be neglected in many applications of diffusion, propagation, epidemics, or rumor spreading.

The aim is to determine whether temporal variations of the graph, and the ensuing modification of the diffusion operator, may loosen the conditions for the emergence of Turing patterns. In the same line of investigation, as a side-question we want to find out whether diffusion-driven instabilities may arise due to another modification of the diffusion operator: the inclusion of a processing delay.

1.4.3. Third question

In certain practical cases the size of the graph - its number of nodes - is extremely large. Network models of the brain or of the internet, where each page is a different node, are two common examples. There have been attempts

to study the continuum limit of dynamical systems on networks of arbitrary large size, but what was needed first was a rigorous mathematical framework to let this idea operate. When graph-limit theory emerged approximately a decade ago, it started filling that gap neatly. This theory provides a concept of continuum limit for sequences of graphs of increasing size. Not long after it appeared, graph-limit theory was put to work to establish the validity of the continuum-limit approximation to fundamental linear and nonlinear evolution equations on graphs, including the heat equation and the scalar reaction-diffusion equation. There are however two natural topics to address in this work, that have remained untouched so far.

The aim of the first topic is to prove the convergence of the continuum limit scheme for the most basic random walk model, the so-called node-centric walk, and to study the resulting problem on a continuum. Concerning the second topic, we want to determine the conditions for Turing instabilities in a RD system, applied directly to the problem on the continnum. Keep in mind that the questions of convergence have already been affirmatively answered for various problems where diffusion is encoded by the combinatorial Laplacian, including the (non)linear heat equation and the reaction-diffusion equation (or semilinear heat equation). This second point is two-faceted: it comprises the determination the spectral properties of the Laplacian operator and the formulation of a principle of linearized stability, which will be affected by the infinite-dimensional nature of the problem.

1.5. Structure and contributions

The goal of the section is to outline the content of each individual chapter. This will allow the reader to make connections between the chapters and the different research questions, as represented by fig. 1.3. The descriptions of the chapters also give a quick summary of the contributions, which will eventually be reviewed in the conclusion in greater detail.

As the coloring of the nodes on fig. 1.3 shows, the first part of the thesis is about linear models of diffusion on (in)finite graphs, and comprises chapters 2 to 4.

Chapter 2 : We consider random walks on dynamical networks where edges appear and disappear possibly during finite time intervals. The walks are grounded on up to three independent stochastic processes: one determining the walker's waiting-time, associated to the edges. We determine the transition kernels and then compare the mean resting times and stationary states of the different models, and pay particular attention to the emergence of non-Markovian behavior, even when all dynamical entities are governed by memoryless distributions.

Chapter 3 : We first propose a comprehensive analytical and numerical treatment of a generic three-timescales walk introduced in chapter 2 on directed acyclic graphs. Once cycles are allowed in the network, memory may emerge, remarkably even if the walker and the evolution

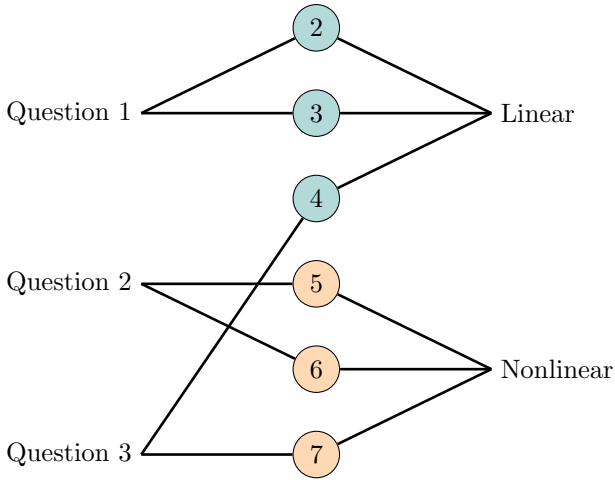


FIGURE 1.3. The thesis at a glance. Each node labeled 2 to 7 refers to a chapter. The diagram indicates the research question and the type of equation addressed by each chapter.

of the edges are governed by Poisson processes. We then introduce a general analytical framework to characterize such non-Markovian walks and validate our findings with numerical simulations.

Chapter 4 : We explore the continuum limit of random walks by adopting a methodology based on graph-limit theory. Through an approximation procedure, the standard system of coupled ordinary differential equations is replaced by a nonlocal evolution equation on the unit interval. We focus on Markov chains on dense weighted graphs whose dynamics are encoded in the transition matrix and the random-walk Laplacian. After having established the well-posedness of the continuum problem, we prove the convergence of the approximation procedure in the case of a convergent sequence of dense graphs. We also consider the case when the problem on a graph results from two different types of discretization of the continuum version. We then apply the spectral theory of operators to characterize the relaxation time of the process in the continuum limit. We finally show that our results can straightforwardly be extended to the discrete time walk.

The second part, made up of chapters 5 to 7, is devoted to nonlinear models.

Chapter 5 : Turing and Hopf bifurcations in networked RD systems are introduced. A analytical expression of all possible stable steady states in a two-component system is obtained in the neighborhood of the bifurcation point. The effect of a processing delay in the diffusive coupling is briefly discussed in the scalar and vector cases. We demonstrate the emergence of instabilities seeded by the delayed-diffusion operator.

Chapter 6 : We study Markovian diffusion on periodic temporal networks when it is combined with local reactions to yield a reaction-diffusion on a time varying graph. In particular, we determine conditions for the emergence of diffusion-driven instabilities in this scenario. Our main result shows that spatial instabilities may unexpectedly arise in RD systems on time-varying networks, in the sense that the linear stability analysis would conclude on the impossibility of such phenomenon if the topology were to be frozen at any random time. An instability threshold on the timescale of the evolution of the graph is obtained, both for switched and continuously-varied networks.

Chapter 7 : This chapter mirrors the graph-limit approach to Markovian random walks in chapter 4. The validity of the continuum-limit approximation being well-established, we concentrate our effort on achieving a principle of linear stability for reaction-diffusion equations on graphons, by applying results of spectral to the combinatorial Laplacian. Illustrating the theory with examples, we prove that the conclusions about stability with the linearized problem may or may not hold depending on the underlying space of the initial perturbation. The conclusions require only mild differentiability conditions on the generator of the nonlinear semigroup associated to the continuum RD equation.

Finally, chapter 8 outlines the conclusions and is an opportunity to elaborate on future research directions. Appendices A and B complement chapters 4 and 7 by providing background material surrounding spectral measures and stability-related properties of linear semigroups and their generators.

1.6. Publications

This work is partially based on a list of papers listed below and (co)written since the start of this thesis. The main sources for chapters 2 and 3 are papers [5] and [6] below. We would like to acknowledge that part of the analytical derivation in [5] is the result of close collaboration with Martin Gueuning. The content of Chapter 4 matches that of [7] intimately. Chapter 5 comprises some material from [1,2], and the content of chapter 6 originates in [3]. This paper subsequently lead to the variation in [4], for which my co-authors deserve the most credit. Chapter 7 contains only unpublished work.

Journal Publications

[1] Julien Petit, Timoteo Carletti, Malbor Asllani, and Duccio Fanelli. Delay-induced turing-like waves for one-species reaction-diffusion model on a network. Europhysics Letters, 111(5):58002, 2015.

[2] Julien Petit, Malbor Asllani, Duccio Fanelli, Ben Lauwens, and Timoteo Carletti. Pattern formation in a two-component reaction-diffusion system with

delayed processes on a network. Physica A: Statistical Mechanics and its Applications, 462:230249, 2016.

[3] Julien Petit, Ben Lauwens, Duccio Fanelli, and Timoteo Carletti. Theory of turing patterns on time varying networks. Physical review letters, 119(14):148301, 2017.

[4] Maxime Lucas, Duccio Fanelli, Timoteo Carletti, and Julien Petit. Desynchronization induced by time-varying network. Europhysics Letters, 121(5):50008, 2018.

[5] Julien Petit, Martin Gueuning, Timoteo Carletti, Ben Lauwens, and Renaud Lambiotte. Random walk on temporal networks with lasting edges. Physical Review E, 98(5):052307, 2018.

[6] Julien Petit, Renaud Lambiotte and Timoteo Carletti. Classes of random walks on temporal networks with competing timescales. Applied Network Science 4, 72, 2019.

Preprint

[7] Julien Petit, Renaud Lambiotte and Timoteo Carletti. Random walks on dense graphs and graphons, arXiv:1909.11776, 2019.

CHAPTER 2

Mesoscopic foundations of diffusion on temporal networks

2.1. Introduction

The mesoscopic foundations of diffusion rely on random walk models made of a set of microscopic motion rules at the level of individual particles. The mathematical properties of random walks on static networks are overall well-established [**83**] and, under assumptions described section 1.3.2, are essentially equivalent to those of a Markov chain. However, in real world scenarios a core assumption that the network is a static entity, is often disputed. Empirical evidence shows instead that the network should be regarded as a dynamical entity, with edges appearing and disappearing in the course of time [**57, 88, 56**].

As part of our first research question, in this chapter we introduce and analyze models of random walks that interconnect the dynamics of the walker, namely the diffusing entity, and of the network, each contributing to the problem with their own timescale(s). In particular, we focus on the mostly unexplored scenario when the edges remain active - meaning they are available to the jumper - over finite time intervals.

2.1.1. Temporal networks

The underlying object of a temporal network is a dynamical graph.

Definition 2.1 (Dynamical graph)**.** Let $\mathcal{G} = (V, E)$ be a (static) underlying graph and let

$$\mathcal{A} = \{A_\ell\}_{\ell=1}^{m}\,, \tag{2.1}$$

with $m = 2^{|E|}$, be the set of adjacency matrices A_k corresponding to the possible configurations or subgraphs $\mathcal{G}' = (V, E')$, $E' \subset E$, allowed by $\mathcal{G}$. A dynamical graph is a random marked point process

$$(A_{\ell_k}, \tau_k)_{k\in\mathbb{Z}}, \quad \ell_k \in \{1, \dots, m\} \tag{2.2}$$

with mark space $\mathcal{A}$ and (τ_k) a sequence of rewiring times, which is symmetric (resp. directed) if $\mathcal{G}$ is symmetric (resp. directed).

Using the notation of the definition, the time-dependent adjacency matrix is written as

$$A(t) = \sum_{k\in\mathbb{Z}} A_{\ell_k} \chi_{[\tau_k,\tau_{k+1})}(t), \tag{2.3}$$

with χ the indicator function. We denote $k_i^{\text{out}}(t) = \sum_{j=1}^{N} A_{ij}(t)$ the out-degree and $k_i^{\text{in}}(t) = \sum_{j=1}^{N} A_{ji}(t)$ the in-degree of node i at time t.

We consider directed, model-driven temporal networks[1] where the underlying graph is strongly connected and where the duration of the up-times meaning the edge is available for transport (resp. down-times, edge unavailable) is randomly distributed.

Assumption 2.2 (Distributed up-times and down-times)**.** The time evolution of the edges is modeled by independent and identically distributed random variables. The up-time (resp. down-time) of an edge is the random variable $X_u(t)$ (resp. X_d) with PDF $\mathsf{U}(t)$ (resp. $\mathsf{D}(t)$) with support in $\mathbb{R}^+$ and finite mathematical expectation.

The benefit of this restriction is simplified calculations without jeopardizing the intended range of the conclusion.

2.1.2. Single-timescale random walks

We start with single-timescale random walks, where $\mathsf{U}(t)$ and $\mathsf{D}(t)$ may actually be dirac distributions in zero. We will explain two types of binary classifications, and present the three standard models relying on these classifications.

2.1.2.1. *Classification*

The impact of the temporal properties of networks on diffusive processes have been explored by means of numerical simulations [**64, 128, 107**] and analytical tools [**33**]. In those works, the distinction between the dynamics on the network and the dynamics of the network is reflected by a standard, two-faceted classification of random-walk processes [**89**]:

Node-centric *vs* edge-centric: It is generally relevant to distinguish between node-based and edge-based dynamics [**112**]. In a node-centric walk, a stochastic process occurring at the level of the node determines the duration before the next jump. In an edge-centric model, also known under the intuitive denomination of fluid model, the links are the driving units. They become available for transport, then vanish, according to their stochastic process.

Active *vs* passive: One draws a line between active and passive models. The walk is active when the waiting time before the next move is reset by each jump of the walker. Active walks are common models for animal trajectories. On the other hand, in passive models the

[1] As opposed to data-driven temporal networks.

motion of the walker is instead constrained by the temporal patterns of (typically the edges of) the network. An example would be that of a person randomly exploring a public transportation network, taking every available ride.

Put differently, there are clocks, either on the nodes or on the edges (node-centric vs edge-centric models), and either the clocks are reset following each jump of the walker or they evolve in an independent manner (active vs passive models).

2.1.2.2. *Standard models*

Let us first briefly review the three classical models of continuous-time random walks. Their respective sets of microscopic motion rules are described by the three panels of figure 2.1.

Model 1 : active node-centric model where the walker resets the clock of a node upon arrival on the node. One may think of the clock as being attached to the walker, and as obeying the walker's dynamics. The waiting time on a node corresponds to the random variable X_w with PDF $\psi(t)$. In this work, we will concentrate on the simpler case that the renewal counting process of the jumps is a Poisson process with rate μ, allowing us to bring out effects due to walker-network interaction. The variable X_w represents the active, node-level feature in all models.

Model 2: active edge-centric model. The walker accordingly resets the down-time of the edges leaving a node, upon arrival on that node. This down-time is the random variable X_d and determines the period of unavailability of the edge.

Model 3: again an edge-centric model with a passive walker, who passively follows edge activations.

The fourth combination corresponding to the two classification criteria, the passive node-centric walk, appears to be irrelevant for practical applications [**89**]. We won't discuss it further.

2.1.3. Beyond standard models

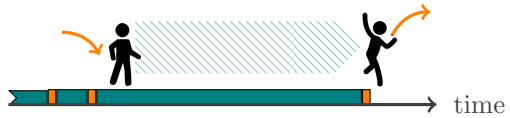
The first point of this section is to elaborate on the motivations for introducing new models. The second point adds to the argument by giving details about possible applications. The third point is a quick overview of the method.

2.1.3.1. *Decoupled walker and transport layer*

The standard models come short when the dynamics of the walker is decoupled from the transport layer represented by a dynamical network, such that the diffusive process is effectively a combination of active and passive diffusion with independent node-centric and edge-centric processes combined. Hence a first objective of this chapter is to formulate natural extensions of existing random

(A) Model 1. Active node-centric

(B) Model 2. Active edge-centric



(C) Model 3. Passive edge-centric

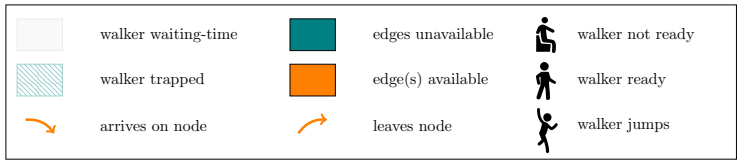


FIGURE 2.1. Microscopic motion rules for the three classical single-timescale models

walks, that arise when the walker does not necessarily jump through an available link. Instead, it is constrained by its own waiting time after each jump.

On top of that, the network is such that the duration of the links between the nodes is not instantaneous [**45, 124**]. This is in contrast with a majority of approaches which assume that the network evolution can be modeled as a point process [**54**]. By doing so, we allow for the possibility to investigate the importance of a timescale associated to edge duration, which has plenty of applications in real-life systems. Take contact networks and their impact on epidemic or information spreading as a canonical example. For instance, recent research on empirical face-to-face network data collected via Bluetooth was performed by the authors of [**124**]. They looked into the predictability of the interactions between a large group of individuals taking part in the study. It is representative that their work would mostly rely on data, and doesn't go as far as spreading processes, where our modeling would come in handy. In many cases, the durations of availability timespans of the edges are inevitably finite and, as results from [**10**] and a wealth of publications relying on data collected by the SocioPatterns initiative [**25, 120**], even feature a long-tailed distribution. This heterogeneity prohibits a well-defined timescale for the interactions, but can be captured by the proposed models. In engineering, practical applications

include peer-to-peer and proximity networks of mobile sensors with wireless connections.

2.1.3.2. *Applications*

Different properties such as recurrence and transience, and quantities such as hitting times and mean first passage time help characterize random walks. We choose to focus on the mean resting time, that is, the expected total waiting time on the nodes, with the motivation that this is an appropriate measure for the speed of spreading processes in the different models. Because random walks can be used to determine the most central nodes in a network, a final application will be the analysis of the rankings of the nodes in the steady states resulting from the different models.

Overall, given the central role of random walks in the design of algorithms on networks, our results open the way to generalize standard tools such as PageRank for centrality measures [**21**] and Markov stability for community detection [**34**].

2.1.3.3. *Method*

Our approach is based on the generalized integro-differential master equation obtained for semi-Markovian random walks on temporal networks [**54**]. This equation requires to determine the distribution of the resting time on the nodes. By direct integration, we then obtain the mean resting time. We also compare the different models, and evaluate the dominating timescales in regimes of extreme values for the dynamical parameters. Finally, combining an asymptotic analysis of the master equation and the resting times leads to the steady state for the different models. The analysis up to that point does not take into account a possible memory effect arising from walker-network interaction, which is the subject of the next chapter.

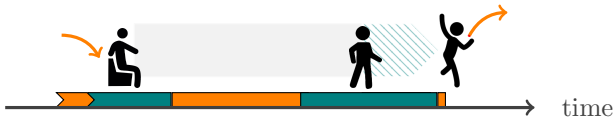
2.2. Mesoscopic foundations

We introduce and classify new random walk models with up to three timescales represented by fig. 2.2, and discuss their master equation.

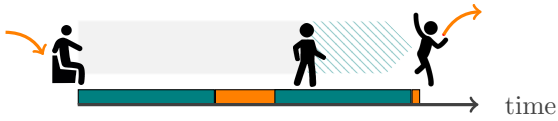
2.2.1. Multiple-timescales random walks

Following assumption 2.2, the timescales are well-defined and are correctly represented by the expectation of the random variables. This assumption holds for exponential distributions, but wouldn't for example apply for power-law distributions. Each model features an active walker and we consider different behaviors at edge level, as depicted by fig. 2.2.

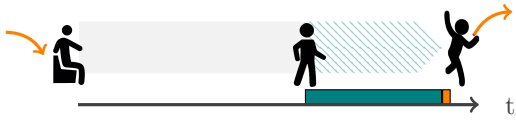
Model 4 : The edge-level dynamics is passive and each edge cycles through states of availability followed by periods of unavailability. The duration of the up-times is a random variable X_u, and the duration of the down-times is random variable X_d, with already-introduced



(A) Model 4. Active at node level and passive at edge level



(B) Model 5. Active at node and edge level



(C) Model 6. Active at node and edge level

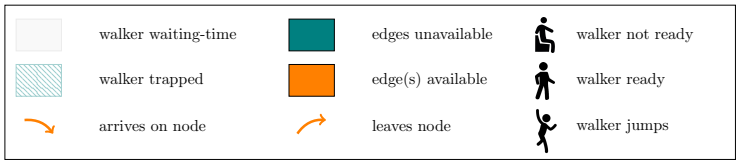


FIGURE 2.2. Microscopic motion rules for natural extensions of the classical random walks, with three (models 4 and 5) or just two (model 6) competing timescales.

PDF's $\mathsf{U}(t)$ and $\mathsf{D}(t)$. When the walker is ready to jump, there are two possibilities. If one or more outgoing edges are available, a new destination node is selected randomly without bias. The total waiting time on the node is thus X_w. If no outgoing edge is available, then the walker is trapped on the node, and waits until the next activation of an edge[2].

Model 5 : This time the model is active at node and edge level. At the beginning of the walker's waiting time, the state of each outgoing edge is reset to unavailable, for a duration X_d. Then follows a period of

[2]Another behavior would have been that the ready-to-jump but trapped walker waits for another period drawn again from the distribution of X_w, before attempting another jump. In this scenario, the induced delay before the jump also depends on the dynamics of the walker and not only on that of the network through availability of edges at the end of the prolonged stay [**42**]. The analysis of the model would go along the lines presented here.

availability X_u, then a down-time, and so on, until after a duration X_w the walker is ready and an edge is eventually available.

Model 6 : The third declination is similar to the previous one in the sense that the walker also actively resets the transport layer, but this reset occurs at the end of the walker's waiting time X_w. Therefore, the walker is always trapped for a duration X_d and the dynamics only has two timescales. This model relatively slows down the walker and will serve as a baseline to compare the previous two models.

The mathematical modeling is expectedly more involved than with the single-timescale models 1, 2 or 3, and is presented in the next section.

2.2.2. Master equation

Starting from the microscopic motion rules, the mesoscopic modeling produces a master equation for the evolution of the walker's position across the network. This position is encoded in the row vector of residence probabilities on the N nodes of the graph,

$$n(t) = (n_1(t), n_2(t), \dots, n_N(t))$$

with

$$n_i(t) = P\{Z(t) = i\}, \quad i = 1, \dots, N, \tag{2.4}$$

where $Z(t)$ is the location of the walker at time t. Once the equation is obtained, the analysis of various properties of the walk becomes accessible.

2.2.2.1. *Markovian random walks*

The active node-centric walk, model 1, is arguably the simplest to study among continuous-time random walks. It is Markovian when the waiting time on the node is exponentially distributed. If the rate on node j is μ_j, the walk is governed by the well-known master equation [**4**]:

$$\dot{n}_i(t) = \sum_{j=1}^{N} \mu_j n_j \frac{1}{k_j} A_{ji} - \mu_i n_i, \tag{2.5}$$

If the rate is the same on all nodes, $\mu_j = \mu$ for all j, and after scaling of the time variable $t \mapsto \mu t$, equation (2.5) reads

$$\dot{n} = nD^{-1}(A - D) = nL^{rw}, \tag{2.6}$$

where D denotes the diagonal matrix of the out-degrees. The matrix

$$L^{rw} = D^{-1}A - I \tag{2.7}$$

is the so-called random walk Laplacian of the graph.

An alternative modeling for the process of model 1 is to consider that the walker is always ready to jump, and that the edges activate like in the edge centric walk of model 2. To illustrate this fact, let us start from model 1 where we assume that the rates on the nodes are proportional to degree: $\mu_j = \lambda k_j$, so

that the rate of jump across each edge of the graph is the same, independently of the degrees of the nodes. The resting time on node j is the random variable

$$\mathbb{X}_d^{(j)} = \min\left\{ X_d^{\ell \leftarrow j} \,\middle|\, \ell \in V_j \right\}, \tag{2.8}$$

where the superscript in the right-hand side explicitly refers to the edges. Equation (2.8) refers to the minimum of k_j exponential distributions with rate λ, which again follows an exponential distribution with rate $k_j\lambda$. Equation (2.5) becomes

$$\dot{n}_i(t) = \sum_{j=1}^{N} n_j \lambda A_{ji} - \lambda k_i n_i. \tag{2.9}$$

In matrix form, recalling that n is a row vector we have $\dot{n} = \lambda n(A - D)$, or again, after scaling of time with respect to $1/\lambda$,

$$\dot{n} = nL, \tag{2.10}$$

where this time

$$L = A - D \tag{2.11}$$

is known as the graph or combinatorial Laplacian. Equation 2.10 is well known to correspond to model 2, where λ would be the rate of the exponential distribution governing the time before activation of each edge. This shows that the same process - in terms of trajectories - can be seen as happening atop of a static graph, or a temporal graph. The walk on a static graph generates a temporal network where jumps across edges are considered as edges activation.

Remark 2.3. Note again that this freedom of the choice of modeling exists thanks to the restrictive framework of exponential distributions. Also observe that eqs. (2.6) and (2.10) emerge as particular cases of the general eq. (2.21), which effectively differentiates between the node-centric and edge-centric equations when the exponential assumption is relaxed.

Remark 2.4 (On the time-dependent Laplacian)**.** In previous works dealing with synchronization [**130, 12**], desynchronization [**86**], but also in reaction-diffusion systems in chapter 6 and other works about dynamical systems on time-varying networks [**136, 56**], a time-dependent Laplacian $L(t)$ has replaced the usual graph Laplacian L in the equations such that (2.10) becomes

$$\dot{n} = nL(t). \tag{2.12}$$

In our framework, we consider this equation as associated to a passive edge-centric walk on switched networks, where the underlying network of possible links varies in time. The rewiring occurs for several edges simultaneously at discrete time steps, as opposed to the continuous-time process we have considered so far, where almost surely no two edges change states at the same time. The adjacency matrix is

$$A(t) = A_{\xi(t)} \tag{2.13}$$

where $\xi : \mathbb{R}^+ \to I \subset \mathbb{N}$ selects one possible graph configurations in the set $\{A_i\}_{i\in I}$. The Laplacian is then given by

$$L(t) = A(t) - D(t), \tag{2.14}$$

where $D(t)$ contains the time-dependent degrees on its diagonal. Note that for simplicity, we have again assumed that the rate λ is the same for all edges of all configurations of the underlying graph, allowing us to use the timescaled equation (2.10) between any two switching times. This remark applies in the context of discrete switching, but can be extended to a continuously-varying, weighted adjacency matrix, where $L(t)$ is no longer a piecewise constant matrix function.

2.2.2.2. *Semi-Markovian random walks*

In general, when the random variables X_w, X_u, X_d are no longer exponentially distributed, the Markov property is lost[3]. However when the process is semi-Markovian and Markovianity in trajectory is preserved, the differential equations (2.6) and (2.10) are replaced by a generalized Montroll-Weiss equation. Such generalizations have been developed from a node-centric perspective for instance in [**4**], and for the edge-centric approach in [**54**], ending up in essentially the same mesoscopic equation, the only difference being the underlying microscopic mechanism regulating the resting times on the nodes. This lets us choose between the two approaches. We will mainly follow [**54**], in which the generalized master equation valid for arbitrary distributions for X_d in the passive edge-centric walk is derived. This master equation will be an important ingredient in the next two sections and is worth presenting.

The building quantity in the models is the resting time on a node j, namely the duration between the arrival-time on the node, and a jump to any other node. This duration is a random variable $X^{(j)}$ with PDF $\mathbf{\Psi}_j(t)$. This resting time density (also known as transition density) satisfies the normalization condition

$$\int_0^\infty \mathbf{\Psi}_j(\tau)d\tau = 1, \tag{2.15}$$

meaning that a jump will eventually occur since the out-degree in the underlying graph is positive. The diagonal matrix of the resting time densities is $D_T(t)$, so that the elements are given by

$$[D_T(t)]_{ij} = \mathbf{\Psi}_j(t)\delta_{ij}. \tag{2.16}$$

The PDF of the resting time is written as the sum

$$\mathbf{\Psi}_j(t) = \sum_{i\in V_j} T_{ij}(t), \tag{2.17}$$

where $T_{ij}(t)$ refers to a jump across edge $j \to i$. If $j \to i$ is an edge of $\mathcal{G}$, the integral

$$\int_0^\infty T_{ij}(\tau)d\tau = \frac{1}{k_j} =: \mathbf{T}_{ij} \tag{2.18}$$

[3]In this work, we are interested in the loss of the Markov property resulting from the interaction between dynamical entities, some or all of them following exponential distributions. Therefore, we will not elaborate on the idea that conversely, Markovianity could emerge out of the interaction between non-Markovian walker and edges.

is thus the probability that at the time of the jump, the walker located on node j selects node i as destination. Here $\mathbf{T}$ is the (effective) transition matrix. Let the time-dependent functions $T_{ij}(t)$ be the entries of the matrix function $\mathbb{T}(t)$. Finally, recall that the Laplace transform of the function $f(t)$ is the map $s \mapsto \int_0^\infty f(t)e^{-st}dt$, and is written $\widehat{f}(s)$ or $\mathcal{L}\{f(t)\}$.

We are now in position to write the master equation which in the Laplace domain reads

$$\mathcal{L}\{\dot{n}(t)\} = \left(\widehat{\mathbb{T}}(s)\widehat{D}_T^{-1}(s) - I\right) K(s)\widehat{n}(s), \tag{2.19}$$

where

$$\widehat{K}(s) = \frac{s\widehat{D}_T(s)}{1 - \widehat{D}_T(s)} \tag{2.20}$$

is the nonlocal-in-time memory kernel. The time-domain version of (2.19) is

$$\dot{n}(t) = \left(\mathbb{T}(t) * \mathcal{L}^{-1}\left\{\widehat{D}_T^{-1}(s)\right\} - \delta(t)\right) * K(t) * n(t), \tag{2.21}$$

where $*$ denotes a convolution. When the analysis is pushed further in the s-domain, one obtains a compact expression for the steady state of the walk. This fact is discussed in section 2.5.

2.3. Resting time density

The work ahead is now to compute the resting time densities $\mathbf{\Psi}_j(t)$, from which the average time spent on a node for a given model follows directly, as is shown in the next section. We compute and interpret this quantity starting from the agent- and edge-level rules of the different models, to obtain a mesoscopic interpretation. Our analysis is restricted to:

Assumption 2.5. The random variables X_w, X_u and X_d are exponentially distributed.

This will allow to shed light on the effect of having up to three timescales, and not on complications arising from otherwise possibly fat-tailed distributions for these three random variables. As a result, the density for model 1 is exponential and the two edge-centric models 2 and 3 generate statistically equivalent trajectories, and hence the analysis for one model holds for the other.

Models 2 and 3

Under assumption 2.5, the instantaneous activation times follow a Poisson process and we have

$$T_{ij}(t) = \mathsf{D}(t)\left[\int_t^\infty \mathsf{D}(\tau)d\tau\right]^{k_j-1}.$$

The interpretation is that the edge $j \to i$ must activate after a time t, whereas all competing edges must remain unavailable at least up to that point. Performing the integration and multiplying by k_j gives

$$\mathbf{\Psi}_j(t) = k_j \lambda e^{-k_j \lambda t}, \tag{2.22}$$

a result already found in [**54**]. This is again an exponential distribution with rate $k_j\lambda$.

Model 6

The walker residing in node j will jump along edge $j \to i$ at time t if all competing edges are unavailable at least up until then - that is, their period of unavailability will last for at least $t - x$, where x marks the time the walker is ready to jump. Moreover, edge $j \to i$ needs to activate exactly after the duration $t - x$. With this in mind,

$$T_{ij}(t) = \int_0^t \psi_j(x) \left[\int_{t-x}^{\infty} \mathsf{D}(s)ds\right]^{k_j-1} \mathsf{D}(t-x)dx,$$

and noting that $\left[\int_{t-x}^{\infty} \mathsf{D}(s)ds\right]^{k_j-1}$ simplifies to $e^{-\lambda(k_j-1)(t-x)}$,

$$= \int_0^t \mu e^{-\mu x} \lambda e^{-k_j\lambda(t-x)} dx \tag{2.23}$$

$$= \mu\lambda e^{-k_j\lambda t} \int_0^t e^{(-\mu+k_j\lambda)x} dx. \tag{2.24}$$

Note that (2.23) is merely the convolution between the waiting time of the walker and the minimum of k_j independent down-times for the edges, reflecting the fact that the process results in an addition of random variables. To proceed, we observe that the integral in (2.24) depends on whether $\mu = k_j\lambda$ or $\mu \neq k_j\lambda$. In the former case, the integral is equal to t and multiplying (2.24) by k_j yields

$$\mathbf{\Psi}_j(t) = \mu k_j \lambda e^{-k_j\lambda t} t. \tag{2.25}$$

In the second case, $\mu \neq k_j\lambda$ and eq. (2.24) gives

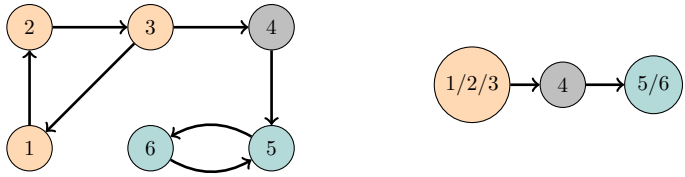
$$\mathbf{\Psi}_j(t) = \frac{\lambda\mu}{k_j\lambda - \mu}\left(e^{-\mu t} - e^{-k_j\lambda t}\right). \tag{2.26}$$

Model 4

This model is generally non-Markovian, but the following derivation assumes a directed acyclic graph (DAG). The reason is twofold.

(1) When there are cycles, the walker may be influenced by the statistical information left at the previous passage, which may induce biases in the walker's trajectory [**127, 48**]. Working with DAGs restores Markovianity in the trajectories, and equation (2.19) is valid.

(2) DAGs include directed trees and find many applications, see for instance [**95**]. Every undirected graph possesses an acyclic orientation. Moreover, by contracting each strongly connected component, every directed graph can be mapped to a DAG, a process illustrated by fig. 2.3. The material presented in this section therefore provides tools to analyze a random walk on a coarse-grained model obtained by condensation of a given graph into a DAG. The acyclic predictions are expected to remain good approximations when edges along paths can be considered statistically independent, in the sense that no correlation exists between their states. For this to hold, the process on the nodes should be slow with respect to the edges dynamics, cycles should be long or locally, nodes should have high degree.



(A) Directed graph with cycles

(B) Directed acyclic graph after condensation

FIGURE 2.3. Mapping of a directed network with cycles (A) to a DAG (B) through a condensation process. Strongly connected components are transformed into super-nodes.

Let us now determine the resting time density. In model 4, two possible scenarios face the walker when ready to jump: either an edge is available with probability r,

$$r = P\{\text{edge } j \to i \text{ is available at a random time}\} \tag{2.27}$$

that is,

$$r = \frac{\langle U \rangle}{\langle U \rangle + \langle D \rangle} = \frac{\lambda}{\lambda + \eta}, \tag{2.28}$$

or an extra wait period is needed before an outgoing edge turns available, with probability

$$1 - r = \frac{\eta}{\lambda + \eta}. \tag{2.29}$$

We decompose $T_{ij}(t)$ in two terms,

$$T_{ij}(t) = (\mathbf{1}) + (\mathbf{2}). \tag{2.30}$$

The first term corresponds to the case that the walker does not get trapped,

$$(\mathbf{1}) = \psi_j(t) \sum_{\ell=1}^{k_j} \frac{1}{\ell} \binom{k_j - 1}{\ell - 1} r^\ell (1-r)^{k_j - \ell}. \tag{2.31}$$

Edge $j \to i$ has to be available, and needs to be chosen among the $k_j - 1$ other edges which are also active at time t. To proceed, we use:

Fact 2.6 (Combinatorial identity)**.** *For $n \in \mathbb{N}_0$ and $0 \leq p \leq 1$, it holds that*

$$\sum_{k=1}^{n} \frac{1}{k} \binom{n-1}{k-1} p^k (1-p)^{n-k} = \frac{1}{n} \left(1 - (1-p)^n\right). \tag{2.32}$$

Proof. Using Newton's binomial theorem, we have

$$\begin{aligned}
\sum_{k=1}^{n} \frac{1}{k} \binom{n-1}{k-1} p^k (1-p)^{n-k} &= \sum_{k=1}^{n} \frac{(n-1)!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \frac{1}{n} \sum_{k=1}^{n} \binom{n}{k} p^k (1-p)^{n-k} \\
&\quad + \frac{1}{n} \sum_{k=0}^{0} \binom{n}{k} p^k (1-p)^{n-k} - \frac{1}{n}(1-p)^n \\
&= \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} - \frac{1}{n}(1-p)^n \\
&= \frac{1}{n} \left(p + (1-p)\right)^n - \frac{1}{n}(1-p)^n \\
&= \frac{1}{n} \left(1 - (1-p)^n\right).
\end{aligned}$$

□

Fact 2.6 allows to rewrite equation (2.31) as

$$(\mathbf{1}) = \psi_j(t) \frac{1}{k_j} \left[1 - (1-r)^{k_j}\right]. \tag{2.33}$$

The quantity between square brackets is the probability that at least one edge is available. The factor $1/k_j$ appears because all outgoing edges are treated indifferently, and so the probability to be chosen is distributed uniformly among all edges including $j \to i$.

In the second case represented by figure 2.4, the jump occurs after the walker got trapped. If this happens on node j, then for a given $i \in V_j$ the time w before $j \to i$ becomes available has the PDF

$$\mathscr{D}(t) = \frac{1}{\langle \mathsf{D} \rangle} \int_t^\infty \mathsf{D}(\nu) \mathrm{d}\nu \tag{2.34}$$

as follows from the so-called bus paradox. Under assumption 2.5, this is again an exponential distribution,

$$\mathscr{D}(t) = \mathsf{D}(t). \tag{2.35}$$

Edge $j \to i$ is selected by the trapped walker to perform the jump a time t if (i) the waiting-time expires before t, (ii) at that moment all other edges are

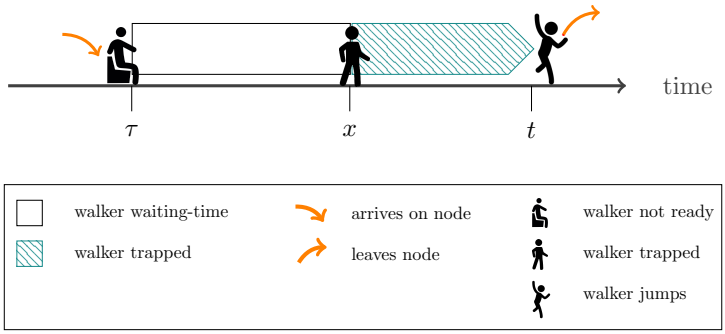


FIGURE 2.4. Resting time of the trapped walker. The walker arrives at time τ on node j. At the end of the waiting-time $x-\tau$ determined by $\psi_j(t)$, none of the k_j outgoing edges is active. The walker needs to wait a subsequent duration $w = t - x$ before a link becomes available. By time-homogeneity, the resting time needs only be computed for $\tau = 0$.

not active and will remain inactive at least until t, and (iii) edge $j \to i$ was also unavailable but becomes active exactly at time t. It results that

$$(\mathbf{2}) = \int_0^t \psi_j(x)\big[(1-r)P\{w > t-x\}\big]^{k_j-1}(1-r)\mathscr{D}(t-x)dx \tag{2.36}$$

or more compactly,

$$(\mathbf{2}) = (1-r)^{k_j}\int_\tau^t \psi_j(x-\tau)\mathscr{D}(t-x)\Big[\int_{t-x}^\infty \mathsf{D}(s)ds\Big]^{k_j-1}dx \tag{2.37}$$

where we have used eq. (2.35) to obtain

$$P\{w > t-x\} = \int_{t-x}^\infty \mathsf{D}(s)ds. \tag{2.38}$$

We have thus found

$$T_{ij}(t) = C_j^{(1)}\psi_j(t) + C_j^{(2)}\int_0^t \psi_j(x)\left[\int_{t-x}^\infty \mathsf{D}(s)ds\right]^{k_j-1}\mathsf{D}(t-x)dx, \tag{2.39}$$

where the constants

$$C_j^{(1)} := \frac{1}{k_j}\left[1-(1-r)^{k_j}\right] \tag{2.40}$$

$$C_j^{(2)} := (1-r)^{k_j} \tag{2.41}$$

depend only on $\langle\mathsf{U}\rangle$, $\langle\mathsf{D}\rangle$, and on the topology through k_j. The distribution of U only matters through its mean that appears in r. On the other hand, the distribution of D does matter beyond its mean, because the jump of a trapped walker occurs directly at the end of a down-time. The resting-time density on node j then reads

$$\mathbf{\Psi}_j(t) = \left[1-(1-r)^{k_j}\right]\psi_j(t) + (1-r)^{k_j}k_j\lambda\mu e^{-k_j\lambda t}\int_0^t e^{(-\mu+k_j\lambda)x}dx. \tag{2.42}$$

The two terms of this PDF reflect a weighted combination of models 1 and 6, where the weight $(1-r)^{k_j}$ is the probability that all outgoing edges are unavailable at a random time.

Model 5

When the walker is ready to jump, whether network edges are available or not depends on the duration since the walker arrived on the node. That makes the analysis somewhat more involved. Assume the walker is ready after s time units. Let $p^*(s)$ be the probability that an edge is in the same state it was at time $t=0$, namely

$$p^*(s) = P\left\{j \to i \text{ is unavailable at time } s \,|\, \text{it is unavailable at time } 0\right\}, \tag{2.43}$$

where $s>0$ is random. Let also

$$q^*(s) = 1 - p^*(s) \tag{2.44}$$

be the probability the edge is available for transport at s.

Remark 2.7 (On the independence from the degree)**.** The fact that p^* does not depend on the degree of the node is due to the memorylessness of the exponential distribution. In general, smaller degrees favor the probability that the move happened after the walker was trapped, that is, at the beginning of an up-time, hence in turn having a reducing impact on p^*.

The two quantities $p^*(s)$ and $q^*(s)$ can be computed by accounting for all possible on-off switches of the edge in the interval $[0,s]$. For the sake of clarity, we make a simplifying assumption, resulting in a compact form for eqs. (2.43) and (2.44).

Assumption 2.8. The rate parameters of the exponential random variables U and D are the same.

Fact 2.9. *Under assumption 2.8, for all $s>0$ it holds that*

$$p^*(s) = \frac{1}{2}(1+e^{-2\lambda s}), \tag{2.45}$$

$$q^*(s) = \frac{1}{2}(1-e^{-2\lambda s}), \tag{2.46}$$

where λ is the exponential rate of the common exponential distribution.

Proof. Edge $j \to i$ is unavailable at time s in case it switched stated exactly $2n$ times over that interval, $n=0,1,\dots$ Hence we have

$$p^*(s) = \int_s^\infty \mathsf{D}(r)dr + \int_0^s (\mathsf{D}*\mathsf{U})(r) \int_{s-r}^\infty \mathsf{D}(t)\,dtdr + \int_0^s (\mathsf{D}*\mathsf{U}*\mathsf{D}*\mathsf{U})(r) \int_{s-r}^\infty \mathsf{D}(t)\,dtdr + \dots \tag{2.47}$$

Introducing the notation for repeated convolutions

$$f^{*k_1} * g^{*k_2} = \underbrace{f*\dots*f}_{k_1 \text{ factors}} * \underbrace{g*\dots*g}_{k_2 \text{ factors}}, \quad k_1,k_2 \in \mathbb{N}, \tag{2.48}$$

eq. (2.47) has the compact form

$$p^*(s) = \int_s^\infty \mathsf{D}(r)dr + \sum_{k=0}^{\infty} \int_0^s \left(\mathsf{D} * \mathsf{U}^{*(k+1)} * \mathsf{D}^{*k}\right)(r) \int_{s-r}^\infty \mathsf{D}(t)\, dt dr. \tag{2.49}$$

The convolution $\mathsf{U}^{*k} * \mathsf{D}^{*k}$ is the PDF of the sum of the random variables

$$X_{\mathsf{U}}^{(k)} + X_{\mathsf{D}}^{(k)} \tag{2.50}$$

where $X_{\mathsf{U}}^{(k)}$ (resp. $X_{\mathsf{D}}^{(k)}$) is the sum of k exponential random variables with parameter η (resp. λ). It is well known that

$$X_{\mathsf{U}}^{(k)} \sim \text{Erlang}(k, \eta) \tag{2.51}$$

$$X_{\mathsf{D}}^{(k)} \sim \text{Erlang}(k, \lambda). \tag{2.52}$$

Using [**62**] for the convolution of Erlang densities, we find that the density of (2.50) is given by

$$\left(\mathsf{U}^{*k} * \mathsf{D}^{*k}\right)(t) = \frac{(\eta\lambda)^k}{(\lambda-\eta)^{2k}} \tag{2.53}$$
$$\times \sum_{j=1}^{k} \left[\frac{(-1)^{k-j}}{(j-1)!}\binom{2k-j-1}{k-j}(\lambda-\eta)^j \left\{e^{-\eta t} + (-1)^j e^{-\lambda t}\right\}\right] t^{j-1}\chi_{\mathbb{R}^+}(t).$$

Introducing (2.53) in (2.49) and using assumption 2.8, after a computation one comes to

$$p^*(s) = e^{-\lambda s}\cosh(\lambda s) = \frac{1}{2}\left(1 + e^{-2\lambda s}\right), \tag{2.54}$$

and the expression for $q^*(s)$ follows directly. □

If the walker is ready after a short time s, the edge will probably still be down, $p^*(0) = 1$, while for large s, the state of the edge is up or down with equal probability, $\lim_{s\to\infty} p^*(s) = \frac{1}{2}$. The second term of eq. (2.45) is the increase with respect to $r = \frac{1}{2}$, and it is smaller for a higher rate λ and for larger $s - \nu$. This is because more switches will decrease the memory effect on the state of the edge. These observations are depicted by fig. 2.5.

So now, we have an expression similar to (2.42) except that r and $1 - r$ are essentially replaced by the time-dependent q^* and p^*. Let us begin by first writing an expression for $T_{ij}(t)$:

$$\begin{aligned} T_{ij}(t) &= \frac{1}{k_j}\left[1 - p^*(t)^{k_j}\right]\psi_j(t) \\ &\quad + \int_0^t \psi_j(x)\left[p^*(x)\int_{t-x}\mathsf{D}(s)ds\right]^{k_j-1} p^*(x)\mathsf{D}(t-x)dx \\ &= \frac{1}{k_j}\left[1 - p^*(t)^{k_j}\right]\psi_j(t) + \lambda\mu e^{-k_j\lambda t}\int_0^t p^*(x)^{k_j} e^{(-\mu+k_j\lambda)x}dx \\ &= \frac{1}{k_j}\left[1 - p^*(t)^{k_j}\right]\psi_j(t) + \frac{\lambda\mu}{2^{k_j}} e^{-k_j\lambda t}\int_0^t \left(1 + e^{-2\lambda x}\right)^{k_j} e^{(-\mu+k_j\lambda)x}dx, \end{aligned}$$

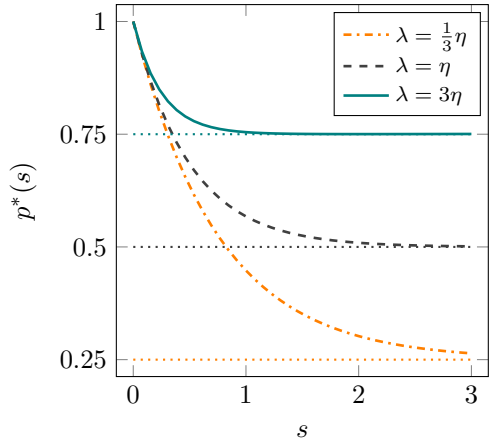


FIGURE 2.5. Evolution of $p^*(s)$, the probability for an edge to be unavailable at time s knowing it was down at 0, in the all-exponential case and for various ratios of η/λ. The teal and orange curves come from eq. (2.49), whereas the gray series corresponds to eq. (2.54). In all cases, the horizontal dotted series indicate the corresponding values of $r = \lambda/(\lambda + \eta)$.

and using Newton's binomial formula in both terms,

$$= \frac{1}{k_j}\left[1 - \frac{1}{2^{k_j}}\sum_{m=0}^{k_j}\binom{k_j}{m}e^{-2m\lambda t}\right]\psi_j(t) + \frac{\lambda\mu}{2^{k_j}}e^{-k_j\lambda t}\sum_{m=0}^{k_j}\binom{k_j}{m}\frac{1}{\beta_m}\left(e^{\beta_m t} - 1\right) \tag{2.55}$$

where we have set $\beta_m = -\mu + k_j\lambda - 2m\lambda$. The resting time density therefore reads:

$$\begin{aligned}\boldsymbol{\Psi}_j(t) = \mu e^{-\mu t} - \frac{\mu}{2^{k_j}}\sum_{m=0}^{k_j}\binom{k_j}{m}e^{-(\mu+2m\lambda)t} \\ + \frac{\mu}{2^{k_j}}k_j\lambda\sum_{m=0}^{k_j}\binom{k_j}{m}\frac{1}{\beta_m}e^{-(\mu+2m\lambda)t} - \frac{\mu}{2^{k_j}}k_j\lambda\sum_{m=0}^{k_j}e^{-k_j\lambda t}\binom{k_j}{m}\frac{1}{\beta_m}.\end{aligned} \tag{2.56}$$

2.4. Mean resting time

The mathematical expectation the resting time $X^{(j)}$ on node j is given by

$$E\{X^{(j)}\} = \int_0^\infty t\boldsymbol{\Psi}_j(t)dt, \tag{2.57}$$

and will also be referred to by $\langle\boldsymbol{\Psi}_j\rangle$. It is naturally called the mean resting time (or mean residence time, MRT) and is relevant in many scenarios, as it will for instance directly determine the relaxation time on tree-like structures.

2.4.1. Derivation of the mean resting time

In this section we handle the models by increasing level of complexity.

Models 1, 2 and 3

In the active node-centric walk, one writes directly $\langle \mathbf{\Psi}_j \rangle_{\text{model 1}} = \frac{1}{\mu}$. It follows from the resting time density of model 2 in eq. (2.22) that

$$\langle \mathbf{\Psi}_j \rangle_{\text{model 2}} = E\{\mathbb{X}_d^{(j)}\}, \tag{2.58}$$

where $\mathbb{X}_d^{(j)}$ was introduced by equation (2.8).

Model 6

Consider first the case $\mu = k_j\lambda$. Following eq. (2.25), the mean resting time reads

$$\langle \mathbf{\Psi}_j \rangle_{\text{model 6}} = \int_0^\infty \mu t^2 k_j \lambda e^{-k_j\lambda t} dt. \tag{2.59}$$

Recalling that the n-th moment of an exponential distribution with rate λ is $E(X^n) = n!/\lambda^n$, we have

$$\langle \mathbf{\Psi}_j \rangle_{\text{model 6}} = 2/\mu = 1/\mu + 1/(k_j\lambda). \tag{2.60}$$

The same expression is obtained in the case when $\mu \neq k_j\lambda$. Indeed, from (2.26),

$$\langle \mathbf{\Psi}_j \rangle_{\text{model 6}} = \frac{\lambda\mu}{k_j\lambda - \mu} \int_0^\infty \left(e^{-\mu\tau} - e^{-k_j\lambda\tau}\right) d\tau = \frac{\lambda\mu}{k_j\lambda - \mu} \left(\frac{1}{\mu^2} - \frac{1}{(k_j\lambda)^2}\right),$$

that is,

$$\langle \mathbf{\Psi}_j \rangle_{\text{model 6}} = \frac{1}{\mu} + \frac{1}{k_j\lambda} = E\{X_w\} + E\{\mathbb{X}_d^{(j)}\}, \tag{2.61}$$

justifying again to consider (Mod. 6) an additive model.

Model 4

We have observed that the resting time PDF (2.42) is a weighted combination of the PDFs of models 1 and 6, such that

$$\begin{aligned}
\langle \mathbf{\Psi}_j \rangle_{\text{model 4}} &= \langle \mathbf{\Psi}_j \rangle_{\text{model 1}} + \langle \mathbf{\Psi}_j \rangle_{\text{model 6}} \\
&= \left[1 - (1-r)^{k_j}\right] E\{X_w\} + (1-r)^{k_j} \left(E\{X_w\} + E\{\mathbb{X}_d^{(j)}\}\right) \\
&= E\{X_w\} + (1-r)^{k_j} E\{\mathbb{X}_d^{(j)}\}. \qquad (2.62)
\end{aligned}$$

Under this form, we see the model is conditionally (depending on r) additive.

Model 5

The mean resting time follows from the PDF (2.56) as

$$\langle \boldsymbol{\Psi}_j \rangle_{\text{model 5}} = E\{X_w\} - \frac{\mu}{2^{k_j}} \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{(\mu + 2m\lambda)^2} + \frac{\mu}{2^{k_j}} k_j \lambda \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\beta_m} \frac{1}{(\mu + 2m\lambda)^2} - \frac{\mu}{2^{k_j}} \frac{1}{k_j \lambda} \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\beta_m}. \tag{2.63}$$

Regrouping the terms, we get

$$\begin{aligned} \langle \boldsymbol{\Psi}_j \rangle_{\text{model 5}} &= E\{X_w\} + \frac{\mu}{2^{k_j}} \sum_{m=0}^{k_j} \binom{k_j}{m} \left(\frac{1}{(\mu + 2m\lambda)^2} \left[\frac{k_j \lambda}{\beta_m} - 1 \right] - \frac{1}{\beta_m} \frac{1}{k_j \lambda} \right) \\ &= E\{X_w\} + \frac{\mu}{2^{k_j}} \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\beta_m} \left(\frac{1}{\mu + 2m\lambda} - \frac{1}{k_j \lambda} \right) \\ &= E\{X_w\} + \frac{\mu}{2^{k_j}} \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\mu + 2m\lambda} E\{\mathbb{X}_d^{(j)}\}. \end{aligned} \tag{2.64}$$

2.4.2. Discussion

All models have a mean resting time that we cast under the form

$$\langle \boldsymbol{\Psi}_j \rangle = \boldsymbol{\alpha} E\{X_w\} + \boldsymbol{\beta}(k_j, \mu, \lambda) E\{\mathbb{X}_d^{(j)}\}, \tag{2.65}$$

where $\boldsymbol{\alpha}$ depends only on the walker and $\boldsymbol{\beta}(k_j, \mu, \lambda)$ is topology driven since it represents the probability that all outgoing edges are unavailable when the walker is ready. Summing up the results of this section, we have

$$\boldsymbol{\alpha}_{\text{model 1}} = 1 \qquad \boldsymbol{\beta}_{\text{model 1}}(k_j, \mu, \lambda) = 0 \tag{2.66}$$

$$\boldsymbol{\alpha}_{\text{model 2}} = 0 \qquad \boldsymbol{\beta}_{\text{model 2}}(k_j, \mu, \lambda) = 1 \tag{2.67}$$

$$\boldsymbol{\alpha}_{\text{model 3}} = 0 \qquad \boldsymbol{\beta}_{\text{model 3}}(k_j, \mu, \lambda) = 1 \tag{2.68}$$

$$\boldsymbol{\alpha}_{\text{model 4}} = 1 \qquad \boldsymbol{\beta}_{\text{model 4}}(k_j, \mu, \lambda) = (1 - r)^{k_j} \tag{2.69}$$

$$\boldsymbol{\alpha}_{\text{model 5}} = 1 \qquad \boldsymbol{\beta}_{\text{model 5}}(k_j, \mu, \lambda) = \frac{1}{2^{k_j}} \mu \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\mu + 2m\lambda} \tag{2.70}$$

$$\boldsymbol{\alpha}_{\text{model 6}} = 1 \qquad \boldsymbol{\beta}_{\text{model 6}}(k_j, \mu, \lambda) = 1. \tag{2.71}$$

Recall that (2.70) was derived under assumption 2.8, $\eta = \lambda$, for which $r = \frac{1}{2}$. To allow a fair comparison between models we write

$$\boldsymbol{\beta}_{\text{model 4}}(k_j, \mu, \lambda) = \frac{1}{2^{k_j}}. \tag{2.72}$$

Using standard algebra, it is straightforward to check that

$$0 = \boldsymbol{\beta}_{\text{model 1}} < \boldsymbol{\beta}_{\text{model 4}} < \boldsymbol{\beta}_{\text{model 5}} < \boldsymbol{\beta}_{\text{model 6}} = 1, \tag{2.73}$$

for all $k_j \in \mathbb{N}_0$ and all positive reals μ and λ. The smaller this coefficient, the larger the expected number of jumps along the trajectories of the walk, all other parameters being chosen equal.

We want to compare the three where there is a dynamical walker-network interaction. To this end, let us define the ratios of mean resting times

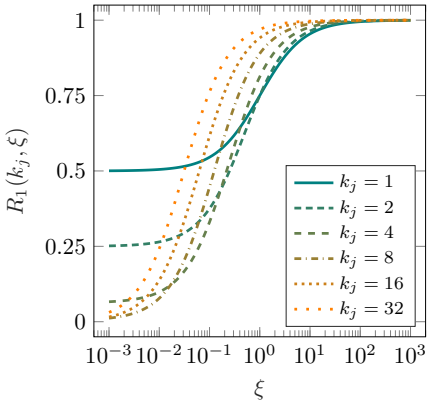
$$R_1 := \frac{\langle \mathbf{\Psi}_j \rangle_{\text{model 4}}}{\langle \mathbf{\Psi}_j \rangle_{\text{model 6}}} , \tag{2.74}$$

$$R_2 := \frac{\langle \mathbf{\Psi}_j \rangle_{\text{model 5}}}{\langle \mathbf{\Psi}_j \rangle_{\text{model 6}}}. \tag{2.75}$$

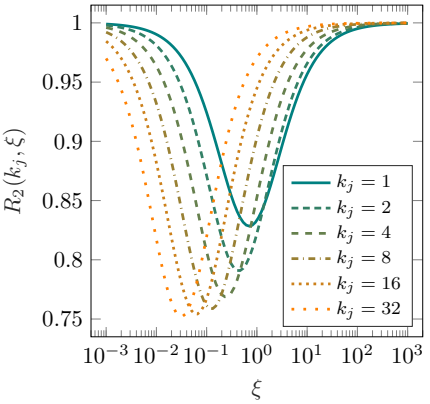
These quantities depend only on the degree k_j, and on a new variable $\xi := \frac{\lambda}{\mu}$. Indeed, we write

$$R_1(k_j, \xi) = \frac{k_j 2^{k_j} \xi + 1}{k_j 2^{k_j} \xi + 2^{k_j}} , \tag{2.76}$$

$$R_2(k_j, \xi) = \frac{k_j 2^{k_j} \xi + \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{1+2m\xi}}{k_j 2^{k_j} \xi + 2^{k_j}}. \tag{2.77}$$



(A) Model 4, eq. (2.76). The reduction in MRT is substantial for small ξ, especially combined with larger degrees such that the relatively slow network timescales have less impact.



(B) Model 5, eq. (2.77). The reduction in MRT is less pronounced than in the comparatively faster model 4. It never goes below $\frac{3}{4}$ and even vanishes for extreme values of ξ.

FIGURE 2.6. Reduction in mean resting times of models 4 and 5 with respect to the additive model 6. The ratios R_1 and R_2 compare the three models with two timescales given by μ and $\lambda = \eta$, for which there is walker-edges interaction. Going from walk (1) to (4) to (5) to (6), the MRT increases - all dynamical and topological parameters being equally chosen. Diffusion on a tree topology would be slower. Notice that only the degree k_j and $\xi = \frac{\lambda}{\mu}$ determine these comparisons.

The above expressions are plotted in figure 2.6 for various values of the degree. That $R_i(k_j, \xi \to \infty) = 1$ for $i = 1, 2$ and $R_2(k_j, 0) = 1$ is a consequence of the asymptotic behavior of the resting time densities.

2.5. Steady state

One core application of random walks is their ability to rank nodes of a network according to the long-term residence probability of the walker, an information directly accessible from the steady state (SS) of the preceding section. We have mentioned that the steady state is essential to Google's PageRank algorithm, but obviously goes beyond the ranking of webpages. It can notably help identify the most prominent or threatening figures in a network of individuals, a motivation for what follows.

2.5.1. Derivation of the steady state

Let $D_{\langle T\rangle}$ be the diagonal matrix of the MRT,

$$\left[D_{\langle T\rangle}\right]_{ij} = \langle \mathbf{\Psi}_j \rangle \delta_{ij}. \tag{2.78}$$

A small-s analysis of the generalized master equation (2.19) showed in [**54**] the steady-state to be

$$n(\infty) := \lim_{t\to\infty} n(t) \propto D_{\langle T\rangle} \boldsymbol{v}, \tag{2.79}$$

where $\propto$ is the proportionality symbol and $\boldsymbol{v} = (v_1, \ldots, v_N)$ is the eigenvector associated to the unit eigenvalue of the effective $N \times N$ transition matrix $\mathbf{T}$. Recall eq. (2.18),

$$\mathbf{T}_{ij} = \int_0^\infty T_{ij}(t)dt = \frac{1}{k_j} A_{ji}. \tag{2.80}$$

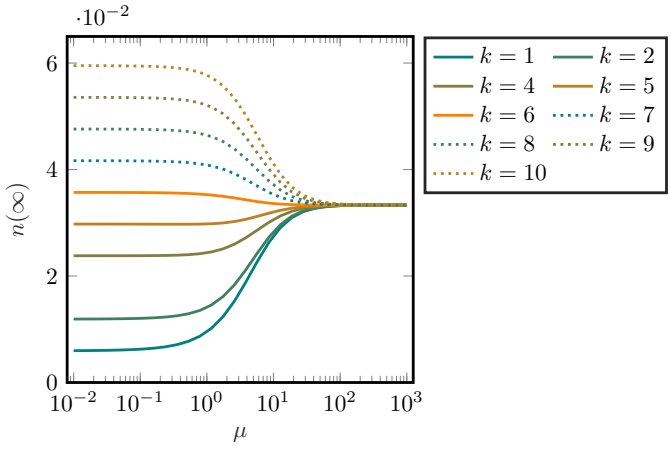
One straightforwardly checks that $v_j = k_j$ satisfies

$$(\mathbf{T}\boldsymbol{v})_i = \sum_{\ell=1}^{N} A_{\ell i} = k_i, \tag{2.81}$$

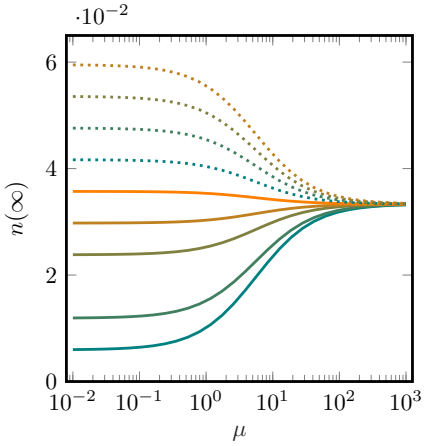
where the last equality assumes that the network is balanced, namely, the in-degree of node i is equal to its out-degree k_i. In other words, when the underlying network is balanced, the steady state is [**54**]

$$n_j(\infty) \propto D_{\langle T\rangle} k_j. \tag{2.82}$$

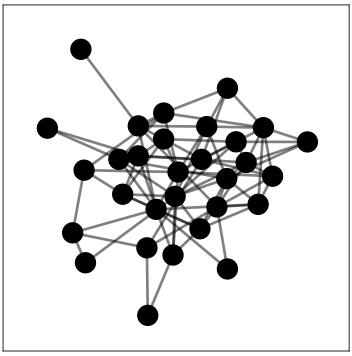
Our goal is to understand how the SS depends on the modeling scheme and hence how the ranking process changes accordingly. We will compute those steady states and report the result graphically on figs. 2.7 and 2.8.



(A) Model 5, eq. (2.90)



(B) Model 6, eq. (2.87)



(C) Directed Erdös-Rényi graph. Each edge represents a pair of reciprocal links.

FIGURE 2.7. Steady states for models 5 and 6. Those models behave in the limit for the fast walker, $\mu \to \infty$, as a Markovian passive edge-centric random walk, where the steady state is uniform across the network. In the small-μ limit, the steady state probability is proportional to the degree, like in the Markovian active node-centric walk. This behavior does not hold true for model 4. The rate parameters are $\eta = 1 = \lambda = 1$.

Models 1, 2 and 3.

For later comparisons, let us briefly recall that with the node-centric model 1, $\langle \mathbf{\Psi}_j \rangle = \frac{1}{\mu}$ and the steady-state is proportional to degree,

$$n_j^{(Mod.\ 1)}(\infty) = \frac{k_j}{\sum_{\ell=1}^{N} k_l}, \tag{2.83}$$

whereas with the active edge-centric model 2 we have

$$n_j^{(Mod.\ 2)}(\infty) = \frac{1}{N}. \tag{2.84}$$

As pointed out, the same expression is valid for the passive edge-centric walk of model 3 when the down-time distribution is exponential. One easily verifies that the right-hand side of $\dot{n} = nL^{rw}$ vanishes for $n = n^{(Mod.\ 1)}(\infty)$, while the same holds true with $\dot{n} = nL$ for $n = n^{(Mod.\ 2)}(\infty)$. Uniformity of the steady-state (2.84) is due to the fact that the edges are all selected just as frequently. Well-connected nodes are more likely to be visited by the walker, but will be left sooner.

Model 6

Based on eq. (2.82) we have

$$n_j^{(Mod.\ 6)}(\infty) \propto \left(\frac{1}{\mu} + \frac{1}{k_j\lambda}\right) k_j = \frac{1}{\lambda\mu}(\mu + k_j\lambda), \tag{2.85}$$

and after normalisation we get

$$n_j^{(Mod.\ 6)}(\infty) = \frac{k_j\lambda + \mu}{\sum_{\ell=1}^{N} k_\ell\lambda + \mu} \tag{2.86}$$

or under a different form, after division by $k_j\lambda\mu$,

$$n_j^{(Mod.\ 6)}(\infty) = \frac{\frac{1}{\mu} + \frac{1}{k_j\lambda}}{\sum_{\ell=1}^{N} \frac{k_\ell}{k_j}\frac{1}{\mu} + \frac{1}{k_j\lambda}} = \frac{E\{X_w\} + E\{\mathbb{X}_d^{(j)}\}}{\sum_{\ell=1}^{N} \frac{k_\ell}{k_j} E\{X_w\} + E\{\mathbb{X}_d^{(j)}\}} \tag{2.87}$$

We recover the expressions of the active node-centric (Mod. 1) and edge- centric (Mod. 4) walks in the respective limits $\lambda \to \infty$ and $\mu \to \infty$.

Model 4

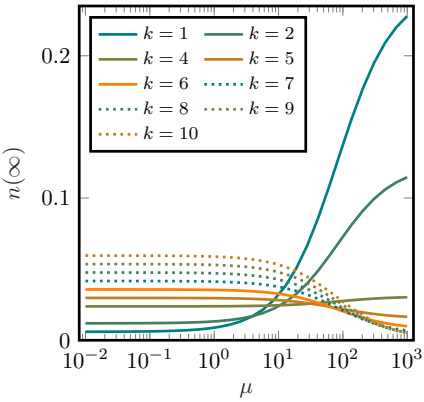
This model necessitates a preliminary observation concerning the method. The transition density derived in the preceding section results to be only an approximation when the graph has cycles. On the other hand, the steady state formula (2.82) assumes balanced networks. Hence, the steady state we will obtain with eq. (2.89) is an approximation if the balanced network has cycles, for instance when the network has reciprocal links.

Let us now proceed with the analysis. We have

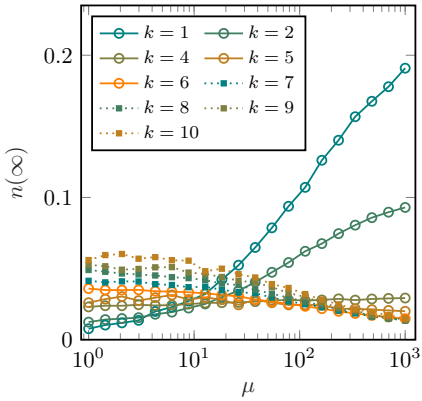
$$n_j^{(Mod.\ 4)}(\infty) \propto \frac{k_j}{\mu} + (1-r)^{k_j}\frac{1}{\lambda} \tag{2.88}$$

and through normalization we obtain

$$n_j^{(Mod.\ 4)}(\infty) = \frac{\frac{k_j}{\mu} + (1-r)^{k_j}\frac{1}{\lambda}}{\sum_{\ell=1}^{N}\left[\frac{k_\ell}{\mu} + \frac{(1-r)^{k_\ell}}{\lambda}\right]}. \tag{2.89}$$



(A) Analytic, eq. (2.89). Remarkably, when $\mu \to \infty$, the steady-state probability is high when degree is low.



(B) Monte Carlo simulation of 10^4 trajectories of a walker, in the range of values of μ corresponding to the reversed degree-based ranking

FIGURE 2.8. Steady state for model 4. The analytical prediction of panel (A), which does not account for the presence of cycles in the network, is qualitatively confirmed by Monte Carlo simulation of the actual process on panel (B). Kendall's Tau coefficient on fig. 2.9 captures the inversion with respect to the degree-based ranking. The rate parameters and network are those of fig. 2.7.

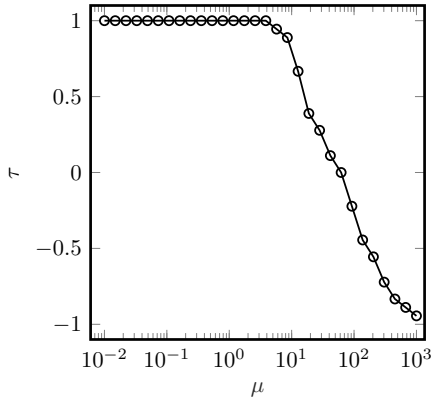


FIGURE 2.9. Kendall's Tau correlation coefficient for the steady state of model 4 relative to fig. 2.8b. When Kendall's τ is one , the ranking is degree-based; when τ is -1, it is reversed. There is no link between the two when τ is 0.

As expected, $n_j^{(Mod.\ 4)}(\infty)$ tends to $n_j^{(Mod.\ 1)}(\infty)$ when $\lambda \to \infty$. But more importantly, in the limit of a very fast walker we have

$$\lim_{\mu\to\infty} n_j^{(Mod.\ 4)}(\infty) = \frac{(1-r)^{k_j}}{\sum_{\ell=1}^{N}(1-r)^{k_\ell}}.$$

It results that smaller residence probabilities are associated with nodes with larger degree. At variance for a relatively slow walker, larger degree means larger probability:

$$\lim_{\mu\to 0} n_j^{(Mod.\ 4)}(\infty) = \frac{k_j}{\sum_{\ell=1}^{N} k_\ell}.$$

It was reported before that fat-tailed resting times on a portion of the nodes of a network could lead to accumulation on these nodes in spite of their low degree [**41**]. In our case, the renewal process ruling the jump times arises from interaction between walker and network, without explicitly reverting to long-tailed distributions of the resting time on certain nodes.

Model 5

Resulting directly from the transition density given by (2.64), we have

$$n^{(Mod.\ 5)}(\infty) \propto \frac{1}{\mu} + \frac{\mu}{2^{k_j}\lambda} \sum_{m=0}^{k_j} \binom{k_j}{m} \frac{1}{\mu + 2m\lambda} \tag{2.90}$$

and the normalization factor is given by

$$\left[\frac{N}{\mu} + \frac{\mu}{\lambda} \sum_{\ell=1}^{N} \left(\frac{1}{2^{k_\ell}} \sum_{m=0}^{k_\ell} \binom{k_\ell}{m} \frac{1}{\mu + 2m\lambda} \right) \right]^{-1}. \tag{2.91}$$

2.5.2. Discussion

In this discussion we compare the introduced models based on a numerical experiment, and then give details about the Monte-Carlo random walk simulator.

2.5.2.1. *Comparison between models*

For our simulations, we selected an Erdös-Rényi graph with 30 nodes and connection probability $\frac{1}{5}$ pictured on fig. 2.7c. All reciprocal links are present, implying the graph is not cycle-free. Figures 2.7a and 2.7b indicate no departure for models 5 and 6 from the degree-based ranking, at variance with the peculiar behavior of model 4 on fig. 2.8. The figure shows that the steady state computed with the theoretical formula is in good agreement with Monte-Carlo simulations describe below. The formula proved mostly accurate throughout our numerical investigation, certainly in terms of ranking of the nodes.

The discrepancy between fig. 2.8a and fig. 2.8b is not a finite-size effect of our sample, and is confirmed by a 95% confidence interval, although it was not reported on the figure for clarity. The network topologies and range of values for the dynamical parameters which preserve the agreement between analytical approximation and simulation, together with confidence intervals, are further discussed in chapter 3.

2.5.2.2. *Numerical simulation*

The numerical method relies on Monte-Carlo simulation to determine the probabilities $n(t)$ by averaging over a large number M of realizations. Each trajectory of the walker corresponds to new samples of the stochastic processes regulating the walker's waiting-times and the up- and down-times of the edges. The time interval $[0, T]$ of each realization is discretized according to a uniform partition

$$0 = t_0 < t_1 < \cdots < t_m = T, \quad \Delta t = t_k - t_{k-1}. \tag{2.92}$$

For a fixed $t \in [0, T]$ and $j \in V$, the probability $n_j(t)$ is given by

$$n_j(t) = P\{Z(t) = j\} = E\left\{\chi_{Z(t)=j}\right\}, \tag{2.93}$$

which is approximated by its estimator $\overline{\chi}_{Z(t)=j}$. Here the averaging is performed over the sample of the walker's position at time t,

$$\left\{Z^{\omega_1}(t), \ldots, Z^{\omega_M}(t)\right\}, \tag{2.94}$$

which results form the simulation of M independent trajectories $\omega_1, \ldots, \omega_M$. Since M is large, it results from the central limit theorem that $\overline{\chi}_{Z(t)=j}$ is approximately normally distributed. Therefore,

$$\sqrt{M}\frac{\overline{\chi}_{Z(t)=j} - E\left\{\chi_{Z(t)=j}\right\}}{s_M} \sim t_{M-1}, \tag{2.95}$$

where s_M^2 denotes the unbiased sample variance and t_{M-1} is a Student's t-distribution with $M-1$ degrees of freedom. A confidence interval readily follows from (2.95) noticing the convergence in distribution of the t-distribution to a normal standard law when $M \to \infty$.

Since the jumps occur at irregular times and the probabilities $n_j(t)$ are computed on the partition (2.92), averaging is performed of the intervals $(t_{k-1}, t_k]$, and each sample is given by

$$\chi_{Z^{\omega_\ell}(t_k)=j} \approx \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} \chi_{Z^{\omega_\ell}(t)=j} dt \tag{2.96}$$

for $k = 1, \ldots, m$ and $\ell = 1, \ldots, M$. In essence the probabilities $n_j(t_k)$ are approximated by the fraction of time spent by the walker on node j over the time interval $(t_{k-1}, t_k]$ [**54**].

2.6. Conclusion

This chapter proposes models of random walks where the walker does not have a memory and is unbiased, and offers a look into their trajectories when defined on temporal networks. As we have discussed, there exist different ways to inter-connect the dynamics of the walker and of the network, and this interplay may break the Markovianity of the system, even in purely active models or passive models without cycles. A long-tailed walker waiting time on (a subset of) the node(s), such as in [**41**], is not required to observe a dramatic departure from the steady-states of the classical random walk models. We have revealed that

the mean resting time may be impacted, resulting in a slowed-down diffusion on tree-like structures.

Overall, our work underlines the importance of the different timescales associated to random walks on temporal networks, and unveils the importance of the duration of contacts on diffusion. In this paradigm, the walk that most naturally combines the classical active node-centric and passive edge-centric walks, is described by model 4. In this model the interaction of a memoryless walker and a network governed by memoryless distributions leads to the loss of the trajectory Markov property and of time-homogeneity. In the next chapter we develop a mathematical framework further enabling an analytic treatment of this finding, and of the influencing factors such as the rates of the distributions and the topology of the graph.

CHAPTER 3

Emergence of memory with competing timescales

3.1. Introduction

Mathematical analysis of dynamics on temporal networks often relies on the assumption that links activate during an infinitesimal duration [**54**]. We have seen in chapter 2 that in the case of continuous-time random walks, this framework naturally reduces to Markovian and semi-Markovian process, and the dynamics then exhibits interesting properties including the so-called waiting time paradox [**127, 48**]. The emergence of non-Markovian trajectories is even more pronounced in situations when the activations of the edges are correlated, often requiring the use of higher-order models for the data [**122, 78**].

However, this whole stream of research neglects an important aspect of the edge dynamics, the non-zero duration of their availability, which has been observed and characterized in a variety of real-life systems, including sensor data [**45, 138, 121**]. In the previous chapter, microscopic models have been introduced and analyzed, in order to capture this specificity of the edges and some of its applied consequences. In particular, we proposed a natural combination of the fundamental active node-centric, and passive edge-centric random walks. This lead to a model which is neither node-centric nor edge centric, but where the active feature of the walker-level dynamics, and the passive character of its decoupled edge-level counterpart are both preserved.

We have seen with fig. 2.8 that this new walk, model 4, is a semi-Markovian process described by the master equation eq. (2.21) only when the underlying graph does not have cycles. Our main objective is therefore to develop an analytical framework that also applies to non-Markovian random walks on temporal networks with finite up-times. This will finalize our contributions with respect to our first research question.

In addition to the presence of cycles, the competition between the three timescales makes the problem particularly rich. Figure 3.1 summarizes possible scenarios of timescale separation. It conveys the message that in two limit regimes the regular single-timescale CTRWs are essentially recovered. In one regime, the trajectory of the walker is solely determined by the waiting time on the nodes, since the edges are nearly always active. In the second regime, the time spans between two instantaneous up-times of the edges dominate the dynamics, as it is the slowest component. But generally speaking, not only one

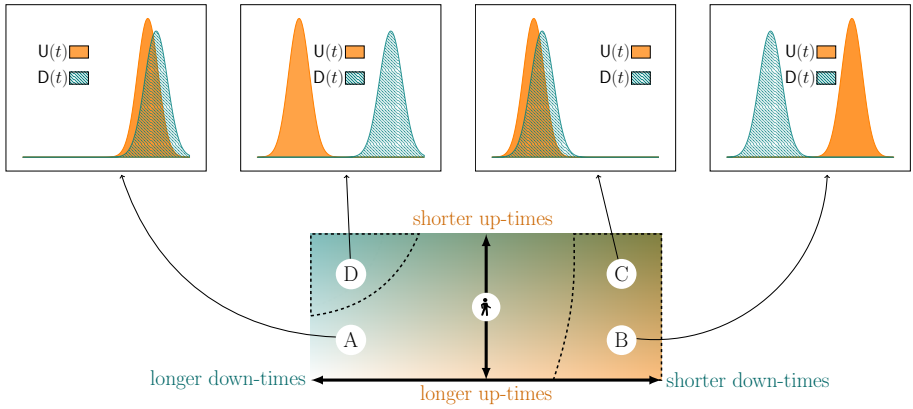


FIGURE 3.1. Modeling depending on timescale separation. The bottom panel represents the characteristic up-time and down-time durations with respect to the walker's own timescale. The corners labeled from A to D correspond to the four cases where there is a clear timescale separation between up-times, down-times and waiting time of the walker. Each of these situations is described by one of the top panels, representing typical PDFs $\mathsf{U}(t)$ and $\mathsf{D}(t)$. At the right of the domain, in the region ranging from B to C, the network dynamics has less influence. In the region around D, the modeling may possibly only retain the down-times, and the walker's waiting time. Our focus is on the center of the domain, between the dotted regions, where the full complexity of the model needs to be accounted for.

timescale prevails and neither of the former two asymptotic regimes can capture the full dynamics, thereby motivating this chapter.

This chapter is organized as follows. In section 3.2, we consider the semi-Markovian case that arises when the graph is a DAG. We present equivalent node-centric and edge-centric formulations for the walk. Then we derive the propagator and a transport equation without using the Laplace transform. We discuss in section 3.3 the two sources, dynamical and topological, for the emergence of memory in the trajectories. In section 3.4 we then apply our time-domain method initially developed for DAGs, to graphs with cycles. We explicitly compute corrections to the transition kernel accounting for short cycles of length 2. The analytical predictions are confronted with numerical simulations in section 3.5, which also contains details about the numerical implementation of our formalism. We finally conclude and give perspectives in section 3.6.

3.2. The semi-Markovian case

As a first step, we consider the trajectory of a random walker as defined by model 4 on a DAG. The reason for that is twofold. First, we will show that the three different timescales still allow for purely node-centric, or edge-centric

models with statistically equivalent trajectories. Secondly, we will show how to find the propagator which maps the initial condition to the residence probabilities at a subsequent time, without reverting to the Laplace transform. Although not necessary at this stage, our approach will come in handy when we deal with the non-Markovian case.

3.2.1. Equivalent node- and edge-centric models

Even beyond Poisson processes, the model on DAGs can be cast into a node-level-only process, or into an edge-level-only variant with instantaneous durations of edges availability, and a walker who has no own waiting time. We will discuss both approaches, which enable the use of the formalism valid form models 1 and 3 of chapter 2.

3.2.1.1. *Node-centric model*

The goal is to compute the PDF $\mathbf{\Psi}_j$ for the resting time $X^{(j)}$ of the walker, based on ψ, U and D. Observe that $\mathbf{\Psi}_j$ depends on the node only through its out-degree. Let $\mathscr{D}_{(1),j}$ be the PDF of the random variable $w_{(1),j}$ defined as the minimum of k_j independent random variables w with density $\mathscr{D}$ given by the bus paradox, eq. (2.34). Using independence we have

$$\begin{aligned} \mathscr{D}_{(1),j}(t) &= \frac{d}{dt}\left(1 - P\left\{w_{(1),j} > t\right\}\right) \\ &= -\frac{d}{dt}\left(P\{w > t\}\right)^{k_j} \qquad (3.1) \\ &= k_j \left(1 - F_w(t)\right)^{k_j - 1} \mathscr{D}(t), \qquad (3.2) \end{aligned}$$

where $F_w(t)$ is the distribution function of w. The contribution of the edges to the total waiting time has density given by

$$\mathcal{E}_{(1),j}(t) = (1-r)^{k_j}\mathcal{D}_{(1),j}(t) + (1-(1-r)^{k_j})\delta(t). \qquad (3.3)$$

Observe that it is properly normalized,

$$\int_0^\infty \mathcal{E}_{(1),j}(t)dt = (1-r)^{k_j}\int_0^\infty \mathcal{D}_{(1),j}(t)dt + 1 - (1-r)^{k_j} = 1. \qquad (3.4)$$

By additivity, the resting-time of the walker in node j is given by

$$\mathbf{\Psi}_j(t) = (\psi * \mathcal{E}_{(1),j})(t) \qquad (3.5)$$

such that

$$\mathbf{\Psi}_j(t) = (1-r)^{k_j}(\psi * \mathcal{D}_{(1),j})(t) + (1-(1-r)^{k_j})\psi(t). \qquad (3.6)$$

When all densities are exponential, $w_{(1),j} \equiv \mathbb{X}_d^{(j)}$ as given by eq. (2.8) is exponentially distributed with rate $k_j\lambda$ and we know that

$$(\psi * \mathcal{D}_{(1),j})(t) = \frac{k_j\lambda\mu}{\mu - k_j\lambda}\left(e^{-k_j\lambda t} - e^{-\mu t}\right), \qquad (3.7)$$

which may be plugged in (3.6) and we recover $\mathbf{\Psi}_j(t)$ of eq. (2.62).

3.2.1.2. *Edge-centric model*

When opting for an edge-centric model, one has to determine the distribution of the down-times on the edges such that $\mathbf{\Psi}_j$ be given by eq. (3.6). Let $X_d^{(j)}$ be the random variable associated in this edge-centric model to the down-time of an edge originating from node j. This is the usual notation, except for the superscript indicating the dependence on the (degree of the) node. Denoting as always by F_A the cumulative distribution of a random variable A, we have

$$F_{X^{(j)}}(t) = 1 - P\left\{X^{(j)} > t\right\} = 1 - \left(1 - F_{X_d^{(j)}}(t)\right)^{k_j} \tag{3.8}$$

which is then solved for $F_{X_d^{(j)}}(t)$. The PDF of the down-times is then given by differentiation,

$$\frac{d}{dt}F_{X_d^{(j)}}(t) = \frac{1}{k_j}\big(1 - F_{X^{(j)}}(t)\big)^{1/k_j - 1}\mathbf{\Psi}_j(t). \tag{3.9}$$

3.2.2. Time-domain derivation of the propagator

Consider a random walker initially located at node $Z(0) = j_0$. The propagator of the walk maps the initial condition $\boldsymbol{n}(0)$ given by

$$n_j(0) = P\{Z(0) = j\} = \delta_{jj_0}, \quad j = 1, \dots, N, \tag{3.10}$$

to $\boldsymbol{n}(t)$, the vector of residence probabilities at a later time t, namely $n_j(t) = P\{Z(t) = j\}$. Let $q_j^{(k)}(t)$ be the conditional PDF of the arrival time on node j in k jumps, conditionally on the initial condition, obeying the normalization condition

$$\sum_{j=1}^{N}\int_0^\infty q_j^{(k)}(t)dt = 1, \quad \forall k \in \mathbb{N}. \tag{3.11}$$

In particular observe that for $k = 0$, we have

$$\sum_{j=1}^{N}\int_0^\infty q_j^{(0)}(t)dt = \sum_{j=1}^{N}\int_0^\infty n_j(0)\delta(t)dt = \sum_{j=1}^{N}\delta_{jj_0} = 1, \tag{3.12}$$

and for $k = 1$,

$$\sum_{j=1}^{N}\int_0^\infty q_j^{(1)}(t)dt = \sum_{j=1}^{N}\int_0^\infty T_{jj_0}(t)dt = \int_0^\infty \mathbf{\Psi}_{j_0}(t)dt = 1. \tag{3.13}$$

Therefore,

$$q_j(t) := \sum_{k=0}^{\infty} q_j^{(k)}(t) \tag{3.14}$$

represents the conditional probability density of the arrival time on node j in any number of jumps. Observe that $q_j(t)$ does not satisfy a normalization condition such as eq. (3.11). Let also $\Phi_j(t, \tau)$ be the survival probability, *i.e.*

the probability to stay on node j at least up to $t > \tau$ with τ the arrival time on node i, such that

$$n_j(t) = \int_0^t \Phi_j(t-\tau) q_j(\tau) d\tau. \tag{3.15}$$

Here we have used time-homogeneity, $\Phi_j(t,\tau) \equiv \Phi_j(t-\tau)$, which also applies for the transition kernel, $T_{ij}(t,\tau) \equiv T_{ij}(t-\tau)$. We have the relationship

$$\Phi_j(t-\tau) = 1 - \int_\tau^t \mathbf{\Psi}_j(\nu-\tau) d\nu = 1 - \int_\tau^t \sum_{i \in V_j} T_{ij}(\nu-\tau) d\nu. \tag{3.16}$$

Remark 3.1 (Survival probability with exponential distributions)**.** Under assumption 2.5, following a direct calculation based on the expression for $T_{ij}(t)$, eq. (2.39), the survival probability reads

$$\begin{aligned}\Phi_j(t) = 1 - (1-r)^{k_j} \left(1 - \frac{1}{\lambda k_j - \mu}\left(\lambda k_j e^{-\mu t} - \mu e^{-\lambda k_j t}\right)\right) \\ - \left(1 - (1-r)^{k_j}\right)\left(1 - e^{-\mu t}\right),\end{aligned} \tag{3.17}$$

a fact that will be used throughout our numerical simulations.

We want to write the column vector $\boldsymbol{n}(t) = (n_1(t), \ldots, n_N(t))^T$ in terms of the transition kernel T_{ij} and of the initial condition $\boldsymbol{n}(0)$. Combining (3.15) and (3.16) shows that we have to determine $q_j(t)$. We have

$$q_j(t) = \sum_{k=1}^{\infty} q_j^{(k)}(t) + q_j^{(0)}(t), \tag{3.18}$$

where we recall $q_j^{(0)}(t) = n_j(0)\delta(t)$. Equivalently,

$$q_j(t) = \sum_{k=0}^{\infty} q_j^{(k+1)}(t) + q_j^{(0)}(t), \tag{3.19}$$

where

$$q_j^{(k+1)}(t) = \sum_{i=1}^{N} \int_0^t T_{ji}(t-\nu) q_i^{(k)}(\nu) d\nu. \tag{3.20}$$

Summing on both sides over $k \geq 0$ and adding $q_j^{(0)}(t)$ yields

$$q_j(t) = \sum_{i=1}^{N} \int_0^t T_{ji}(t-\nu) q_i(\nu) d\nu + q_j^{(0)}(t). \tag{3.21}$$

In vector form, with $\boldsymbol{q}(t) = (q_1(t), \ldots, q_N(t))^T$, this becomes

$$\boldsymbol{q}(t) = \mathcal{T}\boldsymbol{q}(t) + \boldsymbol{q}^{(0)}(t) \tag{3.22}$$

where $\mathcal{T}$ is the finite-rank linear integral operator acting on $\boldsymbol{q}(t)$ defined by:

$$\mathcal{T}\boldsymbol{q}(t) = \int_0^t T(t-\nu)\boldsymbol{q}(\nu) d\nu, \tag{3.23}$$

and where the kernel $T(t)$ is a matrix function with component (i,j) given by $T_{ij}(t)$. Due to the acyclic nature of the graph, we obtain with eq. (3.23) a

convolution and applying a Laplace transform allows to solve (3.22) for $\boldsymbol{q}(t)$, as was done in [**54**].

It is however not mandatory to use the Laplace transform to solve the integral equations (3.22) for $\boldsymbol{q}(t)$ and obtain a Laplace-domain Montroll-Weiss equation for $\boldsymbol{n}(t)$. We can proceed directly in the time domain and solve the equation relying on the acyclic nature of the graph. From eq. (3.22) we find

$$\boldsymbol{q}(t) = (I - \mathcal{T})^{-1}\boldsymbol{q}^{(0)}(t) = \sum_{k=0}^{\infty} \mathcal{T}^k \boldsymbol{q}^{(0)}(t), \tag{3.24}$$

where the convergence of the Neumann series follows directly from an argument in [**47**, Appendix A.2], that is based on the Perron-Frobenius theorem for non-negative matrices. In terms of the initial condition, we have

$$\boldsymbol{q}(t) = \sum_{k=0}^{\infty} \mathcal{T}^k \delta(t)\boldsymbol{n}(0). \tag{3.25}$$

The successive terms $\mathcal{T}^k\boldsymbol{n}(0)\delta(t)$ account for the probability to arrive on a given node at time t, starting from the initial node in exactly k steps.

Remark 3.2. In general, the Neumann series does not offer a practical way for computing $(I - \mathcal{T})^{-1}$ since it involves an infinite number of terms. Because we make the assumption that the underlying graph $\mathcal{G}$ has no cycles, the series can be cut after d terms, where d is the diameter of the graph.

Once $\boldsymbol{q}(t)$ is found, we consider equation (3.15), which can be cast under the form

$$\boldsymbol{n}(t) = \mathcal{P}\boldsymbol{q}(t) \tag{3.26}$$

where $\mathcal{P}$ is another finite-rank, diagonal operator with components given by

$$\left(\mathcal{P}\boldsymbol{q}(t)\right)_j = \int_0^t \Phi_j(t-\tau)q_j(\tau)d\tau, \quad j = 1, \ldots, N. \tag{3.27}$$

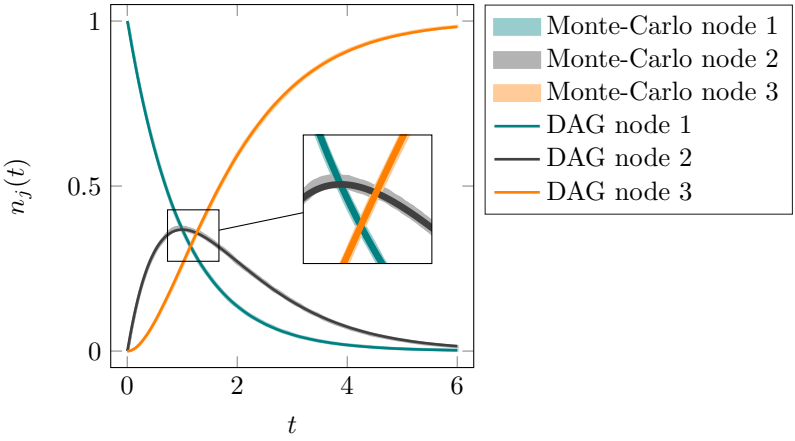
The right-hand-side of (3.27) can be computed directly in the time-domain, or through a Laplace transform and we have thus found the propagator. An integration of the analytical model is compared against Monte-Carlo simulation on figure 3.2 under assumption 2.5 that all random variables are exponentially distributed.

3.2.2.1. *Time-domain transport equation*

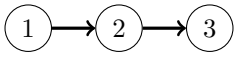
From (3.15) we can compute $\dot{\boldsymbol{n}}(t)$ in terms of the transition kernel and of the initial condition. Applying Leibniz's rule for differentiation under the integral sign, we obtain

$$\dot{n}_j(t) = q_j(t) - \int_0^t q_j(\tau) \sum_{i \in V_j} T_{ij}(t-\tau)d\tau = q_j(t) - \int_0^t q_j(\tau)\boldsymbol{\Psi}_j(t-\tau)d\tau. \tag{3.28}$$

The interpretation is that the rate of variation of $n_j(t)$ is given by a sum of all arrivals minus the departures, with each departure resulting from a previous



(A) The solid curves represent the analytical probabilities $n_j(t)$. Monte-Carlo simulation based on 10^5 independent trajectories determined shaded areas representing a confidence interval of one standard deviation above and below the mean. The mean was not drawn to preserve readability.



(B) The walker is initially located on node 1.

FIGURE 3.2. Validation of the analytical model on a DAG with exponential distributions with equal rates, $\mu = \eta = \lambda = 1$.

arrival at any point in time. In matrix form, with the obvious definition for the diagonal operator $\mathcal{Q}$, eq. (3.28) reads

$$\dot{\boldsymbol{n}}(t) = (I - \mathcal{Q})\boldsymbol{q}(t) = (I - \mathcal{Q}) \sum_{k=0}^{\infty} \mathcal{T}^k \boldsymbol{q}^{(0)}(t) \tag{3.29}$$

and using eq. (3.26), we formally obtain the generalized master equation

$$\dot{\boldsymbol{n}}(t) = (I - \mathcal{Q})\mathcal{P}^{-1}\boldsymbol{n}(t). \tag{3.30}$$

It should be remarked that the operator $(I - \mathcal{Q})\mathcal{P}^{-1}$ is nonlocal in time, and that this equation is generally non-Markovian, or put differently, that the process it describes has memory, as we discuss in the next section.

3.3. Emergence of memory

There are two causes for the loss of the Markov properties in time and trajectory. The first one is dynamical and due to the different timescales, whereas the second one is topological with the existence of cycles. We will illustrate both contributions numerically. Our approach is to neglect these sources in the modeling, and to compare the result against Monte-Carlo simulation.

3.3.1. Neglecting timescales

Figure 3.1 hints towards the existence of three possible scenarios: one where the dynamics of the walker dominates, one where the down-times dominate, and a third one where the modeling should clearly account for all the three processes. This claim is sustained by an illustrative example on fig. 3.3 where Monte-Carlo simulation over $[0, T]$ compares the error in the modeling when keeping only one timescale,

$$e_{Mod.1}(T) = \int_0^T \left\| \boldsymbol{n}^{(Mod.1)}(t) - \boldsymbol{n}(t) \right\|_2 dt \tag{3.31}$$

$$e_{Mod.3}(T) = \int_0^T \left\| \boldsymbol{n}^{(Mod.3)}(t) - \boldsymbol{n}(t) \right\|_2 dt \tag{3.32}$$

for a range of values of the exponential rates (η, λ) on the edges, while the walkers timescale is normalized, $\mu = 1$. Here $\boldsymbol{n}^{(Mod.1)}(t)$ and $\boldsymbol{n}^{(Mod.3)}(t)$ are respectively given by the node-centric eq. (2.6) and the edge-centric eq. (2.10), whereas $\boldsymbol{n}(t)$ results from Monte-Carlo simulation. The existence of the two regimes (low error, teal-colored regions) corresponding to the dotted areas on fig. 3.1 is hereby confirmed. Keep in mind that the error levels displayed on the figure obviously dependent on the various topological and dynamical parameters, but also on the duration of the simulation, and that this plot hardly provides any rule to decide when timescales may be neglected.

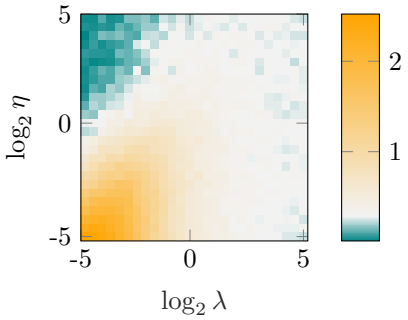
Section 3.4 aims at providing the necessary modeling framework to cover the full domain of this plot, including the region where both errors are large and where the full interplay of the walker's and edges' behaviors should be accounted for.

3.3.2. Neglecting the cycles in the graph

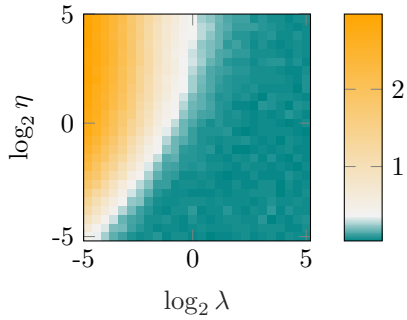
The results derived thus far relied on an assumption of independence between events, edges appearing or vanishing, encountered by the random walker. This assumption ceases to hold true when the underlying network has cycles. In chapter 2, we have seen with fig. 2.8 that despite the qualitative agreement of the steady-state-based ranking between analytical modeling on DAGs and Monte-Carlo simulation, there was some amount of quantitative discrepancy between the two. The purpose of fig. 3.4 is to highlight the significant deviations between the acyclic approximation and the numerical simulation of the walk, even in situations when each of the three processes is Poisson, on the small graph with cycles visible on fig. 3.3c. The semi-Markov property is visibly lost, and the memory effect is visibly stronger when the walker's timescale is faster. These observations justify the corrective framework of the next section.

3.4. Mesoscopic modeling with cycles

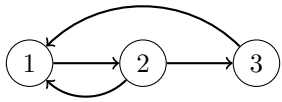
In general, if cycles are present in the network, the state space is essentially the full trajectory of the random walk, which makes the problem intractable



(A) Error of the passive edge-centric model, eq. (3.32).



(B) Error of the active node-centric model, eq. (3.31)



(C) Directed graph with a 2-cycle and a 3-cycle

FIGURE 3.3. Neglecting timescales. This figure provides a comparison of the passive edge-centric (A) and active-node centric (B) models with Monte-Carlo simulation involving $4 \cdot 10^3$ independent trajectories on the graph in (C). The error between the predictions of those single-timescale models and the actual (Monte-Carlo) probabilities $n_j(t)$ is plotted for various combinations of the rates η and λ, with $\mu = 1$. The time interval of the simulation is $[0, 5]$.

analytically. We therefore propose a method estimating the corrections due to cycles of a given length. Although the proposed framework is general, we restrict the following discussion to contributions of cycles of length two. This will speed up the numerical simulations, as the incorporation of long cycles comes with increased computational cost. Also note that longer cycles are associated to weaker corrections, as more time between two passages tends to wash out footprints left by the walker.

3.4.1. Propagator with corrections for 2-cycles

In order to allow a correction for 2-cycles, we will basically view the walk as a second order Markov chain. Accordingly, we condition the arrival-times on a node based on the time of the previous move. Let us define $q_{jmm'}(\tau, \nu)$ to be the arrival time density for the couple (τ, ν) on nodes $m' \to m \to j$. Observe that almost surely, $0 < \nu < \tau$. As depicted by figure 3.5, let $T_{i|jmm'}(t|\tau, \nu)$ be the conditional transition kernel across edge $j \to i$ at time t, taking the two previous jumps into account: from m' to m at time ν and from m to j at time τ.

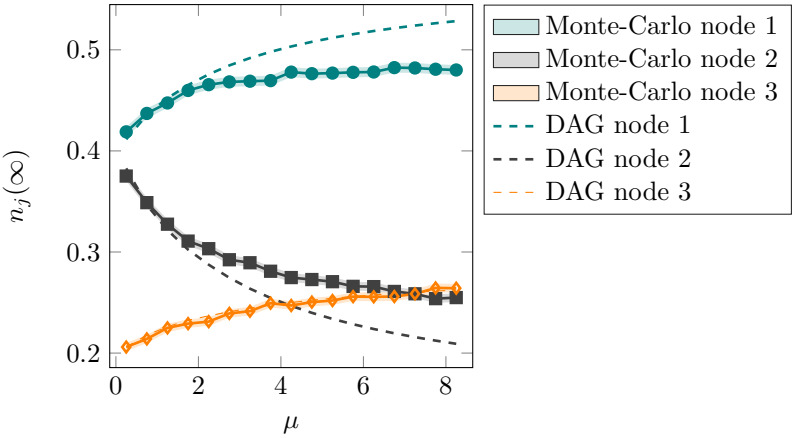


FIGURE 3.4. Neglecting the cycles. The stationary state $n_j(\infty)$ for several values of the rate μ of the walker is plotted resulting from Monte-Carlo simulation (solid lines with filled markers) and the analytical model on DAGs (dashed lines). The width of the shading around the Monte-Carlo curves corresponds to twice the standard deviation of the mean computed from $4 \cdot 10^4$ independent trajectories. The up- and down-time rates are $\eta = 1 = \lambda$ and the initial condition of the walk is $\boldsymbol{n}(0) = (1, 0, 0)^T$. This figure indicates that the memory effect is more pronounced when the walker's rate increases. We used the graph of fig. 3.3c.

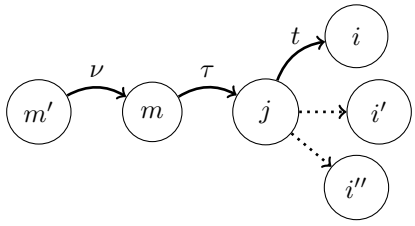


FIGURE 3.5. Jump times and nodes in the definition of the transition kernel $T_{i|jmm'}(t|\tau,\nu)$ and of the survival probability $\Phi_{jmm'}(t|\tau,\nu)$. Here, nodes m' and j and nodes m and i are not necessarily different nodes.

By the limited amount of memory we take into account, this conditional kernel actually depends only on the durations $t-\nu$ and $\tau-\nu$:

$$T_{i|jmm'}(t|\tau,\nu) = T_{i|jmm'}(t-\nu|\tau-\nu,0) =: \widetilde{T}_{i|jmm'}(t-\nu|\tau-\nu). \tag{3.33}$$

Here we use the tilde symbol $\widetilde{\ }$ to indicate explicitly that this is the kernel after time-homogeneity was used. Let also $\Phi_{jmm'}(t|\tau,\nu)$ be the survival probability on node j, having made the two previous jumps at times $\nu \leq \tau$. We have

$$\Phi_{jmm'}(t|\tau,\nu) = 1 - \sum_{i\in V_j} \int_\tau^t T_{i|jmm'}(s|\tau,\nu)\, ds. \tag{3.34}$$

The normalization condition reads

$$\lim_{t\to\infty} \Phi_{jmm'}(t|\tau,\nu) = 0, \quad \forall\ 0 \le \nu \le \tau, \tag{3.35}$$

or equivalently, in terms of transitions

$$\sum_{i\in V_j} \int_\tau^\infty T_{i|jmm'}(s|\tau,\nu)\, ds = 1 \tag{3.36}$$

for all $0 \le \nu \le \tau$ and $1 \le j \le N$. In the remainder of this section the computations assume the conditional transition density to be known. Its exact form will be determined in section 3.4.2.

Using the same steps as for acyclic graphs, let us first write the probability that the walker is on node j at time t as

$$n_j(t) = n_j^{(0)}(t) + n_j^{(1)}(t) + n_j^{(k\ge 2)}(t), \tag{3.37}$$

where the superscript refers to the number of jumps performed up to time t. The first two terms are not impacted by the memory effect, and can be computed based on the transition kernel for DAGs:

$$n_j^{(0)}(t) = \int_0^t \Phi_j(t,\tau) q_j^{(0)}(t) d\tau = \Phi_j(t,0) n_j(0), \tag{3.38}$$

and

$$n_j^{(1)}(t) = \int_0^t \Phi_j(t,\tau) q_j^{(1)}(\tau) d\tau = \sum_{m\in V_j'} n_m(0) \int_0^t \Phi_j(t,\tau) T_{jm}(\tau,0) d\tau. \tag{3.39}$$

It remains to compute $n_j^{(k\ge 2)}(t) = \sum_{k\ge 2} n_j^{(k)}(t)$. Note that in $n_j^{(k)}(t)$ we also need the transition density of the $(k+1)$-th jump which determines the survival probability on node j after k jumps. For all $k \ge 2$ we write

$$n_j^{(k)}(t) = \sum_{m'\to m\to j} \sum \iint_{0\le\nu\le\tau} \Phi_{jmm'}(t|\tau,\nu) q_{jmm'}^{(k,k-1)}(\tau,\nu)\, d\nu d\tau, \tag{3.40}$$

where again the superscript in $q_{jmm'}^{(k,k-1)}$ gives the number of jumps. In order to determine $n_j^{(k\ge 2)}(t)$ we will need

$$q_{jmm'}(\tau,\nu) = \sum_{k\ge 2} q_{jmm'}^{(k,k-1)}(\tau,\nu). \tag{3.41}$$

Once we have computed this quantity, then the third term in (3.37), $n_j(t) = n_j^{(0)}(t) + n_j^{(1)}(t) + n_j^{(k\ge 2)}(t)$, will indeed follow as

$$n_j^{(k\ge 2)}(t) = \sum_{m'\to m\to j} \sum \iint_{0\le\nu\le\tau} q_{jmm'}(\tau,\nu) \Phi_{jmm'}(t|\tau,\nu)\, d\nu d\tau, \tag{3.42}$$

and we have the probability $n_j(t)$ as a function of the initial condition $\boldsymbol{n}(0)$.

Let us therefore determine the arrival-times density in a given number of jumps, $q^{(k,k-1)}_{jmm'}(\cdot,\cdot)$. Let us write equation (3.41) by splitting the sum as

$$q_{jmm'}(\tau,\nu) = \sum_{k=2}^{\infty} q^{(k+1,k)}_{jmm'}(\tau,\nu) + q^{(2,1)}_{jmm'}(\tau,\nu). \tag{3.43}$$

In this expression, for all $k \geq 2$,

$$q^{(k+1,k)}_{jmm'}(\tau,\nu) = \sum_{m''\in V'_{m'}} \int_0^{\nu} T_{j|mm'm''}(\tau|\nu,\nu') q^{(k,k-1)}_{mm'm''}(\nu,\nu')\, d\nu' \tag{3.44}$$

and using again (3.41), equation (3.43) becomes

$$q_{jmm'}(\tau,\nu) = \sum_{m''\in V'_{m'}} \int_0^{\nu} T_{j|mm'm''}(\tau|\nu,\nu') \times q_{mm'm''}(\nu,\nu')\, d\nu' + q^{(2,1)}_{jmm'}(\tau,\nu). \tag{3.45}$$

The extended initial condition of arrival times for the first two jumps is given by

$$\begin{aligned} q^{(2,1)}_{jmm'}(\tau,\nu) &= \widetilde{T}_{im}(\tau-\nu) \int_0^{\nu} \widetilde{T}_{mm'}(\nu-\nu') q^{(0)}_{m'}(\nu') d\nu' \\ &= \widetilde{T}_{im}(\tau-\nu)\widetilde{T}_{mm'}(\nu) n_{m'}(0) \end{aligned} \tag{3.46}$$

where $\widetilde{T}_{ji}(t) := T_{ij}(t,0)$ is the transition kernel for the acyclic case.

Equation (3.45) is a linear Volterra integral equation of the second kind, with kernel given by the conditional transition kernel that we determine hereafter. We have a vector of unknown functions where each component $q_{jmm'}(\cdot,\cdot) : [0,\infty)^2 \to [0,\infty)$ corresponds to a path of length 2 in the underlying graph. As will become clear, this equation cannot be cast under the form of a convolution. Consequently, the Laplace-transform-based method cannot be applied.

3.4.2. Transition kernel with correction for 2-cycles

We want to compute the conditional transition kernel $T_{i|jmm'}(t|\tau,\nu)$, where the trajectory before ν is not taken into account, that is, we need to determine

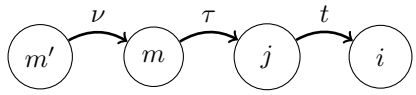
$$\widetilde{T}_{i|jmm'}(x|y) = T_{i|jmm'}(x|y,0), \quad 0 \leq y \leq x. \tag{3.47}$$

Assumption 2.5 still applies again throughout this section. Note that the letters t,τ,ν will indicate absolute times, whereas x and y are durations. We will keep both in order to avoid having to assume a jump a time 0. There are three cases represented on fig. 3.6, depending on whether $m' \to m \to j$ is a cycle or not. In the first case there is no memory effect due to 2-cycles and the density reads as before

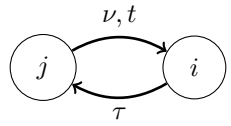
$$\widetilde{T}_{i|jmm'}(x|y) = \widetilde{T}_{ij}(x-y) \tag{3.48}$$

where the right-hand side comes from the modeling for DAGs. In the second and third cases the kernel cannot be written in terms of the one obtained for acyclic graphs. As before, in these cases we will write

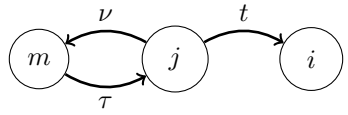
$$\widetilde{T}_{i|jmm'}(x|y) = \mathbf{(1)} + \mathbf{(2)}, \tag{3.49}$$



(A) Case 1.



(B) Case 2.



(C) Case 3.

FIGURE 3.6. Three cases for the transition kernel $T_{i|jmm'}(t|\tau,\nu)$ with corrections for cycles. The arrows are labeled by the jump time. Not all possible out-neighbors of node j are represented, although they would influence the transition kernel.

where the first term corresponds to a jump at the end of the waiting time on the node, whereas the second term is for the jump of a trapped walker. The computation of both terms requires first to determine the probability for an edge to be (un)available some time after exact knowledge of its state.

3.4.2.1. *Two corrections on r*

When the walker returns to a node after completion of a 2-cycle, the next destination node depends on the choice previously made from the same location. Firstly, the outgoing edge that was selected at the beginning of the cycle has an increased probability with respect to $r = \langle \mathsf{U} \rangle / (\langle \mathsf{U} \rangle + \langle \mathsf{D} \rangle)$, to still be available. The smaller the time $y = \tau - \nu$ to go through the cycle and the subsequent walker's waiting time, the more pronounced this effect. Secondly, the converse is also true: any other edge that wasn't selected is more likely to have been and still be unavailable some short time later.

3.4.2.2. *First correction with p^**

The proof of fact 2.9 presents the derivation for the first correction,

$$p^*(s,\nu) = P\left\{ j \to i \text{ is up at } s \,|\, j \to i \text{ was up at } \nu \right\}, \tag{3.50}$$

where the knowledge of the state of the edge stems from a move from j to i at time ν and where $p^*(s,\nu) \equiv p^*(s-\nu)$. This probability does not depend on the number of nodes under assumption 2.5 that the distributions are exponential, see remark 2.7. Observe that in chapter 2, we have actually computed the probability

for an edge to still be unavailable, knowing is was originally unavailable. But it suffices to swap U and D in the reasoning, and moreover this operation is transparent when both distributions have the same rate, $\eta = \lambda$. In that case it was found that $p^*(x) = \frac{1}{2}(1 + e^{-2\lambda x})$, see fig. 2.5.

3.4.2.3. *Second correction with* $p_j^\dagger$

Consider two consecutive jumps of the walker following a cycle $j \to i \to j$ of the underlying graph $\mathcal{G}$, where node j has at least one neighbor i' other than i. Let ν be the time of the first jump of the walker trough $j \to i$. The aim is to determine the second correction captured by $p_j^\dagger(s, \nu)$, namely

$$p_j^\dagger(s,\nu) = P\left\{j \to i' \text{ is up at } s \,|\, j \to i \text{ selected at } \nu\right\}, \tag{3.51}$$

where we write $p_j^\dagger(s,\nu) \equiv p_j^\dagger(s-\nu)$.

For simplicity we consider that the random variables X_u and X_d follow the same exponential distribution, assumption 2.8. The expression of $p_j^\dagger$ will be given in fact 3.4 after a preliminary identity is established in fact 3.3.

Fact 3.3 (Combinatorial identity)**.** *For $n \in \mathbb{N}$, $n \geq 2$ and $0 \leq p \leq 1$, it holds that*

$$\sum_{k=1}^{n} \binom{n-1}{k-1} \frac{1}{k+1} p^k (1-p)^{n-k} = (1-p)^{n+1} + (n+1)p - 1. \tag{3.52}$$

Proof. First note that for all $1 \leq k \leq n-1$ we have

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{3.53}$$

Then, still for $n \geq 2$, write the left-hand side of eq. (3.52) as

$$\begin{aligned}
\sum_{k=1}^{n} \binom{n-1}{k-1} \frac{1}{k+1} p^k (1-p)^{n-k} &= \sum_{k=1}^{n-1} \binom{n-1}{k-1} \frac{1}{k+1} p^k (1-p)^{n-k} \\
&\qquad + \binom{n-1}{n-1} \frac{1}{n+1} p^k \\
&= \underbrace{\sum_{k=1}^{n-1} \binom{n}{k} \frac{1}{k+1} p^k (1-p)^{n-k}}_{(a)} \\
&\qquad - \underbrace{\sum_{k=1}^{n-1} \binom{n-1}{k} \frac{1}{k+1} p^k (1-p)^{n-k}}_{(b)} + \frac{1}{n+1} p^k,
\end{aligned} \tag{3.54}$$

where we have used eq. (3.53) to obtain the second equality. Let us consider term (a) and write

$$\begin{aligned}(a) &= \sum_{k=1}^{n-1} \frac{n!}{(n-k)!(k+1)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n-1} \frac{1}{n+1} \frac{(n+1)!}{(n+1-k-1)!(k+1)!} p^k (1-p)^{n-k} \\ &= \frac{1}{n+1} \sum_{k=1}^{n-1} \binom{n+1}{k+1} p^k (1-p)^{n-k} \\ &= \frac{1}{n+1} \sum_{k=1}^{n-1} \binom{n+1}{k+1} \frac{1}{p} p^{k+1} (1-p)^{n+1-k-1},\end{aligned}$$

translating the indexes k,

$$= \frac{1}{(n+1)p} \sum_{k=2}^{n} \binom{n+1}{k} p^k (1-p)^{n+1-k},$$

then letting $\tilde{n} := n+1$

$$\begin{aligned}&= \frac{1}{(n+1)p} \sum_{k=2}^{\tilde{n}-1} \binom{\tilde{n}}{k} p^k (1-p)^{\tilde{n}-k} \\ &= \frac{1}{(n+1)p} \left(\sum_{k=0}^{\tilde{n}} \binom{\tilde{n}}{k} p^k (1-p)^{\tilde{n}-k} - \binom{\tilde{n}}{0} (1-p)^{\tilde{n}} \right. \\ &\qquad \left. - \binom{\tilde{n}}{1} p (1-p)^{\tilde{n}-1} - \binom{\tilde{n}}{\tilde{n}} p^{\tilde{n}} \right),\end{aligned}$$

and reverting to the variable n

$$= \frac{1}{(n+1)p} \left(1 - (1-p)^{n+1} - (n+1)p(1-p)^n - p^{n+1} \right).$$

Next, consider (b) in eq. (3.54). We find

$$\begin{aligned}(b) &= \sum_{k=1}^{n-1} \frac{(n-1)!}{(n-1-k)!k!} \frac{1}{k+1} p^k (1-p)^{n-k} \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \binom{n}{k+1} p^k (1-p)^{n-k},\end{aligned}$$

translating the index k,

$$\begin{aligned}&= \frac{1}{np}(1-p) \sum_{k=2}^{n} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{1}{np}(1-p) \left(\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} - \binom{n}{0} (1-p)^n - \binom{n}{1} p (1-p)^{n-1} \right) \\ &= \frac{1}{np}(1-p) \left(1 - (1-p)^n - np(1-p)^{n-1} \right).\end{aligned}$$

Inserting the expressions for (a) and (b) in eq. (3.54) yields

$$\begin{aligned}\sum_{k=1}^{n}\binom{n-1}{k-1}\frac{1}{k+1}p^k(1-p)^{n-k} &= \frac{1}{n(n+1)p}\Big(n-n(1-p)^{n+1}\\ &\quad -n(n+1)p(1-p)^n-np^{n+1}\\ &\quad -(n+1)(1-p)+(n+1)(1-p)^{n+1}\\ &\quad +n(n+1)p(1-p)^n+np^{n+1}\Big)\\ &= (1-p)^{n+1}+(n+1)p-1,\end{aligned}$$

which is the desired form. □

Fact 3.4. *Under assumption 2.8, for every $j\in 1,\dots,N$ such that $k_j>2$, the length-two-cycle correction to $r=\frac{1}{2}$ is given by*

$$p_j^\dagger(s,\nu)=(2a_j-1)p_i^*(s,\nu)-a_j+1, \tag{3.55}$$

where $a_j:=\frac{k_jr+(1-r)^{k_j}-1}{k_j-1}$.

Proof. Let E_s and E_ν denote respectively the events that $j\to i$ is up at time s and at time ν. Let E'_s and E'_ν be the corresponding events for edge $j\to i'$ and let also F_ν be the event that edge $j\to i$ was selected by the walker for the move performed at time ν. Using the law of total probabilities for conditional probabilities we have[1]

$$\begin{aligned}p_j^\dagger(s,\nu) &= P\{E'_s|F_\nu\}\\ &= P\{E'_s\cap E'_\nu|F_\nu\}+P\left\{E'_s\cap\overline{E'_\nu}|F_\nu\right\}\\ &= P\{E'_s|E'_\nu\cap F_\nu\}\,P\{E'_\nu|F_\nu\}+P\left\{E'_s|\overline{E'_\nu}\cap F_\nu\right\}P\left\{\overline{E'_\nu}|F_\nu\right\}.\end{aligned}$$

Now, using the assumption that the up- and down-times follow the same distribution, $P\{E'_s|E'_\nu\cap F_\nu\}=p^*(s,\nu)$ and $P\left\{E'_s|\overline{E'_\nu}\cap F_\nu\right\}=1-p^*(s,\nu)$. Also observe that $P\left\{\overline{E'_\nu}|F_\nu\right\}=1-P\{E'_\nu|F_\nu\}$. So it only remains to compute

$$a_j:=P\{E'_\nu|F_\nu\}, \tag{3.56}$$

the probability for an edge to be available at some time, knowing a jump was performed through a competing edge at the same time. This will yield the final expression

$$p_j^\dagger(s,\nu)=(2a_j-1)p_i^*(s,\nu)-a_j+1. \tag{3.57}$$

Let H_ν be the event that the jump at time ν happened after the walker was trapped. Since we assume second order time-homogeneity, namely no knowledge of the past trajectory up to time ν, we have

$$P\{H_\nu\}=(1-r)^{k_j}=:b_j. \tag{3.58}$$

Using again the law of total probabilities,

$$a_j=\underbrace{P\{E'_\nu|F_\nu\cap H_\nu\}}_{=0}P\{H_\nu|F_\nu\}+P\left\{E'_\nu|F_\nu\cap\overline{H_\nu}\right\}\underbrace{P\left\{\overline{H_\nu}|F_\nu\right\}}_{=1-b_j}. \tag{3.59}$$

[1]We write $\bar{A}$ the complement of event A, such that $P\{A\cup\bar{A}\}=1$ and $P\{A\cap\bar{A}\}=0$.

In the second term,

$$P\left\{E'_\nu | F_\nu \cap \overline{H_\nu}\right\} = \frac{P\left\{E'_\nu \cap F_\nu \cap \overline{H_\nu}\right\}}{P\left\{F_\nu \cap \overline{H_\nu}\right\}} \tag{3.60}$$

where the denominator is decomposed as

$$P\left\{F_\nu \cap \overline{H_\nu}\right\} = P\left\{F_\nu \cap \overline{H_\nu} | E'_\nu\right\} P\left\{E'_\nu\right\} + P\left\{F_\nu \cap \overline{H_\nu} | \overline{E'_\nu}\right\} P\left\{\overline{E'_\nu}\right\} \tag{3.61}$$

with $P\{E'_\nu\} = r = 1 - P\left\{\overline{E'_\nu}\right\}$. Moreover, let $E_\nu^{(\ell)}$ be the event that ℓ out of $k_j - 2$ out-neighbors of node j are reachable at time ν, so that

$$\begin{aligned}
P\left\{F_\nu \cap \overline{H_\nu} | E'_\nu\right\} = P\{F_\nu | E'_\nu\} &= \sum_{\ell=0}^{k_j-2} P\left\{F_\nu | E_\nu^{(\ell)} \cap E'_\nu\right\} P\left\{E_\nu^{(\ell)} | E'_\nu\right\} \\
&= \sum_{\ell=0}^{k_j-2} r \frac{1}{\ell+2} \times \binom{k_j-2}{\ell} r^\ell (1-r)^{k_j-2-\ell} \\
&= \sum_{\ell=0}^{k_j-2} \binom{k_j-2}{\ell} \frac{1}{\ell+2} r^{\ell+1} (1-r)^{k_j-2-\ell} \\
&= \sum_{\ell=1}^{k_j-1} \binom{k_j-2}{\ell-1} \frac{1}{\ell+1} r^\ell (1-r)^{k_j-1-\ell}.
\end{aligned} \tag{3.62}$$

Using fact 3.3 the right-hand side of (3.62) reads

$$P\left\{F_\nu \cap \overline{H_\nu} | E'_\nu\right\} = \frac{k_j r + (1-r)^{k_j} - 1}{k_j(k_j-1)r}, \quad k_j \geq 2. \tag{3.63}$$

Similarly, for the remaining factor of (3.61) we have

$$\begin{aligned}
P\left\{F_\nu \cap \overline{H_\nu} | \overline{E'_\nu}\right\} &= \sum_{\ell=0}^{k_j-2} P\left\{F_\nu \cap \overline{H_\nu} | E_\nu^{(\ell)} \cap E'_\nu\right\} P\left\{E_\nu^{(\ell)} | E'_\nu\right\} \\
&= \underbrace{P\left\{F_\nu \cap \overline{H_\nu} | E_\nu^{(0)} \cap \overline{E'_\nu}\right\}}_{=r} \times \underbrace{P\left\{E_\nu^{(0)} | \overline{E'_\nu}\right\}}_{=(1-r)^{k_j-2}} \\
&\quad + \sum_{\ell=1}^{k_j-2} \underbrace{P\left\{F_\nu \cap \overline{H_\nu} | E_\nu^{(\ell)} \cap \overline{E'_\nu}\right\}}_{=\frac{1}{\ell+1} r} \times \underbrace{P\left\{E_\nu^{(\ell)} | \overline{E'_\nu}\right\}}_{=\binom{k_j-2}{\ell} r^\ell (1-r)^{k_j-2}} \\
&= \sum_{\ell=0}^{k_j-2} \binom{k_j-2}{\ell} \frac{1}{\ell+1} r^{\ell+1} (1-r)^{k_j-2-\ell} \\
&= \sum_{\ell=1}^{k_j-1} \binom{k_j-2}{\ell-1} \frac{1}{\ell} r^\ell (1-r)^{k_j-1-\ell},
\end{aligned} \tag{3.64}$$

and relying again on fact 2.6,

$$P\left\{F_\nu \cap \overline{H_\nu} | \overline{E'_\nu}\right\} = \frac{1-(1-r)^{k_j-1}}{k_j-1}, \quad k_j \geq 2. \tag{3.65}$$

Inserting (3.63) and (3.65) in (3.61) leads to writing (3.60) as

$$P\left\{E'_\nu | F_\nu \cap \overline{H_\nu}\right\} = \frac{k_j r + (1-r)^{k_j} - 1}{(1-k_j)\left((1-r)^{k_j}-1\right)}, \tag{3.66}$$

and eventually (3.56) becomes

$$a_j = \frac{k_j r + (1-r)^{k_j} - 1}{k_j - 1}, \quad k_j \geq 2. \tag{3.67}$$

The expression of $p_j^\dagger(s,\nu)$ results from inserting (3.67) into (3.57). □

Under assumption 2.8 that U and D share the same rate λ, it follows directly from (3.55) that

$$p_j^\dagger(s,\nu) = \frac{1}{2} - \frac{1}{4}e^{-2\lambda(s-\nu)}, \tag{3.68}$$

when we have set $k_j = 2$, and hence $a_j = \frac{1}{4}$, a choice that maximizes the importance of this correction. The second term represents the difference with respect to $r = \frac{1}{2}$, and is such that $p_j^\dagger(s,\nu) \to \frac{1}{4}$ if $s - \nu \to 0^+$ and $p_j^\dagger(s,\nu) \to \frac{1}{2}$ if $s - \nu \to +\infty$, see fig. 3.7.

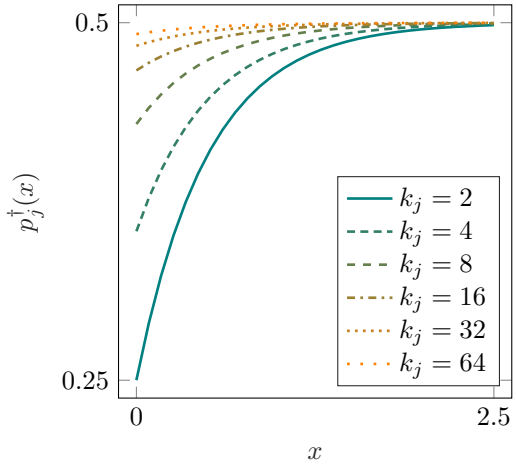


FIGURE 3.7. Second correction on r. Contrary to the first correction determined by p^*, here there is dependence not only on the dynamical parameters, but also on the topology through the local connectivity: the larger the degree, the weaker the correction $p^\dagger(x) - r$, which vanishes in the limit $k_j \to \infty$. The rates for the edges are $\eta = 1 = \lambda$.

3.4.2.4. *The second case:* $T_{i|jij}(t|\tau,\nu)$

Having computed the necessary corrections on r, we are now in position to further develop equation (3.49). Assumption 2.5 allows us to neglect the bus paradox, and hence sensibly simplifies the derivation. The first term, when the walker was not trapped before the move, reads

$$\begin{aligned}\mathbf{(1)}_{(i|jij)} &= \psi_j(t-\tau)\sum_{\ell=1}^{k_j}\frac{1}{\ell}p^*(t,\nu)\binom{k_j-1}{\ell-1}\times(p_j^\dagger(t,\nu))^{\ell-1}(1-p_j^\dagger(t,\nu))^{k_j-\ell}\\ &= \frac{p^*(t,\nu)}{p_j^\dagger(t,\nu)}\psi_j(t-\tau)\left[\frac{1-(1-p_j^\dagger(t,\nu))^{k_j}}{k_j}\right].\end{aligned} \tag{3.69}$$

We notice that this expression is the same as for the acyclic graphs, up to a correction factor $p^*(t,\nu)/p_j^\dagger(t,\nu)$, and with r replaced by $p_j^\dagger(t,\nu)$.

Using the same approach as for p^* in section 2.3, we obtain the second term of $T_{i|jij}(t|\tau,\nu)$ corresponding to a walker who was trapped before moving:

(3.70)
$$\begin{aligned}\mathbf{(2)}_{(i|jij)} = \int_\tau^t \psi_j(s-\tau)\times\left[\sum_{k=0}^{\infty}\int_0^{s-\nu}\left(\mathsf{U}*\mathsf{D}^{*k}*\mathsf{U}^{*k}\right)(r)\times\mathsf{D}(t-\nu-r)\,dr\right]\\ \times\left[(1-p_j^\dagger(s,\nu))P\{w>t-s\}\right]^{k_j-1}ds.\end{aligned}$$

Relying on the previous computation of $p^*(s,\nu)$, expression (3.70) simplifies to

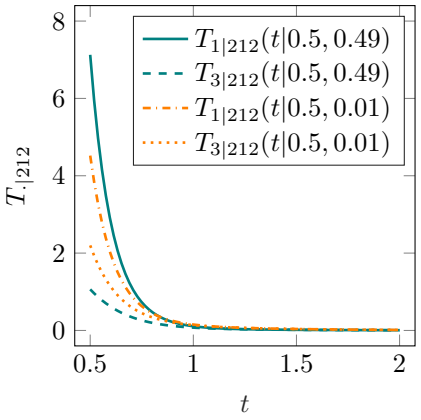
$$\begin{aligned}\mathbf{(2)}_{(i|jij)} = \int_\tau^t \psi_j(s-\tau)(1-p^*(s,\nu))\mathsf{D}(t-s)\\ \times\left[(1-p_j^\dagger(s,\nu))P\{w>t-s\}\right]^{k_j-1}ds.\end{aligned} \tag{3.71}$$

In this expression, $(1-p^*(s,\nu))\mathsf{D}(t-s)$ refers to the probability that edge $j\to i$ is down at time s, and will remain so exactly until time t when it becomes available to the jumper again.

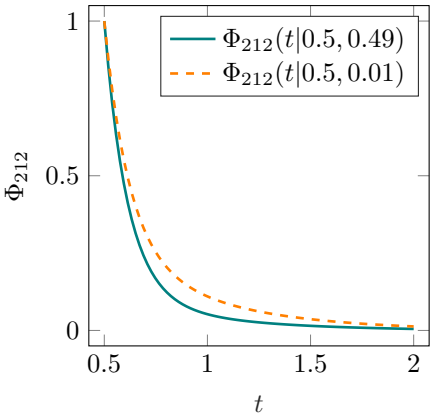
3.4.2.5. *The third case:* $T_{i|jmj}(t|\tau,\nu)$ *with* $m\neq j$

The first term of the transition kernel in the case of fig. 3.6c, when the walker was not trapped, is given by

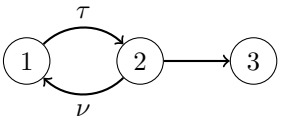
$$\begin{aligned}\mathbf{(1)}_{(i|jmj)} = \psi_j(t-\tau)\Big[p^*(t,\nu)\times P\{\text{choose } i\mid j\to m \text{ is up}\}+\\ (1-p^*(t,\nu))\times P\{\text{choose } i\mid j\to m \text{ is down}\}\Big],\end{aligned} \tag{3.72}$$



(A) The differentiation between the transition kernel towards nodes 1 and 3 is more pronounced when the walker goes faster around the cycle $2 \to 1 \to 2$, than when it takes longer, $\tau - \nu = 0,01$ (teal) *vs* $\tau - \nu = 0,49$ (orange). The bias decreases with time $t - \nu$.



(B) After completion of the cycle $2 \to 1 \to 2$, the memory effect translates into a probability to stay put on node 2 that is lower when $\tau - \nu$ is small (teal), as compared to when $\tau - \nu$ is large (orange, weaker memory).



(C) The walker initialy moves from node 2 to 1 at ν then back at τ. The rates are $\mu = 8$, $\eta = 1 = \lambda$.

FIGURE 3.8. Stronger (teal) vs weaker (orange) memory effect on the transition kernel and survival probability depending on the time to go through a cycle of the graph.

where the two still undetermined probabilities are for events at time t. They are determined by

$$P\{\text{choose } i \mid j \to m \text{ is up}\}$$
$$= p_j^\dagger(t,\nu) \sum_{\ell=0}^{k_j-2} \binom{k_j-2}{\ell} \frac{1}{\ell+2} (p_j^\dagger(t,\nu))^\ell (1-p_j^\dagger(t,\nu))^{k_j-\ell-2}$$
$$= \frac{k_j p_j^\dagger(t,\nu) + (1-p_j^\dagger(t,\nu))^{k_j-1}}{k_j(k_j-1)p_j^\dagger(t,\nu)} \tag{3.73}$$

and

$$
\begin{aligned}
&P\{\text{choose } i \mid j \to m \text{ is down}\} \\
&\qquad = p_j^\dagger(t,\nu) \sum_{\ell=0}^{k_j-2} \binom{k_j-2}{\ell} \frac{1}{\ell+1} (p_j^\dagger(t,\nu))^\ell (1-p_j^\dagger(t,\nu))^{k_j-\ell-2} \\
&\qquad = \frac{1-(1-p_j^\dagger(t,\nu))^{k_j-1}}{k_j-1},
\end{aligned}
\tag{3.74}
$$

where the final forms (3.73) and (3.74) were obtained as in the derivation of section 3.4.2.3.

The second term $\mathbf{(2)}_{(i|jmj)}$ associated to a trapped walker can be shown to have the same expression as in (3.71).

Combining the effects of p^* and $p_j^\dagger$ results in figure 3.8 where it appears clearly that a shorter time to go around a cycle induces a stronger bias in favor of repeating a previous jump, hence inducing a positive correlation between the successive moves.

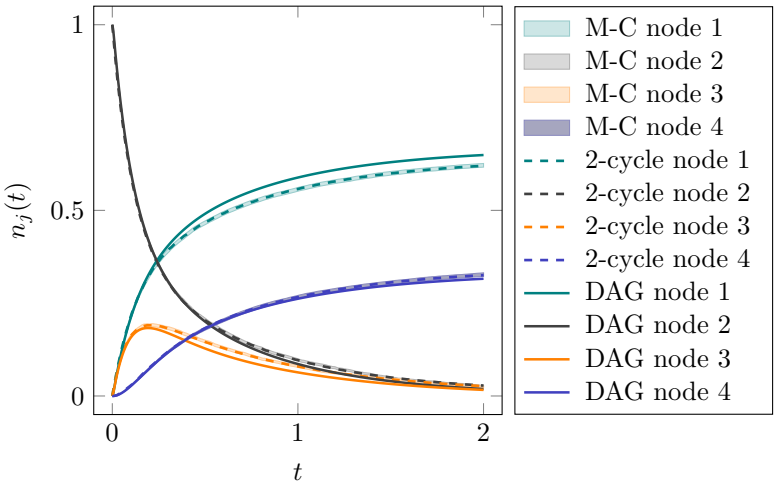
3.5. Numerical validation

We solved the Volterra vector integral equations (3.22) of the modeling for the DAGs, and (3.45) of the modeling with corrections, by applying a trapezoidal scheme for discretization of the integrals [**35**]. In the case of (3.45) this resulted in a linear system of equations, and we relied on its block-triangular structure due to causality of events, to solve it with a limited need of memory.

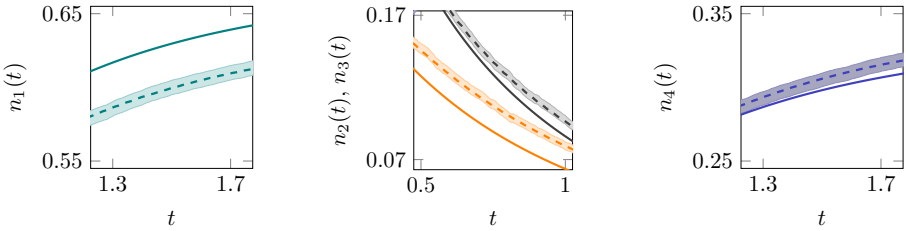
The initial condition $\boldsymbol{q}^{(0)}(t) = \boldsymbol{n}(0)\delta(t)$ arising in these equations was approximated using a positive normal distribution with probability density function $\delta_\epsilon(t)$ parametrized by a small parameter ϵ, such that

$$
\boldsymbol{q}^{(0)}(t) \approx \boldsymbol{n}(0)\delta_\epsilon(t), \quad \int_0^\infty \delta_\epsilon(t)dt = 1. \tag{3.75}
$$

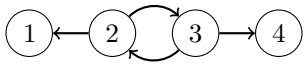
A validation of the comprehensive analytical framework through a simple numerical example is the purpose of figure 3.9, which we interpret as follows. Due to the cycles, the increase of $n_1(t)$ for node 1 is much slower when compared with the curve resulting from the transition kernel valid for acyclic graphs. Indeed, the memory effect comes into play only if and after the walker has followed the path $2 \to 3 \to 2$. The effect then acts in favor of increased probability in node 4. The memory tends to bring the curves corresponding to the two nodes 2 and 3 belonging to the cycle closer together. By the same mechanism, the relaxation of $n_2(t)$ and $n_3(t)$ to 0 is notably slower. The dashed curves resulting from the modeling with corrections fall in line with the Monte-Carlo simulation, which shows the effectiveness of the developed framework in this simple case.



(A) Monte-Carlo simulation based on $4 \cdot 10^4$ independent trajectories (shading) *vs* analytical predications for DAGs (solid) *vs* analytical predictions accounting for cycles (dashed). The shaded areas determine an interval of width equal to twice the standard deviation and centered around the mean.



(B) node 1 (C) nodes 2 and 3 (D) node 4



(E) The initial position of the walker is node 2. The rates are $\mu = 8$, $\eta = 1 = \lambda$.

FIGURE 3.9. Validation of the analytical framework accounting for second-order corrections.

3.6. Conclusion

A very common assumption in the study of dynamical processes on networks is to take only the direction of the edges and their weights into account. Accordingly, one often assumes that temporal events on the edges occur as a Poisson process. An important contribution of the field of temporal networks is to question this assumption and to propose more complex temporal models, including renewal processes with arbitrary event-time distributions. Yet in a

majority of works, one considers implicitly or explicitly instantaneous interactions. The main purpose of this chapter was to incorporate the finite duration of those interactions in stochastic models of temporal networks, and to estimate its impact on random walk processes.

We have derived the analytical expression of the propagator in the case of directed acyclic graphs. In general, graphs may have cycles and then, as we have showed with numerical (counter-)examples, the trajectories of the walker deviate from Markov or semi-Markov processes. We have discussed qualitatively and backed numerically the existence of regimes of the dynamical parameters where the quality of predictions for DAGs is somewhat preserved, even when there are cycles. Still, the need for analytically addressing the problem was established. With this in mind we have presented corrections to encompass in the modeling the presence of cycles in the underlying network.

This work is mostly theoretical but it has plenty of potential applications in real-life systems, potentially every time links activate for finite times. Yet in our view, the key message of chapters 2 and 3 that brings an answer to the overall objective of the first research question, is the now established and better understood importance of the three timescales to characterize diffusion on temporal networks with finite duration of edge activation: one for the moving entity and two for the edges. Future research directions include a more thorough investigation on when certain timescales can be neglected over other ones, hence leading to simplified mathematical treatment, and models including a fourth timescale, associated to the possible non-stationarity of the evolution of the graph, for instance due to circadian rhythms.

CHAPTER 4

Random walks on graphons

4.1. Introduction

Temporal properties of networks have led to intricate microscopic models of diffusion in chapters 2 and 3, which notably helped refine our understanding of network-theoretic methods in practical applications such as ranking. This chapter focuses on another oftentimes critical attribute of networks, that is not related to time fluctuations but lies at the core of our research theme centered on the continuum-limit : their potentially very large size.

As huge graphs become increasingly common in scientific research and real-world applications, a range of algorithms and computational problems indeed face scalability issues. An elegant workaround is to consider the continuum limit of graphs, defined when the number of nodes goes to infinity. This approach has for instance been used for network identification [**43**], spectral clustering [**53**], but also to study different classes of diffusion-based problems in networked systems [**91, 94**]. As we will see in that case, the dynamics on a large graph is rightfully approximated by the dynamics on its continuous limit.

In this chapter, we first revisit existing results for the continuum limit of the discrete heat equation and some nonlinear variants. This limit was the subject of a series of recent papers [**91, 93, 94**]. We then concentrate on the continuum limit of the node-centric case, hence considering the limit of the random-walk Laplacian operator. In general, for non-regular graphs, this operator differs from the combinatorial Laplacian, which is often preferred in algorithmic implementations such as spectral clustering because it properly accounts for the heterogeneous degree distributions observed in real-life networks. The random-walk operator in this work shouldn't be confused with a another operator[1] common in the machine learning community, sometimes also called random-walk Laplacian, which has an established convergence to the Laplace-Beltrami operator [**53, 11**].

Our approach is based on graph-limit theory [**84**], which does not rely on the assumption that the data generating the graphs is sampled from a distribution on a manifold [**46, 117**]. Our main contribution that answers part of the second research question, concerns the convergence of the space-discrete problem to a continuous problem in some appropriate setting. The problem on the continuum then falls in the realm of nonlocal evolution equations. More precisely, it is a

[1]We further discuss the different Laplacian integral operators in sections 7.1.1 and 7.1.2.

volume-constrained diffusion problem [**32**], and its analysis is voluntarily limited to some immediate consequences of spectral theory applied to our operators. Importantly, graph-limit theory defines a framework for the convergence of graphs of increasing size, but it may as well be seen as a possibly random graph-generating method. From that perspective, our work demonstrates that one may analyze the continuum model, to draw valid conclusions regarding the dynamics on the graphs generated by that model.

The chapter is organized as follows. Section 4.2 contains preliminary background material before it concentrates on graph-limit theory and introduces graphons as the limit objects of dense graph sequences. A random walk interpretation of the continuum limit of the heat equation on graphs opens section 4.3. We then focus on our main concern, the continuous-time node-centric walk. Well-posedness of the continuum problem is the subject of section 4.4. The main convergence results are presented in section 4.5. These results apply to dense graphs, and follow from a semigroup approach. We distinguish between different scenarios: first the discrete problem on graphs is sampled from the continuum version, and then the other way around. We then proceed with an analysis of the relaxation of the process using spectral theory in section 4.6. In section 4.7 we comment on the application of our method to the discrete-time random-walk, before coming to a conclusion in section 4.8.

4.2. Preliminaries

This section establishes the notation for concepts from operator and semigroup theory and identifies some known results used throughout the text, following [**115, 38**]. For the sake of self-consistency we then introduce the necessary basics of graph-limit theory and graphons [**84**].

4.2.1. Related to functional analysis

Let $\mathbb{L}(X, Y)$ denote the space of operators between the Banach spaces X and Y with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, which is in turn a Banach space with operator norm $||T|| = \sup_{||x||_X=1} ||Tx||_Y$. We let $\ker T$ and $\operatorname{ran} T$ denote respectively the kernel and range of T. When $X = Y$, we let $\mathbb{L}(X) := \mathbb{L}(X, Y)$. Further if $X = Y$ is a Hilbert space H, then $(\cdot,\cdot)_H$ denotes the scalar product, and we drop the subscript when there is no risk of confusion. The Hilbert space adjoint of T is denoted T^* and $\mathcal{I}$ is the identity operator.

Definition 4.1 (Resolvent set, spectrum)**.** Let $T \in \mathbb{L}(X)$. Then the complex number λ is in the resolvent set $\rho(T)$ of T if $\lambda\mathcal{I} - T$ is a bijection with a bounded inverse. The resolvent of T at λ is the operator $R_\lambda := (\lambda\mathcal{I} - T)^{-1}$. If $\lambda \notin \rho(T)$, then λ is in the spectrum $\sigma(T)$ of T.

The spectrum of an operator can be partitioned in various ways, that we will use in this chapter and in chapter 7. Since not all authors follow the same conventions, we will give the definition we use or explicitly refer the reader to appendix B.

Definition 4.2 (Eigenvalue, eigenvector and point spectrum)**.** Let $T \in \mathbb{L}(X)$. A nonzero vector x satisfying $Tx = \lambda x$ for some $\lambda \in \mathbb{C}$ is an eigenvector and λ is a corresponding eigenvalue. If λ is an eigenvalue, then $\lambda\mathcal{I} - T$ is not injective and so $\lambda \in \sigma(T)$. The point spectrum of T is the set of all eigenvalues of T.

From a spectral analysis perspective, the following theorem is an important result regarding those operators of $\mathbb{L}(X, Y)$ which map bounded sequences to sequences with a convergent subsequence, namely the compact operators. Next we give a definition that encompasses one of the central operators of this chapter.

THEOREM 4.3 (The Hilbert-Schmidt theorem [**115**, Theorem VI.16]). *Let A be a self-adjoint compact operator on a Hilbert space H. Then, there is a complete orthonormal basis of eigenvectors $\{\phi_n\}$ such that $A\phi_n = \lambda_n\phi_n$ and $\lambda_n \to 0$ as $n \to \infty$.*

Definition 4.4 (Hilbert-Schmidt operator)**.** Consider the Hilbert space $H = L^2(M, d\mu)$. Let $K(\cdot,\cdot) \in L^2(M \times M, d\mu \otimes d\mu)$. The compact integral operator defined by $f \to \int_0^1 K(\cdot, y)f(y)d\mu(y)$, for all $f \in H$ is called a Hilbert-Schmidt operator.

4.2.2. Related to semigroups

Definition 4.5 (One-parameter semigroup)**.** A family $(T(t))_{t\geq 0}$ of bounded linear operators on a Banach space X is called a (one-parameter) semigroup on X if the following functional equation is satisfied:

$$T(t + s) = T(t)T(s), \quad \forall t, s \geq 0, \tag{4.1a}$$

$$T(0) = \mathcal{I}. \tag{4.1b}$$

Definition 4.6 (Uniformly continuous semigroup)**.** A one-parameter semigroup $(T(t))_{t\geq 0}$ on a Banach space X is called uniformly (or norm) continuous if for $t \in \mathbb{R}^+$ the map $t \to T(t) \in \mathbb{L}(X)$ is continuous with respect to the uniform operator topology on $\mathbb{L}(X)$, that is, $\lim_{t\downarrow 0} T(t) - \mathcal{I} = 0$ uniformly.

Definition 4.7 (Strongly continuous semigroup)**.** A one-parameter semigroup $(T(t))_{t\geq 0}$ on a Banach space X is called strongly continuous, noted C_0-semigroup, if the orbit maps $t \to T(t)x$ are continuous from $\mathbb{R}^+$ into X for every $x \in X$, namely $\lim_{t\downarrow 0} T(t) - \mathcal{I} = 0$ strongly.

Definition 4.8 (Infinitesimal generator)**.** The infinitesimal generator A of a strongly continuous semigroup $(T(t))_{t\geq 0}$ is defined by $Ax := \lim_{t\downarrow 0} \frac{T(t)x-x}{t}$ for every x in its domain given by $\operatorname{dom} A := \{x \in X : t \to T(t)x \text{ is differentiable}\}$.

In the following pages and in chapter 7 we more than once rely on the fact that for a C_0-semigroup on X with generator A, the following are equivalent [**38**, Corollary II.5.5]: (a) the generator is bounded; (b) $\operatorname{dom} A = X$; (c) the domain is closed in X; (d) the semigroup is uniformly continuous.

4.2.3. Related to graphons

We now introduce graphons as the limit object of convergent sequences of graphs of increasing size. Then we show that graphons can actually be used to generate deterministic and random graphs of any size. Finally we consider them as kernels of associated integral operators. The reference works cited within this section should be referred to for a complete treatment of the subject.

4.2.3.1. *Graphons as limits of dense graph sequences*

Recent research [**85, 20, 19, 84**] provides a theoretical framework to study convergence of symmetric[2] dense graphs sequences. As a starting point, the so-called cut (or rectangular) metric allows to define the notion of Cauchy sequence of graphs of increasing number of nodes. Their limit object, called graphon, is a symmetric Lebesgue-measurable function[3] $W : [0,1]^2 \to [0,1]$. Therefore, the space of graphons is essentially the completion of the set of finite graphs seen as step functions (see Section 4.2.3.2), endowed with the so-called cut metric[4] which we introduce hereafter in its graphon version. Let us review the main concepts, as exposed also in [**84, 91, 44**]. The cut norm for graphons is given by

$$\|W\|_\square = \sup_{S,T \subset \mathfrak{M}[0,1]} \iint_{S \times T} W(x,y)dxdy,$$

where the supremum is over measurable subsets of $[0,1]$. The notation $\|W\|_p$ refers to the usual L^p norm of function defined on $[0,1]^2$, for $1 \leq p \leq \infty$. The following inequalities are immediate consequences of this definition, and of the inclusion theorem of L^p spaces on finite-measure spaces:

$$\|W\|_\square \leq \|W\|_1 \leq \|W\|_2 \leq \|W\|_\infty \leq 1. \tag{4.2}$$

Graphons are unique up to a composition with an invertible measure preserving mapping $\phi : [0,1] \to [0,1]$, which amounts to invariance of the limit graphon with respect to a relabeling of the nodes of the graphs. The graphons W^ϕ defined by $W^\phi(x,y) = W(\phi(x), \phi(y))$ and W are in the same equivalence class. The cut metric $\delta_\square$ between two graphons U and W is therefore defined by

$$\delta_\square(U,W) = \inf_{\phi \in \mathfrak{L}} \|U^\phi - W\|_\square \tag{4.3}$$

where $\mathfrak{L}$ is the space of the Lebesgue measurable bijections on the unit interval. The definition is similar for the $\delta_p(\cdot,\cdot)$ metrics based on the L^p norms, $1 \leq p \leq \infty$. Since two different graphons U, W can satisfy $\delta_\square(U,W) = 0$, strictly speaking $\delta_\square$

[2]This chapter is restricted to graphs with symmetric adjacency matrices, but when it comes to a random walks, edges are considered oriented, and *symmetric* effectively means *directed with reciprocal links*.

[3]Note that this choice of domain and range is somehow restrictive by comparison with other works where for instance $W : [0,1]^2 \to \mathbb{R}$. However, we will work with the standard definition because it achieves the desired degree of generality.

[4]There is a different though equivalent notion of convergence for dense graph sequences. It is called subgraph convergence, and is defined via associated sequences of induced subgraph densities [**20**].

is a metric only when we identify such graphons U and W [**20**]. Let us denote by $\mathbb{W}$ the space of graphons after this identification.

It holds that the metric space $(\mathbb{W}, \delta_\square)$ is compact, namely sequences of graphons possess at least one convergent subsequence in the cut metric. Unless explicitly mentioned, in this work we assume convergence of graphons in the L^2 norm topology. Hence by completeness, Cauchy sequences in $(\mathbb{W}, \|\cdot,\cdot\|_2)$ converge in the L^2 metric, and thus also in the δ_2 and $\delta_\square$ metrics, the limit being the same.

Many attributes of graphs already have natural counterparts in the realm of graphons[5]. A prominent example is the notion of degree, or strength, which plays a key role in this chapter. For a given graphon we let

$$k(x) := \int_0^1 W(x,y)dy \tag{4.4}$$

denote the (generalized) degree function. Since in this work graphons are bounded function $W : [0,1]^2 \to [0,1]$, the degree function is bounded, $0 \leq k(x) \leq 1$ for all $x \in [0,1]$.

Before we close this section, we should point out two important remarks. The first one relates to an alternate, though equivalent approach to graph-limit theory, whereas the second one shows that graph-limit theory is general enough to accommodate for random graph sequences.

Remark 4.9 (Convergence of graph sequences via homomorphism density)**.** Convergence of dense graph sequences (G_n) can be equivalently introduced in terms of convergence of an associated sequence $t(H, G_n)$, where H is any simple graph and

$$t(H, G_n) = \frac{\hom(H, G_n)}{|V_G|^{|V_H|}}. \tag{4.5}$$

Here, V_H is the vertex set for H (and similarly for G). For two unweighted graphs, $\hom(H, G)$ denotes the number of homomorphisms from H into G. This definition generalizes to weighted graphs [**85**]. Each term $t(H, G_n)$ in the sequence therefore represents the density of edge-preserving maps (or homomorphisms) $V_H \to V_{G_n}$, that is, the density of H as a subgraph of G_n, and the limit of the sequence reads

$$t(H, W) := \int_{[0,1]^{|V_H|}} \prod_{i \sim_H j} W(x_i, x_j)dx_1 \dots dx_{|V_H|}, \tag{4.6}$$

where $\sim_H$ denotes the symmetric adjacency relation in H.

Remark 4.10 (Convergence of random sequences)**.** From the beginning, graph-limit theory was developed to also provide a concept of limit for sequences of dense random graphs, namely those sequences where each graph is randomly generated based on the same set of rules. Rather than going into details, we refer the interested reader to [**19**] and to example 4.11, which we re-visit in chapter 7.

[5]There is however work left in defining graphon equivalents of graph related concepts, see page 177.

Example 4.11 (Uniform attachment graphon)**.** The graphon given by $W(x, y) = 1 - \max(x, y)$ is the limit of a randomly grown graph sequence, where each of the graphs follows from a uniform attachment scheme [**19**], see fig. 4.1. The construction starts with a single node, and nodes are then added one node at a time. After the i-th node was added, every pair of non-adjacent nodes at that point is connected with probability $1/i$. We give details to build intuition around the resulting expression of $W(x, y)$. Let $\{v_0, v_1, \ldots, v_{n-1}\}$ be the nodes after n steps, in order of appearance given by the indices. Consider two different nodes v_i, v_j and assume $0 \leq i < j \leq n$. Then after the n-th step, the probability that they are not connected is given by

$$P\{v_i \nsim v_j\} = \left(1 - \frac{1}{j+1}\right) \cdot \left(1 - \frac{1}{j+2}\right) \cdot \ldots \cdot \left(1 - \frac{1}{n}\right) = \frac{j}{n}, \tag{4.7}$$

and thus

$$P\{v_i \sim v_j\} = 1 - \frac{j}{n} = 1 - \frac{\max(i, j)}{n}. \tag{4.8}$$

Now choosing x, y such that $i = xn$ and $j = yn$, we have $P\{v_i \sim v_j\} = 1 - \max(x, y)$. That $W(x, y)$ given above is with probability 1 the limit of the sequence in the cut metric is established by [**19**, Theorem 3.1].

4.2.3.2. *Graphs as step graphons and graphs from graphon models*

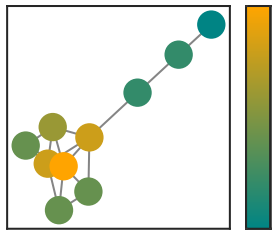
The connection between graphs and graphons is a two-way street. First, graphs can be mapped to the graphon space through a step function representation of their adjacency matrix. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a uniform partition of $[0, 1]$, where $P_i = \left[\frac{i-1}{n}, \frac{i}{n}\right)$ for $i = 1, \ldots, n-1$, and $P_n = \left[\frac{n-1}{n}, 1\right]$. Then let $\eta : \mathbb{M}_n \to \mathbb{W}$ be a mapping such that

$$\eta(G)(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \chi_{P_i}(x) \chi_{P_j}(y), \tag{4.9}$$

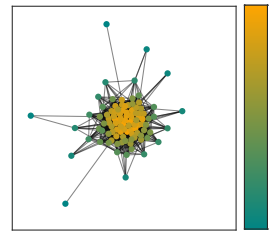
where χ_S is the indicator function of set S and A the adjacency matrix of graph G. The mapping thus defines the step (or empirical) graphon $\eta(G)$ associated to G. Similarly, η maps vectors $\boldsymbol{u} = (u_1, \ldots, u_n)$ to piecewise constant functions on $[0, 1]$, so that

$$\eta(\boldsymbol{u})(x) = \sum_{i=1}^{n} u_i \chi_{P_i}(x). \tag{4.10}$$

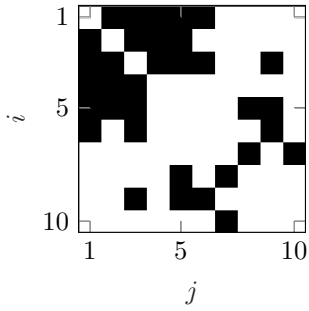
On the other hand, graphons can be considered as deterministic or (exchangeable [**37**]) random graph models. In this work we mainly treat the deterministic setting, and do not go beyond commenting on the random version as in remark 4.12. We leave the topic of random walks on the continuum limit of sparse graphs for future work, see page 95. Let $W \in \mathbb{W}$ be a graphon and let the integer n denote the desired number of nodes in the graph. Then W generates a dense graph by assigning weights to the edges, which can be done in two ways.



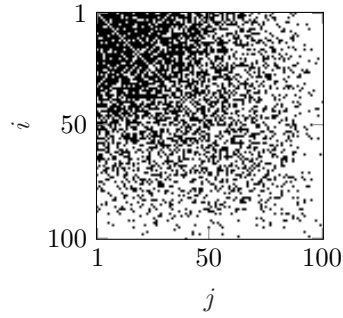
(A) Uniform attachment graph, 10 nodes. Color by degree, from 1 to 5.



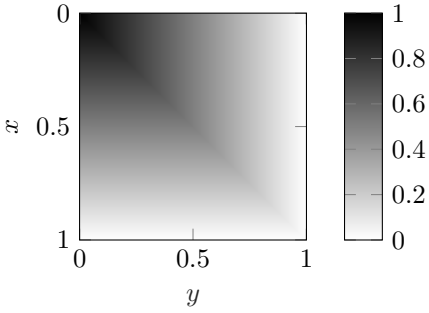
(B) Uniform attachment graph, 100 nodes. Color by degree, from 1 to 57.



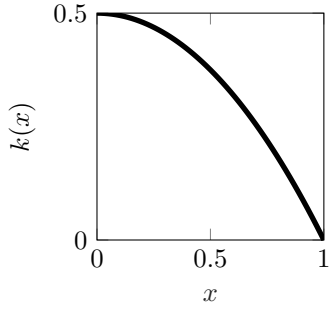
(C) Binary adjacency matrix of the graph in (A), black means 1



(D) Binary adjacency matrix of the graph in (B)



(E) Graphon $W(x, y) = 1 - \max(x, y)$, black means 1



(F) Generalized degree function $k(x) = \frac{1}{2}(1 - x^2)$

FIGURE 4.1. Growing uniform attachment random graph sequence.

Quotient graph: In a first approach, the weight A_{ij} of the edge between two nodes i and j equals the mean value of W on the corresponding cell of the partition of the unit square:

$$A_{ij} = n^2 \int_{P_i} \int_{P_j} W(x, y) dx dy, \quad i, j = 1, \dots, n. \tag{4.11}$$

This results in the so-called quotient graph $W/\mathcal{P}$. One can prove that there is almost everywhere point-wise convergence of the associated step graphon $\eta(W/\mathcal{P})$ to W ([**20**], lemma 3.2).

Sampled graph: A second approach to generate a graph from a given graphon $W \in \mathbb{W}$, is to define

$$A_{ij} = W\left(\frac{i}{n}, \frac{j}{n}\right), \quad i,j = 1,\dots,n, \tag{4.12}$$

in a way that is reminiscent of W-random graphs [**85**]. Let us denote $W_{[n]}$ the corresponding graph. Observe that $\eta\left(W_{[n]}\right) \to W$ point-wise at every point of continuity of W [**91**].

Remark 4.12 (Graphons as sparse graph models)**.** As we know, the bounded graphons $W \in L^\infty[0,1]$ considered in this work emerge in the continuum limit of dense graphs. However, graphs modeling real-life systems tend to be sparse, and graph-limit theory has addressed this fact in [**17, 18**], which requires to work with graphons in $L^p[0,1]^2$ $(p > 1)$. Dynamical systems with diffusion and L^p graphons were notably investigated by [**63, 94**], with obvious connections to the scope of this work. In this construct, L^p graphons generate graphs with a given number of nodes n and a density controlled by a sequence of positive numbers (ρ_n) satisfying $\rho_n \to 0$, $n\rho_n \to \infty$ as $n \to \infty$. For $\ell, m \in \{0,1,\dots,n\}$, $\ell \neq m$, an edge is created between two nodes v_ℓ, v_m, independently of the other pairs, with probability

$$P\{v_\ell \sim v_m\} = \min\left\{1, \rho_n W\left(\frac{\ell}{n}, \frac{m}{n}\right)\right\}. \tag{4.13}$$

Let $\Omega_n = \{0,1\}^n$. A random graph model on n nodes is actually a probability space $\left(\Omega_n, 2^{\frac{n(n+1)}{2}}, P\right)$ and $\Gamma_n(\omega)$ denotes a sampled graph, with $\omega \in \Omega_n$. The edges set of Γ_n is determined by the Bernoulli random variables $\mathcal{E}_{\ell m}(\omega) = \chi_{\ell \sim m}(\omega)$. The expected value of $\mathcal{E}_{\ell m}$ is equal to the probability given by eq. (4.13) and

$$d_\ell := E\{\deg v_\ell\} = E\left\{\sum_{m \neq \ell} \mathcal{E}_{\ell m}\right\} \tag{4.14}$$

denotes the expected degree of node v_ℓ. This quantity has an important role to play when it comes to convergence results related to dynamics in the continuum-limit of sparse random graph sequences, see remark 7.24.

4.2.3.3. *Graphons as kernels of operators*

Every graphon $W \in \mathbb{W}$ can be considered as a kernel, allowing to formally define an integral operator $\mathcal{W}$ on functional spaces on $[0,1]$ through

$$\mathcal{W}f(x) = \int_0^1 W(x,y) f(y) dy \tag{4.15}$$

The composition (product) of two such operators is given by

$$\mathcal{U}\mathcal{W}f(x) = \int_0^1 (U \circ W)(x,y) f(y) dy, \tag{4.16}$$

where $\circ$ is the operator product between the graphon kernels, defined by

$$(U \circ W)(x,y) = \int_0^1 U(x,z)W(z,y)dz, \quad \forall x,y \in [0,1]. \tag{4.17}$$

Observe that in general, $U \circ W$ is not a symmetric function. We denote $W^{\circ n}$ the operator product of the kernel, as opposed to the point-wise product $W^n(x,y) = (W(x,y))^n$, which is associated to the operator W^n. It follows from eq. (4.17) that

$$W^{\circ n}(x,y) = \int W(x,z_1)W(z_1,z_2)\dots W(z_{n-1},y)dz_1dz_2\dots dz_{n-1}. \tag{4.18}$$

4.3. Continuum limit of random walks

The aim of this section is twofold. First, we recall existing results on the continuum limit of the discrete heat equation, namely the edge-centric walk. Secondly, we formally derive the continuum limit of the node-centric random walk. As mentioned in section 4.2.3, our random-walk perspective implies that edges are directed, because they are associated with possible moves of the walker, with an origin and a destination. Therefore, the symmetry of the adjacency matrix indicates there exists a reciprocal to each edge and that both have the same weight.

4.3.1. Continuum limit of the discrete heat equation

The master equation for the edge-centric random walk is the graph version of the heat equation on a continuum. From chapter 2 we know that in terms of microscopic modeling, the trajectories are determined by a Poisson process on the edges with constant rate λ, such that the resting-time on the node follows an exponential distribution with rate μ_j proportional to the degree or strength of the node v_j: $\mu_j = \kappa\,\mathrm{str}(v_j)$. This allows for a constant rate of jump across all edges of the graph. In matrix form, we have that

$$\dot{\boldsymbol{u}} = \lambda \boldsymbol{u}(A - D). \tag{4.19}$$

Here we have written $\boldsymbol{u}$ for the row vector of residence probabilities, and $L = A - D$ is the combinatorial Laplacian of the graph. This model exhibits a homogeneous asymptotic state.

For simplicity assume an unweighted graph. If $\lambda > 0$, then $\lambda \deg(v_j) \to \infty$ if $\deg(v_j) \to \infty$, which will happen for some if not all nodes of a dense graph. The walker would perform jumps at an infinite rate, which is physically unrealistic. Normalizing the rate of the process according to the number of vertices avoids this situation. If λ becomes dependent on n, say $\lambda_n = \frac{1}{n}$, the resulting rate in each node remains bounded, $\lambda_n \deg(v_j) \leq 1$ for all j independently of the number of nodes. This explains the normalization that was required to justify the continuum limit of eq. (4.19) in [**91**].

4.3.2. Continuum limit of the node-centric walk

In contrast with the edge-centric model, no normalization of the rate parameter λ of the node-centric walk is needed when the number of nodes grows to infinity, since the rate does not depend on the structure of graph. The continuum limit therefore directly applies to the unmodified discrete model. For a formal derivation in this case, consider again the vector $\boldsymbol{u}(t)$ satisfying

$$\dot{u}_i(t) = \sum_{j=1}^{N} u_j \frac{1}{\mathrm{str}(v_j)} A_{ji} - u_i, \tag{4.20}$$

and the uniform partition $\mathcal{P} = \{P_1, \dots, P_n\}$ of $[0,1]$, with $u(\cdot, t) := \eta(\boldsymbol{u}(t))$ an associated step function on the interval. Let the degree function k_η of the step graphon $\eta(G)$ be defined by

$$k_\eta(x) = \int_0^1 \eta(G)(x,y)dy \text{ for } x \in [0,1]. \tag{4.21}$$

This degree function is actually the normalized strength (or also degree, when the graph is unweighted) of the nodes in G:

$$\mathrm{str}(v_i) = n \sum_{j=1}^{n} \int_{P_j} A_{ij} dy = n \sum_{j=1}^{n} \int_{P_j} \eta(G)(x,y)dy = n k_\eta(x) \tag{4.22}$$

for all $x \in P_i$. It follows that

$$\begin{aligned} \sum_{j=1}^{n} \frac{A_{ij}}{\mathrm{str}\,(v_j)} u_j(t) &= \sum_{j=1}^{n} n \int_{P_j} \frac{A_{ij}}{\mathrm{str}\,(v_j)} u(y,t)dy \\ &= \sum_{j=1}^{n} n \int_{P_j} \frac{A_{ij}}{n k_\eta(y)} u(y,t)dy = \int_0^1 \frac{\eta(G)(x,y)}{k_\eta(y)} u(y,t)dy, \end{aligned} \tag{4.23}$$

for every $x \in P_i$. Hence, the node-centric walk on the graph has an equivalent continuum domain formulation

$$\frac{\partial}{\partial t} u(x,t) = \int_0^1 \frac{\eta(G)(x,y)}{k_\eta(y)} u(x,t)dy - u(x,t). \tag{4.24}$$

The goal of this work is to prove convergence in the appropriate norm of the solution of eq. (4.24) to the solution of the evolution equation on the continuum

$$\frac{\partial}{\partial t} w(x,t) = \int_0^1 \frac{W(x,y)}{k(y)} w(y,t)dy - w(x,t), \tag{4.25}$$

where W is the limit graphon of $\eta(G)$ in the L^2 metric, and k its degree function.

Observe that similarly, a discrete equation of the form eq. (4.20) is obtained starting from eq. (4.25), when the graph is $W/\mathcal{P}$ or $W_{[n]}$.

4.4. Well-posedness

Before we prove the above-mentioned convergence, let us determine whether the up-to-now formal eq. (4.25), together with initial condition $w(x,0) = g(x)$, defines a well-posed initial-value problem (IVP).

4.4.1. Connectedness and integrability of W/k

Care will be taken first regarding how connectedness in the graph translates to graphons. The following definition follows from [**60, 84**].

Definition 4.13. A graphon W is connected if

$$\int_{S\times([0,1]\setminus S)} W(x,y)dxdy > 0$$

for every $S \in \mathfrak{M}[0,1]$ with lebesgue measure $\mu(S) \in (0,1)$.

Notice at this point that the connectedness (or lack thereof) of the graphs G_n of the sequence does not imply that of their limit [**60**]. Indeed, one could always make all the (otherwise disconnected) graphs of the sequence connected by a adding each time a node connected to all other nodes. This would leave the limit unchanged. And conversely, disconnecting one node in each connected graph of the sequence would not change the limit either. Also note that if a graphon W is (dis)connected, then so are all the kernels in the same equivalence class ([**60**], theorem 1.16). Let us now look into the implications of connectedness of the graphon on the positiveness of the degree function and hence on the definition of the random walk Laplacian operator.

Proposition 4.14. *Let W be a connected graphon, then $k > 0$.*

Proof. For every $x \in [0,1]$, define the neighborhood of x in W as

$$N_x = \{y \in [0,1] \,:\, W(x,y) > 0\}\,.$$

Since W is connected, $\mu(N_x) > 0$ for μ-almost every x ([**60**], lemma 5.1) and therefore,

$$k(x) = \int_{N_x} W(x,y)dy > 0 \quad \text{for} \quad \mu\text{-a.e. } x. \tag{4.26}$$

□

Remark 4.15. The connectedness of the graphon does not imply however that the degree function is bounded away from zero, namely that there exists a constant c such that $0 < c \leq k$ on $[0,1]$. Take for instance $W(x,y) = x^m y^m$ with $m > 0$, for which $k(x) = \frac{x^m}{m+1}$.

That k can be arbitrarily small influences the integrability of the kernel $K(x,y) := \frac{W(x,y)}{k(y)}$ in eq. (4.25), as the following remark explains.

Remark 4.16. The connectedness of the graphon does not imply that the integral kernel $K(x,y)$ is in $L^p[0,1]^2$ for $p>1$. Consider for example the binary graphon $W=\chi_{x^\alpha+y^\alpha\leq 1}$ for $\alpha>0$, where the subscript $x^\alpha+y^\alpha\leq 1$ is short for the set of couples $(x,y)\in[0,1]^2$ such that the inequality is satisfied. By a direct integration, $k(x)=(1-x^\alpha)^{\frac{1}{\alpha}}$. The integral

$$\|K\|_p^p=\iint_{x^\alpha+y^\alpha\leq 1}(1-y^\alpha)^{-\frac{p}{\alpha}}dxdy=\int_0^1(1-y^\alpha)^{\frac{1-p}{\alpha}}dy \tag{4.27}$$

is finite if and only if $p<1+\alpha$. Hence, K is in $L^2[0,1]$ only if $\alpha>1$, and in particular, the kernel K of the threshold graphon [**36**] obtained with $\alpha=1$ is not square-integrable. However, using Fubini-Tonelli it is easy to show that $\|K\|_1=1$ for all connected graphons, such that K is always in $L^1[0,1]^2$.

Based on the preceding remark, in order to ensure that the kernel is square integrable, we will make the following assumption:

Assumption 4.17. There exists a constant c such that $0<c\leq k$ on $[0,1]$.

If W is bounded away from zero, so is k, but graphons with localized support may still fulfill the assumption, as shown by Figure 4.3.

4.4.2. The IVP on $L^2[0,1]$

Resting on the operator in the right-hand side of eq. (4.25), we come to the following definition.

Definition 4.18. Let $W\in\mathbb{W}$ be a connected graphon that verifies assumption 4.17. Let the random-walk Laplacian operator $\mathcal{L}^{rw}:L^2[0,1]\to L^2[0,1]$ be defined by

$$\mathcal{L}^{rw}f(x)=\int_0^1\frac{W(x,y)}{k(y)}f(y)dy-f(x). \tag{4.28}$$

By definition, W is bounded on $[0,1]^2$ and following our hypothesis, $\frac{1}{k(x)}$ is bounded on $[0,1]$. Therefore, $K(x,y)=\frac{W(x,y)}{k(y)}$ is a Hilbert-Schmidt kernel and $\mathcal{K}:L^2[0,1]\to L^2[0,1]$ defined by

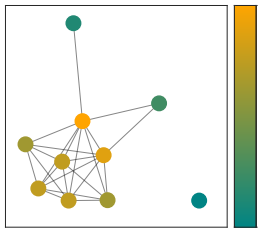
$$\mathcal{K}f(x)=\int_0^1\frac{W(x,y)}{k(y)}f(y)dy,\quad\forall x\in[0,1]\text{ and }f\in L^2[0,1] \tag{4.29}$$

is a compact Hilbert-Schmidt operator. Following definition 4.18, the continuum IVP has the form

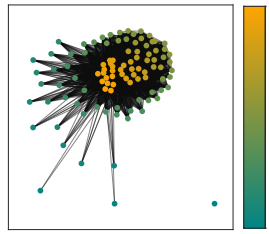
$$\frac{\partial}{\partial t}w(x,t)=\mathcal{L}^{rw}w(x,t) \tag{4.30a}$$

$$w(x,0)=g(x)\in L^2[0,1] \tag{4.30b}$$

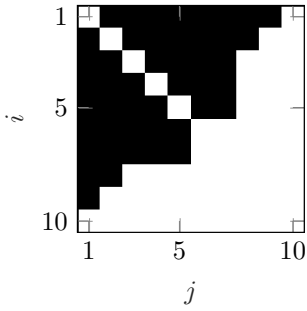
THEOREM 4.19. *Let $W\in\mathbb{W}$ be connected and satisfying assumption 4.17. Then there exists a unique classical solution to the initial-value problem eq.* (4.30).



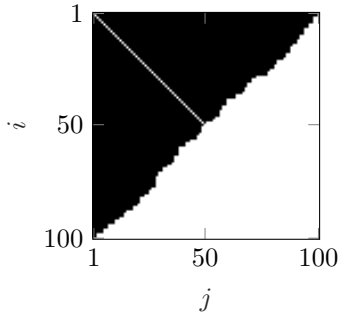
(A) Threshold random graph, 10 nodes. Color by degree, from 0 to 8.



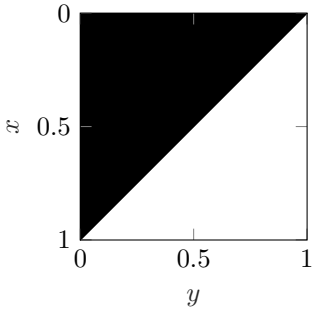
(B) Threshold random graph, 100 nodes. Color by degree, from 0 to 100.



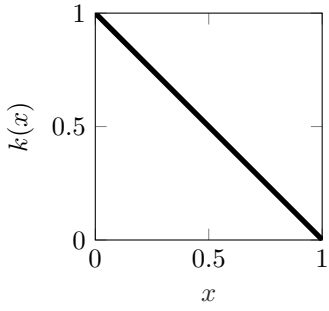
(C) Binary adjacency matrix of the graph in (A), black means 1



(D) Binary adjacency matrix of the graph in (B), black means 1



(E) Graphon $W(x, y) = \chi_{x+y\leq 1}$, black means 1

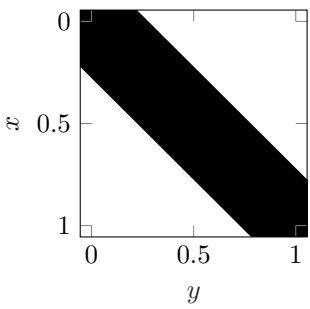


(F) Generalized degree function $k(x) = 1 - x$

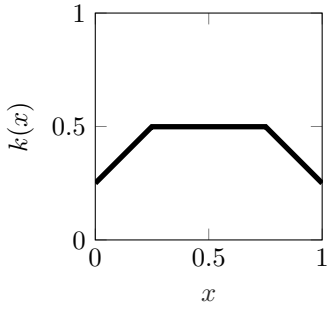
FIGURE 4.2. Threshold random graphs with uniformly distributed weights on the vertices, and their associated graphon. Two vertices v_i, v_j are adjacent if their weights w_i, w_j satisfy $w_i + w_j > t = 1$. The nodes are sorted by their degrees. Details on the construct in [**36**].

PROOF. The operator $\mathcal{K}$ is linear, and continuous hence bounded. It follows that $\mathcal{L}^{rw}$ is linear and bounded. Hence it is closed[6] and therefore $\mathcal{L}^{rw}$ is the

[6]The linear transformation $T : H \to H$ is closed if its graph $\Gamma(T)$, namely the set of pairs $\{\langle \varphi, T\varphi \rangle | \varphi \in \operatorname{dom} T\}$ is a closed subset of $H \times H$.



(A) $W(x,y) = \chi_{|x-y|\leq r}(x,y)$ with $r = \frac{1}{4}$, black means 1.



(B) Generalized degree function

FIGURE 4.3. The stripe graphon and its degree function. Observe that the support of W is localized on a subset of the square, but k is bounded away from zero.

infinitesimal generator of the (uniformly and thus) strongly continuous semigroup

$$\mathcal{T}^{rw}(t) = e^{\mathcal{L}^{rw}t} := \sum_{\ell=0}^{\infty} \frac{t^\ell \left(\mathcal{L}^{rw}\right)^\ell}{\ell!}. \tag{4.31}$$

Proposition 6.2 in [**38**] allows to conclude. □

Remark 4.20. (Classical solution) By definition of classical solution of the abstract Cauchy problem eq. (4.30), the orbit maps

$$t \in \mathbb{R}^+ \mapsto w(t,x) \in L^2[0,1]$$

are continuously differentiable. The forthcoming convergence results of section 4.5 are established in norm

$$\|w\|_{C([0,T],L^2[0,1])} = \sup_{t\in[0,T]} \|w(t,\cdot)\|_{L^2[0,1]} \tag{4.32}$$

defined for any positive real T.

Remark 4.21. The asymptotic steady state w_∞ of eq. (4.30) follows from $\mathcal{L}^{rw} w_\infty = 0$ and is proportional to the degree, $w_\infty \propto k$.

4.4.3. Positivity

The continuum IVP eq. (4.30) would loose physical relevance if its solution were to loose the possible positivity of the initial condition, $w(\cdot,0) \geq 0$. Before we proceed to a proof of positivity, let us first introduce a notation. For $g \in L^\infty[0,1]$, and $1 \leq p \leq \infty$, let $\mathcal{M}_g : L^p[0,1] \to L^p[0,1]$ denote the multiplication operator defined by

$$\mathcal{M}_g f(x) = g(x) f(x). \tag{4.33}$$

Proposition 4.22. *Let W be a connected graphon satisfying assumption 4.17 and let $w(\cdot,0) = g \geq 0$ be the initial condition of IVP eq. (4.30). Then the classical solution $w(x,t)$ of the IVP satisfies $w(\cdot,t) \geq 0$ for all $t \geq 0$.*

Proof. Let us define[7] $\mathcal{L}^{cons} := \mathcal{M}_{1/k}\mathcal{L}^{rw}\mathcal{M}_k$, yielding by a direct calculation

$$\mathcal{L}^{cons}f(x,t) = \frac{1}{k(x)}\int_0^1 W(x,y)f(y,t)dy - f(x,t), \quad \forall f \in L^2[0,1]. \tag{4.34}$$

Further let $u = \mathcal{M}_{1/k}w$ with $w(x,t)$ the solution of eq. (4.30) such that

$$\frac{\partial}{\partial t}u = \mathcal{M}_{1/k}\frac{\partial}{\partial t}w = \mathcal{M}_{1/k}\mathcal{L}^{rw}w = \mathcal{M}_{1/k}\mathcal{L}^{rw}\mathcal{M}_k u = \mathcal{L}^{cons}u.$$

Since $w(\cdot,t) \geq 0$ iff $u(\cdot,t) \geq 0$, it remains to prove the positivity of $u(\cdot,t)$. Choose $\epsilon > 0$ arbitrarily and let $v(x,t) = u(x,t) + \epsilon t$. Observe that $\mathcal{L}^{cons}v = \mathcal{L}^{cons}u$, and hence

$$\frac{\partial}{\partial t}v - \mathcal{L}^{cons}v = \frac{\partial}{\partial t}u + \epsilon - \mathcal{L}^{cons}u = \epsilon.$$

Let us show that $v(x,t)$ reaches its minimum at some $(a,0)$, $a \in [0,1]$. Assume by contradiction that there exists $(a,\tau) \in [0,1] \times (0,T)$ for some $T > 0$ such that $v(x,t) \geq v(a,\tau)$ for all x and t. It follows that

$$\begin{aligned}\mathcal{L}^{cons}v(a,\tau) &= \frac{1}{k(a)}\int_0^1 W(a,y)v(y,\tau)dy - v(a,\tau)\\ &\geq \frac{1}{k(a)}\int_0^1 W(a,y)v(a,\tau)dy - v(a,\tau) = 0.\end{aligned}$$

Hence, $\frac{\partial}{\partial t}v(a,\tau) = \mathcal{L}^{cons}v(a,\tau) + \epsilon = \epsilon > 0$ which is in contradiction with the assumption of v attaining its minimum in (a,τ) with $\tau > 0$, so $\tau = 0$. We have thus proved $v(x,t) \geq v(a,0)$, so that

$$u(x,t) + \epsilon t = v(x,t) \geq v(a,0) = u(a,0) = \frac{g(a)}{k(a)} \geq 0.$$

Since ϵ is arbitrary, this allows to conclude. □

4.4.4. The IVP with probability density functions

Let us observe that when $w(\cdot,t)$ in eq. (4.30) is a probability density function, it is natural to consider $w(\cdot,t) \in L^1[0,1]$, and one may define $\mathcal{L}^{rw}$ as a mapping $L^1[0,1] \to L^1[0,1]$. Indeed, as in eq. (4.29) let us still write $\mathcal{K}$ the integral part of $\mathcal{L}^{rw}$ defined on $L^1[0,1]$. By Fubini-Tonelli, the operator norm $\|\mathcal{K}\|_{1,1} := \|\mathcal{K}\|_{L^1[0,1]\to L^1[0,1]}$ satisfies

$$\|\mathcal{K}\|_{1,1} \leq \sup_{||f||_1=1}\int_{[0,1]^2}|K(x,y)f(y)|dxdy = \sup_{||f||_1=1}\int_{[0,1]}|f(y)|dy = 1. \tag{4.35}$$

This, combined with the fact that $\|\mathcal{K}f\|_1 = 1$ if $f = 1$, shows that $\|\mathcal{K}\|_{1,1} = 1$, and so even without assumption 4.17, $\mathcal{L}^{rw}$ is a bounded mapping of $L^1[0,1]$

[7] This Laplacian operator is know as consensus Laplacian, and is further discussed in section 7.1.1 on page 142.

into itself. Additionally, theorem 4.19 about the existence and unicity of a solution to the IVP has a similar formulation and proof in the present case. Further, the positivity established in section 4.4.3 also applies here, and this would still not require assumption 4.17. The only significant change in the proof of proposition 4.22 would be to use the auxiliary operator $\mathcal{L}^{rw}\mathcal{M}_k$ instead of $\mathcal{L}^{cons} = \mathcal{M}_{\frac{1}{k}}\mathcal{L}^{rw}\mathcal{M}_k$. When $w(\cdot,0) \geq 0$ we further have conservation of the L^1 norm:

$$\frac{\partial}{\partial t}\|w(\cdot,t)\|_1 = \frac{\partial}{\partial t}\int_0^1 |w(x,t)|dx = \frac{\partial}{\partial t}\int_0^1 w(x,t)dx = 0. \tag{4.36}$$

In the remainder of the chapter, for the sake of simplicity and in order to benefit from the Hilbert space framework at a later stage, we will however assume that W satisfies assumption 4.17. This allows to define $\mathcal{L}^{rw}$ as an operator acting on $L^2[0,1]$ and we do not use L^1 but rather the stronger L^2 norms also present in other works about dynamics on graphons [**91, 94**].

4.5. Convergence on dense graphs

This section is divided in three parts. In loose terms, the first two show that the solution of the discretized problem on $W/\mathcal{P}$ or $W_{[n]}$ converges to that of the continuum IVP, in the norm of eq. (4.32). The goal of the third part is to prove that the discrete problem can be approximated by its continuum version.

4.5.1. Convergence on the quotient graph $W/\mathcal{P}$

Let us start with two simple lemmas.

Lemma 4.23. *Let $\mathcal{A}_\eta : L^2[0,1] \to L^2[0,1]$ be an integral operator with bounded kernel A_η. Assume that A_η is a.e.-constant on every cell $P_i \times P_j$ of the uniform partition of $[0,1]^2$. Further let $f \in L^2[0,1]$ and define f_η by*

$$f_\eta(x) = n\sum_{i=1}^{n}\int_{P_i} f(y)dy\chi_{P_i}(x), \quad \forall x \in [0,1].$$

Then for all $\ell \in \mathbb{N}_0$, it holds that $\mathcal{A}_\eta^\ell f = \mathcal{A}_\eta^\ell f_\eta$.

Proof. The proof in the case $\ell = 1$ follows from a direct calculation, see for instance [**44**], lemma 3. The claim for $\ell > 1$ is a direct consequence since then

$$\mathcal{A}_\eta^\ell f = \mathcal{A}_\eta^{\ell-1}\mathcal{A}_\eta f = \mathcal{A}_\eta^{\ell-1}\mathcal{A}_\eta f_\eta = \mathcal{A}_\eta^\ell f_\eta.$$

□

Lemma 4.24. *Let $\mathcal{A}, \mathcal{B} : L^2[0,1] \to L^2[0,1]$ be two Hilbert-Schmidt integral operators with respective kernels A and B defined on the unit square, with $A \leq \beta$ for some constant $\beta > 0$. Then, for all $f \in L^2[0,1]$ and $\ell \in \mathbb{N}_0$*

$$\|\mathcal{A}^\ell f - \mathcal{B}^\ell f\|_2 \leq \beta^{\ell-1}\|A - B\|_2\|f\|_2 + \|(\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2.$$

PROOF. Using the Minkowski inequality, we have

$$\begin{aligned}\|\mathcal{A}^\ell f - \mathcal{B}^\ell f\|_2 &= \|\mathcal{A}^{\ell-1}\mathcal{A}f - \mathcal{B}^{\ell-1}\mathcal{B}f\|_2 \\ &= \|\mathcal{A}^{\ell-1}\mathcal{A}f - \mathcal{A}^{\ell-1}\mathcal{B}f + \mathcal{A}^{\ell-1}\mathcal{B}f - \mathcal{B}^{\ell-1}\mathcal{B}f\|_2 \\ &= \|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f) + (\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2 \\ &\leq \|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f)\|_2 + \|(\mathcal{A}^{\ell-1} - \mathcal{B}^{\ell-1})\mathcal{B}f\|_2. \end{aligned} \tag{4.37}$$

Now $\mathcal{A}^{\ell-1}$ is a Hilbert-Schmidt integral operator with kernel $A^{\circ(\ell-1)}$, as defined by eq. (4.18). For such operator, as a product of the Cauchy-Schwarz inequality it is known about the operator norm $\|\cdot\|$ that $\|\mathcal{A}^{\ell-1}\| \leq \|A^{\circ(\ell-1)}\|_2$, or equivalently

$$\|\mathcal{A}^{\ell-1}f\|_2 \leq \|A^{\circ(\ell-1)}\|_2 \|f\|_2. \tag{4.38}$$

The first term in the right hand side of eq. (4.37) therefore satisfies

$$\|\mathcal{A}^{\ell-1}(\mathcal{A}f - \mathcal{B}f)\|_2 \leq \|A^{\circ(\ell-1)}\|_2 \|\mathcal{A}f - \mathcal{B}f\|_2 \leq \beta^{\ell-1}\|\mathcal{A}f - \mathcal{B}f\|_2 \tag{4.39}$$

where we use $\|A^{\circ(\ell-1)}\|_2 \leq \|A\|_2^{\ell-1}$ ([**44**], lemma 6) and $A(x,y) \leq \beta$ for all $0 \leq x, y \leq 1$ to obtain the last inequality. Using again eq. (4.38) with $\ell = 2$, we also have $\|\mathcal{A}f - \mathcal{B}f\|_2 \leq \|A - B\|_2 \|f\|_2$ which, together with eq. (4.37) and eq. (4.39) leads to the conclusion. □

Now we are in a place to formulate the convergence results. The continuous formulation of the discrete problem associated to eq. (4.30) on the quotient graph reads[8]

$$\frac{\partial}{\partial t}u(x,t) = \mathcal{L}_\square^{rw} u(x,t) \tag{4.40a}$$

$$u(x,0) = g_\square(x) \tag{4.40b}$$

where the random walk Laplacian operator on $W/\mathcal{P}$ satisfies

$$\mathcal{L}_\square^{rw} f(x) = \int_0^1 \frac{\eta\,(W/\mathcal{P})\,(x,y)}{k_\square(y)} f(y)dy - f(x), \quad \forall f \in L^2[0,1], \tag{4.41}$$

and the initial condition is averaged on each cell of the partition as

$$g_\square(x) = n\sum_{i=1}^n \int_{P_i} g(y)dy\chi_{P_i}(x), \quad \forall x \in [0,1]. \tag{4.42}$$

Based on the following proposition, operator $\mathcal{L}_\square^{rw}$ is well-defined.

Proposition 4.25. *Let W be a connected graphon satisfying assumption 4.17, then the strength of every node of the quotient graph determined by the uniform partition $\mathcal{P} = \{P_1, \ldots, P_n\}$ of $[0,1]$ is positive.*

[8]The subscript $\square$ refers to fact that the averaging is performed on square cells of $[0,1]^2$. To lighten the notations, we do not refer explicitly to the number of nodes of the graph, so we write $u(x,t)$ instead of for instance, $u^{(n)}(x,t)$.

PROOF. The strength of the i-th node v_i, $i = 1, \dots, n$, is given by $\operatorname{str}(v_i) = nk_\square(x)$, for every $x \in P_i$. We have

$$k_\square(x) = \int_0^1 \sum_{j=1}^n A_{ij}\chi_{P_j}(y)dy = \frac{1}{n}\sum_{j=1}^n A_{ij}, \quad \forall x \in P_i, \tag{4.43}$$

where A_{ij} was defined by eq. (4.11). Hence,

$$\begin{aligned} k_\square(x) &= n\int_{P_i}\sum_{j=1}^n \int_{P_j} W(x', y')dy'dx' \\ &= n\int_{P_i}\int_0^1 W(x', y')dy'dx' \\ &= n\int_{P_i} k(x')dx' \geq c, \end{aligned}$$

where $c > 0$ is the constant from assumption 4.17. □

Remark 4.26. It follows that the finite-dimensional IVP eq. (4.40) on the quotient graph has a unique solution given by $e^{t\mathcal{L}_\square}g_\square$.

THEOREM 4.27 (Convergence with $W/\mathcal{P}$). *Let W be a connected graphon satisfying assumption 4.17, and let $w(x,t)$ be the solution of IVP eq.* (4.30). *Further let $u(x,t)$ be the solution of the associated discrete problem eq.* (4.40). *Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad \text{as} \quad n \to \infty.$$

PROOF. Using remark 4.26, by the Minkowski inequality we have

$$\begin{aligned} \|u(\cdot,t) - w(\cdot,t)\|_2 &= \left\|e^{t\mathcal{L}_\square}g_\square - e^{t\mathcal{L}}g\right\|_2 \\ &= \left\|\sum_{k=0}^\infty \frac{t^k}{k!}\mathcal{L}_\square^k g_\square - \sum_{k=0}^\infty \frac{t^k}{k!}\mathcal{L}^k g\right\|_2 \\ &\leq \|g_\square - g\|_2 + \sum_{k=1}^\infty \frac{t^k}{k!}\underbrace{\left\|\mathcal{L}_\square^k g_\square - \mathcal{L}^k g\right\|_2}_{(*)}. \end{aligned} \tag{4.44}$$

Let us write $\mathcal{L}^{rw} = \mathcal{K} - \mathcal{I}$ where $\mathcal{K}$ is the operator previously defined in eq. (4.29) and $\mathcal{I}$ is the identity operator. We have a similar decomposition $\mathcal{L}_\square^{rw} = \mathcal{K}_\square - \mathcal{I}$ for the Laplacian of the step graphon. For $k \geq 1$ and $0 \leq m \leq k$ let us write $\alpha_{mk} = (-1)^m\binom{k}{m}$, and consider $(*)$ in the right-hand side of eq. (4.44). Using Newton's binomial theorem we have

$$\begin{aligned} \left\|\mathcal{L}_\square^k g_\square - \mathcal{L}^k g\right\|_2 &= \left\|(\mathcal{K}_\square - \mathcal{I})^k g_\square - (\mathcal{K}^k - \mathcal{I})g\right\|_2 \\ &= \left\|\sum_{m=0}^k \alpha_{mk}\mathcal{K}_\square^{k-m}g_\square - \sum_{m=0}^k \mathcal{K}^{k-m}g\right\|_2 \\ &\leq \left\|\sum_{m=0}^{k-1} \alpha_{mk}\left(\mathcal{K}_\square^{k-m}g_\square - \mathcal{K}^{k-m}g\right)\right\|_2 + \|\alpha_{kk}(g_\square - g)\|_2 \end{aligned}$$

with $|\alpha_{mk}| = \binom{k}{m}$ and using lemma 4.23,

$$\leq \sum_{m=0}^{k-1} \binom{k}{m} \underbrace{\left\| \left(\mathcal{K}_\square^{k-m} - \mathcal{K}^{k-m}\right) g \right\|_2}_{(**)} + \|(g_\square - g)\|_2. \tag{4.45}$$

By assumption 4.17 and proposition 4.25 there exists some constant $c > 0$ such that $k_\square(y) \geq c$ for all $y \in [0,1]$. Further, $0 \leq W \leq 1$ on $[0,1]^2$, and so

$$\|K_\square\|_2 = \left\| \frac{\eta(W/\mathcal{P})}{k} \right\|_2 = \|\eta(W/\mathcal{P})\|_2 \left\| \frac{1}{k_\square} \right\|_2 \leq \frac{1}{c} =: \beta_\square, \tag{4.46}$$

where $K_\square$ denotes the integral kernel of $\mathcal{K}_\square$. For $\ell \in \mathbb{N}_0$, let us define $\mathcal{E}_\ell := \mathcal{K}_\square^\ell - \mathcal{K}^\ell$ and $E_\ell := K_\square^\ell - K^\ell$. Then, applying lemma 4.24 successively $\ell - 1$ times to $(**)$ in eq. (4.45) with $\ell = k - m$, we obtain

$$\begin{aligned}
\|\mathcal{E}_\ell g\|_2 &\leq \beta_\square^{\ell-1} \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_{\ell-1} \mathcal{K} g\|_2 \\
&\leq \left(\beta_\square^{\ell-1} + \beta_\square^{\ell-2}\right) \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_{\ell-2} \mathcal{K}^2 g\|_2 \\
&\vdots \\
&\leq \left(\sum_{j=1}^{\ell-1} \beta_\square^{\ell-j} \right) \|E_1\|_2 \|g\|_2 + \|\mathcal{E}_1 \mathcal{K}^{\ell-1} g\|_2,
\end{aligned}$$

and since E_ℓ is the kernel of $\mathcal{E}_\ell$ if $\ell = 1$,

$$\leq \left(\sum_{j=1}^{\ell-1} \beta_\square^{\ell-j} \right) \|E_1\|_2 \|g\|_2 + \|E_1\|_2 \|\mathcal{K}^{\ell-1} g\|_2$$

and with $\|\mathcal{K}^{\ell-1} g\|_2 \leq \|K^{\circ(\ell-1)}\|_2 \|g\|_2 \leq \|K\|_2^{\ell-1} \|g\|_2$,

$$\begin{aligned}
&\leq \left(\|K\|_2^{\ell-1} + \sum_{j=1}^{\ell-1} \beta_\square^{\ell-j} \right) \|E_1\|_2 \|g\|_2 \\
&\leq \ell \beta^{\ell-1} \|E_1\|_2 \|g\|_2,
\end{aligned} \tag{4.47}$$

where the last inequality stems from $\beta := \max\{\|K\|_2, \beta_\square\} \geq 1$. Combining eq. (4.45) and eq. (4.47) yields

$$\begin{aligned}
\left\| \mathcal{L}_\square^k g_\square - \mathcal{L}^k g \right\|_2 &\leq \sum_{m=0}^{k-1} \binom{k}{m} (k-m) \beta^{k-m-1} \|K_\square - K\|_2 \|g\|_2 + \|(g_\square - g)\|_2 \\
&\leq \beta^{k-1} \sum_{m=0}^{k-1} \binom{k}{m} (k-m) \|K_\square - K\|_2 \|g\|_2 + \|(g_\square - g)\|_2.
\end{aligned} \tag{4.48}$$

From eqs. (4.44) and (4.48) we obtain

$$\begin{aligned}\|u(\cdot,t)-w(\cdot,t)\|_2 &\le \|g_\square - g\|_2 + \sum_{k=1}^{\infty}\frac{t^k}{k!}\|(g_\square - g)\|_2 \\ &\quad + \|K_\square - K\|_2\|g\|_2\sum_{k=1}^{\infty}\frac{t^k}{k!}\beta^{k-1}\sum_{m=0}^{k-1}\binom{k}{m}(k-m) \\ &= \|g_\square - g\|_2 e^t + \|K_\square - K\|_2\|g\|_2\sum_{k=1}^{\infty}\frac{t^k}{k!}\beta^{k-1}\sum_{m=0}^{k-1}\binom{k}{m}(k-m)\end{aligned}$$

and with $\sum_{m=0}^{k-1}\binom{k}{m}(k-m) = k2^{k-1}$,

$$\begin{aligned}&= \|g_\square - g\|_2 e^t + \|K_\square - K\|_2\|g\|_2\sum_{k=1}^{\infty}\frac{t^k}{k!}k(2\beta)^{k-1} \\ &\le \|g_\square - g\|_2 e^t + \|K_\square - K\|_2\|g\|_2 t\sum_{k=1}^{\infty}\frac{(2\beta t)^{k-1}}{(k-1)!} \\ &\le \|g_\square - g\|_2 e^t + \underbrace{\|K_\square - K\|_2}_{(***)}\|g\|_2 t e^{2\beta t}. \qquad (4.49)\end{aligned}$$

By the Lebesgue differentiation theorem, $g_\square \to g$ pointwise for almost every $x \in [0,1]$ as $n \to \infty$, so that

$$\|g_\square - g\|_2 \xrightarrow[n\to 0]{} 0 \qquad (4.50)$$

by dominated convergence [**91**]. Let us consider $(***)$ in eq. (4.49):

$$\begin{aligned}\|K_\square - K\|_2^2 &= \int_{[0,1]^2}\left(\frac{\eta(W/\mathcal{P})(x,y)}{k_\square(x,y)} - \frac{W(x,y)}{k(y)}\right)^2 dxdy \\ &\le \operatorname*{ess\,sup}_{y\in[0,1]}\frac{1}{k_\square^2(y)k^2(y)}\int_{[0,1]^2}\big(\eta(W/\mathcal{P})(x,y)k(y) - W(x,y)k_\square(y)\big)^2 dxdy \\ &\le \beta^2\int_{[0,1]^2}\big(\eta(W/\mathcal{P})(x,y)(k(y)-k_\square(y))\big)^2 dxdy \\ &\quad + \beta^2\int_{[0,1]^2}\big((W(x,y)-\eta(W/\mathcal{P})(x,y))k_\square(y)\big)^2 dxdy\end{aligned}$$

and because $\|\eta(W/\mathcal{P})\|_2 \le 1$ and $\|k_\square\|_2 \le 1$,

$$\le \beta^2\left(\|k-k_\square\|_2^2 + \|W-\eta(W/\mathcal{P})\|_2^2\right). \qquad (4.51)$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}\|k-k_\square\|_2^2 &= \int_0^1\left(\int_0^1 (W(y,z)-\eta(W/\mathcal{P})(y,z))\,dz\right)^2 dy \\ &\le \int_0^1\int_0^1 (W(y,z)-\eta(W/\mathcal{P})(y,z))^2\,dzdy \\ &= \|W-\eta(W/\mathcal{P})\|_2^2,\end{aligned}$$

which together with eq. (4.51) yields

$$\|K_\square - K\|_2^2 \leq 2\beta^2 \|W - \eta(W/\mathcal{P})\|_2^2. \tag{4.52}$$

By the same argument leading to eq. (4.50), we have $\|W - \eta(W/\mathcal{P})\|_2 \to 0$ as $n \to \infty$ which with eq. (4.52) implies

$$\|K_\square - K\|_2 \xrightarrow[n\to 0]{} 0. \tag{4.53}$$

Combining eqs. (4.49), (4.50), and (4.53) allows to conclude. □

4.5.2. Convergence on the sampled graph $W_{[n]}$

The case of the discrete problem on $W_{[n]}$ can be handled similarly as the discrete problem on $W/\mathcal{P}$, and the convergence theorem follows mainly from the observation in section 4.2.3.2 that $W_{[n]} \to W$ at every point of continuity of W. The necessary convergence in L^2 will follow from the supplemental assumption that the graphon is almost everywhere continuous. The discrete problem (in its step function form) associated to eq. (4.30) on the sampled graph $W_{[n]}$ reads

$$\frac{\partial}{\partial t} u(x,t) = \mathcal{L}_{[n]}^{rw} u(x,t) \tag{4.54a}$$

$$u(x,0) = g_\square(x) \tag{4.54b}$$

where the random walk Laplacian operator on $W_{[n]}$ satisfies

$$\mathcal{L}_{[n]}^{rw} f(x) = \int_0^1 \frac{\eta\left(W_{[n]}\right)(x,y)}{k_{[n]}(y)} f(y)dy - f(x), \quad \forall f \in L^2[0,1], \tag{4.55}$$

and the initial condition is again averaged on each cell of the partition as in eq. (4.42). One needs to assume sufficiently large n to guarantee $k_{[n]}$ to be bounded away from 0 and so the Laplacian to be well-defined.

THEOREM 4.28 (Convergence with $W_{[n]}$). *Let W be a connected, almost everywhere continuous graphon satisfying assumption 4.17 and let $w(x,t)$ be the solution of IVP eq. (4.30). Further let $u(x,t)$ be the solution of the associated discrete problem eq. (4.54). Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad \text{as} \quad n \to \infty.$$

The proof is similar as for theorem 4.27.

Remark 4.29. The initial condition could have been sampled in a similar fashion as the graphon, to yield the step function

$$g_{[n]} = \sum_{i=1}^{n} g\left(\frac{i}{n}\right) \chi_{P_i}. \tag{4.56}$$

Almost everywhere continuity of g would ensure that $\|g - g_{[n]}\|_2 \to 0$ when $n \to \infty$, and would be part of the hypothesis of a convergence theorem. The proof of theorem 4.28 would only require minor changes, which are similar to those discussed next in the new context of section 4.5.3.

4.5.3. Convergence for a sequence of discrete problems

This time we consider a sequence of problems defined on graphs with increasing number of nodes. We assume the sequence of dense connected graphs, say (G_n), converges to a limit graphon W in the L^2 metric, in the sense that $\|\eta(G_n)-W\|_2 \to 0$ as $n \to \infty$. Let k_n denote the degree function of the empirical graphon $\eta(G_n)$. Consider the family of discrete problems under the mapping η

$$\frac{\partial}{\partial t}u(x,t) = \mathcal{L}_n^{rw}u(x,t) \tag{4.57a}$$

$$u(x,0) = g_n(x) \in L^2[0,1], \tag{4.57b}$$

where the random walk Laplacian operator $\mathcal{L}_n^{rw}$ satisfies

$$\mathcal{L}_n^{rw}f(x) = \int_0^1 \frac{\eta\left(G_n\right)(x,y)}{k_n(y)}f(y)dy - f(x), \quad \forall f \in L^2[0,1]. \tag{4.58}$$

Similarly as before, we write $\mathcal{L}_n^{rw} = \mathcal{K}_n - \mathcal{I}$.

THEOREM 4.30 (Convergence with (G_n)). *Let (G_n) be a sequence of connected graphs that converges to a connected graphon W satisfying assumption 4.17. Let $w(x,t)$ be the solution of the IVP eq. (4.30) associated to W with initial condition $w(\cdot,0) = g \in L^2[0,1]$. Further let $u(x,t)$ be the solution of the corresponding discrete problem eq. (4.57), and assume that $\|g_n - g\|_2 \to 0$ as $n \to \infty$. Then for all $t \in \mathbb{R}^+$ it holds that*

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \to 0 \quad \text{as} \quad n \to \infty.$$

PROOF. The proof follows the same steps as for theorem 4.27. However, using lemma 4.23 to obtain eq. (4.45) is now prohibited due to the initial condition of a discrete problem no longer resulting from an averaging of the continuous IVP. Consider a sufficiently large n such that the degree function of the empirical graphon satisfies $k_n \geq c$ for some constant $c > 0$. Not relying this time on lemma 4.23, we write

$$\begin{aligned}\|\mathcal{K}_n^{\ell-m}g_n - \mathcal{K}^{k-m}g\|_2 &= \|\mathcal{K}_n^{\ell-m}g_n - \mathcal{K}_n^{\ell-m}g + \mathcal{K}_n^{\ell-m}g - \mathcal{K}^{k-m}g\|_2 \\ &\leq \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 + \|(\mathcal{K}_n^{\ell-m} - \mathcal{K}^{\ell-m})g\|_2,\end{aligned}$$

with the first term in the right-hand side newly present. Following the same steps leading to eq. (4.45), we obtain

$$\begin{aligned}\left\|\mathcal{L}_n^k g_n - \mathcal{L}^k g\right\|_2 \leq &\sum_{m=0}^{k-1}\binom{k}{m}\|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 \\ &+ \sum_{m=0}^{k-1}\binom{k}{m}\left\|\left(\mathcal{K}_n^{k-m} - \mathcal{K}^{k-m}\right)g\right\|_2 + \|(g_n - g)\|_2,\end{aligned}$$

where again the first term right of the inequality is new. In fashion similar to the proof of theorem 4.27, with $\beta := \max\left\{\|K\|_2, \left\|\frac{1}{k_n}\right\|_\infty\right\}$ we have

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 + \|g_n - g\|_2 e^t + \|K_n - K\|_2 \|g\|_2 t e^{2\beta t}.$$

Using $\|\mathcal{K}^{k-m}\| \leq \beta^{k-m} \leq \beta^k$ and $\sum_{m=0}^{k-1} \binom{k}{m} \leq 2^k$, we have

$$\sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_{m=0}^{k-1} \binom{k}{m} \|\mathcal{K}_n^{\ell-m}(g_n - g)\|_2 \leq \sum_{k=1}^{\infty} \frac{t^k}{k!} 2^k \beta^k \|(g_n - g)\|_2 = \|(g_n - g)\|_2 e^{2\beta t},$$

leading to

$$\|u(\cdot,t) - w(\cdot,t)\|_2 \leq \|g_n - g\|_2 \left(e^t + e^{2\beta t}\right) + \|K_n - K\|_2 \|g\|_2 t e^{2\beta t}.$$

□

4.6. Relaxation

The evolution of a system towards its asymptotic state w_∞ starting from any initial condition is known as relaxation. The so-called relaxation time characterizes the rate of this evolution. In the continuum limit of the node-centric walk, it is determined by the spectral properties of $\mathcal{K}$, in a way reminiscent of random walks on finite graphs. For the node-centric continuous-time walk, we will show now that this rate can be exponential. Let us define a normalized adjacency operator, which is then used in the definition of a normalized Laplacian.

Definition 4.31 (Graphon normalized adjacency operator)**.** Under assumption 4.17, let the normalized adjacency operator be the integral operator $\mathcal{A}^{norm} : L^2[0,1] \to L^2[0,1]$ defined by

$$\mathcal{A}^{norm} f(x) = \int_0^1 \frac{W(x,y)}{\sqrt{k(x)}\sqrt{k(y)}} f(y) dy, \quad \forall f \in L^2[0,1], \tag{4.59}$$

namely $\mathcal{A}^{norm} = \mathcal{M}_{1/\sqrt{k}} \mathcal{K} \mathcal{M}_{\sqrt{k}}$.

Observe that under assumption 4.17 the kernel is square-integrable and symmetric. Hence $\mathcal{A}^{norm}$ is a compact, self-adjoint Hilbert-Schmidt integral operator and the Hilbert-Schmidt theorem on page 71 applies. Therefore, there exists an orthonormal basis of eigenfunctions $\{\phi_m\}$ with associated eigenvalues λ_m, so that operator $\mathcal{A}^{norm}$ has the canonical form

$$\mathcal{A}^{norm} = \sum_{m=1}^{\infty} \lambda_m \left(\phi_m, \cdot\right) \phi_m. \tag{4.60}$$

The operator $\mathcal{L}^{norm} := \mathcal{A}^{norm} - \mathcal{I}$ is the associated normalized (or sometimes also called symmetric) Laplacian. Note that for $\ell \in \mathbb{N}$, $\left(\mathcal{A}^{norm}\right)^\ell$ has eigenfunctions

ϕ_m and eigenvalues λ_m^ℓ, and that $(\mathcal{L}^{rw})^\ell = \mathcal{M}_{\sqrt{k}} (\mathcal{L}^{norm})^\ell \mathcal{M}_{1/\sqrt{k}}$. Combined with eq. (4.60) this yields the singular value decomposition

$$\begin{aligned} e^{\mathcal{L}^{rw} t} &= \mathcal{M}_{\sqrt{k}} \left(\sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \left(\sum_{m=1}^{\infty} \lambda_m \left(\phi_m, \cdot\right) \phi_m - \mathcal{I} \right)^\ell \right) \mathcal{M}_{1/\sqrt{k}} \\ &= \mathcal{M}_{\sqrt{k}} \left(\sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \left(\sum_{m=1}^{\infty} \left(\lambda_m - 1\right) \left(\phi_m, \cdot\right) \phi_m \right)^\ell \right) \mathcal{M}_{1/\sqrt{k}} \\ &= \mathcal{M}_{\sqrt{k}} \left(\sum_{m=1}^{\infty} \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \theta_m^\ell \left(\phi_m, \cdot\right) \phi_m \right) \mathcal{M}_{1/\sqrt{k}} \\ &= \sum_{m=1}^{\infty} e^{\theta_m t} \left(\frac{\phi_m}{\sqrt{k}}, \cdot \right) \sqrt{k} \phi_m \end{aligned} \tag{4.61}$$

with $\theta_m = \lambda_m - 1$ the eigenvalues of $\mathcal{L}^{norm}$. By letting $\psi_m = \frac{\phi_m}{\sqrt{k}}$ and $\zeta_m = \sqrt{k}\phi_m$, the solution of IVP eq. (4.30) reads

$$w(x,t) = \sum_{m=1}^{\infty} e^{\theta_m t} \left(\psi_m, g\right) \zeta_m(x). \tag{4.62}$$

The following proposition allows for a characterization of the rate of the relaxation towards w_∞.

Proposition 4.32. *Let W be a graphon satisfying assumption 4.17, then the eigenvalues θ_m of $\mathcal{L}^{norm}$ are non-positive reals, and the largest eigenvalue is zero. If moreover W is connected, then the eigenvalue zero has multiplicity one.*

Proof. That the eigenvalues are reals results from $\mathcal{L}^{norm}$ being a self-adjoint operator on $L^2[0,1]$. Let θ be an eigenvalue of $\mathcal{L}^{norm}$ associated to ϕ. Then θ is given by the Rayleigh quotient

$$\theta = \frac{(\theta\phi, \phi)}{(\phi, \phi)} = \frac{(\mathcal{L}^{norm}\phi, \phi)}{(\phi, \phi)}. \tag{4.63}$$

Consider the numerator in the right-hand side of eq. (4.63). For all $f \in L^2[0,1]$ we can write

$$\begin{aligned} (\mathcal{L}^{norm} f, f) &= \int_0^1 \int_0^1 \frac{W(x,y)}{\sqrt{k(x)}\sqrt{k(y)}} f(x) f(y) dx dy - \int_0^1 f^2(x) dx \\ &= \frac{1}{2} \Bigg(2 \int_0^1 \int_0^1 \frac{\sqrt{W(x,y)}}{\sqrt{k(x)}} \frac{\sqrt{W(x,y)}}{\sqrt{k(y)}} f(x) f(y) dx dy \\ &\quad - \int_0^1 \int_0^1 \frac{W(x,y)}{k(x)} f^2(x) dx dy - \int_0^1 \int_0^1 \frac{W(x,y)}{k(y)} f^2(y) dx dy \Bigg) \\ &= -\frac{1}{2} \int_0^1 \int_0^1 \left(\frac{\sqrt{W(x,y)}}{\sqrt{k(x)}} f(x) - \frac{\sqrt{W(x,y)}}{\sqrt{k(y)}} f(y) \right)^2 dx dy, \end{aligned} \tag{4.64}$$

which is non-positive. The claim that zero is an eigenvalue follows from the fact that $\mathcal{L}^{norm}\sqrt{k(x)} = 0$ on $[0,1]$. Finally, let us show that the nullspace of $\mathcal{L}^{norm}$ has dimension one if W is connected. By defining $g = \mathcal{M}_{1/\sqrt{k}} f$ on $[0,1]$, we have to show that if the right-hand side of eq. (4.64) is zero, namely

$$-\frac{1}{2}\iint_{[0,1]^2} W(x,y)\left(g(x)-g(y)\right)^2 dxdy = 0, \tag{4.65}$$

then g has to be a constant function on $[0,1]$. By contradiction, assume that there exists some non-constant function g that verifies eq. (4.65). For simplicity, consider the case that $g = c_1$ on some $S \in \mathfrak{M}[0,1]$ with $\mu(S) \in (0,1)$, and $g = c_2$ on $S^c := [0,1] \setminus S$, with $c_1, c_2 \in \mathbb{R}$, $c_1 \neq c_2$. The reasoning would be similar if g is a piecewise constant function on any other partition of $[0,1]$, and can be extended by density to any not piecewise constant g. Based on eq. (4.65), we can write

$$0 = \iint_{[0,1]^2} W(x,y)\left(g(x)-g(y)\right)^2 dxdy \geq \iint_{S\times S^c} W(x,y)\left(g(x)-g(y)\right)^2 dxdy, \tag{4.66}$$

and the integral in the right-hand side is zero. Since W is connected, we have $\iint_{S\times S^c} W(x,y)dxdy > 0$, and hence there exists a positive-measured subset $E \times F$ of $S \times S^c$ such that $W > 0$ on $E \times F$. Therefore, $g(x) - g(y) = 0$ for almost every $(x,y) \in E \times F$. But then, since $g = c_1$ on $E \subset S$, $g = c_1$ on $F \subset S^c$, a contradiction. □

Remark 4.33 (Spectral gap)**.** The last claim of proposition 4.32 means that the spectral gap of $\mathcal{L}^{norm}$, namely the positive difference between the largest and the second largest eigenvalue, is nonzero when the graphon is connected with degree function bounded away from zero. Observe that if k is not bounded away from zero, we may no longer write $\mathcal{A}^{norm} = \mathcal{M}_{1/\sqrt{k}}\mathcal{K}\mathcal{M}_{\sqrt{k}}$ because $1/\sqrt{k}$ is not bounded. This implies that the spectrum of $\mathcal{L}^{rw}$ can no longer be deduced directly from the spectrum of the compact self-adjoint operator $\mathcal{L}^{norm}$. However, the eigenvalues of $\mathcal{L}^{rw}$ may in some cases be computed directly, see example 4.34. If the graphon is not connected, one can analyze the dynamics on each connected component independently, as follows from the decomposition introduced in [**60**]. Therefore, it only remains open to fully characterize relaxation in the case of a connected graphon where assumption 4.17 is not satisfied, meaning $k(x)$ becomes arbitrarily small on positive measured subsets of $[0,1]$, and such that $\mathcal{K}$ is still well-defined. This situation could lead to a vanishing spectral gap for $\mathcal{L}^{rw}$ whilst in the discrete (or discretized) version, the spectral gap would be positive as a result of connectedness in finite graphs.

Example 4.34 (Eigenvalues of $\mathcal{L}^{rw}$ on a separable graphon)**.** Consider the separable[9] graphon $W(x,y) = xy$. The degree function is $k(x) = x/2$, in which case $\mathcal{L}^{rw} = \mathcal{K} - \mathcal{I}$ with $\mathcal{K}f(x) = 2x\int_0^1 f(y)dy$. Any eigenvalue $\lambda_{\mathcal{K}}$ of $\mathcal{K}$ satisfies

$$2x\int_0^1 \phi_{\mathcal{K}}(y)dy = \lambda_{\mathcal{K}}\phi_{\mathcal{K}}(x), \quad x \in [0,1], \tag{4.67}$$

[9]We say the graphon $W(x,y)$ is separable when it can be written as $W(x,y) = \zeta(x)\zeta(y)$ for some function ζ.

where $\phi_{\mathcal{K}}$ is an eigenfunction. It suffices to subtract one to the eigenvalues of $\mathcal{K}$ and to hold the same eigenfunctions to obtain the eigenpairs of $\mathcal{L}^{rw}$. From (4.67), one finds $\lambda_{\mathcal{K}} = 1$ with the one-dimensional eigenspace $\operatorname{span}\{x\}$, or $\lambda_{\mathcal{K}} = 0$ with the infinite-dimensional eigenspace $\{1\}^{\perp}$. Observe these spaces are not orthogonal, but their sum is $L^2[0,1]$.

Note that this particular example is not representative of all the technicalities of the spectral analysis of the generally non-self-adjoint operator $\mathcal{L}^{rw}$, when the degree function is not bounded away from zero. A more complete picture can be obtained by revisiting the above limited analysis with the tools of spectral theory introduced to study another graphon-based Laplacian operator in chapter 7.

4.7. Extension to the discrete-time walk

The analysis of the node-centric continuous-time walk carries over to the discrete-time version. Recall that the discrete-time walk is actually a Markov chain where the set V of vertices is the state-space and the transition probability from node v_i to v_j is encoded in the matrix

$$T_{ji} = \begin{cases} A_{ij}/\operatorname{str}(v_i) & \text{if } v_j \sim v_i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.68}$$

If we let $p_i(\ell)$ denote the probability that the walker is located on node i after ℓ steps, and $\boldsymbol{p}(\ell) := (p_1(\ell), \dots, p_n(\ell))$, then

$$\boldsymbol{p}(\ell+1) = \boldsymbol{p}(\ell)T, \tag{4.69}$$

where $T = D^{-1}A$, with D the diagonal matrix of the strengths of the nodes. It follows that for any $\ell \in \mathbb{N}$, $\mathbf{p}(\ell) = \mathbf{p}(0)T^{\ell}$. The corresponding IVP on the continuum reads

$$w(\cdot, \ell+1) = \mathcal{K}w(\cdot,\ell), \quad \ell \in \mathbb{N} \tag{4.70a}$$

$$w(\cdot, 0) = w_0 \in L^2[0,1], \tag{4.70b}$$

with solution given by $w(\cdot,\ell) = \mathcal{K}^{\ell} w_0$ for every $\ell \in \mathbb{N}$.

Following the same steps as in sections 4.3.2, 4.4, and 4.5, we obtain similar convergence results on the quotient graph $W/\mathcal{P}$, on the sampled graph $W_{[n]}$ and for a sequence of discrete problems. Analogously as for eq. (4.62), the spectral expansion of the solution of the discrete-time IVP eq. (4.70) is of the form

$$w(\cdot,\ell) = \sum_{m=1}^{\infty} \lambda_m^{\ell} \left(\psi_m, w_0\right) \zeta_m, \quad \ell \geq 0, \tag{4.71}$$

where the notation refers to (4.62).

4.8. Conclusion

The motivation to apply graph-limit theory to random walks as part of our third research question was twofold. On the one hand, random walks and Laplacians play a central role in the study of graphs, and a better understanding of their behavior on graphons has a clear mathematical interest, with theoretical and algorithmic objectives. On the other hand, as large networks become ever more common in numerous fields of research, a rigorous study of the continuum limit of the different types of random walks on graphs was needed.

This chapter was intended as a first step towards a systematic study of classes of random walks on discrete domains, relying on the adequate framework provided by graph-limit theory. We have first shown that the continuum-limit of the discrete heat equation [**91**] could be interpreted as the limit of a rescaled edge-centric continuous-time Poisson random walk. We have then studied the continuum limit of the remaining two fundamental classes of random walks on graphs, which complement the discrete heat equation: the continuous-time node-centric walk and its more basic discrete-time version. A final part of the document was devoted to spectral aspects of the introduced random walk Laplacian operator, thereby allowing for a characterization of the relaxation time of the process.

The world of random walks is a very broad one, and in this respect the scope of this initial work had to be narrowed. Hence, a promising research direction would consist in generalizing the semigroup approach developed here, or the one in [**91, 93, 94**], to the diverse classes of random walk processes omitted here, for instance walks on temporal or directed networks. A second line of research could focus on the case of sparse graphs. Sparsity is indeed known to be the norm rather than the exception in real-life networks. Such extension was already provided for the graph-limit version of the heat equation, using L^p graphons [**17, 93, 63**]. In these works, sparsity follows from randomness in the sampling procedure. This induces some technicalities due to the quotient structure of the kernel of the random walk Laplacian, and the fact that it is normalized by the degree. They are certainly not insurmountable but do require a careful treatment in deriving the continuum limit and obtaining convergence results. Another possible venue of investigation to deal with sparsity could follow from recent works on sparse exchangeable graphs generated via graphon processes or graphexes [**19, 16, 23**].

With that, we conclude the part of the thesis devoted to linear diffusion processes. In what comes next, we will build on the foundation of the first chapters to study stability and self-organization in nonlinear problems with diffusive coupling on graphs.

CHAPTER 5

Diffusion- and delay-driven instabilities on graphs

5.1. Introduction

Reaction-diffusion (RD) equations as introduced in chapter 1 arise out of the combination of two phenomena : local reactions and diffusion. When the domain is a graph, diffusion is commonly in the line of the process studied in chapter 2. In all models we will consider, the reaction term of the equation is nonlinear. Abandoning the linearity of the system has several dramatic implications. To start with, new classes of problems such as epidemics, and spreading processes in general can be modeled. Secondly, nonlinearity naturally comes with the cost of a generally more involved analysis. Thirdly, a distinctive feature of RD processes is that the number of agents (particles, reactants, or any other described quantity) is not necessarily preserved. We will also have to keep a check on the boundedness and positivity of solutions when we introduce delay systems. And finally, in terms of dynamical behavior, patterns and fronts are intimately linked to RD systems and even occupy the central position in this chapter. The overall objective, as part of our second research question, is to gain insight into the mechanism that governs the emergence of these patterns, and to determine whether time delay may increase the likelihood to observe instability seeded by diffusion in a given system.

In this introduction, we start with definitions of stability in section 5.1.1, which then allows to introduce and discuss diffusion-driven instabilities in section 5.1.2. Section 5.1.3 is a follow-up devoted to the limitations of the approach, and closes the introduction. Next in section 5.2 we will make a first step beyond the classical linear stability analysis, by using a bifurcation argument to describe the asymptotic steady state analytically. In section 5.3 we introduce delay systems, and present our findings related to the emergence of phenomena that could not arise without the delay. The final section summarizes the results and offers some conclusions.

5.1.1. Definitions of stability

Contrary to the modeling and analysis of the asymptotic steady state of the linear systems in the previous chapters, our dealings with nonlinear system will now revolve around stability. This requires some standard definitions we

will mention now, and also reuse in the following chapters in finite and infinite dimensions.

Let us consider a possibly nonlinear dynamical system on a Banach space X with norm $\|\cdot\|_X$ given by

$$\dot{z}(t) = \mathcal{F}(z(t)), \quad t \geq 0, \tag{5.1a}$$

$$z(0) = z_0, \tag{5.1b}$$

where $\mathcal{F} : \operatorname{dom}\mathcal{F} \subset X \to X$ is densely defined and $z_0 \in \operatorname{dom}\mathcal{F}$ is the initial condition. It is assumed that eq. (5.1) is well-posed and admits a unique solution $z(t) = S(t)z_0$, where $S(t)$ is the (possibly nonlinear[1]) C_0-semigroup of operators on X generated by $\mathcal{F}$. Let z_e be an equilibrium of eq. (5.1), $\mathcal{F}(z_e) = 0$. The equilibrium is not necessarily unique; there may be an infinite number of equilibrium points. So let E be a set of such equilibrium points and define

$$d(\cdot, E) := \inf \{\|\cdot - x\|_X \, : \, x \in E\}. \tag{5.2}$$

Definition 5.1 (Lyapunov stability)**.** The equilibrium set E is (Lyapunov) stable if for every $\epsilon > 0$, there exists $\delta > 0$ such that if $d(z_0, E) < \delta$, then $d(z_0, E) < \epsilon$ for all $t \geq 0$. The equilibrium is globally stable if $d(z_0, E) < \epsilon$ for all $t \geq 0$ and all $z_0 \in \operatorname{dom}\mathcal{F}$. Otherwise it is called (globally) unstable.

Definition 5.2 (Asymptotic stability)**.** The equilibrium set E is asymptotically stable if it is stable and there exists $\delta > 0$ such that if $d(z_0, E) < \delta$, then $\lim_{t\to\infty} d(z(t), E) = 0$. The equilibrium is globally asymptotically stable $\lim_{t\to\infty} d(z(t), E) = 0$ for all $z_0 \in \operatorname{dom}\mathcal{F}$.

Definition 5.3 (Exponential stability)**.** The equilibrium set E is exponentially stable if there exists $\delta, \alpha, \beta > 0$ such that if $d(z_0, E) < \delta$, then $d(z(t), E) \leq \alpha e^{-\beta t} d(z_0, E)$ for all $t \geq 0$. The equilibrium is globally exponentially stable if there exists $\alpha, \beta > 0$ such that $d(z(t), E) \leq \alpha e^{-\beta t} d(z_0, E)$ for all $t \geq 0$ and $z_0 \in \operatorname{dom}\mathcal{F}$.

Remark 5.4 (On the stability of points)**.** One may particularize definitions 5.2 to 5.3 to the stability of points by replacing in those definitions the expressions $d(\cdot, E) := \inf \{\|\cdot - x\|_X \, : \, x \in E\}$ by $\|\cdot - z_e\|_X$.

We will be interested in exponential stability and not asymptotic stability, because the former is the one needed to apply the principle of linearized stability in (in)finite dimensions.

5.1.2. Diffusion-driven instability

Did Turing ever expect that the 1952 paper [**131**] where he proposed his diffusion-driven mechanism for pattern formation in reaction-diffusion systems would have such dramatic and lasting impact? Patterns, namely the patchy motifs we come across in nature, for instance on the coating of some animals or on whorled leaves[2], are both common and compelling. The amount of mathematics

[1] See [**2**, Definition 2.2] for a definition of nonlinear strongly continuous semigroup.

[2] To use an example of Turing's.

behind his explanation of the destabilizing effect of diffusion on the otherwise stable point of the local reaction kinetics was relatively contained, which certainly helped his idea propagate. And he proposed a broad investigation, by considering RD on both continuous and discrete domains, which demonstrated stationary and oscillatory behaviors alike. Still, that the so-called Turing instabilities evolved into a field of research of their own is remarkable. This may have to do with the debate that surrounds the actual validity of the Turing approach, as we discuss in the next section. But let us first present the basics about Turing's diffusion-driven instability (DDI), focusing on systems on graphs [**102, 103**].

5.1.2.1. *On graphs*

We consider a two-species RD system evolving on a connected symmetric network with N nodes and no self-loops. The adjacency matrix is A and the degrees are denoted by k_j. The combinatorial Laplacian matrix L is given by $L_{ij} = A_{ij} - k_i\delta_{ij}$, and is the matrix operator for diffusion. The levels of the species in node i read respectively $u_i(t)$ and $v_i(t)$. The reactions in each node are modeled via two smooth nonlinear functions $f, g \in C^\infty(\mathbb{R})$, which are taken identical in all the nodes following an assumption of spatial homogeneity. The diffusion coefficients D_u and D_v characterize the mobility of the two species across the links. Recall that they can be interpreted as the exponential rate for the activation of the links in the edge-centric random-walk. The model is the discrete version of (1.27) and reads

$$\dot{u}_i(t) = f(u_i(t), v_i(t)) + D_u \sum_{j=1}^{N} L_{ij} u_j(t) \tag{5.3a}$$

$$\dot{v}_i(t) = g(u_i(t), v_i(t)) + D_v \sum_{j=1}^{N} L_{ij} v_j(t) \tag{5.3b}$$

for all $i = 1, \dots, N$ and $t > 0$, with initial conditions $u_i(0) = u_{i,0}$, $v_i(0) = v_{i,0}$ given in every node.

First one assumes there exists a uniform constant solution $(u_i(t), v_i(t)) = (u_e, v_e)$ to the local problem associated to (5.3),

$$\dot{u}_i(t) = f(u_i(t), v_i(t)), \tag{5.4a}$$

$$\dot{v}_i(t) = g(u_i(t), v_i(t)), \tag{5.4b}$$

with the same initial conditions as the original problem, such that

$$f(u_e, v_e) = 0 = g(u_e, v_e). \tag{5.5}$$

Assume further that the steady-state (u_e, v_e) is linearly stable, meaning the Jacobian matrix given by

$$\boldsymbol{J} = \begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix} \tag{5.6}$$

where the partial derivatives are evaluated in (u_e, v_e), has eigenvalues with negative real parts. This condition is expressed as

$$\operatorname{tr} \boldsymbol{J} < 0, \quad \det \boldsymbol{J} > 0. \tag{5.7}$$

Linearizing (5.3) around (u_e, v_e) and dropping the dependence on t in the notation, one obtains

$$\begin{pmatrix} \dot{u}_i \\ \dot{v}_i \end{pmatrix} = \boldsymbol{J} \begin{pmatrix} u_i \\ v_i \end{pmatrix} + \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix} \begin{pmatrix} (Lu)_i \\ (Lv)_i \end{pmatrix}, \tag{5.8}$$

for $i = 1, \ldots, N$. This system is diagonalized using an orthonormal basis $\{\varphi_0, \ldots, \varphi_{N-1}\}$ of eigenvectors of the Laplacian matrix, associated to eigenvalues ordered as

$$\lambda_{N-1} \leq \ldots \leq \lambda_1 < \lambda_0 = 0, \tag{5.9}$$

where the last strict inequality is due to the connectedness of the graph. Note that

$$\varphi_m \circ \mathbf{1} = \begin{cases} \sqrt{n} & \text{if } m = 0 \\ 0 & \text{if } m > 1 \end{cases}, \tag{5.10}$$

where all entries of $\mathbf{1} \in \mathbb{R}^N$ are equal to 1 and $\circ$ denotes the usual scalar product. The solution of (5.8) is written as[3]

$$u(t) = \sum_{m=1}^{N} a_m e^{\boldsymbol{\lambda}_m t} \varphi_m, \tag{5.11}$$

$$v(t) = \sum_{m=1}^{N} b_m e^{\boldsymbol{\lambda}_m t} \varphi_m. \tag{5.12}$$

where the a_m's and b_m's are real constants. It is a matter of inserting these expressions in eq. (5.8) to diagonalize the system and obtain the characteristic equation for each mode m,

$$\det\left(\boldsymbol{\lambda}_m I - \boldsymbol{J}_m\right) = 0, \tag{5.13}$$

where

$$\boldsymbol{J}_m = \boldsymbol{J} + \begin{pmatrix} D_u & 0 \\ 0 & D_v \end{pmatrix} \lambda_m, \quad m = 1, \ldots, N. \tag{5.14}$$

Let us observe that

$$\operatorname{tr} \boldsymbol{J}_m = \operatorname{tr} \boldsymbol{J}_m + (D_u + D_v)\lambda_m, \tag{5.15}$$

and so $\operatorname{tr} \boldsymbol{J}_m < \operatorname{tr} \boldsymbol{J} < 0$ for all $m > 1$. The solutions of (5.13) read

$$\boldsymbol{\lambda}_m^+ = \frac{1}{2}\left(\operatorname{tr} \boldsymbol{J}_m + \sqrt{(\operatorname{tr} \boldsymbol{J}_m)^2 - 4 \det \boldsymbol{J}_m}\right), \tag{5.16}$$

$$\boldsymbol{\lambda}_m^- = \frac{1}{2}\left(\operatorname{tr} \boldsymbol{J}_m - \sqrt{(\operatorname{tr} \boldsymbol{J}_m)^2 - 4 \det \boldsymbol{J}_m}\right). \tag{5.17}$$

The following definition formalizes what is meant by diffusion-driven instability.

Definition 5.5 (Turing instability)**.** An unstable homogeneous fixed point of a reaction-diffusion system is Turing unstable if it is an exponentially stable equilibrium of the associated local problem.

[3]Here bold font is used to distinguish between $\boldsymbol{\lambda}_m$ in these expression and the (topological) eigenvalues λ_m of the Laplacian matrix.

In case of problem (5.3), this means (5.7) applies, and there exists $m \in \{1, \dots, N-1\}$ such that the rightmost eigenvalue $\boldsymbol{\lambda}_m^+$ of $\boldsymbol{J}_m$ has positive real part,

$$\operatorname{tr} \boldsymbol{J}_m + \operatorname{Re} \sqrt{(\operatorname{tr} \boldsymbol{J}_m)^2 - 4 \det \boldsymbol{J}_m} > 0. \tag{5.18}$$

In terms of terminology, $\operatorname{Re} \boldsymbol{\lambda}_m^+$ is called the linear growth rate, and the map $\lambda_m \mapsto \operatorname{Re} \boldsymbol{\lambda}_m^+$ is often referred to as the dispersion relation. Condition (5.18) is the same as $\det \boldsymbol{J}_m < 0$ for some $m > 0$, or equivalently

$$\det \boldsymbol{J} + (f_u D_v + g_{v_\tau} D_u)\lambda_m + D_u D_v \lambda_m^2 < 0. \tag{5.19}$$

Together with the stability condition (5.7) of the local system, this implies $f_u D_v + g_v D_u > 0$. Since $\operatorname{tr} \boldsymbol{J} = f_u + g_v < 0$, from another use of (5.7), we have $f_u > 0$ and $g_v < 0$, or $f_u < 0$ and $g_v > 0$. Assuming u is the activator and v the inhibitor, $f_u > 0$, $g_v < 0$, one has $f_u D_v + g_v D_u > 0$ iff $D_v/D_u > -g_v/f_u > 1$. Summing up, we have the implication

$$\det \boldsymbol{J}_m < 0 \implies \frac{D_v}{D_u} > \frac{-g_v}{f_u} > 1 \tag{5.20}$$

which provides a necessary and practically restrictive condition for Turing instability.

Inequality (5.19) determines a range for the topological eigenvalues:

$$\lambda_m \in \left(\frac{-(f_u D_v + g_v D_u) - \sqrt{\rho}}{2 D_u D_v}, \frac{-(f_u D_v + g_v D_u) + \sqrt{\rho}}{2 D_u D_v} \right), \tag{5.21}$$

where $\rho = (f_u D_v + g_v D_u)^2 - 4 D_u D_v \det \boldsymbol{J}$. The critical ratio for the diffusion coefficients for which the above interval reduces to a single point, $\sigma_c := D_v/D_u|_{\rho=0}$, is given by

$$\sigma_{\mathrm{c}} = \frac{2 \det \boldsymbol{J} - f_u g_v + 2\sqrt{\det \boldsymbol{J}(\det \boldsymbol{J} - f_u g_v)}}{(f_u)^2}. \tag{5.22}$$

For a given diffusion coefficient D_u, the Laplacian eigenvalue corresponding to σ_{c} reads

$$\lambda_{\mathrm{c}} = \frac{(f_u - g_v)\sigma_{\mathrm{c}} - (\sigma_{\mathrm{c}} + 1)\sqrt{-f_v g_u \sigma_{\mathrm{c}}}}{D_u \sigma_{\mathrm{c}}(\sigma_{\mathrm{c}} - 1)}. \tag{5.23}$$

Remark 5.6. The value of D_u can be tuned so that the resulting new λ_{c} corresponds exactly to any nonzero eigenvalue λ_m of the Laplacian, for instance the eigenvalue closest to the original λ_{c}. It suffices to replace D_u by $\frac{\lambda_{\mathrm{c}}}{\lambda_m} D_u$, because then

$$\left(\frac{\lambda_{\mathrm{c}}}{\lambda_m} D_u \right) \lambda_m = \lambda_{\mathrm{c}} D_u.$$

In section 5.2 we will make the assumption that λ_{c} is an eigenvalue with algebraic multiplicity one for $m = m_{\mathrm{c}}$, and we write write $\varphi_{m_{\mathrm{c}}} =: \varphi_c$ the corresponding eigenvector of the Laplacian.

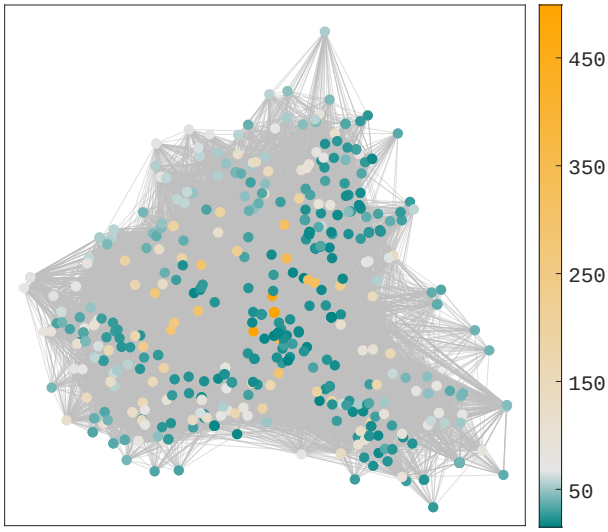


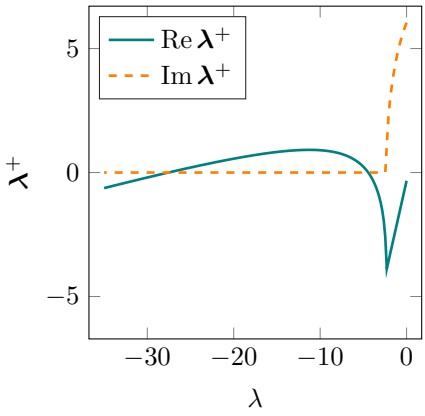
FIGURE 5.1. Network with 500 nodes corresponding to figs. 5.2 and 5.3. Node color by degree.

5.1.2.2. *Example*

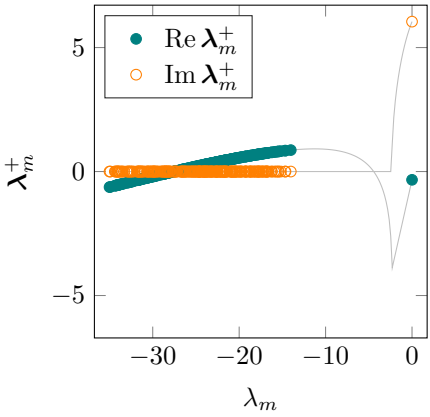
These considerations are best supported by an example. To this end, we selected Mimura-Murray kinetics, eq. (1.17), on top of a network with 500 nodes visible on fig. 5.3. The parameters are $a = 35$, $b = 16$, $c = 9$, $d = \frac{2}{5}$, while the diffusion coefficients are set to $D_u = 0.1$ and $D_v = 3$. The growth rates are plotted on fig. 5.2. Observe that the shape of the curves on figs. 5.2a and 5.2b is determined by the parameters of the Mimura-Murray model and by the diffusion coefficients, but not by the underlying network.

Next we integrated the system subject to a small initial perturbation around the equilibrium given by (1.18a). As predicted by the positive growth rates, the interaction of reactions and diffusion drove the system away from the homogeneous equilibrium, towards a non-uniform steady state written $u(\infty), v(\infty)$. If the term "pattern" is natural for RD systems on a continuum, the steady-states of discrete problems visibly do not come with an obvious shape to them. Finding the best way to represent the resulting pattern is a legitimate quest, and the answer is at least partly topological, as can be guessed from the fact that the Fourier modes correspond to the eigenvectors of the Laplacian matrix of the graph.

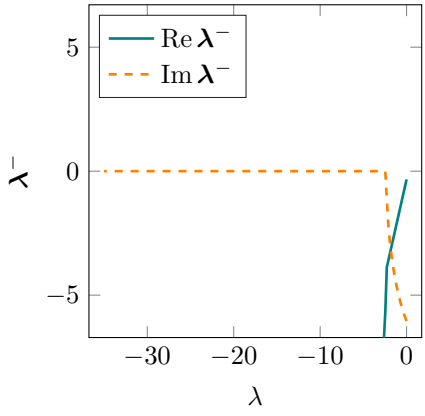
To make sense of that, first note that the graph of our example was generated with a spatial model where the nodes are connected depending on an embedding in the hyperbolic plane [**72**]. As is standard in this construct, each node is referred to by an angular and a radial coordinate (Θ, R) which are two random



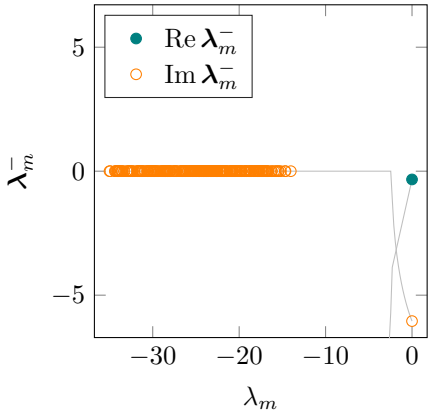
(A) Growth rates $\boldsymbol{\lambda}^+$ valid were the system defined on a continuum. The eigenvalues of the matrix Laplacian would be replaced by those of the differential Laplacian ∇^2.



(B) Actual discrete set of growth rates $\boldsymbol{\lambda}^+$ for the discrete problem. The plot does not display all eigenvalues since it is truncated left and below.



(C) Growth rates $\boldsymbol{\lambda}^-$ valid were the system defined on a continuum



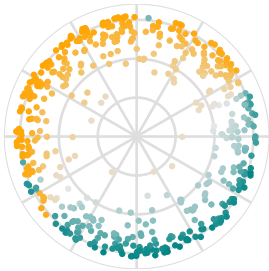
(D) Actual discrete set of growth rates $\boldsymbol{\lambda}^-$ for the discrete problem.

FIGURE 5.2. Linear growth rates for problem (5.3) with Mimura-Murray kinetics. The eigenvalues of the Laplacian of the graph on fig. 5.1 determine how the continuous curves of panel (A) are populated on panel (B). The stability of the system on the graph is determined by the teal-colored series $\operatorname{Re}\boldsymbol{\lambda}_m^+$ on panel (B), since the real part of all growth rates on panel (D) is negative.

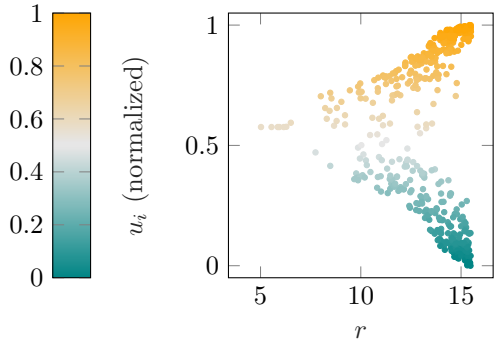
variables with respective PDF's

$$f_\Theta(\theta) = \frac{1}{2\pi}\chi_{[0,2\pi)}(\theta), \tag{5.24}$$

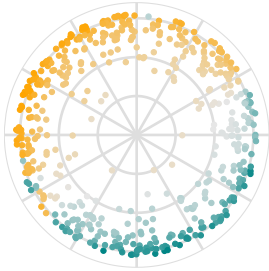
$$f_R(r) = \frac{\alpha\sinh(\alpha r)}{\cosh(\alpha r) - 1}\chi_{[0,\rho)}(r). \tag{5.25}$$



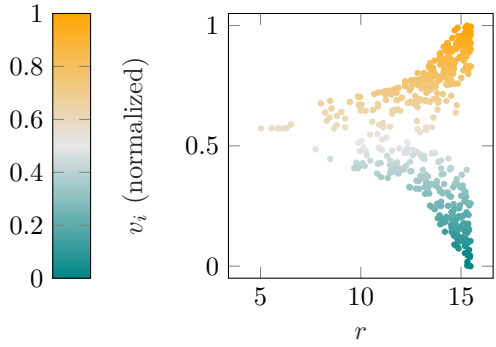
(A) Embedding in the hyperbolic plane, nodes colored according to $u(\infty)$



(B) Activation $u(\infty)$ depending on the radial coordinate of the hyperbolic embedding



(C) Embedding in the hyperbolic plane, nodes colored according to $v(\infty)$



(D) Activation $v(\infty)$ depending on the radial coordinate of the hyperbolic embedding

FIGURE 5.3. Finding order in a Turing pattern on graph, with Mimura-Murray kinetics on top of a spatial network embedded in the hyperbolic plane.

The parameters[4] have been set to $\alpha = 5.5$ and $\rho = 15.5$. Every pair of nodes (θ_1, r_1), (θ_2, r_2) forms an edge if their hyperbolic distance $d_H(\theta_1, r_1, \theta_2, r_2)$ given by

$$\cosh(d_H) = \cosh(r_1)\cosh(r_2) - \sinh(r_1)\sinh(r_2)\cos(\theta_1 - \theta_2) \tag{5.26}$$

is below the deterministic threshold, 1.4ρ in our case. Representing the asymptotic steady state after the nodes were embedded in the hyperbolic plane is

[4]The values of these parameters come from a choice of intermediate variables with physical meaning, which control the type of network.

thus a natural way to try and find structure in the discrete pattern. If for this particular example the procedure is clearly effective, as fig. 5.3 indicates, it suffices to change the parameters in the network generation procedure to see it loose its efficacy. Interpreting the shape of a Turing pattern on a network, let alone making a prediction, is nontrivial and is made even trickier since the pattern generally depends on the initial random perturbation.

5.1.3. Pro's and flaws of the Turing approach

The arguments in favor of Turing's mechanism can be summarized as mathematical simplicity coupled with a wide range of non-uniform steady states, that is, many different patterns which match the types of motifs encountered in real-life. But how meaningful and practical are delay-driven instabilities as a route to pattern formation? The literature on the matter is abundant, and we only aim at giving some general considerations to make sure the reach of this and the forthcoming chapters is adequately apprehended.

When it comes to meaningfulness, we know reaction-diffusion models suffer from limitations inherent to the modeling they arise from. Turing further had criticism about his own work, for instance that his closest examples to our graph-based problems, namely an isolated ring of cells, was "mathematically convenient though biologically unusual". And many years later, there appears to be consensus on the fact that real-world biological systems respond to genetic programming rather than diffusion of morphogenes obeying Fick's law of diffusion. But experimental evidence of Turing instabilities exists, even if rather in the context of a laboratory experiment with chemical reactors.

Overall, any experiment designed to validate the Turing mechanism will have to overcome a significant barrier, be it on continuous or discrete domains, arising from the required difference in diffusion coefficients between the two species, see (5.20). The therefore arguably small region of instability in the parameters space is a common criticism of DDI's. Several workarounds have been proposed, in order to relax the demanding conditions for Turing instability.

The research community has considered stochastic effects, models with cross diffusion, and in the particular case of discrete domains, it has introduced multiplex, directed or non-normal networks to the effect of less stringent instability conditions. In this chapter, we will contribute to this stream with delay models, section 5.3, and with RD on top of a temporal graph in chapter 6.

Moving forward, it is worth noting that recent developments keep flowing in, like in [**50**] where a new paradigm for mass conserving systems is introduced. The proposed analysis is based on so-called moving equilibria of elementary building blocks which partition the system in space. A remarkable feature is that the approach is not bound by the hypothesis that the nonlinear terms be stabilizing, and therefore shows that even if the Turing paradigm certainly offers a convenient framework for the mathematical analysis of non-uniform steady states in RD systems, it should not exclude other approaches.

5.2. Asymptotic steady state

Reaching beyond the conditions for instability, in this section we want to investigate the long-term behavior of the solutions of the reaction-diffusion problem (5.3), by looking for non-constant solutions to the stationary problem

$$0 = f(u_i, v_i) + D_u \sum_{j=1}^{N} L_{ij} u_j \tag{5.27a}$$

$$0 = g(u_i, v_i) + D_v \sum_{j=1}^{N} L_{ij} v_j \tag{5.27b}$$

for $i = 1, \dots, N$.

We first recall the definition of Fréchet derivative, not to be confused with the weaker Gâteaux derivative of definition 7.26 on page 157.

Definition 5.7 (Fréchet derivative)**.** Let X be a normed linear space. Then operator $F : X \to X$ is Frechet differentiable at $x \in X$ if there exists a bounded linear operator $DF(x)$ such that for all $h \in X$,

$$\lim_{h \to 0} \frac{\|F(x+h) - F(x) - DF(x)h\|}{\|h\|} = 0.$$

If F is Fréchet differentiable in every $x \in X$, then F is said to be Fréchet differentiable in X.

Note that in finite dimensions, the Fréchet derivative is represented in coordinates by the Jacobian matrix, and that existence and continuity of all partial derivatives is a sufficient condition for Fréchet differentiability.

5.2.1. Bifurcation from a simple eigenvalue

The approach relies on the use of a theorem in bifurcation theory ([**30**], Theorem 1.7), which was also used in [**75**] to study a growth-diffusion-chemotaxis model on a continuous, rectangular domain.

In the sequel D_u is considered a constant given value and we let $X = \mathbb{R}^N \times \mathbb{R}^N$, endowed with the usual scalar product, and use the notation $\boldsymbol{w} := (u, v) \in X$ where $u = (u_1, \dots, u_N)$ and $v = (v_1, \dots, v_N)$, and $w_e := (u_e, w_e) \in \mathbb{R}^2$. Associated with the stationary problem (5.27) we define the operator $F : \mathbb{R}^+ \times X \to X$ by

$$F_i(\sigma, \boldsymbol{w}) = \begin{cases} f(u_e + u_i, v_e + v_i) + D_u \sum_{j=1}^{N} L_{ij} u_j, & i = 1, \dots, N, \\ g(u_e + u_i, v_e + v_i) + \sigma D_u \sum_{j=1}^{N} L_{ij} v_j, & i = N+1, \dots, 2N. \end{cases} \tag{5.28}$$

The partial Fréchet derivative $D_{\boldsymbol{w}} F$, namely the Fréchet derivative of the map $\boldsymbol{w} \mapsto F(\sigma, w)$ for a fixed σ and evaluated in $\boldsymbol{w} = \mathbf{0}$, is the operator $D_{\boldsymbol{w}=\mathbf{0}} F$

given by

$$(D_{\boldsymbol{w}=\mathbf{0}}F)_i(\sigma,\boldsymbol{w}) = \begin{cases} f_u u_i + f_v v_i + D_u \sum_{j=1}^N L_{ij}u_j, & i=1,\dots,N, \\ g_u u_i + g_v v_i + \sigma D_u \sum_{j=1}^N L_{ij}v_j, & i=N+1,\dots,2N. \end{cases} \tag{5.29}$$

We can now state the main result of the section. The proof will largely consist in verifying the hypotheses of the bifurcation theorem cited above. We will establish the proof under the assumption of an activator-inhibitor system, $\operatorname{sign} J = \left(\begin{smallmatrix} + & - \\ + & - \end{smallmatrix}\right)$, but the result holds for cross activator-inhibitor systems such as the Brusselator where $\operatorname{sign} J = \left(\begin{smallmatrix} + & + \\ - & - \end{smallmatrix}\right)$. The only change is that the signs of b_c and B_c would be reversed. It only matters that the partial derivatives be nonzero.

THEOREM 5.8. *If λ_c is an eigenvalue of the Laplacian matrix with algebraic multiplicity one associated to the eigenvector φ_c, then there exists an interval $(-s,s)$ and a positive constant $b_c = -g_u/(g_v + \sigma_c D_u \lambda_c)$ such that nonconstant solutions of (5.27) near (u_e, v_e, σ_c) can be represented as*

$$\big(u(\alpha), v(\alpha)\big) = \big(u_e, v_e\big) + \alpha\left(\varphi_c, b_c\varphi_c\right) + \alpha^2\left(U(\alpha), V(\alpha)\right), \tag{5.30}$$

$$\sigma(\alpha) = \sigma_c + \alpha\beta(\alpha), \tag{5.31}$$

for all $\alpha \in (-s,s)$, where $\beta : (-s,s) \to \mathbb{R}$ and $U, V : (-s,s) \to \mathbb{R}^n$ are continuous functions, such that $\operatorname{ran} U \times \operatorname{ran} V$ *is a subspace of the orthogonal complement of* $\ker D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c,\cdot)$ *in X.*

PROOF. In order to apply the bifurcation theorem, let us first determine the dimension of $\ker D_{\boldsymbol{w}=\mathbf{0}}F$ for $\sigma = \sigma_c$. For this we have to determine the nonzero solutions of the linearization of (5.27). Let us consider the perturbations $\tilde{\boldsymbol{w}}_i = (\tilde{u}_i, \tilde{v}_i)$ of the equilibrium w_e and consider the linear system

$$0 = f_u\tilde{u}_i + f_v\tilde{v}_i + D_u\sum_{j=1}^N L_{ij}\tilde{u}_j \tag{5.32a}$$

$$0 = g_u\tilde{u}_i + g_v\tilde{v}_i + \sigma D_u\sum_{j=1}^N L_{ij}\tilde{v}_j \tag{5.32b}$$

for $i = 1,\dots N$. We expand the perturbation in the basis of the eigenvectors of the Laplacian with constant coefficients, $\tilde{u} = \sum_{m=1}^N a_m\varphi_m$, $\tilde{v} = \sum_{m=1}^N b_m\varphi_m$. Using the orthogonality of the chosen basis and the relation $L\varphi_m = \lambda_m\varphi_m$, we come to

$$\begin{pmatrix} f_u + D_u\lambda_m & f_v \\ g_u & g_v + \sigma D_u\lambda_m \end{pmatrix}\begin{pmatrix} a_m \\ b_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.33}$$

for $m = 1, \dots, N$. Notice for later use that this equation shows that $D_{\boldsymbol{w}=\mathbf{0}}F$ is also defined by

$$D_{\boldsymbol{w}=\mathbf{0}}F(\sigma, \sum_{m=1}^{N} a_m\varphi_m, \sum_{m=1}^{N} b_m\varphi_m) = \sum_{m=1}^{N} (a_m, b_m) \begin{pmatrix} f_u + D_u\lambda_m & g_u \\ f_v & g_v + \sigma D_u\lambda_m \end{pmatrix} \varphi_m. \tag{5.34}$$

The linear problem (5.32) thus yields the characteristic equation $\det \boldsymbol{J}_m = 0$. If we take $\sigma = \sigma_\mathrm{c}$ then $\det \boldsymbol{J}_m = 0$ only for the value of m such that $\lambda_m = \lambda_\mathrm{c}$, and it remains to solve

$$\begin{pmatrix} f_u + D_u\lambda_\mathrm{c} & f_v \\ g_u & g_v + \sigma_\mathrm{c} D_u\lambda_\mathrm{c} \end{pmatrix} \begin{pmatrix} a_c \\ b_c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{5.35}$$

for a_c and b_c. Let us normalize the solution assuming $a_c = 1$. It follows that $b_c = -g_u/\left(g_v + \sigma_\mathrm{c} D_u\lambda_\mathrm{c}\right)$, and one has

$$\ker D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot) = \operatorname{span}(\varphi_\mathrm{c}, b_c\varphi_\mathrm{c}), \tag{5.36}$$

which is one-dimensional. Applying the rank-nullity theorem,

$$\dim\left(\operatorname{ran} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot)\right) = \dim(X) - \dim\left(\ker D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot)\right) = 2N - 1, \tag{5.37}$$

and $\dim\left(X \setminus \operatorname{ran} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot)\right) = 1$.

Next we need to verify that the partial derivative with respect to σ of $D_{\boldsymbol{w}=\mathbf{0}}(\sigma, \tilde{\boldsymbol{w}})$ evaluated in $\sigma = \sigma_\mathrm{c}$ is not in the range of $D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot)$:

$$\frac{\partial}{\partial\sigma} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \varphi_\mathrm{c}, b_c\varphi_\mathrm{c}) \notin \operatorname{ran} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \cdot). \tag{5.38}$$

By the Fredholm alternative theorem, this is equivalent to

$$\frac{\partial}{\partial\sigma} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \varphi_\mathrm{c}, b_c\varphi_\mathrm{c}) \notin \ker D_{\boldsymbol{w}=\mathbf{0}}F^*(\sigma_\mathrm{c}, \cdot)^\perp, \tag{5.39}$$

where $D_{\boldsymbol{w}=\mathbf{0}}F^*$ denotes the adjoint operator of $D_{\boldsymbol{w}=\mathbf{0}}F$ defined for $\sigma = \sigma_\mathrm{c}$ by

$$D_{\boldsymbol{w}=\mathbf{0}}F^*(\sigma_\mathrm{c}, \sum_{m=1}^{N} a_m\varphi_m, \sum_{m=1}^{N} b_m\varphi_m) = \sum_{m=1}^{N} (a_m, b_m) \begin{pmatrix} f_u + D_u\lambda_m & f_v \\ g_u & g_v + \sigma_\mathrm{c} D_u\lambda_m \end{pmatrix} \varphi_m. \tag{5.40}$$

Using (5.34) the left-hand-side of (5.39) becomes

$$\begin{aligned} \frac{\partial}{\partial\sigma} D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c, \varphi_\mathrm{c}, b_c\varphi_\mathrm{c}) &= \frac{\partial}{\partial\sigma}(1, b_c) \begin{pmatrix} f_u + D_u\lambda_\mathrm{c} & g_u \\ f_v & g_v + \sigma D_u\lambda_\mathrm{c} \end{pmatrix} \varphi_\mathrm{c} \\ &= (1, b_c) \begin{pmatrix} 0 & 0 \\ 0 & D_u\lambda_\mathrm{c} \end{pmatrix} \varphi_\mathrm{c} \\ &= (0, b_c D_u\lambda_\mathrm{c}\varphi_\mathrm{c}) \,. \end{aligned} \tag{5.41}$$

To determine $\ker D_{\boldsymbol{w}=\mathbf{0}}F^*(\sigma_\mathrm{c},\cdot)$ in the right-hand side of (5.39), we again first have to solve

$$\det\begin{pmatrix} f_u + D_u\lambda_m & f_v \\ g_u & g_v + \sigma_\mathrm{c} D_u\lambda_m \end{pmatrix} = 0 \tag{5.42}$$

for λ_m. The solution is $\lambda_m = \lambda_\mathrm{c}$ and $\ker D_{\boldsymbol{w}=\mathbf{0}}F^*(\sigma_\mathrm{c},\cdot) = \operatorname{span}(A_\mathrm{c}\varphi_\mathrm{c}, B_\mathrm{c}\varphi_\mathrm{c})$ where the constants $A_\mathrm{c}, B_\mathrm{c}$ are given by

$$\begin{pmatrix} f_u + D_u\lambda_\mathrm{c} & g_u \\ f_v & g_v + \sigma_\mathrm{c} D_u\lambda_\mathrm{c} \end{pmatrix}\begin{pmatrix} A_\mathrm{c} \\ B_\mathrm{c} \end{pmatrix} = 0. \tag{5.43}$$

Normalizing by taking $\mathrm{A_c} = 1$, we get $\ker D_{\boldsymbol{w}=\mathbf{0}}F^*(\sigma_\mathrm{c},\cdot) = \operatorname{span}(\varphi_\mathrm{c}, B_\mathrm{c}\varphi_\mathrm{c})$ with $B_\mathrm{c} = -f_v/\left(g_v + \sigma_\mathrm{c} D_u\lambda_\mathrm{c}\right) < 0$. Combining with (5.41) allows to write condition (5.39) as

$$(0, b_c D_u\lambda_\mathrm{c}\varphi_\mathrm{c}) \notin \operatorname{span}(\varphi_\mathrm{c}, B_\mathrm{c}\varphi_\mathrm{c})^\perp. \tag{5.44}$$

We will use interchangeably the notations $\circ$ or $\langle\cdot,\cdot\rangle$ to denote the scalar product in X and in $\mathbb{R}^n$. Due to the fact that $b_c > 0$, $B_\mathrm{c} < 0$, one observes that

$$(0, b_c D_u\lambda_\mathrm{c}\varphi_\mathrm{c}) \circ (\varphi_\mathrm{c}, B_\mathrm{c}\varphi_\mathrm{c}) = b_c B_\mathrm{c} D_u\lambda_\mathrm{c}\varphi_\mathrm{c} \circ \varphi_\mathrm{c} = b_c B_\mathrm{c} D_u\lambda_\mathrm{c} > 0, \tag{5.45}$$

which shows that condition (5.44) holds true.

Let us now denote Z the orthogonal complement of $\ker D_{\boldsymbol{w}=\mathbf{0}}(\sigma_c,\cdot)$ in X. For some positive s, from the bifurcation theorem we have that nonconstant solutions of the stationary problem (5.27) near (σ_c, w_e) can be represented as

$$\big(u(\alpha), v(\alpha)\big) = w_e + \alpha\left(\varphi_c, b_c\varphi_c\right) + \alpha\left(\psi_u(\alpha), \psi_v(\alpha)\right) \tag{5.46}$$

and $\sigma(\alpha) = \sigma_\mathrm{c} + \varphi(\alpha)$ for all $\alpha \in (-s, s)$, where $\varphi : (-s, s) \to \mathbb{R}$ and $\psi_u, \psi_v : (-s, s) \to \mathbb{R}^n$ are once continuously differentiable functions such that $\operatorname{ran}\psi_u \times \operatorname{ran}\psi_v$ is in the orthogonal complement of $\ker D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c,\cdot)$ in X. Moreover from the theorem we have that $\varphi(0) = 0$ and $\psi_u(0) = 0 = \psi_v(0)$, allowing us to write $\varphi(\alpha) = \alpha\beta(\alpha)$ and $\psi_u(\alpha) = \alpha U(\alpha)$, $\psi_v(\alpha) = \alpha V(\alpha)$ with β, U, V continuous functions, so that we have

$$\beta(0) = \lim_{\alpha\to 0}\beta(\alpha) = \varphi'(0), \tag{5.47}$$

$$U(0) = \lim_{\alpha\to 0} U(\alpha) = \psi_u'(0), \tag{5.48}$$

$$V(0) = \lim_{\alpha\to 0} V(\alpha) = \psi_v'(0). \tag{5.49}$$

Now it remains to check that $\operatorname{ran} U \times \operatorname{ran} V$ is a subspace of the orthogonal complement of $\ker D_{\boldsymbol{w}=\mathbf{0}}F(\sigma_c,\cdot)$ in X. Since this is the case for $\operatorname{ran}\psi_u \times \operatorname{ran}\psi_v$ we can write

$$\left\langle\left(\psi_u(\alpha), \psi_v(\alpha)\right), (\varphi_\mathrm{c}, b_c\varphi_\mathrm{c})\right\rangle = 0 \tag{5.50}$$

for all $\alpha \in (-s, s)$, which implies that

$$\left\langle\left(U(\alpha), V(\alpha)\right), (\varphi_\mathrm{c}, b_c\varphi_\mathrm{c})\right\rangle = 0, \tag{5.51}$$

for all $0 < |\alpha| < s$. The functions U, V being continuous, we further have

$$\left\langle\left(U(0), V(0)\right), (\varphi_\mathrm{c}, b_c\varphi_\mathrm{c})\right\rangle = \lim_{\alpha\to 0}\left\langle\left(U(\alpha), V(\alpha)\right), (\varphi_\mathrm{c}, b_c\varphi_\mathrm{c})\right\rangle = 0, \tag{5.52}$$

and the proof is complete. □

5.2.2. Next-order characterization of the solution

The values of $U(\alpha)$, $V(\alpha)$ and $\beta(\alpha)$ in $\alpha = 0$ provided by the next proposition partially characterize the solutions of (5.27) in the neighborhood of the bifurcation. It is helpful to write these vectors in the basis of the φ_m's, such that

$$U(0) = \sum_{m=1}^{N} \mathsf{a}_m \varphi_m, \quad V(0) = \sum_{m=1}^{N} \mathsf{b}_m \varphi_m, \tag{5.53}$$

for some to-be-determined constants[5] $\mathsf{a}_m, \mathsf{b}_m$. In particular, for $m = m_c$ such that $\lambda_m = \lambda_c$ given by (5.23), we write $\mathsf{a}_c := \mathsf{a}_{m_c}$ and $\mathsf{b}_c := \mathsf{b}_{m_c}$. We further define the constants k_1, k_2 by

$$2k_1 = f_{uu} + (f_{uv} + f_{vu})b_c + f_{vv}b_c^2, \tag{5.54}$$

$$2k_2 = g_{uu} + (g_{uv} + g_{vu})b_c + g_{vv}b_c^2, \tag{5.55}$$

with the second partial derivatives of f and g evaluated in (u_e, v_e). We then let

$$\ell_m := \sum_{i=1}^{N} \varphi_{\mathrm{c},i}^2 \varphi_{m,i} \tag{5.56}$$

and in particular $\ell_c := \ell_{m_c}$.

Based on this premises, we have the following result.

Proposition 5.9. *If λ_c associated to the Laplacian eigenvector φ_c has algebraic multiplicity one, and if β given in (5.31) is one time differentiable in zero, then*

$$\begin{pmatrix} U(0) \\ V(0) \end{pmatrix} = -\sum_{m \neq m_\mathrm{c}} \ell_m \, \boldsymbol{J}_m^{-1}\big|_{\sigma=\sigma_\mathrm{c}} \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} \varphi_m - k_1 \ell_c \begin{pmatrix} f_u + D_u\lambda_\mathrm{c} & f_v \\ 1 & b_c \end{pmatrix}^{-1} \begin{pmatrix} \varphi_\mathrm{c} \\ 0 \end{pmatrix}, \tag{5.57}$$

and

$$\beta(0) = -\frac{k_2\ell_c + g_u\mathsf{a}_c + (g_v + \sigma_\mathrm{c} D_u \lambda_\mathrm{c})\,\mathsf{b}_c}{D_u b_c \lambda_c} \tag{5.58}$$

where $\mathsf{a}_c = \left(f_v(b_c^2+1)\right)^{-1} b_c\ell_c k_1 = -b_c\mathsf{b}_c$.

Proof. To begin with, since $(U(0), V(0))$ is in the orthogonal complement of $\ker D_{\boldsymbol{w}=\boldsymbol{0}} F(\sigma_c, \cdot)$ in X, we have:

$$\begin{aligned} 0 &= \langle (U(0), V(0)), (\varphi_\mathrm{c}, b_c\varphi_\mathrm{c}) \rangle \\ &= \sum_{m=1}^{N} \mathsf{a}_m \langle \varphi_m, \varphi_\mathrm{c} \rangle + \sum_{m=1}^{N} \mathsf{b}_m b_c \langle \varphi_m, \varphi_\mathrm{c} \rangle \\ &= \mathsf{a}_c + b_c \mathsf{b}_\mathrm{c} \end{aligned} \tag{5.59}$$

[5]Note the sans serif font to distinguish between a_m and b_m from a_m and b_m appearing in the expressions of $\tilde{u}(t)$ and $\tilde{v}(t)$ introduced on page 107.

Since $(u(\alpha), v(\alpha))$ is a solution of (5.27) with $D_v = \sigma(\alpha) D_u$, we also have

$$0 = g(u_i(\alpha), v_i(\alpha)) + \sigma(\alpha) D_u \sum_{j=1}^{N} L_{ij} v_j(\alpha), \tag{5.60}$$

for $i = 1, \dots N$. Differentiating and omitting to write the dependency on α we find

$$0 = g_u(u_i, v_i) u_i' + g_v(u_i, v_i) v_i' + \sigma' D_u \sum_{j=1}^{N} L_{ij} v_j + \sigma D_u \sum_{j=1}^{N} L_{ij} v_j'. \tag{5.61}$$

The second derivative reads

$$\begin{aligned} 0 = {} & g_{uu}(u_i, v_i) \left(u_i'\right)^2 + g_{uv}(u_i, v_i) u_i' v_i' + g_u(u_i, v_i) u_i'' \\ & + g_{vu}(u_i, v_i) u_i' v_i' + g_{vv}(u_i, v_i) \left(v_i'\right)^2 + g_v(u_i, v_i) v_i'' \\ & + \sigma'' D_u \sum_{j=1}^{N} L_{ij} v_j + 2\sigma' D_u \sum_{j=1}^{N} L_{ij} v_j' + \sigma D_u \sum_{j=1}^{N} L_{ij} v_j'', \end{aligned} \tag{5.62}$$

for $i = 1, \dots, N$. Notice that $\sigma' = \beta + \alpha\beta'$, $\sigma'' = 2\beta' + \alpha\beta''$, and that

$$(u', v') = (\varphi_c, b_c \varphi_c) + 2\alpha \left(U, V\right) + \alpha^2 \left(U', V'\right), \tag{5.63}$$

$$(u'', v'') = 2 \left(U, V\right) + 4\alpha (U', V') + \alpha^2 (U'', V''). \tag{5.64}$$

Inserting $\alpha = 0$ into the previous four equations, we get

$$\sigma'(0) = \beta(0) \tag{5.65}$$

$$\sigma''(0) = 2\beta'(0) \tag{5.66}$$

$$(u'(0), v'(0)) = (\varphi_c, b_c \varphi_c) \tag{5.67}$$

$$(u''(0), v''(0)) = 2(U(0), V(0)). \tag{5.68}$$

Hence for $\alpha = 0$, (5.62) becomes

$$\begin{aligned} 0 = {} & g_{uu} \varphi_{c,i}^2 + (g_{uv} + g_{vu}) b_c \varphi_{c,i}^2 + g_{vv} b_c^2 \varphi_{c,i}^2 + 2 g_u U_i(0) + 2 g_v V_i(0) \\ & + 2\beta'(0) D_u \sum_{j=1}^{N} L_{ij} v_e + 2\beta(0) D_u b_c \sum_{j=1}^{N} L_{ij} \varphi_{c,j} + 2\sigma_c D_u \sum_{j=1}^{N} L_{ij} V_j(0) \end{aligned} \tag{5.69}$$

for $i = 1, \dots, N$, where we notice that $\sum_{j=1}^{N} L_{ij} v_e = 0$ and $\sum_{j=1}^{N} L_{ij} \varphi_{c,j} = \lambda_c \varphi_{c,i}$. If we write the elements of $\mathbb{R}^n$ as column vectors, and use the constant k_2 defined by (5.55), then (5.69) reads

$$0 = k_2 \operatorname{diag}(\varphi_c) \varphi_c + g_u U(0) + g_v V(0) + \beta(0) D_u b_c \lambda_c \varphi_c + \sigma_c D_u L V(0), \tag{5.70}$$

where $\operatorname{diag}(x)$ is a square diagonal matrix with the elements of x on the main diagonal. We are to project equation (5.70) on the chosen eigenbasis of the Laplacian. First for all $m = 1, \dots, N$, we have

$$\langle L V(0), \varphi_m \rangle = \sum_{j=1}^{N} \mathsf{b}_j \langle L \varphi_j, \varphi_m \rangle = \mathsf{b}_m \lambda_m. \tag{5.71}$$

Taking the inner product of the right-hand side of (5.70) with φ_m yields

$$0 = k_2\ell_m + g_u\mathsf{a}_m + (g_v + \sigma_c D_u\lambda_m)\mathsf{b}_m, \qquad \text{for } m \neq m_c, \tag{5.72}$$

$$0 = k_2\ell_c + g_u\mathsf{a}_c + (g_v + \sigma_c D_u\lambda_c)\mathsf{b}_c + D_u b_c\lambda_c\beta(0), \qquad \text{for } m = m_c. \tag{5.73}$$

Let us repeat the same procedure that led to (5.72) and (5.73), this time with the first N equations of the stationary problem (5.27), producing N new equations with unknowns $\mathsf{a}_m, \mathsf{b}_m$. For all $\alpha \in (-s, s)$ we have:

$$0 = f(u_i(\alpha), v_i(\alpha)) + D_u \sum_{j=1}^{N} L_{ij} u_j(\alpha) \tag{5.74}$$

for $i = 1, \dots, N$. Differentiating twice with respect to α and taking $\alpha = 0$ yields

$$0 = \left(f_{uu} + (f_{uv} + f_{vu})b_c + f_{vv}b_c^2\right)\varphi_{c,i}^2 + 2f_u U_i(0) + 2f_v V_i(0) + 2D_u \sum_{j=1}^{N} L_{ij} U_j(0) \tag{5.75}$$

for $i = 1, \dots, N$. Noticing the constant k_1 given by (5.54), in vector form the above equation reads

$$0 = k_1 \operatorname{diag}(\varphi_c)\varphi_c + f_u U(0) + f_v V(0) + D_u L U(0). \tag{5.76}$$

Similarly to (5.71), $\langle LU(0), \varphi_m\rangle = \mathsf{a}_m\lambda_m$ and the scalar product of the right-hand side of (5.76) with φ_m yields

$$0 = k_1\ell_m + (f_u + D_u\lambda_m)\,\mathsf{a}_m + f_v\mathsf{b}_m \tag{5.77}$$

for $m = 1, \dots, N$. We are left with $2N + 1$ equations (5.59), (5.72), (5.73) and (5.77) which we solve for $\beta(0), \mathsf{a}_m, \mathsf{b}_m$, $m = 1, \dots, N$. Combining (5.72) and (5.77) yields

$$\boldsymbol{J}_m|_{\sigma=\sigma_c} \begin{pmatrix} \mathsf{a}_m \\ \mathsf{b}_m \end{pmatrix} = -\ell_m \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}, \quad m \neq m_c\,, \tag{5.78}$$

and (5.59) and (5.77) give

$$\begin{pmatrix} f_u + D_u\lambda_c & f_v \\ 1 & b_c \end{pmatrix} \begin{pmatrix} \mathsf{a}_c \\ \mathsf{b}_c \end{pmatrix} = -\ell_c \begin{pmatrix} k_1 \\ 0 \end{pmatrix}. \tag{5.79}$$

or

$$\begin{pmatrix} \mathsf{a}_c \\ \mathsf{b}_c \end{pmatrix} = -\left(b_c(f_u + D_u\lambda_c) - f_v\right)^{-1} \ell_c k_1 \begin{pmatrix} b_c \\ -1 \end{pmatrix}. \tag{5.80}$$

Since b_c satisfies (5.35), we have $f_u + D_u\lambda_c + f_v b_c = 0$, and we finally get

$$\begin{pmatrix} \mathsf{a}_c \\ \mathsf{b}_c \end{pmatrix} = \left(f_v\left(b_c^2 + 1\right)\right)^{-1} \ell_c k_1 \begin{pmatrix} b_c \\ -1 \end{pmatrix}. \tag{5.81}$$

It remains to introduce this expression in (5.73) to determine $\beta(0)$, which concludes the proof. □

This proposition articulates a procedure that can be repeated any number of times to yield the derivatives of all orders of U, V and β. Doing so, we obtain a Taylor formula for the solutions of the stationary problem. However, we should point out that theorem 5.8 provides no estimate on the basin of attraction of such steady states with respect to the original problem (5.3). Therefore, we

do not know whether a given solution of the stationary problem may actually emerge as a pattern of the RD system for some initial condition. Further, the neighborhood of the bifurcation point where eqs. (5.30) and (5.31) apply is not explicit, and moreover, the emerging pattern and thus its deviation from the uniform steady state depend on the random perturbation. And thirdly, theorem 5.8 does not apply to the case that the eigenvalue is not simple, as can happen in the case of a network-defined system. These considerations limit the practical use of the approach in predicting the pattern(s) of a system.

Before proceeding with diffusion-driving instabilities in chapters 6 and 7, we will examine in the next section another mechanism inherent to delay models.

5.3. Delay-driven instabilities

A Hopf bifurcation occurs in a system when a fixed point looses stability because a pair of eigenvalues of the linearized system (around the fixed point) cross the imaginary axis of the complex plane, when a parameter of the system is varied. The bifurcation predictably gives birth to an oscillating behavior, such as a limit cycle, provided the solution remains bounded.

Reaction-diffusion systems are not immune to Hopf bifurcations. Take for instance [**52, 6**] where sufficient conditions were obtained for wave instabilities in three-component reaction-diffusion systems. Keep in mind that these two works address systems evolving on a continuum, but the conclusion that three species is a minimal requirement for sustained oscillations originating from the Turing mechanism holds on networks. If the number of species is down to two, stochastic amplification of finite-size effects can however produce oscillations even with only one diffusing species [**22**]. But this phenomenon of noise amplification occurs because the key fact that each discrete portion of space has a finite carrying capacity, an assumption that deviates from our mesoscopic point of view. Our contribution in this section is to prove that the minimal number of components can be relaxed when the dynamics is affected by a time delay, without having to revert to stochastic formulations.

Time delays, also called time lags, keep making their way into more and more mathematical models. They come into play in applications of classical mechanical engineering, for load balancing in parallel computing, in traffic flow models and in numerous fields belonging to applied network theory. Time delays are part of our understanding of the interactions between neurons in biology, they naturally belong to processes with distributed, cooperative or remote control, and appear when using networks of sensors. Roughly speaking, delays are inherent to virtually all systems where the time needed for transport, propagation, communication, reaction or decision-making cannot be neglected [**97, 8**].

Introducing some time delay in the modeling can be a very reasonable way to improve the models and avoid unnecessary or complex variants of delay-free approaches to refine the match between predictions and observations. In the case of reaction-diffusion systems, this challenge was addressed by some previous

work related to delay-driven instabilities, notably [**125, 137**] regarding two-component systems, and more recently [**104**] devoted to one-component systems with constant and evolving delay. In these works the reaction kinetics features the delay, and not the diffusion part of the equation. We will examine two alternative ways a delay may find its way in the dynamics.

5.3.1. Local reactions

To start with, we will consider a two following single-component RD system on a graph,

$$\dot{u}_i(t) = f(u_i(t)) + D\sum_{j=1}^{N} L_{ij}u_j(t-\tau), \qquad t > 0, \tag{5.82a}$$

$$u_i(t) = \varphi_i(t), \qquad t \in [-\tau, 0], \tag{5.82b}$$

for $i = 1, \dots, N$, where every φ_i is a continuous function that encodes the history of the single-component, infinite-dimensional system. The delay in the second term shows the move of a particle is delayed with respect to the difference in concentration that triggered the move per Fick's law of diffusion. This retarded scheme was introduced in [**129**] from the standpoint of synchronization of a network of dynamical units. The delay models the time needed to exchange and process information. A prototypical example is that of coupled neurons, where instantaneous interaction is impossible due to potential of the membrane having finite propagating speed [**70**]. Another practical example is the delayed reaction of a car driver trigger by a perceived variation of distance or relative velocity with respect the vehicle ahead.

As was the case with the regular two-component Turing instability, we will hold the following.

Assumption 5.10 (Steady-state of the local kinetics)**.** The local system given by the reaction term of (5.82), namely

$$\dot{u}(t) = f(u(t)) \tag{5.83}$$

$$u(0) = u_0 \tag{5.84}$$

possesses a positive stable fixed point.

Recall that if $\tau = 0$, there can be no bifurcation in (5.82), and u_e is stable for the coupled system (5.82). The following proposition formulates a topological condition for a delay-driven Hopf bifurcation, but requires a remark on well-posedness of problem (5.82).

Remark 5.11 (Boundedness and positivity of the solution of (5.82))**.** The next proposition characterizes the bifurcation from the homogeneous steady state, in the linear regime. But in order to make predictions on the long-term behavior of the system, $t \to \infty$, the solutions of eq. (5.82) need to be bounded. It is well know that this boundedness cannot be established in general, but depends on the systems parameters; the restriction holds true for positivity. Results on the boundedness of the solution of (5.82) in particular, and general methods for well-posedness of DDE's can be found in [**129**, section 2] and references therein.

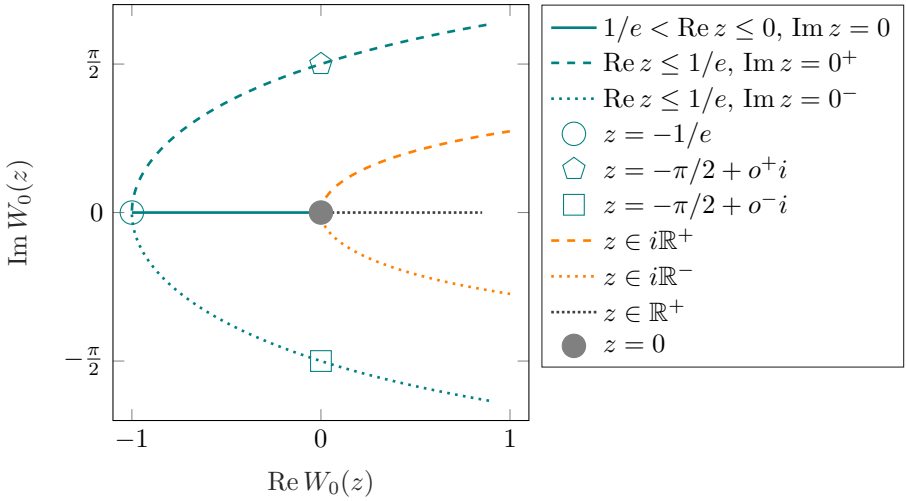


FIGURE 5.4. Principal branch of the Lambert W function. The teal colored curves correspond to the image of the negative real axis. The image of the positive real axis appears in dotted gray. The orange curves represent the image of imaginary axis.

Proposition 5.12 (Delay-driven instability)**.** *Let* (5.82) *be defined on a graph with N nodes where $\lambda_{N-1} \leq \lambda_{N-2} \leq \ldots \leq \lambda_1 \leq \lambda_0 = 0$ denote the ordered eigenvalues of the combinatorial Laplacian. Under assumption 5.10, let $u(t) = u_e$ be the stable fixed point of the local system, $f(u_e) = 0$, and let $\eta \in \left(\frac{\pi}{2}, \pi\right)$ be the solution of $\eta \cos \eta - \tau f'(u_e) \sin \eta = 0$. Then u_e is linearly globally exponentially stable if $\tau D \lambda_{N-1} > \frac{\eta}{\sin \eta}$, and the system undergoes a Hopf bifurcation if $\tau D \lambda_{N-1} < \frac{\eta}{\sin \eta}$.*

PROOF. The solution of the linearized system

$$\dot{w}_i(t) = f'(u_e) w_i + D \sum_{j=1}^{N} L_{ij} w_j(t - \tau), \qquad t > 0, \tag{5.85a}$$

$$w_i(t) = \varphi_i(t) - u_e, \qquad t \in [-\tau, 0], \tag{5.85b}$$

can be written under the form $w(t) = \sum_{m=1}^{N} c_m e^{\mu_m t} \varphi_m$, where φ_m is the eigenvector of the Laplacian associated to λ_m, chosen as before such that $\{\varphi_0, \ldots, \varphi_{N-1}\}$ form an orthogonal basis. The characteristic equation reads

$$\mu_m = f'(u_e) + \lambda_m e^{-\tau \mu_m}, \tag{5.86}$$

and rearranging the terms we get

$$\tau \left(\mu_m - f'(u_e)\right) e^{\tau \left(\mu_m - f'(u_e)\right)} = \tau D \lambda_m e^{-\tau f'(u_e)}. \tag{5.87}$$

To solve this equation for μ_m we use the Lambert W function, defined for $z \in \mathbb{C}$ by [**29**]

$$W(z) z^{W(z)} = z. \tag{5.88}$$

This function has an infinite number of branches W_k indexed by $k \in \mathbb{Z}$. It is partially real-valued only for $k = -1$ and for the index $k = 0$ which corresponds to the principal branch satisfying [**126**, Lemma 3]

$$\max_{k\in\mathbb{Z}} \operatorname{Re} W_k(z) = \operatorname{Re} W_0(z), \quad \forall z \in \mathbb{C}. \tag{5.89}$$

Combining eqs. (5.87) and (5.88) yields

$$\mu_{m,k} = \frac{1}{\tau} W_k\left(\tau D \lambda_m e^{-\tau f'(u_e)}\right) + f'(u_e), \quad k \in \mathbb{Z}, \tag{5.90}$$

where the countable number of solutions is a consequence of problem (5.82) being infinite-dimensional. Due to (5.89), we further have

$$\operatorname{Re} \mu_{N-1,0} \geq \max_{m=0,\dots,N-1,k\in\mathbb{Z}} \operatorname{Re} \mu_{m,k}. \tag{5.91}$$

As illustrated by the teal-colored curves on fig. 5.4, the image of $\left(-\infty, -\frac{\pi}{2}\right)$ by W_0 has positive real part and nonzero imaginary part, and that the image of $\left(-\frac{\pi}{2}, 0\right)$ has negative real part. Therefore stability of u_e is equivalent to

$$\operatorname{Re} W_0\left(\tau D \lambda_{N-1} e^{-\tau f'(u_e)}\right) < -f'(u_e)\tau, \tag{5.92}$$

and the Hopf bifurcation occurs when the equality is reversed. This expression is transcendental, and does not yield the critical value of τ or λ_{N-1} explicitly. However if we let $\xi + i\eta = W(x + iy)$ with $x, y, \xi, \eta \in \mathbb{R}$, from (5.88) we get

$$x = e^{\xi}\left(\xi \cos\eta - \eta \sin\eta\right), \tag{5.93}$$

$$y = e^{\xi}\left(\eta \cos\eta + \xi \sin\eta\right), \tag{5.94}$$

which in the case of (5.92) allows to write

$$\tau D \lambda_{N-1} > -\tau f'(u_e)\cos\eta - \eta \sin\eta, \tag{5.95}$$

$$0 = \eta \cos\eta - \tau f'(u_e) \sin\eta, \tag{5.96}$$

where $\eta := \operatorname{Im} W_0\left(\tau D \lambda_{N-1} e^{-\tau f'(u_e)}\right)$. Since the function $x \mapsto \frac{\tan x}{x}$ is a one-to-one correspondence from $\left(\frac{\pi}{2}, \pi\right)$ to $\mathbb{R}_0^-$, eq. (5.96) defines a unique value of η in $\left(\frac{\pi}{2}, \pi\right)$. Using this value of η and plugging (5.96) into (5.95) yields the stability condition

$$\tau D \lambda_{N-1} > \frac{-\eta\cos\eta}{\sin\eta}\cos\eta - \eta\sin\eta = -\frac{\eta}{\sin\eta}, \tag{5.97}$$

which determines a threshold value for the smallest eigenvalue of the Laplacian, allowing to conclude. □

With fig. 5.5, we provide an example of a wave instability, where the system and parameters are such that the solutions remain bounded. Increasing the delay, raising the relative importance of the diffusive coupling by selecting a larger diffusion coefficient, or augmenting the size of the network sufficiently will break the boundedness displayed by the example.

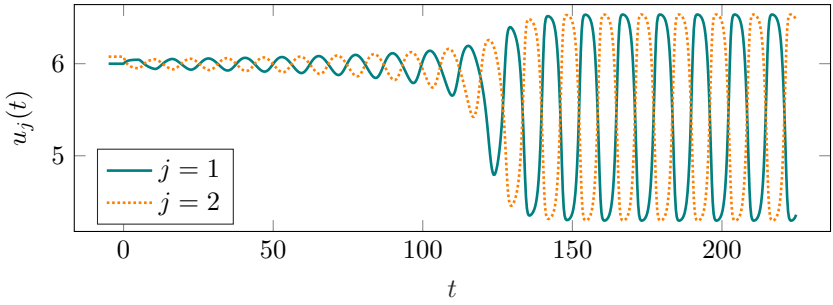


FIGURE 5.5. Wave instability in system (5.82). The graph has only two nodes, and the Laplacian eigenvalues are $-2 = \lambda_1 < \lambda_0 = 0$, therefore going lower than the threshold value of proposition 5.12 for $\lambda_{N-1} = \lambda_1$ and given by -1.9. This value results from $\eta = 2.57$ obtained with eq. (5.96) for the following set of parameters: $f(x) = \frac{1}{5}(x-5)(26-5x)(x-6)$, $u_e = 6$, $\tau = 5$ and $D = 0.05$. The history of the system is determined by two constant functions set to independent random values close to u_e.

Remark 5.13 (Delay-driven *vs* diffusion-driven instability)**.** The mechanism that is at play in proposition 5.12 differs in nature from the diffusion-driven Turing instability. Indeed, the delay-free diffusive coupling has an homogenization effect, and only the interplay with the otherwise stable reaction dynamics turns diffusion into the driving force of instability. This is not the case with delayed diffusion in a one-component setting. Indeed if we select τ such that the delayed diffusion dynamics

$$\dot{u}_i(t) = D\sum_{j=1}^{N} L_{ij}u_j(t-\tau), \qquad t > 0, \tag{5.98a}$$

$$u_i(t) = \varphi_i(t), \qquad t \in [-\tau, 0], \tag{5.98b}$$

admits $u_i(t) \equiv 1$ for $t \geq 0$ as a stable equilibrium, meaning delayed diffusion also has an homogenization effect, then it cannot drive the RD system towards instability. Indeed, assume

$$\operatorname{Re} W_0(\tau D\lambda_m) < 0 \tag{5.99}$$

for all λ_m, or equivalently

$$-\frac{\pi}{2} < \tau D\lambda_m \tag{5.100}$$

such that the homogeneous fixed point of eq. (5.98) is exponentially stable. We have to show that $\operatorname{Re}\mu_{N-1,0} < 0$, that is,

$$\operatorname{Re} W_0\left(\tau D\lambda_{N-1}e^{-\tau f'(u_e)}\right) + f'(u_e)\tau < 0 \tag{5.101}$$

where we have used the stability condition (5.92). As depicted by fig. 5.4, observe that

$$\operatorname{Re} W_0\left(\tau D\lambda_{N-1}e^{-\tau f'(u_e)}\right) < \operatorname{Re} W_0\left(-\frac{\pi}{2}e^{-\tau f'(u_e)}\right), \tag{5.102}$$

which combined with (5.101), implies we have to prove that

$$\operatorname{Re} W_0\left(-\frac{\pi}{2}e^{\alpha}\right) - \alpha < 0, \quad \forall \alpha > 0. \tag{5.103}$$

Let $\xi, \eta \in \mathbb{R}$ be given by $\xi + i\eta = W_0\left(-\frac{\pi}{2}e^{\alpha}\right)$, with $\xi > 0$ because $-\frac{\pi}{2}e^{\alpha} < -\frac{\pi}{2}$ and where we can select $\frac{\pi}{2} < \eta < \pi$ for the same reason. Condition (5.103) is then written as $\xi - \alpha < 0$. Equations (5.93) and (5.94) take the form

$$-\frac{\pi}{2}e^{\alpha} = e^{\xi}\left(\xi\cos\eta - \eta\sin\eta\right), \tag{5.104}$$

$$0 = e^{\xi}\left(\eta\cos\eta + \xi\sin\eta\right), \tag{5.105}$$

which combined, noticing that $\frac{\pi}{2} < \eta < \pi$, yield

$$\xi - \alpha = \ln\left(\frac{\pi}{2}\frac{\sin\eta}{\eta}\right). \tag{5.106}$$

It suffices to observe that

$$\frac{\pi}{2}\frac{\sin\eta}{\eta} < \frac{\pi}{2}\frac{1}{\frac{\pi}{2}} = 1 \tag{5.107}$$

to conclude that in spite of the interplay with the reaction, (5.82) cannot be destabilized by the delayed diffusive coupling.

In [**110**], we studied variants of such system where the delay also affects the reaction terms, and a follow-up on the single species models was the subject of [**109**] where we showed that depending on the size of the time delay, stationary or oscillatory patterns could emerge in a two-component setting. From a technical point-of-view these cases bring no novelty, except that the stability domain for the possibly complex topological eigenvalues is no longer a subset of the negative real axis as was the case with only one species, but extends to a portion of the complex plane. The interested reader is referred to the above references.

5.3.2. Global feedback

With an example, in this section we want to make clear that the delay needs not necessarily appear in the diffusion to induce a bifurcation to oscillations. Indeed, during the numerical exploratory phase of a collaborative work we found out global delay feedback could drive the system towards a wave instability. Consider the following variant of (5.82),

$$\dot{u}_i(t) = u_i(t)(h - u_i(t))(u_i(t) - 1) + D\sum_{j=1}^{N} L_{ij}u_j(t), \tag{5.108}$$

for $i = 1, \ldots, N$, with a bistable reaction term, as can readily be seen on fig. 5.6. Here $h : \mathbb{R}^+ \to \mathbb{R}$ is used as a feedback of the form

$$h(t) = \begin{cases} h_0 & \text{if } 0 \le t \le \tau, \\ h_0 + \mu\left(s(t-\tau) - s_0\right) & \text{if } t > \tau, \end{cases} \tag{5.109}$$

where $s(t) := \sum_{j=1}^{N} u_j(t)$ is called the total activation of the network.

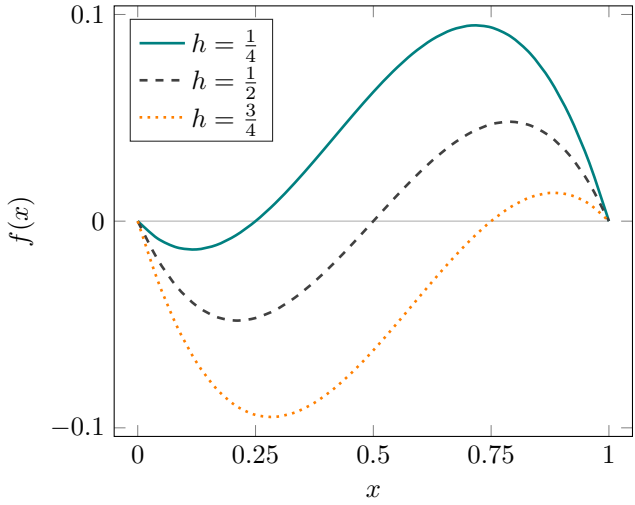


FIGURE 5.6. Bistable reaction kinetics in (5.108), $f(x) = x(h - x)(x - 1)$

The solution to (5.108) proved sensitive to the initial activation and choice of parameters, and a diversity of responses was obtained. With fig. 5.7 we demonstrate that up to a careful selection of the dynamical and topological parameters, sustained oscillations with positive activation in all nodes, also known as breathing, are obtained. Observe that such behavior was long known to occur in RD systems, but required two species[6].

Starting from breathing such as in fig. 5.7, reducing the delay would typically induce pinning, that is, a stationary pattern, while increasing it would lead to a stable limit cycle where $s(t)$ would periodically reach zero. The total activation fades out, $s(t) \equiv 0$ for sufficiently large t, when the delay gets even larger.

5.4. Conclusion

One of the goals of this first chapter, referring to our second research question, was to pave the way for what comes next: a new chapter devoted to the stability analysis of RD systems with two species. Further, we wanted to determine how the Turing mechanism behaves with delayed diffusion. To achieve that goal, we have followed a path starting from the classical finite-dimensional systems on graphs with diffusion driven by the Laplacian matrix, to eventually study variants with time delay.

[6]For instance, the authors of [**49**] revealed the emergence of breathing spots in a 1 spatial dimension model given by:

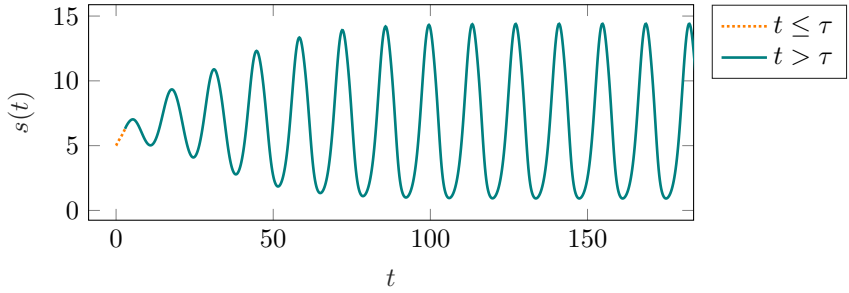
$$u_t = u - u^3 - v + u_{xx}$$
$$v_t = \epsilon(u - a_1 v - a_0) + \delta v_{xx}$$

where ϵ controls the difference of time scale between the reaction kinetics of the two species.

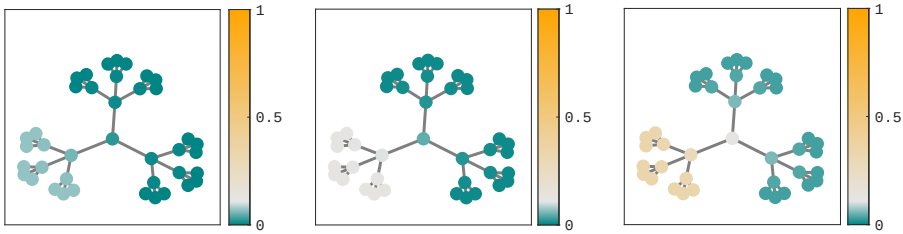
Firstly, in the classical setting we went beyond the well-know conditions for Turing bifurcations. We applied a bifurcation argument to study the structure of the solutions of the stationary problem that the steady state needs to solve. Much remains to be done however to obtain the same degree of understanding as for systems on continuous domains.

Secondly we have shown with proof-of-concept examples that the number of interacting species is no longer critical for the emergence of a wave instability, provided time delay affects the dynamics, be it on the side of the reactions through global feedback, or on the side of diffusion under the form of a processing delay. In all cases we presented, the nature of the bifurcation changed to Hopf-type, and oscillations emerged in selected examples.

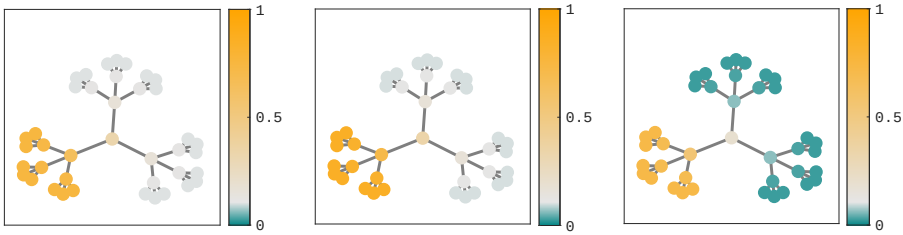
Although a beautifully simple explanation of pattern formation, the Turing paradigm suffers from stringent applicability conditions. Existing workarounds have been commented on in this chapter, but the eventuality that the temporality of the graph could have a facilitating impact was left unexplored. This will no longer be the case with the next chapter.



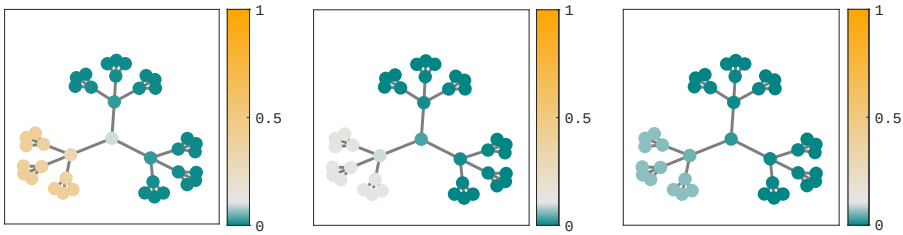
(A) Oscillations of the total activation $s(t) = \sum_{j=1}^{N} u_j(t)$ in case of breathing. That $s(t) > 0$ for all $t \geq 0$ is result of all nodes of the network remaining above zero activation at all times, $u_j(t) > 0$. Global feedback was activated only after $t = \tau$.



(B) $t = 92$ (C) $t = 94$ (D) $t = 96$



(E) $t = 98$ (F) $t = 100$ (G) $t = 102$



(H) $t = 103$ (I) $t = 104$ (J) $t = 106$

FIGURE 5.7. Breathing in a bistable RD system with global time delay feedback, eq. (5.108). The snapshots in panels (C) to (J) illustrate the behavior of the system in regime over one period. The parameters are $\tau = 2.85$, $h_0 = 0.1$, $\mu = 0.15$, $D = 0.11$, $N = 40$, with $s(0) = 5$ nodes initially activated, one at the root and four in the lower left branch.

CHAPTER 6

Diffusion-driven instability on temporal networks

6.1. Introduction

That time-varying graphs could facilitate diffusion-driven instabilities is a natural chapter 2-chapter 5 crossover idea, and the central point of the second research question. Even if we have introduced somewhat intricate models of diffusion on temporal graphs, it makes sense to begin with the simple case of switched networks outlined in remark 2.4. In that setting, diffusion is actually a passive edge-centric random walk, where a new set of edges is substituted to the current one at every discrete time step. Some or even all edges may then be rewired at the same time, so that the adjacency matrix has the form

$$A(t) = A_{\xi(t)}, \tag{6.1}$$

where the switching function $\xi : \mathbb{R}^+ \to I \subset \mathbb{N}$ selects the adjacency matrix of the graph at time t, out of a predetermined set of matrices A_i indexed in I. The Laplacian becomes a matrix valued function of time,

$$L(t) = A(t) - D(t), \tag{6.2}$$

with $D(t)$ the diagonal matrix of the degrees at time t.

Dynamical systems on time-varying networks have evidently already been studied, and the systems and control literature on the subject is abundant, see for instance [**130**] and the references therein. Numerous works have contributed to different interpretations for the time-dependent diffusion operator $L(t)$, and not all of them are based on random-walks. For example, the authors of [**65**] have introduced so-called function dynamics, where the network evolution is self-determined by the dynamical system. And in the same line of thoughts, the dynamics may revolve around activity-driven temporal networks [**108, 82**].

The list goes on, but what sets us apart is the angle of approach. According to our work, the sustained rewiring of the edges clears the way for a variant of diffusion-driven instability. The main effect in that respect will be that even if Turing instability is impeded in the system on any of the static configurations A_i, this will generally no longer hold true with $A(t)$, provided that the timescale of the network is fast enough. This fact contributes a method based on the topology to help overcome a limitation of the Turing mechanism identified in the previous chapter, namely its tight applicability conditions.

Our approach results from the following considerations. With well-posedness in mind, it is convenient to start first with continuously varying edge weights, such that $L(t)$ is a continuous function. In that case the concept of solution remains the classical one, and standard averaging theorems can be used out-of-the-box. We further assume periodicity, so that Floquet theory is applicable. Those two restrictions will be relaxed to encompass switched networks and generalized periodicity, in a sense defined below and which includes quasi-periodic topologies.

The layout of this chapter is as follows. Section 6.2 deals with the case of continuous $L(t)$, and establishes the main stability result of the chapter. In section 6.3 we extend the study to switched systems with solutions in the weaker sense of Caratheodory. Finally, section 6.4 contains a discussion about the relevance of the underlying hypothesis of the chapter, that the timescale for the evolution of the graph is sufficiently fast with respect to the timescale of the local dynamics. We present an example for both situations, first when the assumption does not hold, and second when it arguably applies.

6.2. The continuous case

This section is designed to first build intuition through the use of the method of averaging in a tailored environment, where the evolution of the temporal network is sufficiently smooth and is T-periodic. We then discuss the stability of a critical point of the reaction dynamics. We will also discuss the computation of the threshold for the timescale separation between local dynamics and diffusion.

The reaction-diffusion equations under consideration in this chapter model a two-species system. They are similar to (5.3), and all variables and parameters carry over their meaning of chapter 5. Yet one should notice the time dependency of the Laplacian in the following nonautonomous system:

$$\dot{u}_i(t) = f(u_i(t), v_i(t)) + D_u \sum_{j=1}^{N} L_{ij}(t/\varepsilon) u_j(t) \tag{6.3a}$$

$$\dot{v}_i(t) = g(u_i(t), v_i(t)) + D_v \sum_{j=1}^{N} L_{ij}(t/\varepsilon) v_j(t) \tag{6.3b}$$

for all $i = 1, \dots, N$ and $t > 0$, with as before the initial conditions in $t = 0$. The newly introduced positive parameter ε controls the timescale associated to the temporal graph. The cornerstone is to relate the stability behavior of (6.3) with a time-averaged counterpart. Let us first define the average Laplacian associated to $L(t)$,

$$\langle L \rangle := \frac{1}{T} \int_0^T L(t) dt, \tag{6.4}$$

and introduce a new set of autonomous ODE's

$$\dot{\mathsf{u}}_i(t) = f(\mathsf{u}_i(t), \mathsf{v}_i(t)) + D_u \sum_{j=1}^{N} \langle L \rangle_{ij} \mathsf{u}_j(t) \tag{6.5a}$$

$$\dot{\mathsf{v}}_i(t) = g(\mathsf{u}_i(t), \mathsf{v}_i(t)) + D_v \sum_{j=1}^{N} \langle L \rangle_{ij} \mathsf{v}_j(t) \tag{6.5b}$$

for all $i = 1, \dots, N$ and $t > 0$, called the average system of (6.3), which shares the same initial conditions.

We use the notation of the previous chapter for the variables of problem (6.3), $\boldsymbol{w} = (u_1, \dots, u_N, v_1, \dots, v_N)$ and we write $\mathsf{w} = (\mathsf{u}_1, \dots, \mathsf{u}_N, \mathsf{v}_1, \dots, \mathsf{v}_N)$ for the variables of (6.5). Letting $\boldsymbol{w}_0$ denote the initial condition, we have $\boldsymbol{w}(0) = \boldsymbol{w}_0 = \mathsf{w}(0)$. Note that the solution of (6.3) is parametrized by ε and we write $\boldsymbol{w}(\cdot, \varepsilon)$ only when explicit notation is needed. Here the nonlinear vector function $F : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ is given by

$$F_i(\boldsymbol{w}) = \begin{cases} f(u_i, v_i), & i = 1, \dots, N, \\ \\ g(u_i, v_i), & i = N+1, \dots, 2N. \end{cases} \tag{6.6}$$

The Jacobian of F is the matrix-valued function written as

$$\frac{\partial F}{\partial \boldsymbol{w}} := \begin{pmatrix} \frac{\partial F_1}{\partial w_1} & \cdots & \frac{\partial F_1}{\partial w_{2N}} \\ \vdots & & \vdots \\ \frac{\partial F_{2N}}{\partial w_1} & \cdots & \frac{\partial F_{2N}}{\partial w_{2N}} \end{pmatrix}. \tag{6.7}$$

To further lighten the notation we write $L_\otimes(t) := \begin{pmatrix} D_u L(t) & 0 \\ 0 & D_v L(t) \end{pmatrix}$ and $\langle L_\otimes \rangle$ denotes its time average.

6.2.1. The method of averaging

In this section we establish the proximity of the solutions of eq. (6.3) and of its average (6.5) on finite time intervals. We use the Landau notation for order functions[1]. In particular for a function $h : A \times (0, \varepsilon^\star) \subset \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$ and θ an order function, we say $h(x, \varepsilon) = \mathcal{O}(\theta(\varepsilon))$ when $\varepsilon \searrow 0$ if there exists a constant $C > 0$ such that

$$\sup_A \|h(\cdot, \varepsilon)\| < C\theta(\varepsilon) \tag{6.8}$$

on $(0, \varepsilon)$, where $\|\cdot\|$ is any (equivalent) norm on $\mathbb{R}^m$. Unless stated otherwise, to fix ideas we work with the supremum norm, and $B \subset \mathbb{R}^{2N}$ is a bounded domain.

Proposition 6.1 (Averaging in the continuous case)**.** *Consider IVP's* (6.3) *and* (6.5) *and assume the functions* F, $\frac{\partial F}{\partial \boldsymbol{w}}$ *are continuous and bounded by a constant independent of* ε *on* $B \subset R^{2N}$*, and assume* L *is continuous and bounded independently of* ε *on* $[0, \infty)$*. If* L *is* T*-periodic with average* $\langle L \rangle$ *with*

[1] An order function θ in a is a positive continuous function which decreases monotonically when ε goes to the limit point a, and for which the corresponding limit $\varepsilon \to a$ exists.

T independent of ε, then there exists $t^\star > 0$ such that (6.3) *and* (6.5) *have respectively unique solutions $\boldsymbol{w}(t)$ and $\mathsf{w}(t)$ on $[0, t^\star)$ which satisfy*

$$\boldsymbol{w}(t) - \mathsf{w}(t) = \mathcal{O}(\varepsilon) \tag{6.9}$$

on that interval.

PROOF. Let $\tau := t/\varepsilon$ denote a scaled time variable and define the functions $\boldsymbol{W}(\tau) := \boldsymbol{w}(t)$ and $\mathsf{W}(\tau) := \mathsf{w}(t)$ on $[0, \infty)$, so that (6.3) rewrites

$$\frac{d\boldsymbol{W}(\tau)}{d\tau} = \varepsilon \left(F(\boldsymbol{W}) + L_\otimes(t)\boldsymbol{W}\right), \tag{6.10}$$

and similarly for (6.5),

$$\frac{d\mathsf{W}(\tau)}{d\tau} = \varepsilon \left(F(\mathsf{W}) + \langle L_\otimes \rangle \mathsf{W}\right), \tag{6.11}$$

with initial conditions $\boldsymbol{W}(0) = \mathsf{W}(0) = \boldsymbol{w}_0$. The continuity and boundedness assumptions of the proposition ensure that there exist unique solutions to (6.10), (6.11) by [**133**, Theorem 1.1] on a finite interval $[0, t^\star/\varepsilon)$ for some constant $t^\star > 0$ independent of ε. Moreover the estimate

$$\boldsymbol{W}(\tau) - \mathsf{W}(\tau) = \mathcal{O}(\varepsilon) \tag{6.12}$$

is valid on that interval by the averaging theorem [**133**, Theorem 11.1]. It suffices to go back to the original variable to conclude. □

Remark 6.2 (Higher-order approximation)**.** The averaging theorem does not limit the order of the approximation of the original system given by (6.10) in the τ variable, to only first order like in (6.11). It is possible to construct second-order, and even higher order approximations. For instance, a second order approximation would result in an $\mathcal{O}(\varepsilon^2)$ estimate instead of the $\mathcal{O}(\varepsilon)$ bound in (6.12). However, the length of the validity interval would still scale as $\frac{1}{\varepsilon}$, and is therefore independent of the order of the approximation when expressed back in the original time variable t. This means the validity domain of the estimate does not scale as the order of the approximating system. Therefore we will carry on with a first order approximation, which simplifies the analysis.

Remark 6.3 (Generalized periodicity)**.** The periodicity assumption in the averaging theorem may be relaxed by the weaker existence condition of the generalized average

$$\langle L \rangle_\infty := \lim_{T \to \infty} \frac{1}{T} \int_0^T L(t) dt. \tag{6.13}$$

Therefore, proposition 6.1 encompasses the case of quasi-periodic networks, where the Laplacian $L(t)$ can be written as the sum of $L_1(t)$ and $L_2(t)$ with respective non commensurable periods T_1, T_2 with $T_1/T_2 \notin \mathbb{Q}$.

If proposition 6.1 indicates the responses of systems (6.3), (6.5) to the same random perturbation of a fixed point are similar in the short run, in the next section we seek at giving a spectral characterization of stability that would consolidate the agreement between the two systems.

6.2.2. Stability via averaging

Consider $w_e = (u_e, v_e)$ a critical point of the reaction dynamics of (6.3) and of its average (6.5) and let

$$\boldsymbol{w}_e := (\underbrace{u_e, \dots, u_e}_{N \text{ times}}, \underbrace{v_e, \dots, v_e}_{N \text{ times}}). \tag{6.14}$$

In our case, the linearization of (6.3) around $\boldsymbol{w}_e$ features a time-dependent Jacobian, and we will use Floquet theory to link its stability to the linearization of the average problem. We will then discuss the extension to the nonlinear problems.

6.2.2.1. *Linear stability*

Let $\boldsymbol{z} := \boldsymbol{w} - \boldsymbol{w}_e$ and consider the linearization of (6.3):

$$\dot{\boldsymbol{z}} = M(t, \varepsilon)\boldsymbol{z} \tag{6.15a}$$

$$\boldsymbol{z}(0) = \boldsymbol{w}_0 - \boldsymbol{w}_e =: \boldsymbol{z}_0, \tag{6.15b}$$

where

$$M(t, \varepsilon) := \frac{\partial}{\partial \boldsymbol{w}} F(\boldsymbol{w}_e) + L_{\otimes}(t/\varepsilon) \tag{6.16}$$

is εT-periodic. Under the notation of chapter 5 we have $\frac{\partial}{\partial \boldsymbol{w}} F(\boldsymbol{w}_e) = \boldsymbol{J} \otimes I_N$ with $\otimes$ the Kronecker product, $\boldsymbol{J}$ given by (5.6) and I_N the $N \times N$ identity matrix. Similarly with $\mathsf{z} := \mathsf{w} - \boldsymbol{w}_e$ the linearization of (6.5) is the autonomous problem

$$\dot{\mathsf{z}} = \langle M \rangle \mathsf{z} \tag{6.17a}$$

$$\mathsf{z}(0) = z_0, \tag{6.17b}$$

where this time

$$\langle M \rangle = \boldsymbol{J} \otimes I_N + \langle L_{\otimes} \rangle. \tag{6.18}$$

The spectral approach to linear stability requires that we determine the general form of the solution of the system. For simplicity, we start with the case $\varepsilon = 1$ and we write $M(\cdot) := M(\cdot, 1)$. First notice that the solution of (6.15) can be written as

$$\boldsymbol{z}(t) = \exp\left(\int_0^t M(t)\, dt\right) \boldsymbol{z}(0) \tag{6.19}$$

if and only if the matrices $M(t)$ and $M(t')$ commute for any pair t, t'. This assumption is generally not satisfied even if the Laplacian matrices commute, because the commutation between the Laplacians and the Jacobian of the reaction part is also needed. However we have

$$\boldsymbol{z}(t) = \exp\left(\Omega(t)\right) \boldsymbol{z}_0 \tag{6.20}$$

where $\Omega(t) = \sum_{n=1}^{\infty} \Omega_n(t)$ is known as the Magnus series or Magnus expansion [**87, 14**]. The series is convergent for all $t \in [0, T)$ provided that [**98**]

$$\int_0^T \|M(\tau)\|_2 \, d\tau < \pi. \tag{6.21}$$

Here we use the induced matrix norm $\|A\|_2 := \sup_{v\neq 0} \frac{\|Av\|_2}{\|v\|_2}$ where the norm of v is the 2-norm for vectors. If $[A,B] := AB - BA$ denotes the matrix commutator, indicatively the first terms of the series read

$$\Omega_1(t) = \int_0^t M(\tau)d\tau \tag{6.22}$$

$$\Omega_2(t) = \frac{1}{2}\int_0^t \int_0^\tau [M(\tau), M(\tau')]\, d\tau' d\tau \tag{6.23}$$

$$\Omega_3(t) = \frac{1}{6}\int_0^t \int_0^\tau \int_0^{\tau'} \big([M(\tau), [M(\tau'), M(\tau'')]] + [M(\tau''), [M(\tau'), M(\tau)]] \big) d\tau'' d\tau' d\tau. \tag{6.24}$$

Remark 6.4. (Explicit formulas for the Magnus series) It should be noted for later use in this section that the above expressions for Ω_1, Ω_2 and Ω_3 are obtained after integrating the identity

$$\frac{d\Omega(t)}{dt} = \sum_{n=0}^{\infty} \frac{B_n}{n!} \mathrm{ad}_\Omega^n M(t), \tag{6.25}$$

where the B_n's are the first Bernoulli numbers with $B_1 = -1/2$, and $\mathrm{ad}_X M = [X, M]$ with $\mathrm{ad}_X^0 M = M$, $\mathrm{ad}_X^n M = [X, \mathrm{ad}_X^{n-1} M]$, $n \geq 1$. Working out the recursive expression (6.25) is known to yield

$$\Omega_n(t) = \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\substack{k_1+\ldots+k_j=n-1 \\ k_1\geq 1,\ldots,k_j\geq 1}} \int_0^t \mathrm{ad}_{\Omega_{k_1}(\tau)} \mathrm{ad}_{\Omega_{k_2}(\tau)} \ldots \mathrm{ad}_{\Omega_{k_j}(\tau)} M(\tau) d\tau, \tag{6.26}$$

for $n \geq 2$. This will serve in proving lemma 6.5.

Alternatively, under the periodicity assumption for $L(t)$, the Floquet theorem ensures thus that the solution of (6.15) can be written as

$$\boldsymbol{z}(t) = P(t) \exp(tF)\, \boldsymbol{z}_0, \tag{6.27}$$

where P is T-periodic and bounded, and F does not depend on time. The eigenvalues ρ_m of the so-called monodromy matrix e^{TF} are the characteristic multipliers of the system. Each associated λ_m such that $\rho_m = e^{T\lambda_m}$ is a Floquet or characteristic exponent. These exponents are determined up to a term $2i\ell\pi/T$, $\ell \in \mathbb{Z}$, but can simply be chosen to coincide with the eigenvalues of F. Their real part characterizes the stability of the null solution in the usual way.

Combining the expressions of the solution in terms of the Magnus series and in the Floquet form leads to the Floquet-Magnus expansion [**24**]

$$\boldsymbol{z}(t) = \exp(\Lambda(t)) \exp(tF)\, \boldsymbol{z}_0. \tag{6.28}$$

Here the involved matrices each are obtained as the sum of series[2] $\Lambda(t) = \sum_{n=1}^{\infty} \Lambda_n(t)$ and $F = \sum_{n=1}^{\infty} F_n$. The matrix $\Lambda(t)$ is T-periodic, with $\Lambda(0) = 0$.

[2]The convergence for the series for F follows from the above quoted condition $\int_0^T \|M(\tau)\|_2\, d\tau < \pi$, whereas for the series related to $\Lambda(t)$ we have the more restrictive sufficient condition $\int_0^T \|M(\tau)\|_2\, d\tau < 0.20925$, see [**24**].

Since $\boldsymbol{z}(T) = \exp(\Omega(T))\,\boldsymbol{z}_0$, we have the identity

$$TF_n = \Omega_n(T), \quad n \geq 1\,. \tag{6.29}$$

Let us now consider system (6.15) in the case $\varepsilon < 1$. We write $\Omega(t,\varepsilon)$ for the Magnus expansion corresponding to $M(t,\varepsilon)$, and apply where applicable the same notation to the other functions and variables. We prove a short lemma before proceeding with a result on stability.

Lemma 6.5 (Homogeneity of the terms of $\Omega(\cdot,\cdot)$)**.** *For $t, \varepsilon > 0$ such that the series converge, it holds that*

$$\Omega(\varepsilon t, \varepsilon) = \sum_{n=1}^{\infty} \varepsilon^n \Omega_n(t).$$

Proof. It suffices to prove homogeneity for each term of the series, that is $\Omega_n(\varepsilon t,\varepsilon) = \varepsilon^n \Omega_n(t)$ for all $n \geq 1$. For $n = 1$, from (6.22)

$$\Omega_1(\varepsilon t,\varepsilon) = \int_0^{\varepsilon t} M(\tau,\varepsilon)d\tau = \int_0^t M(\nu)\varepsilon d\nu = \varepsilon\Omega_1(t), \tag{6.30}$$

by virtue of the change of variable $\nu = \tau/\varepsilon$. Take $n \geq 2$ and assume by induction that $\Omega_k(\varepsilon t,\varepsilon) = \varepsilon^k \Omega_k(t)$ for all $1 \leq k \leq n-1$. Then

$$\mathrm{ad}_{\Omega_k(\varepsilon t,\varepsilon)} = [\Omega_k(\varepsilon t,\varepsilon),\cdot] = \varepsilon^k\,[\Omega_k(t),\cdot] = \varepsilon^k \mathrm{ad}_{\Omega_k(t)}. \tag{6.31}$$

From eq. (6.26) with $\vartheta(k_1,\ldots,k_j) = \left\{k_1,\ldots,k_j \in \mathbb{N}_0 : \sum_{\ell=1}^{j} k_\ell = n-1\right\}$, we have

$$\begin{aligned}
&\Omega_n(\varepsilon t,\varepsilon)\\
&= \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\vartheta(k_1,\ldots,k_j)} \int_0^{\varepsilon t} \mathrm{ad}_{\Omega_{k_1}(\tau,\varepsilon)} \mathrm{ad}_{\Omega_{k_2}(\tau,\varepsilon)} \cdots \mathrm{ad}_{\Omega_{k_j}(\tau,\varepsilon)} M(\tau,\varepsilon)d\tau\\
&= \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\vartheta(k_1,\ldots,k_j)} \varepsilon^{k_1+\ldots+k_j} \int_0^{\varepsilon t} \mathrm{ad}_{\Omega_{k_1}\left(\frac{\tau}{\varepsilon}\right)} \mathrm{ad}_{\Omega_{k_2}\left(\frac{\tau}{\varepsilon}\right)} \cdots \mathrm{ad}_{\Omega_{k_j}\left(\frac{\tau}{\varepsilon}\right)} M(\tau,\varepsilon)d\tau\\
&= \varepsilon^{n-1} \sum_{j=1}^{n-1} \frac{B_j}{j!} \sum_{\vartheta(k_1,\ldots,k_j)} \int_0^{\varepsilon t} \mathrm{ad}_{\Omega_{k_1}(\nu)} \mathrm{ad}_{\Omega_{k_2}(\nu)} \cdots \mathrm{ad}_{\Omega_{k_j}(\nu)} M(\nu)\varepsilon d\nu\\
&= \varepsilon^n \Omega_n(t),
\end{aligned}$$

where we have used (6.31) and the change of variable $\nu = \tau/\varepsilon$. □

Proposition 6.6 (Linear stability and averaging)**.** *Assume the hypotheses of proposition 6.1 to hold and let $\boldsymbol{w}_e$ be a fixed point of* (6.3) *for a given $\varepsilon > 0$. Then there exists $\varepsilon^* > 0$ such that if $\varepsilon \in (0,\varepsilon^*)$ and the null solution of the average linear problem* (6.17) *is exponentially stable (resp. unstable), then the null solution of* (6.15) *is exponentially stable (resp. unstable).*

Proof. The Floquet-Magnus expansion for the linear problem (6.15) reads

$$\boldsymbol{z}(t,\varepsilon) = \exp(\Lambda(t,\varepsilon)) \exp(tF(\varepsilon))\,\boldsymbol{z}_0, \tag{6.32}$$

where $\Lambda(t,\varepsilon) = \sum_{n=1}^{\infty} \Lambda_n(t,\varepsilon)$ is εT-periodic, and following (6.29) each term of the series $F(\varepsilon) = \sum_{n=1}^{\infty} F_n(\varepsilon)$ satisfies

$$\varepsilon T F_n(\varepsilon) = \Omega_n(\varepsilon T, \varepsilon), \quad n \geq 1. \tag{6.33}$$

The stability of the null solution of (6.15) is determined by the characteristic exponents rightfully chosen as the eigenvalues of $F(\varepsilon)$, or equivalently by the eigenvalues of $\Omega(\varepsilon T, \varepsilon)$ due to (6.33). By lemma 6.5,

$$\Omega(\varepsilon T, \varepsilon) = \sum_{n=1}^{\infty} \varepsilon^n \Omega_n(T) = \varepsilon \left(\Omega_1(T) + \mathcal{O}(\varepsilon)\right). \tag{6.34}$$

Let $\mathsf{s}(\cdot)$ denote the spectral abscissa[3]. There exists ε^* such that if $\varepsilon \in (0, \varepsilon^*)$ then

$$\operatorname{sign} \mathsf{s}\left(\Omega(\varepsilon T, \varepsilon)\right) = \operatorname{sign} \mathsf{s}\left(\Omega_1(T)\right)). \tag{6.35}$$

Now observe that $\Omega_1(T) = \int_0^T M(\tau) d\tau = \langle M \rangle T$ to conclude. □

We will first comment on the stability of $\boldsymbol{w}_e$ in the nonlinear system, and refer to the upcoming section 6.2.3 for practical implications of this result.

6.2.2.2. *Nonlinear stability*

In nonautonomous ODE's there is no direct equivalent of the Hartman-Grobman theorem and the application of the principle of linearized stability is tied to the problem at hand. Hence we will not draw conclusions with regard to the stability of the $\boldsymbol{w}_e$ in the nonlinear system, and limit the discussion to the following remarks which could inspire further reading.

Remark 6.7 (Fixed point *vs* periodic solutions)**.** Under mild conditions applicable to our case, [**133**, Theorem 11.5] shows that there exists a periodic solution $\phi(t,\varepsilon)$ to (6.3) in the neighborhood of $\boldsymbol{w}_e$. Further, [**133**, Theorem 11.6] states that $\phi(t,\varepsilon)$ is asymptotically stable (resp. unstable) if the eigenvalues of the critical point $\boldsymbol{w}_e$ in the average system all have negative real part (resp. have, for at least one of them, postive real part). Hence if the perturbation of the fixed point is the initial condition of this periodic solution, the perturbation will vanish in (6.5) but not with (6.3).

This remark shows that a principle of linearity will require stringent conditions. A classical theorem in this sense is brought up next.

Remark 6.8 (Palmer's theorem)**.** A standard results for stability of nonautonomous equations is Palmer's linearization theorem, which an extension of the Hartman-Grobman, see [**105**] and following papers, aimed at weakening the applicability conditions of the theorem. Note that recent works focus on variants of these theorems applicable to Caratheodory-type differential equations, as shortly discussed in the forthcoming section 6.3.

[3]The spectral abscissa maps the set of complex matrices to the real numbers and is defined as the maximum of the real parts of the eigenvalues. We use the same notation for the spectral bound, definition B.12 on page 185.

Finally, we conclude with a remark which establishes a bridge between autonomous and nonautonomous equations.

Remark 6.9 (Autonomous formalism)**.** Consider a nonautonomous system such as $\dot{u} = f(u,t)$, $u \in \mathbb{R}^m$ and $u(0) = u_0$ with $f(u,\cdot)$ a T-periodic function for every u. This system may be rendered autonomous by the addition of a new variable ζ which may be seen as an angular variable modulo T, such that

$$\dot{u} = f(u,\zeta) \tag{6.36a}$$

$$\dot{\zeta} = 1, \tag{6.36b}$$

a now formally autonomous problem. The system is said to be suspended in the extended (or autonomous) phase space $\{u,\zeta\}$. For the details on the approach, see [**27**, Section 2.4].

6.2.3. Critical network timescale

Consider the instance that $\boldsymbol{w}_e$ is a linearly exponentially stable fixed point of eq. (6.3) when $\varepsilon = 1$, but that simultaneously, $\boldsymbol{w}_e$ is also linearly unstable with respect to the average system. Proposition 6.6 essentially indicates that accelerating the network timescale $\frac{1}{\varepsilon}$, namely letting $\varepsilon \searrow 0$ will destabilize the fixed point. As a result, it makes sense to determine the critical value[4] $\varepsilon^* \in (0,1)$ for which there is linear marginal stability, namely the solution of

$$\mathsf{s}\left(\Omega(\varepsilon T, \varepsilon)\right) = 0. \tag{6.37}$$

Solving the equation numerically requires to compute the Magnus series. Observe that the direct computation of the terms $\Omega_n(t)$ based on (6.26) is a numerically demanding task since it involves multiple integrals of nested commutators. In this section we introduce an efficient time-stepping method to compute an approximation of the series in the case $\varepsilon = 1$, see [**69, 15**]. We then extend the technique to arbitrary $\varepsilon < 1$.

6.2.3.1. *Time-stepping*

The time-stepping method partitions the interval $[0,T]$ uniformly in ℓ sub-intervals of length h, with ℓ large enough to satisfy to the convergence condition. The discretization points are $0 = t_0 < t_1 < \ldots < t_\ell = T$. For each sub-interval[5] $\Delta_k = [t_{k-1}, t_k)$ a Taylor series for $M(t)$ centered on the midpoint $\tau_k = (k-1/2)h$ is computed such that

$$M(t) = \sum_{k=1}^{\ell} \chi_{\Delta_k}(t) \sum_{j=0}^{\infty} m_j(\Delta_k)(t-\tau_k)^j \tag{6.38}$$

with

$$m_j(\Delta_k) = \frac{1}{j!} \left.\frac{d^j M(t)}{dt^j}\right|_{t=\tau_k}. \tag{6.39}$$

[4]Here we assume unicity of such value of ε.

[5]The last sub-interval $\Delta_\ell = [t_{\ell-1}, T]$ is closed.

The series (6.38) is then truncated to any even order $2s$ of h, and plugged in the recursive formulas (6.26) for the terms of the Magnus series. It results that the state transition matrix from $z(t_{k-1})$ to $z(t_k)$ is given by

$$\psi(\Delta_k) = \exp\left(\sum_{n=1}^{2s-2} \omega_n(\Delta_k) + \mathcal{O}\left(h^{2s+1}\right)\right), \tag{6.40}$$

with as an indication, up to order $2s = 6$ we have the following approximation for terms of the series [**15**]

$$\omega_1(\Delta_k) = h m_0(\Delta_k) + h^3 \frac{1}{12} m_2(\Delta_k) + h^5 \frac{1}{80} m_4(\Delta_k) \tag{6.41}$$

$$\begin{aligned}\omega_2(\Delta_k) &= h^3 \frac{-1}{12} [m_0(\Delta_k), m_1(\Delta_k)] \\ &\quad + h^5 \left(\frac{-1}{80} [m_0(\Delta_k), m_3(\Delta_k)] + \frac{1}{240} [m_1(\Delta_k), m_2(\Delta_k)] \right)\end{aligned} \tag{6.42}$$

$$\begin{aligned}\omega_3(\Delta_k) = h^5 \Bigg(&\frac{1}{360} [m_0(\Delta_k), m_0(\Delta_k), m_2(\Delta_k)] \\ &- \frac{1}{240} [m_1(\Delta_k), m_0(\Delta_k), m_1(\Delta_k)] \Bigg)\end{aligned} \tag{6.43}$$

$$\omega_4(\Delta_k) = h^5 \frac{1}{720} [m_0(\Delta_k), m_0(\Delta_k), m_0(\Delta_k), m_1(\Delta_k)], \tag{6.44}$$

having used the simplified notation

$$[x_1, x_2, \ldots, x_j] = [x_1, [x_2, [\ldots, [x_{j-1}, x_j] \ldots]]]. \tag{6.45}$$

The Baker-Campbell-Hausdorff formula eventually facilitates the computation of a limited-order development of the monodromy matrix

$$\Psi(T) = \prod_{k=1}^{\ell} \psi(\Delta_k). \tag{6.46}$$

Note that here no integration of $M(t)$ is required.

6.2.3.2. *Computing the critical timescale*

Using the time-stepping method, we obtain the monodromy matrix $\Psi(\varepsilon T)$ for a given $\epsilon < 1$. We then compute the characteristic exponents, or leading to the same conclusion on stability, the Floquet multipliers. This procedure needs to be repeated for varied values of ε in order to determine the value ε^* that achieves marginal linear stability.

As one may expect, we only need to approximate the terms of the Magnus series for the original system, that we then multiply by the appropriate integer power of ε. Indeed, let us fix $\varepsilon < 1$ and choose again ℓ sub-intervals in $[0, \varepsilon T]$, each of length $\varepsilon T/\ell = \varepsilon h$ and with midpoints $\varepsilon\tau_k$, $k = 1, \ldots, \ell$. Observe that the convergence condition for the series will automatically be satisfied.

The monodromy matrix reads $\Psi(\varepsilon T,\varepsilon)=\prod_{k=1}^{\ell}\psi\left(\varepsilon\Delta_k,\varepsilon\right)$, where each state transition matrix on the scaled interval $\varepsilon\Delta_k$ is given by

$$\psi(\varepsilon\Delta_k,\varepsilon)=\exp\left(\sum_{n=1}^{2s-2}\omega_n(\varepsilon\Delta_k,\varepsilon)+\mathcal{O}\left(h^{2s+1}\right)\right). \tag{6.47}$$

To support our claim that only the terms of the series for $\varepsilon=1$ need to be computed, observe that

(6.48)

$$m_j(\varepsilon\Delta_k,\varepsilon):=\frac{1}{j!}\left.\frac{d^jM(t/\varepsilon)}{dt^j}\right|_{t=\varepsilon\tau_k}=\frac{1}{j!}\left.\frac{d^jM(t)}{dt^j}\right|_{t=\tau_k}\frac{1}{\varepsilon^{kkj}}=\varepsilon^{-j}m_j(\Delta_k).$$

It easily follows that the truncated terms of the series also have the homogeneity property $\omega_n(\varepsilon\Delta_k,\varepsilon)=\varepsilon^n\omega_n(\Delta_k)$. Hence (6.47) becomes

$$\psi(\varepsilon\Delta_k,\varepsilon)=\exp\left(\sum_{n=1}^{2s-2}\varepsilon^n\omega_n(\Delta_k)+\mathcal{O}\left(h^{2s+1}\right)\right). \tag{6.49}$$

Every aspect developed in this section is revisited next, as we relax the continuity assumption on the evolution of the weights of the edges.

6.3. Extension to switched networks

Two key differences distinguish the case of switched networks as described by eqs. (6.1) and (6.2). To start with, solutions exist only in the weaker sense of Caratheodory, due in this case to the discontinuities of $L(t)$. A Caratheodory solution to eq. (6.3) is absolutely continuous and hence almost everywhere (a.e.) differentiable on its existence domain, and satisfies (6.3a)-(6.3b) a.e. on that interval. For existence and uniqueness results, the reader is referred to the well-established literature on the matter. We will rather present a behavioral description of the solution when the timescale parameter is varied. The second main difference is a simplified analytical treatment, by our initial assumption that the evolution of the network is periodic.

6.3.1. Linear stability on switched networks

We start with the simplest case of a periodically switched network with period T. Let $0=t_0<t_1<\ldots<t_\ell=T$ denote the initial time, the points of discontinuity of the piecewise-constant switching function ξ on $[0,T]$, and the final time of the interval. Mimicking the notation of section 6.2, we let $\Delta_k=[t_{k-1},t_k)$ denote the k-th interval of length $h_k=t_k-t_{k-1}$, where $\xi(t)\equiv\xi(t_{k-1})$.

For a fixed $\varepsilon<1$ the state transition matrix over Δ_k is given by

$$\psi(\varepsilon\Delta_k,\varepsilon)=\exp\left(\int_{\varepsilon t_{k-1}}^{\varepsilon t_k}M(t,\varepsilon)dt\right)=\exp\left(\varepsilon h_kM(t_{k-1})\right), \tag{6.50}$$

where we have used $M(\varepsilon t_{k-1}, \varepsilon) = M(t_{k-1})$. It follows from (6.50) that $\psi(\varepsilon\Delta_k, \varepsilon) = \psi(\Delta_k)^\varepsilon$ and the monodromy matrix reads

$$\Psi(\varepsilon T, \varepsilon) = \prod_{k=1}^{\ell} \psi(\Delta_k)^\varepsilon. \tag{6.51}$$

The value of ε^* for which the null solution of the linearized system is marginally stable solves the equation

$$\varrho\left(\Psi(\varepsilon T, \varepsilon)\right) = 1. \tag{6.52}$$

It is assumed to be unique and is less than one, taking for granted that the null solution is stable when $\varepsilon = 1$.

If the interval between any two consecutive discontinuities of $\xi(t)$ is on average above a given threshold called the dwell time of the system, then the zero solution of (6.17) remains stable. We will now consider examples where the fixed point of the nonlinear system is destabilized by a sufficient increase of the frequency of the switching. Let us first make two brief remarks.

Remark 6.10 (Algebraic observation)**.** Consider again eq. (6.51) yielding the monodromy matrix for a given ε. Observe that in the case that the state transition matrices $\psi(\Delta_k)$ would commute, we would be able to write $\Psi(\varepsilon T, \varepsilon) = \left(\prod_{k=1}^{\ell} \psi(\Delta_k)\right)^\varepsilon = \Psi(T)^\varepsilon$. It would follow that $\varrho\left(\Psi(\varepsilon T, \varepsilon)\right) = \varrho\left(\Psi(T)\right)^\varepsilon$, showing that letting $\varepsilon \searrow 0$ would induce no instability. Therefore, an algebraic quantity based on the commutators between those matrices could hint at the necessary timescale separation to induce instability.

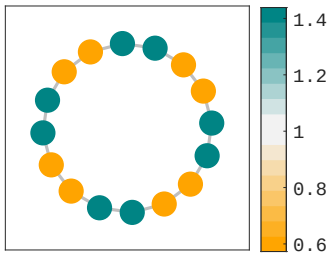
Remark 6.11 (Beyond periodic switching)**.** The averaging theorem used in proposition 6.1 may be given a proof for Caratheodory-type solutions. Hence we may extend the approach to generalized periodicity in the sense of the existence of (6.13), see remark 6.3. Note however that when the spectral radius needs to be replaced by the joint spectral radius in the above formulas, to account for the loss of periodicity, it generates numerical difficulties even with only two possible graph configurations.

6.3.2. Examples

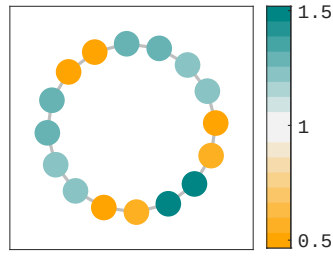
In an attempt to illustrate the above discussion in simple terms, we selected a model of periodically switched network described as follows.

Example 6.12 (The twin model)**.** Consider two static graphs on a set of N nodes regularly arranged on a ring according to their label. The number of nodes is even. In the first configuration with adjacency matrix A_1, the edges link each pair of nodes with labels $2k-1$, $2k$, where $k = 1, \ldots, N/2$. This configuration has a twin with adjacency matrix A_2, the edges connect the pairs with labels $2k$, $2k+1$ for $k = 1, \ldots, N/2-1$, with the addition of the pair labeled by 1 and N. A switched network is created by a periodic switching function on $[0, T]$,

$$\xi(t) = \chi_{[0,\gamma T)}\big(t \bmod T\big) + \chi_{[\gamma T, 1)}\big(t \bmod T\big) \tag{6.53}$$



(A) Global limit cycle



(B) Local limit cycles

FIGURE 6.1. Two qualitatively different solutions of (6.3) plotted at $t = 100$. Only the value of $\boldsymbol{u}$ is shown. In the regime of panel (A) all nodes synchronize, whereas on panel (B) synchronization is achieved only within some pairs of nodes, see fig. 6.2. We recall that we use Brusselator kinetics on the twin model, with parameters given in the main text.

where a mod b is a modulo b. Here $\gamma \in (0, 1)$ is a parameter that controls the fraction of time spent on each configuration. By letting L_1, L_2 denote the Laplacian matrices associated with A_1, A_2, we readily have

$$\langle L \rangle = \gamma L_1 + (1 - \gamma) L_2, \tag{6.54}$$

which corresponds to a weighted ring network. Both static Laplacians have the same spectrum: the zero eigenvalue has multiplicity $N/2$ corresponding to the number of connected components. Next there is the eigenvalue -2 with the same multiplicity.

We selected the Brusselator kinetics introduced in section 1.3.3.1, written under the following form,

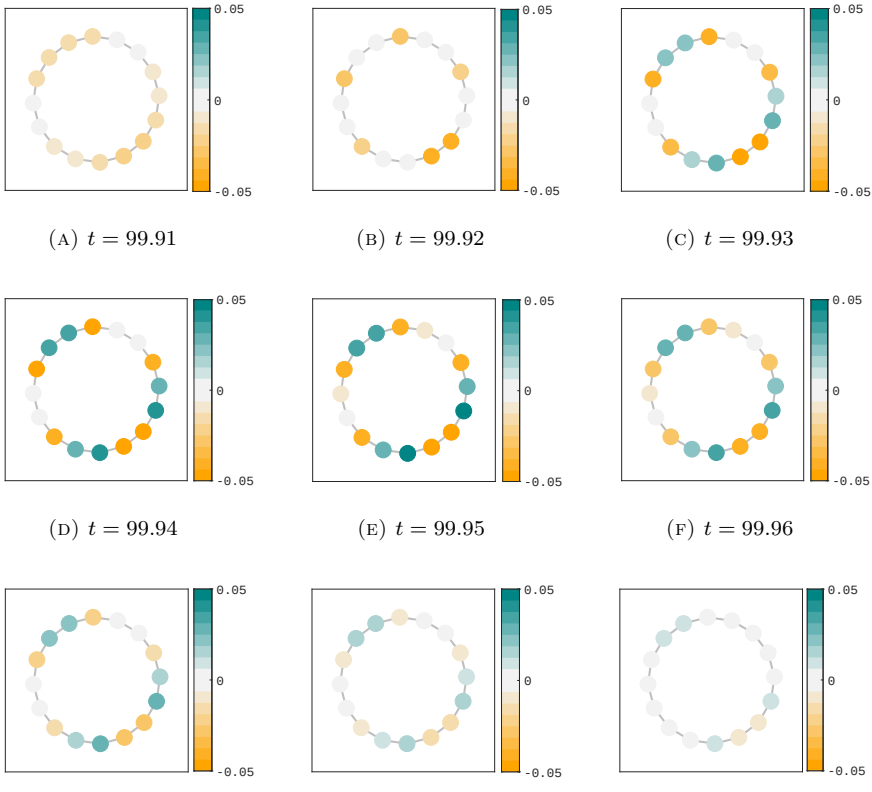
$$f(u, v) = 1 - (b + 1)u + cu^2 v \tag{6.55a}$$

$$g(u, v) = bu - cu^2 v, \tag{6.55b}$$

where the free parameters are chosen as $(b, c) = (8, 10)$. The diffusion coefficients are set to $D_u = 3$, $D_v = 10$. Resulting from this choice of parameter, $w_e = (1, b/c)$ is stable for the system on any of the two static networks with $\varepsilon = 1$, but Turing-unstable on the average graph.

If we select $\varepsilon = 0.15$ and integrate (6.3), qualitatively two behaviors emerge, as visible on fig. 6.1. Either the system evolves towards a limit cycle, or only some nodes synchronize in pairs. In both cases, the system settles in a period regime. The small amplitude oscillations around the reference value $\boldsymbol{w}(100)$ of fig. 6.1b are represented on fig. 6.2. Although not pictured here, the numerical simulations also show that $\boldsymbol{w}(t)$ is not everywhere differentiable, as expected from the discussion on Caratheodory solutions.

We then repeatedly integrated the system on the interval $[0, t_\infty]$ for decreasing values of ε. We computed the amplitude of the deviation, averaged on a small time window $[t_\infty - \Delta_t, t_\infty]$ to account for the low-amplitude oscillations



(A) $t = 99.91$ (B) $t = 99.92$ (C) $t = 99.93$

(D) $t = 99.94$ (E) $t = 99.95$ (F) $t = 99.96$

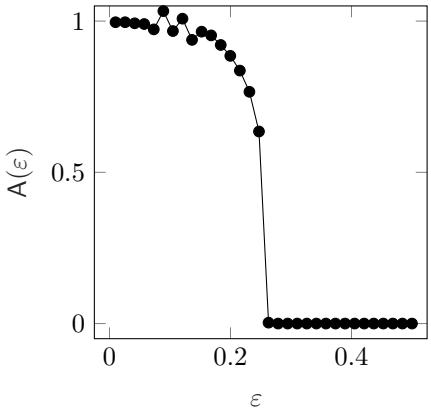
(G) $t = 99.97$ (H) $t = 99.98$ (I) $t = 99.99$

FIGURE 6.2. Oscillations in the local limit cycle regime of fig. 6.1b. Panels (A)-(I) show $\boldsymbol{u}(t) - \boldsymbol{u}(100)$, the deviation with respect to the reference value $\boldsymbol{u}(100)$. The time-interval corresponds to approximately one period of the oscillations, such that the deviations on panels (A) and (I) are relatively smaller. The behavior for $\boldsymbol{v}$ is similar.

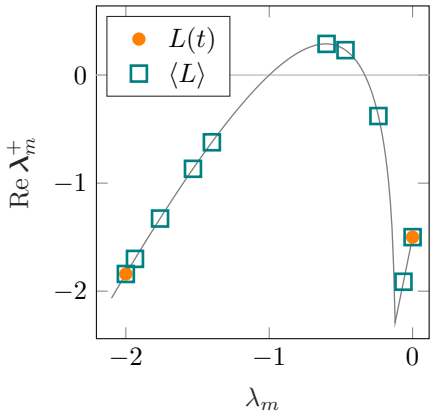
observed on fig. 6.2. Finally, we normalized the result relatively to the pattern obtained in the average system for the same initial perturbation. In formula, we let

$$\mathsf{A}(\varepsilon) = \frac{\int_{t_\infty - \Delta_t}^{t_\infty} \|\boldsymbol{w}(\tau) - \boldsymbol{w}_e\|_2 d\tau}{\Delta_t \|\mathsf{w}(t_\infty) - \boldsymbol{w}_e\|_2} \tag{6.56}$$

denote the normalized pattern amplitude, and plot it on fig. 6.3a. That $\mathsf{A}(\varepsilon) > 0$ for ε sufficiently small was expected based on fig. 6.3b, which shows the linear growth rates of the RD system on any of the twin configurations, and on the average graph. The critical network timescale given by ε^* is visibly located in the interval $[0.2, 0.3]$, and is in agreement with the analytical prediction of the linear stability analysis, see fig. 6.4.

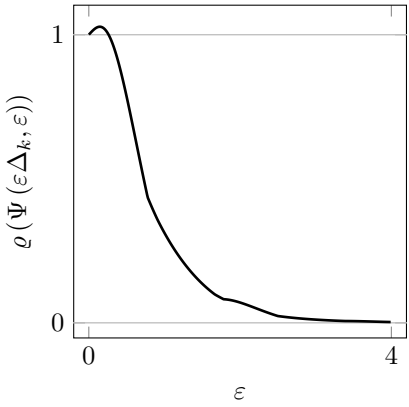


(A) Normalized pattern amplitude with $N = 16$, eq. (6.56)

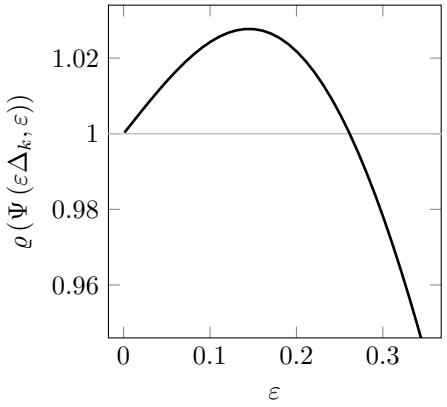


(B) Linear growth rates on the static and average networks.

FIGURE 6.3. Diffusion-driven instability from fast switching. Recall that both static configurations of the twin model have the same spectrum. The parameters are those of fig. 6.1.



(A) Spectral radius of the monodromy matrix, eq. (6.51)

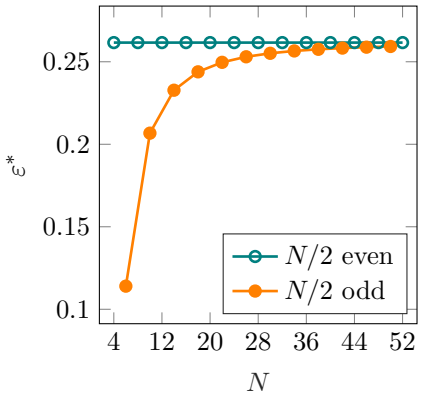


(B) Closer view of the crossing of the axis $\varrho = 1$ of marginal stability

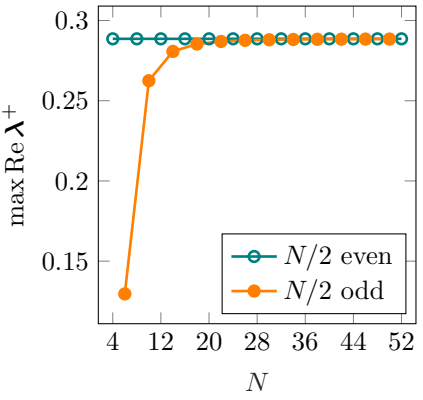
FIGURE 6.4. Spectral radius of the monodromy matrix of the linearized problem. The value ε^* read on the plot corresponds to the behavior on fig. 6.3a. When $\varepsilon \to 0$, the period $\varepsilon T \to 0$ and the monodromy matrix tends to the identity matrix. Otherwise if $\varepsilon \to \infty$, the dynamics is driven by the stable reactions and the spectral radius nears zero.

Observe the peculiar behavior of the critical network timescale in function of the number of nodes of the model, as reported by fig. 6.5a. There is qualitative agreement with the evolution of the largest linear growth rate, as depicted by

fig. 6.5b. Complementing this observation are simulations run in [**111**], where various factors influencing the magnitude of ε^* are discussed. We have considered the case of quasiperiodic temporal networks. We have also examined an example where the switching function $\xi(t)$ is a stochastic process. Loosely speaking, the consistent observation was that the larger the departure from periodicity, the larger the timescale separation needed to make $\boldsymbol{w}_e$ Turing-unstable.



(A) Critical network timescale obtained from eq. (6.52). These analytical curves were confirmed with numerical simulations similar to fig. 6.3a.



(B) Largest linear growth rate depending the number of nodes

FIGURE 6.5. The critical network timescale (A) *vs* largest linear growth rate (B), depending on the parity of half of the number of nodes.

6.4. Conclusion

We have considered the instability that conditions the emergence of patterns in RD systems, when the underlying network varies over time. Specifically we proved that a symmetry-breaking instability can be incited by properly tuning the timescale of the network's dynamics, thereby answering affirmatively the key question of the second aim of the thesis. We would like to refer the interested reader to [**111**, Supplementary material, section V] for our extension to randomly varying networks, and to a co-authored piece of work focused on desynchronization [**86**], which goes the same way in terms of approach[6].

That we have been able to address pattern formation and (de)synchronization underlies the general relevance of including the temporality of the network in the modeling. Our study includes not only switched networks, but also networks where the links may vary smoothly over time. The computed threshold for the

[6]We do not discuss it for it does not address fundamentally different technical difficulties and because the credit for the findings should rest with our co-authors.

timescale of the graph to instigate instability was confirmed through numerical simulation in both cases, and an example based on a ring topology was presented here.

The local dynamics needs to be sufficiently slow relatively to the timescale of the network rearrangements, for the effects of temporality of the graph to manifest. Take Hebbian theory in neuroscience, which models synaptic plasticity as a slow process comparatively with the neuronal dynamics, to convince yourself that our main assumption is certainly a limitation. Still, our setting remains relevant in, for instance, modeling virus spreading mediated by pairwise contacts, or in the same line of thought, for the spread of an airline-carried disease in the United States as in [**28**]. The pathogens will take time to develop while in the meantime, infected individuals will go about their normal life of social interactions. We regard accounting for the temporal variations of the underlying graph in such practical scenario as a natural next step.

If chapter 2 was about models of diffusion on dynamical graphs, chapter 4 considered graphs so huge that a continuum-limit approach to diffusion was justified. In the next chapter, we will discuss the stability properties of RD on massive graphs through the same continuum limit formalism, and RD on graphon models of finite graphs.

CHAPTER 7

Stability in reaction-diffusion equations on graphons

7.1. Introduction

The nonlinear part of the thesis is about to end the same way our tour of random walks did : with a chapter on the continuum limit. And just like chapter 4 pertains to the simple node-centric walk, we will specifically look into reaction-diffusion on static graphs, leaving the more challenging continuum limit of RD on temporal networks aside[1]. Mirroring our previous approach, the analysis will exclusively include dense graphs, making bounded graphons the appropriate limit object. We suggest the non-chronological reader to revert to section 4.2.3 on page 72 for a short recapitulation on the topic.

In this chapter we study diffusion-driven instabilities on graphons, as suggested by our third research question. There is a shift in balance in the targeted achievements, compared to chapter 4. In that chapter, the goal was first and foremost to prove the validity of the approximation procedure of the continuum limit, and the analysis of the continuum problem was kept to some basic facts around relaxation. When it comes to the (semilinear) heat equation, recent works have already validated the graph-limit approach, and we will provide citations accordingly. But it remains to analyze the graphon-based models, and that is where our contribution stands.

The chapter puts the emphasis on the interplay between stability and nonlinearity, but as far as diffusion-driven instability is concerned, the first step is once again to linearize. This fact is reflected by the layout of the chapter which is as follows. To complete this introduction, there are two sections designed to contextualize the focus on the normalized combinatorial Laplacian, and to properly motivate two key assumptions. Next in section 7.2 we cover the way leading to the spectral properties of the combinatorial Laplacian that are then needed in section 7.3. In that section, as a first step we discuss the graph-limit version of the heat equation. We then carry out a linear stability analysis of single-variable and two-variable RD equations. The passage to conclusions on the stability in the nonlinear problem is taken care of in section 7.4, where we distinguish between two representative cases, first with very little and then with significant difference with respect to the finite-dimensional analysis.

[1]Working towards a graph-limit version of dynamical systems on time-varying graphs could start with a look at [**31**].

Some parts of the chapter necessitate preliminary material in spectral theory. To help keep this work reasonably self-contained, the existence of appendices A and B is advertised to the reader. References will be made when appropriate in the main text.

7.1.1. On graphon Laplacian operators

A common feature of the works [**91, 93, 63, 94, 106**] in contexts ranging from synchronization to game theory is that they relate to diffusion processes driven by Laplacian integral operators. These operators generalize the finite-rank matrix operators on networks. In chapter 4 we have already encountered the random walk Laplacian $\mathcal{L}^{rw} = \mathcal{K} - \mathcal{I}$, eqs. (4.28) and (4.29), but also the consensus Laplacian $\mathcal{L}^{cons}$, eq. (4.34). When there was need for a symmetric kernel adjacency operator, $\mathcal{A}^{norm}$ was introduced, (4.59). It corresponds to the normalized Laplacian

$$\mathcal{L}^{norm} = \mathcal{A}^{norm} - \mathcal{I}, \tag{7.1}$$

which was shown to possess a spectrum of non-positive eigenvalues. The focus of the chapter lands on the combinatorial Laplacian $\mathcal{L}$ given by

$$\mathcal{L}f(x) = \int_0^1 W(x,y)(f(y) - f(x))dy, \tag{7.2}$$

acting on to-be-defined function spaces. We will name adjacency operator and write $\mathcal{A}$ for the integral operator associated to W and given by eq. (4.15). If we let $\mathcal{D}$ be the degree multiplication operator, $\mathcal{D} = \mathcal{M}_k$, then we have the decomposition

$$\mathcal{L} = \mathcal{A} - \mathcal{D}. \tag{7.3}$$

As we know from the discussion in section 4.3.1, this operator has a central role as the nonlocal, continuum analogue of the finite-dimensional matrix operator $\frac{1}{n}(A - D)$, which is nothing but the combinatorial Laplacian on graphs, normalized by the number of nodes[2]. This integral Laplacian emerges in previous works such as those cited above, but also in power network dynamics [**73**], or voting models [**81**].

7.1.2. On our assumptions

The extent to which the chapter is relevant mainly results from assumptions that we will review quickly, since they match those of chapter 4.

[2]From chapters 2 and 4, we know the combinatorial Laplacian on a graph encodes the dynamics of the continuous-time, edge-centric random walk. The normalization of the Laplacian by $\frac{1}{n}$ ensures that the rate of the exponentially distributed waiting time on the nodes remains finite when the number of nodes goes to infinity.

7.1.2.1. *Density*

Convergence of graph sequences is limited to dense graphs, restricting our needs to bounded graphons. However, building on remark 4.12 about graphons viewed as random models of sparse graphs, we make the following remark linking Laplacian operator and sparsity, and clarifying some consequences on the portability of our reasoning to the sparse setting.

Remark 7.1 (Graphon Laplacian operators with sparse graphs)**.** Consider again a random graph model on n nodes $\left(\Omega_n, 2^{n(n+1)/2}, P\right)$. For $\omega \in \Omega_n$ and $W \in L^p[0,1]$, $p > 1$, define the finite rank Laplacian operator $L_{\rho_n,\omega} : \mathbb{R}^n \to \mathbb{R}^n$ by [**63, 94**]

$$\left(L_{\rho_n,\omega}(u)\right)_\ell = \frac{1}{d_\ell} \sum_{m \neq \ell} \mathcal{E}_{\ell m}(\omega)\,(u_m - u_\ell) \tag{7.4}$$

for all $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $\ell = 1, \ldots, n$. Recall that here $\mathcal{E}_{\ell m}$ denotes the Bernouilli random variable $\chi_{\ell \sim m}(\omega)$. Consider the deterministic operator $\bar{L}_{\rho_n} : \mathbb{R}^n \to \mathbb{R}^n$ obtained by averaging over all possible realizations of Γ_n as defined by

$$\left(\bar{L}_{\rho_n}(u)\right)_\ell = \frac{1}{d_\ell} \sum_{m \neq \ell} E\left\{\mathcal{E}_{\ell m}(\omega)\right\}(u_m - u_\ell)\,, \tag{7.5}$$

for $\ell = 1, \ldots, n$. In the sense of [**63**], the continuum limit version of the averaged operator is given in our terms by the consensus Laplacian. Since the integral kernel of $\mathcal{L}^{cons}$ is generally not symmetric, the ensuing operator on a Hilbert space is not normal. Therefore, it is not characterized by the general spectral theorem that extends theorem 7.8 to normal operators. As mentioned in section 4.1 this consensus Laplacian was also studied in the machine learning community [**53**], and is known as Laplace-Beltrami operator. This very Laplacian was also called reactive Laplacian in the finite-dimensional setting of [**26**]. Observe that diffusion equations where the Laplacian operator is scaled with the expected degree as in eq. (7.4) are not mass-conserving. Further, note that if the scaling $\frac{1}{n\rho_n}$ is used instead of $\frac{1}{d_\ell}$ in the same equation, then the operator in the continuum limit is formally given by the combinatorial Laplacian of eq. (7.2). However, the degree function is no longer bounded, as the following example shows, which bears important consequences, notably when it comes to the existence of classical solutions of problems where diffusion is based on this Laplacian. This aspect is discussed in remark 7.24.

Example 7.2 (Power-law graphon)**.** Let $0 < \alpha < \gamma < 1$ and consider the sequence (ρ_n) with $\rho_n = n^{-\gamma}$, and the graphon given by $W(x,y) = (xy)^{-\alpha}$. The associated random graph model generates sparse graphs with power-law expected degree distribution [**63**, Lemma 2]. The degree function is given by $k(x) = \frac{x^{-\alpha}}{1-\alpha}$ and is unbounded.

7.1.2.2. *Connectedness*

Recall definition 4.13 for connected graphons, which is the subject of the second assumption. When we consider dynamical systems with diffusion ruled

by the graphon Laplacian operator, and the graphon is not connected, the system can be split in a family of problems each corresponding to one of the so-called connected components of the graphon. Therefore, we will again assume connectedness to hold.

Observe that we avoid altogether our previous discussion about the degree function of the graphon being bounded away from zero.

7.2. Spectral analysis of the combinatorial Laplacian

In the finite-dimensional setting of dynamical systems on symmetric graphs, the study of diffusion-based problems is facilitated by the ability to diagonalize the symmetric Laplacian. Thanks to the existence of a complete set of orthogonal eigenvectors, spectral methods are well suited to address questions such as relaxation time or stability.

In this section we study the spectral properties of the graphon combinatorial Laplacian, building the analysis based on the decomposition (7.3). Observe that due to the boundedness of k, the degree operator is well-defined.

7.2.1. The adjacency and degree operators

Let us consider first $\mathcal{A}$ and $\mathcal{D}$ separately, since depending on the type of graphon, the spectral properties of $\mathcal{L}$ are mostly determined by either $\mathcal{A}$ or $\mathcal{D}$. Regarding $\mathcal{A}$, we simply note that the Hilbert-Schmidt theorem on page 71 applies. Let us now turn to $\mathcal{D}$, which first requires a definition before its spectrum be given by proposition 7.4 that we recall from [**115**].

Definition 7.3 (Essential range)**.** Let f be a real-valued function on a measure space $\langle M, \mu\rangle$. Then λ is in the essential range of f, written $\operatorname{ess\,ran} f$, if and only if

$$\mu\{m|\lambda - \epsilon < f(m) < \lambda + \epsilon\} > 0, \quad \forall \epsilon > 0.$$

Proposition 7.4 (Spectrum of multiplication operators)**.** *Let f be a bounded real-valued function on a measure space $\langle M, \mu\rangle$. Let $\mathcal{M}_f$ be the multiplication operator by f on $L^2(M, d\mu)$. Then $\sigma(\mathcal{M}_f)$ is the essential range of f.*

Remark 7.5 (On the eigenvalues of D)**.** It is well known that multiplication operators such as $\mathcal{D}$ are not compact. Hence the Hilbert-Schmidt theorem does not apply, and there may exist no eigenvalues. Indeed consider a multiplication operator $\mathcal{M}_f$ and let $f^{-1}(\{\lambda\}) := \{x|f(x) - \lambda = 0\}$. Now, $f(x)\phi(x) = \lambda\phi(x)$ implies $\phi(x) = 0$ on the complement of $f^{-1}(\{\lambda\})$. If for every $\lambda \in \operatorname{ran} f$, $f^{-1}(\{\lambda\})$ has measure zero, then there are no eigenvalues. Otherwise, if $f^{-1}(\{\lambda\})$ has positive measure, then for every measurable function ϕ, the product $\phi\chi_{f^{-1}(\{\lambda\})}$ is an eigenfunction of $\mathcal{M}_f$ with eigenvalue λ.

Before we proceed with the analysis of the Laplacian, let us consider the following example where we analyze the spectral properties of $\mathcal{A}$ and $\mathcal{D}$ for the smooth graphon of fig. 4.1.

Example 7.6 (Uniform attachment graphon)**.** This is a smooth graphon, for which the eigenfunctions of the adjacency operator are obtained by solving the equations

$$\frac{d^n}{dx^n}\int_0^1 W(x,y)\phi(y)dy = \lambda\frac{d^n}{dx^n}\phi(x) \tag{7.6}$$

obtained for $n=1$ and $n=2$. Differentiating twice is needed to transform the integro-differential into a differential equation. The method yields the eigenpairs $\lambda_\ell = \left(\frac{\pi}{2}+\ell\pi\right)^{-2}$ and $\phi_\ell(x) = \sqrt{2}\cos\left(\left(\frac{\pi}{2}+\ell\pi\right)x\right)$ for every $\ell\in\mathbb{N}$, where the normalization of the eigenfunctions ensures $\|\phi_\ell\|_2 = 1$. The degree function is $k(x) = \frac{1}{2}(1-x^2)$. By proposition 7.4, $\sigma(\mathcal{D}) = [0,1/2]$. Because k is monotonic, let us also observe that by remark 7.5, $\mathcal{D}$ has no eigenvalues.

7.2.2. Diagonalization of the Laplacian

In this section, we consider all operators act on the Hilbert space $L^2[0,1]$. A key observation is that $\mathcal{L} = \mathcal{A} - \mathcal{D}$ is not compact. Indeed, if it were compact, then so would be the multiplication operator $\mathcal{L} - \mathcal{A}$. Therefore, $\mathcal{L}$ does not necessarily have a complete set of eigenfunctions like $\mathcal{A}$ does. As a result, diagonalizing $\mathcal{L}$ requires a more general framework, which applies thanks to the following proposition.

Proposition 7.7. *The combinatorial Laplacian $\mathcal{L} = \mathcal{A} - \mathcal{D}$ defined on $L^2[0,1]$ is*

(a) *self-adjoint, bounded and hence continuous;*
(b) *dissipative, that is, $(\mathcal{L}f, f) \le 0$ for every $f \in L^2[0,1]$.*

Proof. That $\mathcal{L}$ is self-adjoint follows from $\mathcal{A}$ being self-adjoint, and from k being a real function. The boundedness of $\mathcal{L}$ is a consequence of the boundedness of $\mathcal{A}$ and $\mathcal{D}$. This proves (a). To show (b), for all $f \in L^2[0,1]$ we write

$$\begin{aligned}
(\mathcal{L}f, f) &= \int_0^1\int_0^1 W(x,y)(f(y)-f(x))dyf(x)dx \\
&= \int_0^1\int_0^1 W(x,y)f(y)f(x)dydx - \int_0^1\int_0^1 W(x,y)f^2(x)dydx \\
&= -\frac{1}{2}\int_0^1\int_0^1 W(x,y)\left(f^2(x) - 2f(x)f(y) + f^2(y)\right)dxdy \\
&= -\frac{1}{2}\int_0^1\int_0^1 W(x,y)\left(f(x)-f(y)\right)^2 dxdy \le 0,
\end{aligned}$$

where we have used the symmetry of W to obtain the third equality. □

Consequently, due to (a), the spectral theorem applies, which essentially says that every bounded self-adjoint operator is a multiplication operator. This theorem generalizes the fact that every symmetric matrix is unitary diagonalizable. The theorem exists in several forms, and we give the one most suitable for our purpose [**115**, Theorem VII.3].

THEOREM 7.8 (Spectral theorem – multiplication operator form). *Let A be a bounded self-adjoint operator on a separable Hilbert space H. Then, there exist measures $\{\mu_n\}_{n=1}^N$, with $N = 1, 2, \dots$ or ∞ on $\sigma(A)$ and a unitary operator*

$$U : H \to \oplus_{n=1}^N L^2(\mathbb{R}, d\mu_n)$$

so that

$$(UAU^{-1}\psi)_n(\lambda) = \lambda\psi_n(\lambda),$$

where an element $\psi \in \oplus_{n=1}^N L^2(\mathbb{R}, d\mu_n)$ is an N-tuple $\langle \psi_1(\lambda), \dots, \psi_N(\lambda)\rangle$.

The measures $d\mu_n$ are called spectral measures. They result from a proper choice of cyclic vectors . These two notions are presented in appendices A.1 and A.2. Generally speaking, appendix A covers the material preliminary to the spectral theorem, including the definition of unitary operator used above.

What changes with the operator when it takes a multiplication form is the measures. By a proper scaling of these measures, one can ensure that the operator is mapped by U to a finite measure space, resulting in the following corollary.

Corollary 7.9. *Let A be a bounded self-adjoint operator on a separable Hilbert space H. Then, there exist a finite measure space $\langle M, \mu\rangle$, a bounded function F on M, and a unitary map $U : H \to L^2(M, d\mu)$ so that*

$$(UAU^{-1}f)(m) = F(m)f(m).$$

Remark 7.10 (Definition of U)**.** As explained in appendix A and illustrated by fig. A.1, for a given choice of a cyclic vector ψ, the operator U is defined by

$$U\phi(f)\psi = f, \tag{7.7}$$

where $f \in C(\sigma(\mathcal{L}))$ and ϕ is the mapping of eq. (A.5). Examples 7.14 and 7.15 make a practical use of this definition.

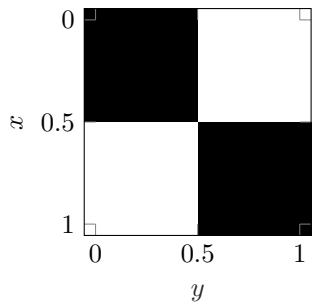
Sections 7.2.2.1 and 7.2.2.2 illustrate the use of the spectral theorem in two different cases, depending on whether the spectral properties of $\mathcal{L}$ are dominated by $\mathcal{A}$ or $\mathcal{D}$. We will make use of appendix B.1 which describes how to decompose the spectrum and hence H, based on spectral measures, in order to apply the spectral theorem.

7.2.2.1. *The case of a constant degree function*

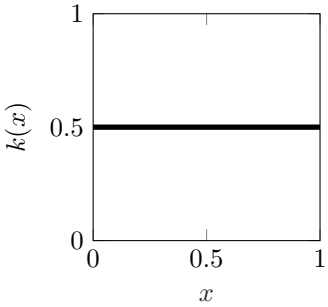
In this section, $\mathcal{A}$ and $\mathcal{D}$ are simultaneously diagonalizable, because the degree function is constant. If we assume there exists some $c \in \mathbb{R}$ such that $k(x) = c$ for a.e. $x \in [0, 1]$, then the spectral properties of $\mathcal{L}$ derive directly from those of $\mathcal{A}$. Indeed, $\mathcal{L}$ admits a complete set of eigenvectors $\{\phi_\ell\}_{\ell=1}^\infty$ associated to the eigenvalues $\lambda_\ell - c$, where the ϕ_ℓ and λ_ℓ are those obtained by application of the Hilbert-Schmidt theorem to $\mathcal{A}$. Therefore $\mathcal{L}$ can be diagonalized without explicitly reverting to the spectral theorem. However, we will illustrate the use of the theorem on an example for which the unitary transform is actually a Fourier transform. In order to do so, we will assume that the graphon is translation invariant, which we define as follows.

Definition 7.11. Let W be a graphon and let the even function $K : [-1, 1] \to [0, 1]$ be defined by $K(x) = W(1, 1-x)$ for a.e. $x \in [0, 1]$. Then W is translation-invariant if it satisfies $W(x, y) = K(x - y)$ for a.e. $x, y \in [0, 1]$.

Remark 7.12. As pointed out with the stripe graphon of fig. 4.3, translation invariance does not imply constant degree. The converse is not true either: constant degree does not imply translation invariance as shown by the example of fig. 7.1.



(A) SBM graphon $W = \chi_{[0,\frac{1}{2}]^2} + \chi_{(\frac{1}{2},1]^2}$



(B) Generalized degree function

FIGURE 7.1. Stochastic block model graphon with two blocks. The graphon is not translation invariant but has constant degree function. Observe that this graphon is not connected.

Under the assumption of translation invariance, the next proposition shows when to treat the adjacency operator as a convolution.

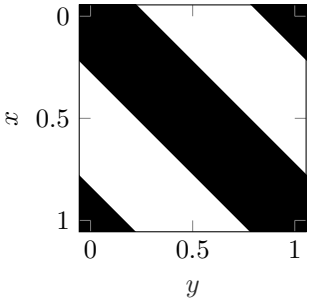
Proposition 7.13. *Let $W(x, y)$ be a translation invariant graphon such that $W(x, y) = K(x - y)$, where K is the even function of definition 7.11. Then $k(x)$ is a constant function on $[0, 1]$ if and only if*

$$K(x + 1) = K(x), \text{ for a.e. } x \in (-1, 0), \tag{7.8}$$

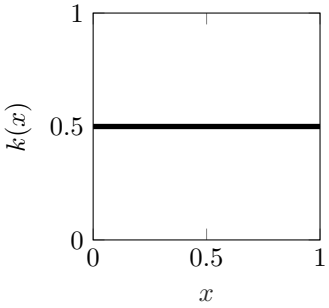
i.e. K has a 1-periodic extension to $\mathbb{R}$.

PROOF. Let us first notice that if W is translation invariant, then the degree function is symmetric with respect to $x = \frac{1}{2}$. Indeed, for a.e. $x \in (0, 1)$,

$$\begin{aligned} k(x) &= \int_0^x W(x, y)dy + \int_x^1 W(x, y)dy \\ &= \int_0^x K(x - y)dy + \int_x^1 K(y - x)dy \\ &= \int_{-x}^0 K(-z)dz + \int_0^{1-x} K(z)dz \\ &= \int_{-x}^{1-x} K(z)dz, \end{aligned} \tag{7.9}$$



(A) Nearest neighbor graphon $W(x,y) = \chi_{d(x,y)<1/4}$



(B) Generalized degree function

FIGURE 7.2. Nearest neighbor graphon, with $d(x,y)$ in the caption of panel (A) given by $d(x,y) = \min\{|x-y|, 1-|x-y|\}$. This is the graphon of example 7.14.

and hence

$$k\left(\frac{1}{2}-x\right) = \int_{-\frac{1}{2}+x}^{\frac{1}{2}+x} K(z)dz = -\int_{\frac{1}{2}-x}^{-\frac{1}{2}-x} K(z)dz = k\left(\frac{1}{2}+x\right)$$

for a.e. $x \in \left(0, \frac{1}{2}\right)$. Let us further assume that the degree function is constant. For a.e. $x \in (0,1)$, we have

$$k(x+\epsilon) - k(x-\epsilon) = 0, \tag{7.10}$$

for any positive ϵ such that $x-\epsilon, x+\epsilon \in (0,1)$. Using (7.9), the left-hand side of eq. (7.10) reads

$$\begin{aligned} k(x+\epsilon) - k(x-\epsilon) &= \int_{-x-\epsilon}^{1-x-\epsilon} K(z)dz - \int_{-x+\epsilon}^{1-x+\epsilon} K(z)dz \\ &= \int_{-x-\epsilon}^{-x+\epsilon} K(z)dz - \int_{1-x-\epsilon}^{1-x+\epsilon} K(z)dz. \end{aligned}$$

By taking the limit $\epsilon \to 0$ in eq. (7.10) and applying the Lebesgue differentiation theorem, it follows that $K(-x) - K(1-x) = 0$ for a.e. $x \in (0,1)$.

Conversely, assume a translation-invariant graphon for which K is 1-periodic. That $k(x)$ is constant follows directly from (7.9). □

As anticipated, we now consider a simple example when the spectral theorem amounts to a Fourier transform.

Example 7.14 (Spectral theorem with the nearest-neighbor graphon)**.** Consider a translation invariant graphon with constant degree, such as the nearest-neighbor graphon of fig. 7.2. The combinatorial Laplacian is given by

$$\mathcal{L}f(x) = \int_0^1 K(x-y)f(y)dy - cf(x),$$

with $c \in [0,1]$, and where according to proposition 7.13, the first term is the convolution $(K * f)(x)$. Let us show that $\mathcal{L}$ has a pure point spectrum, and determine a complete basis of eigenvalues. In this example the hat symbol ($\hat{}$) denotes the Fourier transform, such that

$$\hat{K}(\ell) := \int_0^1 K(x) e^{-2i\pi\ell x} dx$$

for $\ell \in \mathbb{Z}$, and $K(x) = \sum_{\ell\in\mathbb{Z}} \hat{K}(\ell) e^{2i\pi\ell x}$. If we let $\psi_\ell(x) = e^{2i\pi\ell x}$ for $\ell \in \mathbb{Z}$, we have

$$\mathcal{A}\psi_\ell(x) = (K * \psi_\ell)(x) = \hat{K}(\ell)\psi_\ell(x),$$

which shows that $\lambda_\ell := \hat{K}(\ell) - c$ and ψ_ℓ represent an eigenpair for $\mathcal{L}$. Recall that the ψ_ℓ form a Hilbertian basis for $L^2[0,1]$, and we have

$$\mathcal{L} = \sum_{\ell\in\mathbb{Z}} \lambda_\ell(\cdot, \psi_\ell)\psi_\ell. \tag{7.11}$$

The functional calculus introduced in appendix A.2 defines $\phi : C(\sigma(\mathcal{L})) \mapsto \mathbb{L}(L^2[0,1])$ as

$$\phi(f)(\mathcal{L}) := f(\mathcal{L}) = \sum_{\ell\in\mathbb{Z}} f(\lambda_\ell)(\cdot, \psi_\ell)\psi_\ell. \tag{7.12}$$

We may select $\psi := \sum_{\ell\in\mathbb{Z}} \psi_\ell$ as a cyclic vector. The spectral measure μ_ψ follows from the combination of eqs. (7.12) and (A.7):

$$\begin{aligned}
\int_{\sigma(\mathcal{L})} f d\mu_\psi &= (\psi, f(\mathcal{L})\psi) \\
&= \sum_{\ell\in\mathbb{Z}} f(\lambda_\ell)\,(\psi, (\psi, \psi_\ell)\psi_\ell) \\
&= \sum_{\ell\in\mathbb{Z}} f(\lambda_\ell)\,(\psi, \psi_\ell)^2 \\
&= \sum_{\ell\in\mathbb{Z}} f(\lambda_\ell).
\end{aligned}$$

Hence, $\mu_\psi = \sum_{\ell\in\mathbb{Z}} \delta(x - \lambda_\ell)$ where δ is the Dirac measure. The unitary transform U of the spectral theorem (or equivalently of lemma A.3 in this case; lemma A.4 is not needed here) maps $L^2[0,1]$ to $L^2(\sigma(\mathcal{L}), d\mu_\psi)$ which is just $\ell^2(\mathbb{Z})$, the square summable sequences indexed in $\mathbb{Z}$. Consider again the defining equation for U given by (7.7). Take f as the constant function equal to one, namely $f \equiv \{1\}_{\ell\in\mathbb{Z}}$ on $\sigma(\mathcal{L})$. Equation (7.7) yields $U\psi = 1 \equiv \{1\}_{\ell\in\mathbb{Z}}$, and as can be seen on fig. 7.3,

$$U : v = \sum_{\ell\in\mathbb{Z}} \hat{v}(\ell)\psi_\ell \to (\hat{v}(\ell))_{\ell\in\mathbb{Z}}\,, \tag{7.13}$$

which is simply the Fourier transform.

To our best knowledge, linear stability analysis for nonlinear systems with graphons was so far done in the mean-field approximation [**94, 73**] assuming a translation invariant graphon with constant degree, using the Fourier transform. In the following section, we go beyond this particular case.

$$\begin{array}{ccc} 1 \in C(\sigma(A)) \equiv \{1\}_{\ell\in\mathbb{Z}} & \xrightarrow{\quad\mathcal{M}_\lambda\quad} & \lambda\cdot 1 \equiv \left\{\hat{K}(\ell)-c\right\}_{\ell\in\mathbb{Z}} \\ \uparrow U & & \downarrow U^{-1} \\ \psi = \sum_{\ell\in\mathbb{Z}}\psi_\ell & \xrightarrow{\quad\mathcal{L}=\mathcal{A}-\mathcal{D}\quad} & \sum_{\ell\in\mathbb{Z}}\left(\hat{K}(\ell)-c\right)\psi_\ell \end{array}$$

FIGURE 7.3. Illustration of lemma A.3 in the case that U is the Fourier transform, as in example 7.14, with $f = 1$, the constant function equal to one. Here $\mathcal{M}_\lambda$ is the multiplication operator by the identity function $\lambda \to \lambda$.

7.2.2.2. *The case of a nonconstant degree function*

We consider the instance that the eigenvectors of $\mathcal{A}$ no longer help in writing $\mathcal{L}$ under multiplication form, because the degree function is not constant. When $\mathcal{L}$ not longer possesses a basis of eigenfunctions, we have to decompose the Hilbert space $H = L^2[0,1]$ as the direct sum $H = H_{pp} \oplus H_{ac} \oplus H_{sing}$, as follows from theorem B.2. Each subspace is invariant under $\mathcal{L}$, and it remains to find a cyclic vector and a spectral measure for each subspace. To make thinks more explicit, let us consider an example.

Example 7.15 (Spectral theorem with the threshold graphon)**.** The threshold graphon $W(x,y) = \chi_{x+y\leq 1}$ of fig. 4.2 has triangular support in the unit square, and its degree function is $k(x) = 1 - x$. The direct sum decomposition is $L^2[0,1] = \operatorname{span}\{1\} \oplus \{1\}^\perp$, where $\operatorname{span}\{1\} =: H_{pp}$ with basis the constant eigenfunction $\psi(x) = 1$, and $\mu_\psi = \delta(x)$ is the discrete spectral measure. Further, $\{1\}^\perp =: H_{ac}$ admits the cyclic vector $\varphi(x) = x - \frac{1}{2}$. Indeed, $\mathcal{L}^n\varphi(x)$ is a polynomial of degree $n+1$ for all $n \in \mathbb{N}$. To see that, let us assume that for $n \geq 1$, we can write

$$\mathcal{L}^{n-1}\varphi(x) = \alpha_{n-1,n}x^n + \ldots + \alpha_{n-1,1}x + \alpha_{n-1,0}$$

with $\alpha_{n-1,n} \neq 0$. This is the case for $n = 1$ for which $\alpha_{0,1} = 1$. We further have

$$\begin{aligned} \mathcal{L}^n\varphi(x) &= \int_0^{1-x} \mathcal{L}^{n-1}\varphi(y)dy - k(x)\mathcal{L}^{n-1}\varphi(x) \\ &= \sum_{\ell=0}^{n} \alpha_{n-1,\ell}\frac{(1-x)^{\ell+1}}{\ell+1} - (1-x)\sum_{\ell=0}^{n}\alpha_{n-1,\ell}x^\ell, \end{aligned}$$

where the leading coefficient of order $n+1$ is given by

$$\alpha_{n,n+1} := \frac{n+2}{n+1}\alpha_{n-1,n} \neq 0.$$

Recall that the polynomials are dense in $C[0,1]$ and that $L^2[0,1]$ is the completion of $C[0,1]$ with respect to the L^2 norm. As an indication, the first elements of

$\{\varphi, \mathcal{L}\varphi, \mathcal{L}^2\varphi, \ldots\}$ are given by:

$$\begin{aligned}\mathcal{L}\varphi(x) &= \frac{3}{2}x^2 - 2x + \frac{1}{2}\\ \mathcal{L}^2\varphi(x) &= x^3 - 3x^2 + \frac{5}{2}x - \frac{1}{2}\\ \mathcal{L}^3\varphi(x) &= \frac{5}{4}x^4 - 4x^3 + \frac{21}{4}x^2 - 3x + \frac{1}{2}.\end{aligned}$$

As hinted by fig. 7.4, note that for all $n \in \mathbb{N}$, $(\mathcal{L}^n\varphi, 1) = 0$, meaning that $\{1\}^\perp$ is invariant under $\mathcal{L}$ as expected. Indeed, for $n \geq 1$,

$$\begin{aligned}(\mathcal{L}^n\varphi(x), 1) &= \int_0^1 \mathcal{L}^n\varphi(x)dx\\ &= \int_0^1\int_0^1 W(x,y)\mathcal{L}^{n-1}\varphi(y)dydx - \int_0^1 k(x)\mathcal{L}^{n-1}\varphi(x)dx\\ &= \int_0^1 k(y)\mathcal{L}^{n-1}\varphi(y)dy - \int_0^1 k(x)\mathcal{L}^{n-1}\varphi(x)dx = 0.\end{aligned}$$

Let us now determine the spectral measure $d\mu_\varphi$. Following from eq. (A.7) we have

$$\int_{\sigma_{ac}(\mathcal{L})} f(\lambda)d\mu_\varphi = (\varphi, f(\mathcal{L})\varphi)_{H_{ac}} \tag{7.14}$$

for all $f \in C(\sigma_{ac}(\mathcal{L}))$. However eq. (7.14) does not make the measure explicit since we do not have an expression similar to eq. (7.12) for $f(\mathcal{L})$. If we consider eq. (7.14) with the triangular graphon, and we specialize this equation to the case that f is the identity function, we obtain

$$\begin{aligned}\int_{\sigma_{ac}(\mathcal{L})} \lambda d\mu_\varphi &= (\varphi, (\mathcal{L} \restriction H_{ac})\varphi)\\ &= (\varphi, (\mathcal{A} - \mathcal{D})\varphi)\\ &= \int_0^1 \mathcal{A}\varphi(x)\varphi(x)dx - \int_0^1 k(x)\varphi^2(x)dx.\end{aligned}$$

The first integral in the right-hand side is zero, and letting $\theta := -k(x)$ gives

$$\begin{aligned}\int_{\sigma_{ac}(\mathcal{L})} \lambda d\mu_\varphi &= \int_{-1}^0 \theta\varphi^2((-k)^{-1}(\theta))d\theta\\ &= \int_{-1}^0 \theta\varphi^2\left(\theta + 1\right)d\theta.\end{aligned}$$

So we have the candidate

$$d\mu_\varphi = \varphi^2\left(\theta + 1\right)d\theta = \left(\theta + \frac{1}{2}\right)^2 d\theta. \tag{7.15}$$

Further, we observe that for all $n \in \mathbb{N}$, the following identity holds:

$$(\varphi, \mathcal{L}^n\varphi) = \int_0^1 \left(-k(x)\right)^n \varphi^2(x)dx. \tag{7.16}$$

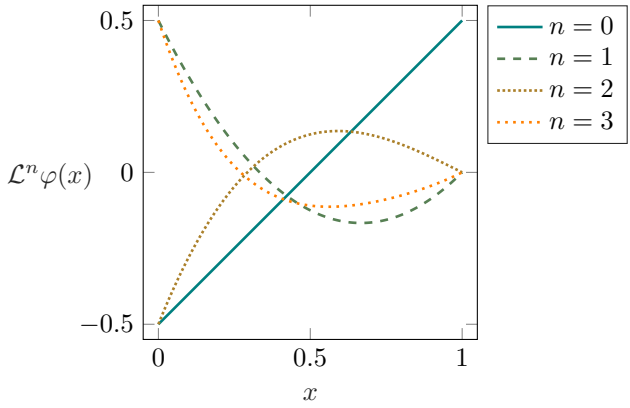


FIGURE 7.4. Representation of the iterates $\mathcal{L}^n\varphi$, $n = 0, \dots, 3$ on the cyclic vector $\varphi(x) = x - \frac{1}{2}$ with the threshold graphon with triangular support of example 7.15.

Letting again $\theta := -k(x)$ in (7.16), it follows that

$$\int_{\sigma_{ac}(\mathcal{L})} \lambda^n d\mu_\varphi = \int_0^1 \theta^n \varphi^2((-k)^{-1}(\theta))d\theta$$

and by density of the polynomials in $L^2[0,1]$, this validates the choice in (7.15).

In the above example, the spectrum of $\mathcal{L}$ was not needed to formally decompose the measure in a discrete and an absolute continuous part. However, the support of the absolute continuous measure is dictated by the knowledge of the spectrum. The next section contains a perturbative approach to obtain the spectrum of $\mathcal{L}$.

7.2.3. Perturbative approach

In appendix B.2, the spectrum is decomposed into a discrete spectrum σ_{disc}, namely the set of isolated eigenvalues of finite multiplicity, and an essential spectrum $\sigma_{disc} = \sigma \setminus \sigma_{disc}$. An interesting property of the essential spectrum resides in it being not sensitive to finite-rank perturbations or limits thereof, that is, compact perturbations. Indeed, we have the following theorem [**114**, Theorem S13]:

THEOREM 7.16 (The classical Weyl theorem). *Suppose that A and B are self-adjoint operators on a Hilbert space, such that $A - B$ is compact. Then $\sigma_{ess}(A) = \sigma_{ess}(B)$.*

From this theorem it results that

$$\sigma_{ess}(\mathcal{L}) = -\sigma_{ess}(\mathcal{D}), \tag{7.17}$$

where $\sigma_{ess}(\mathcal{D})$ follows directly from proposition 7.4 and remark 7.5. This theorem shows that the multiplication operator may dominate the spectral properties of $\mathcal{L}$, as in the next example.

Example 7.17 (Spectrum of $\mathcal{L}$ with the threshold graphon)**.** The pure-point spectrum of $\mathcal{L}$ associated with the graphon $W(x,y) = \chi_{x+y\leq 1}$ of example 7.15 is $\sigma_{pp}(\mathcal{L}) = \{0\}$. The eigenvalue zero is embedded in the essential spectrum, $\sigma_{ess}(\mathcal{L}) = -\sigma_{ess}(\mathcal{D}) = [-1,0]$, and hence $\sigma_{disc}(\mathcal{L}) = \emptyset$ and $\sigma(\mathcal{L}) = \sigma_{ess}(\mathcal{L})$.

7.3. Linear stability of reaction-diffusion on graphons

The spectral analysis in section 7.2 will now serve for a linear stability analysis of reaction-diffusion equations on graphons. To start with, we briefly consider in section 7.3.1 the continuum limit of the heat equation with a single variable. We then move to the linear stability of a critical point in the scalar reaction-diffusion equation in section 7.3.2, and in the vector case with two components in section 7.3.3.

7.3.1. Stability of the heat equation

Consider the discrete heat equation given by

$$\dot{u}_i(t) = \frac{1}{n}\sum_{j=1}^{n} A_{ij}(u_j - u_i), \quad i = 1,\dots,n \tag{7.18a}$$

$$u(0) = u_0, \tag{7.18b}$$

where $A = (A_{ij})$ is a symmetric weight matrix and $u(t) = (u_1(t),\dots,u_n(t))$. The continuum-limit in the $\|\cdot\|_{C([0,T],L^2[0,1])}$ norm[3] for any $T > 0$ was shown in [**91**] to be of the form

$$\dot{\boldsymbol{w}}(t) = \mathcal{L}\boldsymbol{w}(t) \tag{7.19a}$$

$$\boldsymbol{w}(0) = w_0, \tag{7.19b}$$

where $\boldsymbol{w}(t) := w(\cdot,t)$. Here, the matrix A and the graphon W are linked through a discretization step that we do not discuss since our focus is on eq. (7.19). Well-posedness of this equation was proved in [**91**] on the Banach space $L^\infty[0,1]$, with the right-hand side of eq. (7.19a) replaced by the more general nonlinear term

$$\int_0^1 W(x,y)F(w(y,t) - w(x,t))dy \tag{7.20}$$

with $F : \mathbb{R} \to \mathbb{R}$ a Lipschitz continuous function. The proof relies on the Banach contraction mapping principle. For the sake of completeness, we give a proof of well-posedness of this IVP in a Hilbert-space setting using a semigroup argument similar to the one used for theorem 4.19.

Proposition 7.18. *Let W be a graphon and $w_0 \in L^2[0,1]$. Then there exists a unique classical solution $\boldsymbol{u} \in C^1(\mathbb{R}, L^2[0,1])$ of IVP (7.19) with initial condition $\boldsymbol{w}(0) = w_0$.*

[3]The definition of this norm is $\|w\|_{C([0,T],L^2[0,1]} = \sup_{t\in[0,T]}\|w(\cdot,t)\|_{L^2[0,1]}$.

PROOF. Since $\mathcal{L}$ is a bounded linear operator, it is closed and therefore, it is the infinitesimal generator of the uniformly continuous semigroup

$$S_{\mathcal{L}}(t) = e^{\mathcal{L}t} := \sum_{k=0}^{\infty} \frac{t^k \mathcal{L}^k}{k!}. \tag{7.21}$$

Therefore (7.19) is well-posed and admits the unique classical solution

$$\boldsymbol{w}(t) = S_{\mathcal{L}}(t)\boldsymbol{w}(0). \tag{7.22}$$

□

We will mainly examine stability with respect to bounded perturbations, which makes the following remark useful.

Remark 7.19. With initial condition $w_0 \in L^\infty[0,1]$ one alternatively obtains the existence of a solution $\boldsymbol{w} \in C^1(\mathbb{R}, L^\infty[0,1])$ for IVP (7.19). The norm is

$$\|\boldsymbol{w}\|_{C^1(\mathbb{R},L^\infty[0,1])} = \sup_{t\in\mathbb{R}} \|\boldsymbol{w}(t)\|_{L^\infty[0,1]} + \sup_{t\in\mathbb{R}} \|\dot{\boldsymbol{w}}(t)\|_{L^\infty[0,1]}, \tag{7.23}$$

where the derivative $\dot{\boldsymbol{w}}$ of $\boldsymbol{w}$ satisfies the pointwise defining relation

$$\lim_{h\to 0} \left\| \frac{\boldsymbol{w}(t+h) - \boldsymbol{w}(t)}{h} - \dot{\boldsymbol{w}}(t) \right\|_{L^\infty[0,1]} = 0. \tag{7.24}$$

Further, we have the a priori estimate $\|\boldsymbol{w}(t)\|_\infty \leq \|w_0\|_\infty$ for all $t \geq 0$ [**59**, Theorem 3.1].

We will now give a proof of stability with the heat equation, where we expressly apply the spectral theorem and mention all details. The distance $d(\cdot,\cdot)$ in the proof is that of eq. (5.2) with the L^2 norm.

Lemma 7.20. *Let $\boldsymbol{w} \in C^1(\mathbb{R}, L^\infty[0,1])$ be the solution of eq.* (7.19) *with $w_0 \in L^\infty[0,1]$. Then*

$$\frac{\partial}{\partial t} \int_0^1 w(x,t)dx = 0. \tag{7.25}$$

PROOF. The proof can be carried out similarly as in [**93**, Lemma 3.5]. The proof interchanges integration and differentiation as follows

$$\int_0^1 \dot{\boldsymbol{w}}(t)dx = \frac{\partial}{\partial t} \int_0^1 w(x,t)dx,$$

which is permitted because $\dot{\boldsymbol{w}}$ is bounded on $\mathbb{R}$. □

Proposition 7.21 (Stability of the homogeneous equilibrium points)**.** *Let W be a connected graphon such that* $\sup \sigma\left(\mathcal{L} \restriction \{1\}^\perp\right) < 0$. *Then* $\operatorname{span}\{1\} \subset L^\infty[0,1]$ *is a globally exponentially stable set of IVP* (7.19).

PROOF. As in [**93**, Theorem 3.3], consider $L^\infty[0,1]$ as a subspace of $L^2[0,1]$, and write $L^\infty[0,1] = E \oplus E^\perp$ where $E := \operatorname{span}\{1\}$ and $E^\perp := \left\{ f \in L^\infty[0,1] : \int_0^1 f(x)dx = 0 \right\}$. Then according to lemma 7.20, E and $E^\perp$ are invariant sets for (7.19). Let $\boldsymbol{w}(0) = w_\parallel + w_\perp$ with $w_\parallel \in E$ and $w_\perp \in E^\perp$. Observe that

$d(\boldsymbol{w}(0), E) = d(w_\perp, E)$ and $d(\boldsymbol{w}(t), E) = d(\boldsymbol{w}_\perp(t), E)$ with $\boldsymbol{w}_\perp(t) := S_{\mathcal{L}}(t)w_\perp$. Also note that for all $x \in E^\perp$, $d(x, E) = \|x\|_{L^2[0,1]}$. Hence, we have to prove that $\boldsymbol{z}_e \equiv 0$ is a globally exponentially stable equilibrium point of

$$\dot{\boldsymbol{z}}(t) = \mathcal{L}\boldsymbol{z}(t) \tag{7.26a}$$

$$\boldsymbol{z}(0) = z_0 \tag{7.26b}$$

where $z_0 \in E^\perp$. Letting

$$\hat{\boldsymbol{z}}(t) := U\boldsymbol{z}(t) \tag{7.27}$$

where U is the mapping of corollary 7.9, we can also consider

$$\dot{\hat{\boldsymbol{z}}}(t) = F\hat{\boldsymbol{z}}(t) \tag{7.28a}$$

$$\hat{\boldsymbol{z}}(0) = Uz_0. \tag{7.28b}$$

with F the bounded function of corollary 7.9. Making the second variable explicit and with $\hat{z}_0 := Uz_0$,

$$\frac{\partial}{\partial t}\hat{z}(m,t) = F(m)\hat{z}(m,t) \tag{7.29a}$$

$$\hat{z}(m,0) = \hat{z}_0(m) \tag{7.29b}$$

with solution given by

$$\hat{z}(m,t) = e^{F(m)t}\hat{z}_0. \tag{7.30}$$

Since U is unitary, for all $t \geq 0$ we have $\|z(\cdot,t)\|_{L^2([0,1],dx)} = \|\hat{z}(\cdot,t)\|_{L^2(\mathbb{R},d\mu)}$ where μ is the spectral measure with support on $\sigma\left(\mathcal{L} \restriction \{1\}^\perp\right)$. The zero equilibrium is exponentially stable if $\sup\{\operatorname{ess\,ran} F\} < 0$, with the measure having support on the spectrum of $\mathcal{L} \restriction \{1\}^\perp$, or equivalently if

$$\sup\left\{\sigma(U^{-1}(L \restriction \{1\}^\perp)U)\right\} < 0.$$

To conclude, observe that the spectrum is unitary invariant, namely,

$$\sigma(U^{-1}(L \restriction \{1\}^\perp)U) = \sigma(L \restriction \{1\}^\perp).$$

□

Remark 7.22. Following proposition 7.21 we have $\lim_{t\to\infty} \boldsymbol{w}(t) =: w_\infty \in E$ and by lemma 7.20, $w_\infty = \int_0^1 w_0(x)dx = w_\parallel$, where we used the same notation for constant functions and reals.

7.3.2. Linear stability of reaction-diffusion with one variable

In this section we consider the following nonlinear variant of eq. (7.18):

$$\dot{u}_i(t) = \frac{1}{n}\sum_{j=1}^{n} A_{ij}(u_j - u_i) + f(u_i(t)), \quad i = 1, \ldots, n \tag{7.31a}$$

$$u(0) = u_0, \tag{7.31b}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a nonlinear, Lipschitz function. It is mentioned in [**91**, Remark 3.1] that the continuum limit of eq. (7.31) is of the form

$$\frac{\partial}{\partial t} w(x,t) = \int_0^1 W(x,y)\,(w(y,t) - w(x,t))\,dy + f(w(x,t)) \tag{7.32a}$$

$$w(x,0) = w_0. \tag{7.32b}$$

Remark 7.23 (Existence and uniqueness with bounded initial condition)**.** Existence and uniqueness of a classical solution $\boldsymbol{w} \in C^1(\mathbb{R}, L^\infty[0,1])$ to (7.32) with bounded initial condition follows from the same Banach contraction mapping argument that proves the well-posedness of (7.19) in [**92**]. Moreover, for any $T > 0$ we have the a priori estimate

$$\|\boldsymbol{w}\|_{C([0,T],L^\infty[0,1])} \leq C\|\boldsymbol{w}(0)\|_{L^\infty[0,1]} \tag{7.33}$$

where the positive constant C depends only on T [**63**, Theorem 3.2].

Remark 7.24 (Well-posedness of the reaction-diffusion equation in the continuum limit of sparse graphs)**.** Following remark 7.1, in the case of sparse random graphs when the Laplacian of the discrete model is scaled by the expected degree as in eq. (7.4), the continuum limit of the averaged discrete model reads

$$\frac{\partial}{\partial t} w(x,t) = \int_0^1 N(x,y)\,(w(y,t) - w(x,t))\,dy + f(w(x,t)) \tag{7.34a}$$

$$w(x,0) = w_0 \tag{7.34b}$$

where it is assumed that the degree function of the graphon is bounded away from zero. Hence the consensus Laplacian above is well-defined, and $N(x,y) = W(x,y)/k(x)$ is a nonnegative function such that

$$\int_0^1 N(x,y)dy = 1. \tag{7.35}$$

For $p \geq 2$, let q be the conjugated exponent given by

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{7.36}$$

and assume that f is Lipschitz continuous. Existence and uniqueness of a classical solution $\boldsymbol{w} \in C^1(\mathbb{R}, L^q[0,1])$ to (7.34) with $N \in L^p[0,1]$ that satisfies (7.35), and with initial condition $w_0 \in L^q[0,1]$, is given by [**63**, Theorem 3.1]. However, when the Laplacian in the discrete model in scaled by $\frac{1}{n\rho_n}$, the continuum model reads

$$\frac{\partial}{\partial t} w(x,t) = \int_0^1 W(x,y)\,(w(y,t) - w(x,t))\,dy + f(w(x,t)) \tag{7.37a}$$

$$w(x,0) = w_0, \tag{7.37b}$$

and the existence of weak solutions can be proved [**63**, Section 6.2]. Weak solutions are meant in the following sense. Let $T > 0$ and $\boldsymbol{w} \in H^1([0,T], L^2[0,1])$, meaning that for all $t \in [0,T]$, $\boldsymbol{w}(t) \in H^1[0,1]$, that is, $\boldsymbol{w}(t)$ and its weak

derivative are in $L^2[0,1]$. Then $\boldsymbol{w}$ is called a weak solution of (7.37) with $w_0 \in L^2[0,1]$ if

$$\left(\frac{\partial}{\partial t}w(\cdot,t) - \int_0^1 W(\cdot,y)\,(w(y,t) - w(\cdot,t))\,dy + f(w(\cdot,t)), v\right) = 0 \tag{7.38}$$

for all $v \in L^2[0,1]$ and for almost every $t \in [0,T]$.

The existence and uniqueness result [**63**, Theorem 3.1] mentioned in this remark applies to eq. (7.32) where W is a bounded graphon and where we may take $p = q = 2$ in (7.36). Hence we have the following result.

THEOREM 7.25 (Existence and uniqueness on $L^2[0,1]$). *The IVP* (7.32) *with initial condition* $w_0 \in L^2[0,1]$ *and globally Lipschitz* f *has a unique classical solution* $\boldsymbol{w} \in C^1(\mathbb{R}, L^2[0,1])$*, which depends continuously on* w_0*.*

In the rest of this section, we will concentrate on the problem with $w_0 \in L^2[0,1]$; the case with the more restrictive initial condition $w_0 \in L^\infty$ goes along the same line. By writing

$$\boldsymbol{w}(t) = T(t)w_0, \quad \forall t \geq 0, \tag{7.39}$$

we define a nonlinear semigroup $(T(t))_{t\geq 0}$ on $L^2[0,1]$ with infinitesimal generator given by $\mathcal{L}+f$. Let us assume the continuum problem (7.32) possesses a constant equilibrium w_e, such that $(\mathcal{L}+f)(w_e) = f(w_e) = 0$. In order to investigate the stability of w_e with perturbation in $L^2[0,1]$, eq. (7.32a) will be linearized.

We have introduced on page 106 the Fréchet derivative $DF(x)$, definition 5.7. Consider also the following weaker notion of derivative.

Definition 7.26 (Gâteaux derivative)**.** Let X be a Banach space. Then operator $F : \operatorname{dom} F \subset X \to X$ is Gâteaux differentiable at $x \in D$ in the direction $h \in X$, $x, x+h \in \operatorname{dom} F$, if there exists a linear operator $dF(x) : X \to X$ such that for all $h \in X$,

$$\lim_{\epsilon \to 0} \frac{F(x+\epsilon h) - F(x)}{\epsilon} = dF(x)h. \tag{7.40}$$

If (7.40) holds for every direction $h \in X$, then F is said to be Gâteaux differentiable in x. Further if F is Gâteaux differentiable in every $x \in \operatorname{dom} F$, then F is said to be Gâteaux differentiable in $\operatorname{dom} F$.

When the right-hand side of (7.32) is not Fréchet differentiable, as is the case in practical reaction-diffusion models, we rely on the following assumption:

Assumption 7.27. The nonlinear function f is Gâteaux differentiable at w_e, and its Gâteaux derivative $df(w_e)$ is bounded on $L^2[0,1]$.

The Gâteaux derivative is used to linearize the equation around w_e. Equation (7.32) becomes

$$\frac{\partial}{\partial t}z(x,t) = \int_0^1 W(x,y)\,(z(y,t) - z(x,t))\,dy + df(w_e)z(x,t) \tag{7.41a}$$

$$z(x,0) = w_0 - w_e =: z_0. \tag{7.41b}$$

By virtue of the spectral theorem, with $\hat{\boldsymbol{z}}(t)$ and the unitary operator U as in eq. (7.27), we have

$$\frac{\partial}{\partial t}\hat{z}(m,t) = F(m)\hat{z}(m,t) + df(w_e)\hat{z}(m,t), \tag{7.42a}$$

$$\hat{z}(x,0) = \hat{z}_0. \tag{7.42b}$$

Let us assume that the associated local system, namely

$$\frac{\partial}{\partial t}\hat{z}(m,t) = df(w_e)\hat{z}(m,t) \tag{7.43a}$$

$$\hat{z}(x,0) = \hat{z}_0, \tag{7.43b}$$

with solution $\hat{\boldsymbol{z}}(t) = e^{df(w_e)t}\hat{z}_0$ is exponentially stable, that is, $df(w_e) < 0$. Since $F \leq 0$, we further have that w_e is a linearly exponentially stable equilibrium of eq. (7.42) and based on the linear analysis no Turing bifurcation can occur in the one variable case. This is the conclusion we have observed so far for single-species RD systems on networks. Let us consider the two-variable case, where we carry out a linear stability analysis analogously, and then proceed with conclusions on the (in)stability of the nonlinear system.

7.3.3. Linear stability of reaction-diffusion with two variables

In the two-component vector case, with $\boldsymbol{w}_i(t) := w_i(\cdot,t)$, $i \in \{1,2\}$, and $\boldsymbol{w}(t) := w(\cdot,t)$ with $w(\cdot,t) = (w_1(\cdot,t), w_2(\cdot,t))^T$, the reaction-diffusion equations read

$$\frac{\partial}{\partial t}w_i(x,t) = D_i \int_0^1 W(x,y)\,(w_i(y,t) - w_i(x,t))\,dy + f_i(w(x,t)), \quad i \in \{1,2\} \tag{7.44a}$$

$$\boldsymbol{w}(0) = w_0, \tag{7.44b}$$

with $D_1, D_2 > 0$ the diffusion coefficients and with f_1, f_2 nonlinear smooth functions. We will consider the two cases $w_0 \in L^p[0,1] \times L^p[0,1] =: \mathcal{X}_p$ with $p = 2$ or ∞, using the ℓ-norms on $\mathbb{R}^2$,

$$\|x\|_{\mathcal{X}_2} = \left(\|x_1\|^2_{L^2[0,1]} + \|x_2\|^2_{L^2[0,1]}\right)^{\frac{1}{2}} \tag{7.45}$$

$$\|x\|_{\mathcal{X}_\infty} = \max\left\{\|x_1\|_{L^\infty[0,1]}, \|x_2\|_{L^\infty[0,1]}\right\}. \tag{7.46}$$

The choice of the ℓ-norm is not significant since all norms on $\mathbb{R}^n$ are equivalent. We let $\|\cdot\|_{p,p} := \|\cdot\|_{\mathcal{X}_p \to \mathcal{X}_p}$ denote the operator norm.

Existence and uniqueness of classical solutions $\boldsymbol{w} \in C^1(\mathbb{R}, \mathcal{X}_p)$, $p = 2$ or ∞, is obtained as in the scalar case, see remark 7.23 and theorems 7.25 and 7.36. Assuming a constant equilibrium $w_e = (w_{e,1}, w_{e,2})^T$, that is $f_i(w_e) = 0$ for $i \in \{1,2\}$, along the same lines as in the scalar case we linearize around w_e under the following assumption.

Assumption 7.28. For $p = 2$ or ∞, the nonlinear operator $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$ is Gâteaux differentiable in w_e with bounded Gâteaux derivative $df(w_e)$.

We obtain

$$\frac{\partial}{\partial t} z(x,t) = \mathcal{L}_{\otimes} z(x,t) + df(w_e) z(x,t), \tag{7.47a}$$

$$z(x,0) = w_0 - w_e =: z_0. \tag{7.47b}$$

where $\mathcal{L}_{\otimes} := \begin{pmatrix} D_1\mathcal{L} & 0 \\ 0 & D_2\mathcal{L} \end{pmatrix}$ and with the Gâteaux derivative in the forthcoming cases being given by

$$df(w_e) = \begin{pmatrix} \partial_1 f_1(w_e) & \partial_2 f_1(w_e) \\ \partial_1 f_2(w_e) & \partial_2 f_2(w_e) \end{pmatrix}, \tag{7.48}$$

the matrix of partial derivatives $\partial_i := \frac{\partial}{\partial w_i}$, $i \in \{1,2\}$, computed at the equilibrium. As before, via the spectral theorem and the mapping U as in eq. (7.27) we have a problem of the form

$$\frac{\partial}{\partial t} \hat{z}(m,t) = J(m) \hat{z}(m,t) \tag{7.49a}$$

$$\hat{z}(x,0) = \begin{pmatrix} U z_{0,1} \\ U z_{0,2} \end{pmatrix} =: \hat{z}_0. \tag{7.49b}$$

By assumption 7.28 and the bounded perturbation theorem, the solution of eq. (7.49) is given by the C_0-semigroup $\left(\hat{S}(t)\right)_{t \geq 0}$ such that

$$\hat{\boldsymbol{z}}(t) = \hat{S}(t) \hat{z}_0. \tag{7.50}$$

Let us introduce a higher-dimensional generalization of multiplication operators [**55**, Definition 1].

Definition 7.29. Let (X, Σ, η) be a σ-finite measure space, q be a measurable matrix-valued function and $n \in \mathbb{N}_0$. An operator $\mathcal{M}_q$ with domain $\operatorname{dom} \mathcal{M}_q$ defined on $\otimes_{k=1}^n L^p(X, \eta)$, $1 \leq p \leq \infty$, by

$$\mathcal{M}_q f(x) = q(x) f(x) \tag{7.51}$$

for all $x \in X$ and $q \in \operatorname{dom} \mathcal{M}_q = \{\otimes_{k=1}^n L^p(X,\eta) \,:\, qf \in \otimes_{k=1}^n L^p(X,\eta)\}$ is called a matrix multiplication operator.

The generator in the right-hand side of (7.49a) is thus a matrix multiplication operator $\mathcal{M}_J : \mathcal{X}_p \to \mathcal{X}_p$ defined by

$$J(m) = \begin{pmatrix} D_1 F(m) & 0 \\ 0 & D_2 F(m) \end{pmatrix} + df(w_e), \tag{7.52}$$

for all $m \in \mathbb{R}$, and $F \in L^2(\mathbb{R}, d\mu)$ where $d\mu$ is the measure provided by the spectral theorem. In order to obtain the spectrum of a multiplication operator we first need to consider the essential union of the pointwise spectra of the

associated matrix-valued functions. Using the notations of definition 7.29, we have the definition

$$\text{ess}-\bigcup_{x\in X}\sigma\left(q(x)\right):=\bigcap_{p\in[q]}\text{clo}\left(\bigcup_{x\in X}\sigma(p(x))\right), \tag{7.53}$$

where $[q]$ is the equivalence class of all measurable functions which are μ-almost everywhere equal to q. From the definition, one has the practical expression

$$\text{ess}-\bigcup_{x\in X}\sigma\left(q(x)\right)=\left\{z\in\mathbb{C}\,:\,\mu\left\{x\in X\,:\,\sigma(q(x))\cap B_\epsilon(z)\right\}>0,\,\forall\epsilon>0\right\}, \tag{7.54}$$

where $B_r(z_0)=\{z\in\mathbb{C}\,:\,|z-z_0|<\epsilon\}$. The following is well-known about bounded operators, where the operator norm is used.

Fact 7.30 (Nonvoid resolvent)**.** *Let A be a bounded operator. Then*

$$\sigma(A)\subset\left\{\lambda\in\mathbb{C}\,:\,|\lambda|<\|A\|\right\}. \tag{7.55}$$

Proof. It suffices to notice that

$$R(\lambda,A)=\frac{1}{\lambda}\left(1-\frac{1}{\lambda}A\right)^{-1}=\sum_{\ell=0}^{\infty}\frac{A^\ell}{\lambda^{\ell+1}},$$

which exists for all $|\lambda|>\|A\|$. □

As a result, we further have that

$$\varrho(A):=\sup\left\{|\lambda|\,:\,\lambda\in\sigma(A)\right\}\leq\|A\|. \tag{7.56}$$

Here ϱ is obtained by taking the supremum over the whole spectrum, and not only the eigenvalues as in chapter 6.

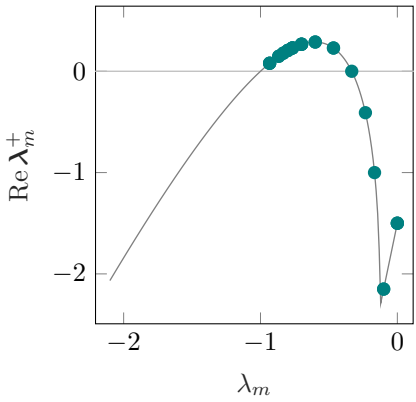
Observe that by assumption 7.28, $\mathcal{M}_J$ is bounded and thanks to fact 7.30, [**55**, Proposition 1] is applicable. We have

$$\sigma(\mathcal{M}_J)=\text{ess}-\bigcup_{m\in\mathbb{R}}\sigma\left(J(m)\right). \tag{7.57}$$

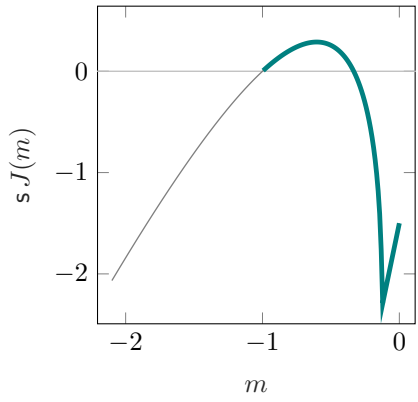
In particular, the spectral bound, definition B.12, is given by

$$\mathsf{s}(\mathcal{M}_J)=\operatorname*{ess\,sup}_{m\in\mathbb{R}}\mathsf{s}(J(m)), \tag{7.58}$$

where $\mathsf{s}(J(m))=\sup_{\lambda\in\sigma(J(m))}\operatorname{Re}\lambda$ for every $m\in\mathbb{R}$ pointwise. Finally, notice that the growth bound, definition B.11, and the spectral bound coincide because $\mathcal{M}_J$ is the generator of a uniformly continuous semigroup. An illustration is provided by fig. 7.5, which shows that just like with graphs, the linear stability of the system on a graphon depends on the location of the spectrum of the Laplacian. It is not tied to the stability of a system with the same local dynamics evolving on a continuum such as a rectangular domain.



(A) Linear growth rates on a random threshold graph with 30 nodes. Two vertices v_i, v_j are adjacent if the sum of their uniform weights is above 1.



(B) Pointwise spectral bounds of the matrices $J(m)$, eq. (7.52), for m belonging to the spectrum $[-1, 0]$ of $\mathcal{L}$ with the threshold graphon $W(x, y) = \chi_{x+y\leq 1}$.

FIGURE 7.5. Linear growth bound for the discrete two-species RD system with Laplacian normalized by the number of node, and for the associated continuum problem eq. (7.44) on the threshold graphon. The parameters for the Brusselator kinetics are those of fig. 6.1.

7.4. On the stability of the nonlinear system

The analysis would not be complete without conclusions on the nonlinear problem in two variables. Emphasis is placed on the analysis of a particular example. In section 7.4.1 we start this section by explaining the whys and wherefores, but also the limits of our choice of Brusselator kinetics. As will emerge from the discussion in section 7.4.2, we need separate treatments depending on the underlying functional space, therefore allocating sections 7.4.3 and 7.4.4 to the two cases $p = 2$ and $p = \infty$ introduced in section 7.3.3.

7.4.1. On the Brusselator

In many models the rate laws depend on the powers of the concentrations, and we will develop our analysis based on the Brusselator with features with polynomial nonlinearity. Recall the equations introduced with (6.55),

$$f_1(x, y) = 1 - (b + 1)x + cx^2y \tag{7.59}$$

$$f_2(x, y) = bx - cx^2y, \tag{7.60}$$

where $w_e = \left(1, \frac{b}{c}\right)$ satisfies $f_i(w_e) = 0$. If $c > b - 1$, $c > 0$, then w_e is a globally exponentially stable equilibrium of the linear systems on $\mathcal{X}_p$. The Gâteaux

derivative of the map $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix}$ in (x,y) is determined by

$$df(x,y) = \begin{pmatrix} -(b+1)+2cxy & cx^2 \\ b-2cxy & -cx^2 \end{pmatrix}. \tag{7.61}$$

Evaluated at the fixed point, this becomes

$$df(w_e) = \begin{pmatrix} b-1 & c \\ -b & -c \end{pmatrix}. \tag{7.62}$$

The Schnakenberg model, a variant of the Brusselator, and the Gray-Scott model which has three variables, also feature polynomial nonlinearity. Hence, our conclusions on the Brusselator apply qualitatively to those models as well. On the other hand, observe that the Gierer-Meinhardt model and the three-variable Epstein model both present negative powers of the concentrations, which makes a direct extension of the conclusions on the Brusselator not straightforward.

Remark 7.31. With the Brusselator, the function f is only locally but not globally Lipschitz continuous. One may partially overcome this problem by working with a radially truncated model [**66**, Remark 1.1] (with $\|\cdot\|$ the norm of the Banach space) :

$$f_r(w) := \begin{cases} f(w) & \text{if } \|x\| \leq r \\ f(rw/\|w\|) & \text{if } \|x\| > r \end{cases}, \tag{7.63}$$

for some $r > 0$. Then f_r is globally Lipschitz, and which is Fréchet differentiable on the open ball $B_r(0)$ if f is Fréchet differentiable.

7.4.2. On the principle of linearized stability

The critical point in applying a principle of linearized stability is the Fréchet differentiablity of the nonlinear semigroup. Indeed, consider [**2**, Theorem 2.5], that we mention in a form that suits our interest.

THEOREM 7.32. *Let w_e be an equilibrium for* (7.44) *on a Banach space, and suppose that the associated nonlinear semigroup $S(t)$ is Fréchet differentiable at w_e, with Fréchet derivative T_{w_e}. If w_e is a globally exponentially stable equilibrium of the linear system given by T_{w_e}, then w_e is a locally exponentially stable equilibrium for* (7.44). *Conversly, if w_e is an unstable equilibrium of the linear system, then w_e is an unstable equilibrium for* (7.44).

When $p = \infty$, we will use a strong condition on the generator of the nonlinear semigroup, namely continuous Fréchet differentiability, that ensures the Fréchet differentiability of the nonlinear semigroup. But we will also exploit the fact that the assumption on the generator can be weakened. Indeed, even if the generator of the nonlinear semigroup is not Fréchet differentiable, the nonlinear semigroup may still be differentiable. This helps us treat the case $p = 2$ in a second step, where only the weaker assumptions are satisfied.

Remark 7.33. The requirement that the fixed point w_e of the linear system is exponentially (un)stable and not asymptotically (un)stable is critical, these two notions of stability being distinct.

7.4.3. The case $p = \infty$

When the reaction-diffusion system is defined on $\mathcal{X}_\infty$, we can show that the nonlinear part f of the generator is continuously Fréchet differentiable on $\mathcal{X}_\infty$ [**66**, Definition 1.1]. This means that f is Fréchet differentiable and further, there exists a continuous increasing function $d : [0, \infty) \to [0, \infty)$ such that

$$\|Df(u) - Df(v)\|_{\infty,\infty} \leq d(r)\|u - v\|_{\mathcal{X}_\infty}, \tag{7.64}$$

for all $u, v \in \mathcal{X}_\infty$ such that $\|u\|_{\mathcal{X}_\infty}, \|v\|_{\mathcal{X}_\infty} \leq r$. Then, we have the following theorem which is a particular case of [**1**, Theorem 3.6].

THEOREM 7.34. *Consider* (7.44) *on a Banach space where f is continuously Fréchet differentiable and let w_e be a globally exponentially stable equilibrium of the linear system* (7.47)*. Then w_e is a locally exponentially stable equilibrium of* (7.44)*. Conversly, if w_e is exponentially unstable in the linear system* (7.47)*, then w_e is an exponentially unstable equilibrium of the nonlinear system* (7.44)*.*

Example 7.35 (Brusselator – Continuous Fréchet differentiability when $p = \infty$)**.** Let us show that $f : \mathcal{X}_\infty \mapsto \mathcal{X}_\infty$, with

$$f(z_1, z_2) = \begin{pmatrix} 1 - (b+1)z_1 + cz_1^2 z_2 \\ bz_1 - cz_1^2 z_2 \end{pmatrix} \tag{7.65}$$

and $c > 0$ is Fréchet differentiable in $z = (z_1, z_2)^T$. A calculation using the Gâteaux derivative (7.61) yields

$$f(z + h) - f(z) - df(z)h = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \tag{7.66}$$

where $h = (h_1, h_2)^T$ and

$$\delta_1 = ch_1(2z_1h_2 + z_2h_1 + h_1h_2) = -\delta_2. \tag{7.67}$$

Hence, writing $\lesssim$ for inequality up to a (positive) constant,

$$\begin{aligned} \left\| \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\|_{\mathcal{X}_\infty} &= \max\{\|\delta_1\|_\infty, \|\delta_2\|_\infty\} \\ &\lesssim \|h_1(2z_1h_2 + z_2h_1 + h_1h_2)\|_\infty \\ &\leq \|z_1\|_\infty\|h_1\|_\infty\|h_2\|_\infty + \|z_2\|_\infty\|h_1\|_\infty^2 + \|h_1\|_\infty^2\|h_2\|_\infty \\ &\leq (\|z_1\|_\infty + \|z_2\|_\infty + \|h_1\|_\infty)\,\|h\|_{\mathcal{X}_\infty}^2, \end{aligned} \tag{7.68}$$

and so

$$\lim_{\|h\|_{\mathcal{X}_\infty} \to 0} \frac{\left\| \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\|_{\mathcal{X}_\infty}}{\|h\|_{\mathcal{X}_\infty}} = 0. \tag{7.69}$$

Let us now show the continuity property. Recall $\|\cdot\|_{\infty,\infty}$ denotes the operator norm $\|\cdot\|_{\mathcal{X}_\infty \to \mathcal{X}_\infty}$. For compactness, we adopt the notation $\|\cdot\|_p$ for $\|\cdot\|_{L^p[0,1]}$.

Fix $r > 0$ and let $u, v \in \mathcal{X}_\infty$ satisfy $\|u\|_{\mathcal{X}_\infty}, \|v\|_{\mathcal{X}_\infty} \leq r$. We compute

$$\begin{aligned}
\|df(u) - df(v)\|_{\infty,\infty} &= \sup_{\|h\|_{\mathcal{X}_\infty}=1} \|df(u)h - df(v)h\|_{\mathcal{X}_\infty} \\
&= \sup_{\|h\|_{\mathcal{X}_\infty}=1} \left\| \begin{pmatrix} 2c(u_1u_2 - v_1v_2)h_1 + c(u_1^2 - v_1^2)h_2 \\ -2c(u_1u_2 - v_1v_2)h_1 - c(u_1^2 - v_1^2)h_2 \end{pmatrix} \right\|_{\mathcal{X}_\infty} \\
&\lesssim \sup_{\|h\|_{\mathcal{X}_\infty}=1} \left(2\|u_1u_2 - v_1v_2\|_\infty \|h_1\|_\infty + \|u_1^2 - v_1^2\|_\infty \|h_2\|_\infty\right) \\
&\leq 2\|u_1 - v_1\|_\infty \|u_2\|_\infty + 2\|v_1\|_\infty \|u_2 - v_2\|_\infty \\
&\qquad + \|u_1 + v_1\|_\infty \|u_1 - v_1\|_\infty \\
&\leq \left(2\|u_2\|_\infty + 2\|v_1\|_\infty + \|u_1 + v_1\|_\infty\right) \|u - v\|_{\mathcal{X}_\infty} \\
&\leq 6r\|u - v\|_{\mathcal{X}_\infty},
\end{aligned}$$

and theorem 7.34 allows to use the principle of linearized stability.

Continous Fréchet differentiablity of f actually ensures that the nonlinear semigroup is everywhere Fréchet differentiable. We have stated in section 7.4.2 that the generator may fail to be Fréchet differentiable while the nonlinear semigroup is still Fréchet differentiable. We then give a sufficient condition based on the Gâteaux differentiability of the generator, which will also allow to address the case $p = 2$ in section 7.4.4.

THEOREM 7.36 (Existence, uniqueness and a priori estimate on $\mathcal{X}_\infty$). *Let $\boldsymbol{w}$ denote the solution of* (7.49) *with initial condition $\boldsymbol{w}(0) = w_0 \in \mathcal{X}_\infty$. Assume $f : \mathcal{X}_\infty \to \mathcal{X}_\infty$ is Lipschitz continuous, and let w_e denote a critical point of f. Then $\boldsymbol{w} \in C(\mathbb{R}, \mathcal{X}_\infty)$ and for any $T > 0$, there exists $C > 0$ depending on T and f but not on W such that*

$$\|\boldsymbol{w} - w_e\|_{L^\infty([0,T],\mathcal{X}_\infty)} \leq C\|\boldsymbol{w}(0) - w_e\|_{\mathcal{X}_\infty}. \tag{7.70}$$

PROOF. For $\tau > 0$, consider $\mathbb{M}_p := C([0, \tau], \mathcal{X}_p)$ with $p = \infty$, and consider the operator K defined on $\mathbb{M}_\infty$ by

$$(K\boldsymbol{w})(t) = w_0 + \int_0^t \left(\mathcal{L}_\otimes w(\cdot, s) + f(w(\cdot, s))\right) ds, \tag{7.71}$$

Then, by a standard calculation similar to [**63**, Theorem 3.1], one can show K is a contraction on $\mathbb{M}_\infty$ for a proper choice of τ. Indeed, for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{M}_\infty$, we

have

$$\begin{aligned}\|K\boldsymbol{u}-K\boldsymbol{v}\|_{\mathbb{M}_\infty} &= \sup_{t\in[0,\tau]}\|K\boldsymbol{u}(t)-K\boldsymbol{v}(t)\|_{\mathcal{X}_\infty}\\ &\le \sup_{t\in[0,\tau]}\left\{\int_0^t \|\mathcal{L}_\otimes(\boldsymbol{u}(s)-\boldsymbol{v}(s))\|_{\mathcal{X}_\infty}ds\right.\\ &\qquad\left.+\int_0^t\|f(\boldsymbol{u}(s))-f(\boldsymbol{v}(s))\|_{\mathcal{X}_\infty}ds\right\}\\ &\le \sup_{t\in[0,\tau]}\left\{(D_1\vee D_2)\|\mathcal{L}\|_{\infty,\infty}\int_0^t\|\boldsymbol{u}(s)-\boldsymbol{v}(s)\|_{\mathcal{X}_\infty}ds\right.\\ &\qquad\left.+L\int_0^t\|\boldsymbol{u}(s)-\boldsymbol{v}(s)\|_{\mathcal{X}_\infty}ds\right\}\\ &\le ((D_1\vee D_2)+L)\,\tau\|\boldsymbol{u}-\boldsymbol{v}\|_{\mathbb{M}_\infty},\end{aligned}$$

where we have used the Minkowski inequality, $\|\mathcal{L}\|_{\infty,\infty}\le 1$, where L is the Lipschitz constant of f, and $a\vee b$ denotes the maximum between a and b. By choosing $\tau=\frac{1}{2}\left((D_1\vee D_2)+L\right)$, this shows that

$$\|K\boldsymbol{u}-K\boldsymbol{v}\|_{\mathbb{M}_\infty}\le\frac{1}{2}\|\boldsymbol{u}-\boldsymbol{v}\|_{\mathbb{M}_\infty}. \tag{7.72}$$

Further, we have that $K\mathbb{M}_\infty\subset\mathbb{M}_\infty$, since (7.72) implies

$$\|K\boldsymbol{u}\|_{\mathbb{M}_\infty}\le\frac{1}{2}\|\boldsymbol{u}\|_{\mathbb{M}_\infty}+\|K0\|_{\mathbb{M}_\infty},$$

where $\|K0\|_{\mathbb{M}_\infty}=\|w_0+tf(0)\|_{\mathbb{M}_\infty}<\infty$. Following the reasoning of [**63**, Theorem 3.1], the unique solution of the IVP on $[0,\tau]$ can be extended to a unique classical solution in $C(\mathbb{R},L^\infty[0,1])$. To prove (7.70), let us conveniently consider the translated system, with $\widetilde{\boldsymbol{w}}:=\boldsymbol{w}-w_e$:

$$\dot{\widetilde{\boldsymbol{w}}}=\mathcal{L}_\otimes\widetilde{\boldsymbol{w}}+f(\widetilde{\boldsymbol{w}}+w_e) \tag{7.73a}$$

$$\widetilde{w}(0)=\widetilde{w}_0:=w_0-w_e. \tag{7.73b}$$

From (7.73a), for all $0\le t\le T$ using again the Lipschitz continuity of f and the Minkowski inequality we have

$$\begin{aligned}\frac{d}{dt}\|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty}&\le\|\mathcal{L}_\otimes\widetilde{\boldsymbol{w}}+f(\widetilde{\boldsymbol{w}}+w_e)\|_{\mathcal{X}_\infty}\\ &\le(D_1\vee D_2)\|\mathcal{L}\|_{\infty,\infty}\|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty}+L\|\widetilde{\boldsymbol{w}}\|_{\mathcal{X}_\infty},\end{aligned}$$

and by the Gronwall inequality, with $\beta:=(D_1\vee D_2)\|\mathcal{L}\|_{\infty,\infty}+L>0$,

$$\|\widetilde{\boldsymbol{w}}\|_{\mathcal{X}_\infty}\le e^{\beta t}\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\le e^{\beta T}\|\widetilde{w}_0\|_{\mathcal{X}_\infty}. \tag{7.74}$$

Since this applies for all $t\le T$, (7.70) follows. □

Remark 7.37. Another way to obtain existence and uniqueness of a classical solution is based on generation theorems for nonlinear semigroups.

Now we formulate a remark that eventually guarantees the formulation of the next lemma makes sense.

Remark 7.38 (Trajectories on finite time intervals)**.** Under its assumptions, the previous theorem shows that for all $T > 0$, the solution $\widetilde{\boldsymbol{w}}$ of (7.73) is such that

$$\begin{aligned}\|f(\widetilde{\boldsymbol{w}} + w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_\infty)} &\leq \Big(L + \|df(w_e)\|_{\infty,\infty}\Big) \sup_{t\in[0,T]} \|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty} \\ &\leq \Big(L + \|df(w_e)\|_{\infty,\infty}\Big) \|\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_\infty)}\end{aligned}$$

as pointed out with a similar argument in [**51**, Lemma 2.2].

Lemma 7.39. *Let $\widetilde{\boldsymbol{w}}$ denote the solution of* (7.73)*, and let $(\widetilde{S}(t))_{t\geq 0}$ be the nonlinear semigroup such that $\widetilde{\boldsymbol{w}}(t) = \widetilde{S}(t)\widetilde{w}_0$. Assume further there exists $w_e \in \mathcal{X}_\infty$ such that $f(w_e) = 0$. If for all $T > 0$*

$$\lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\to 0} \frac{\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_\infty)}}{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}} = 0, \tag{7.75}$$

then the nonlinear semigroup $(\widetilde{S}(t))_{t\geq 0}$ is Fréchet differentiable in 0*, and its Fréchet derivative is the strongly continuous semigroup $(\widetilde{T}(t))_{t\geq 0}$ generated by $\mathcal{L}_\otimes + df(w_e)$, where $df(w_e)$ is the Gâteaux derivative of $\mathcal{L}_\otimes + f(\cdot + w_e)$.*

The proof is an extension of the proof of [**51**, Lemma 3.1] from a Hilbert space setting to a Banach space setting. Details are given for completeness.

Proof. Let $T > 0$ and define

$$\phi(t) := \widetilde{S}(t)\widetilde{w}_0 - \widetilde{T}(t)\widetilde{w}_0.$$

The first step is to bound $\|\phi(t)\|_{\mathcal{X}_\infty}$. Using the fact that $f(w_e) = 0$,

$$\begin{aligned}\dot{\phi}(t) &= \mathcal{L}_\otimes \widetilde{S}(t)\widetilde{w}_0 + f(\widetilde{S}(t)\widetilde{w}_0 + w_e) - \mathcal{L}_\otimes \widetilde{T}(t)\widetilde{w}_0 - df(w_e)\widetilde{T}(t) \\ &= \mathcal{L}_\otimes \phi(t) + df(w_e)\Big(\widetilde{\boldsymbol{w}}(t) - \widetilde{T}(t)\widetilde{w}_0\Big) - df(w_e)\widetilde{\boldsymbol{w}}(t) + f(\widetilde{\boldsymbol{w}}(t) + w_e) \\ &= \mathcal{L}_\otimes \phi(t) + df(w_e)\phi(t) + R(\widetilde{\boldsymbol{w}}(t), w_e)\end{aligned}$$

where

$$R(\widetilde{\boldsymbol{w}}(t), w_e) := f(\widetilde{\boldsymbol{w}}(t) + w_e) - df(w_e)\widetilde{\boldsymbol{w}}(t).$$

Using the triangle inequality and the continuity of the norm to obtain the first inequality,

$$\begin{aligned}\frac{d}{dt}\|\phi(t)\|_{\mathcal{X}_\infty} &\leq \|\dot{\phi}(t)\|_{\mathcal{X}_\infty} \\ &\leq \|\mathcal{L}_\otimes \phi(t)\|_{\mathcal{X}_\infty} + \|df(w_e)\|_{\infty,\infty}\|\phi(t)\|_{\mathcal{X}_\infty} + \|R(\widetilde{\boldsymbol{w}}(t), w_e)\|_{\mathcal{X}_\infty} \\ &= k\|\phi(t)\|_{\mathcal{X}_\infty} + \|R(\widetilde{\boldsymbol{w}}(t), w_e)\|_{\mathcal{X}_\infty}\end{aligned}$$

where $k := \|\mathcal{L}_\otimes\|_{\infty,\infty} + \|df(w_e)\|_{\infty,\infty}$. Multiplying by an integrating factor e^{-kt} we come to

$$\frac{d}{dt}\|\phi(t)\|_{\mathcal{X}_\infty} e^{-kt} \leq \|R(\widetilde{\boldsymbol{w}}(t), w_e)\|_{\mathcal{X}_\infty} e^{-kt}$$

and taking into account $\|\phi(0)\|_{\mathcal{X}_\infty} = 0$, after integration we obtain

$$\begin{aligned}\|\phi(t)\|_{\mathcal{X}_\infty} &= e^{kt}\int_0^t e^{-ks}\|R(\widetilde{\boldsymbol{w}}(s), w_e)\|_{\mathcal{X}_\infty} ds \\ &\leq Te^{kT}\sup_{t\in[0,T]}\|R(\widetilde{\boldsymbol{w}}(s), w_e)\|_{\mathcal{X}_\infty} ds \\ &= Te^{kT}\|R(\widetilde{\boldsymbol{w}}(s), w_e)\|_{L^\infty([0,T],\mathcal{X}_\infty)}.\end{aligned}$$

Finally, relying on the hypothesis (7.75), we have

$$\lim_{\|h\|_{\mathcal{X}_\infty}\to 0}\frac{\|\phi(t)\|_{\mathcal{X}_\infty}}{\|h\|_{\mathcal{X}_\infty}} \leq \frac{Te^{kT}\|R(\widetilde{\boldsymbol{w}}(s), w_e)\|_{L^\infty([0,T],\mathcal{X}_\infty)}}{\|h\|_{\mathcal{X}_\infty}} = 0, \tag{7.76}$$

which yields the conclusion. □

Finally, we formulate the following theorem.

THEOREM 7.40. *Under our working assumptions, and assuming* (7.75) *applies, if w_e is a globally exponentially stable equilibrium of the linear problem* (7.47) *on $\mathcal{X}_\infty$, then it is a locally-exponentially stable equilibrium of the nonlinear problem* (7.44)*. Conversely, if w_e is unstable in the linear system, it is also unstable in the nonlinear system.*

PROOF. It suffices to consider the stability of $\widetilde{\boldsymbol{w}} \equiv 0$ in the translated system (7.73), and to combine lemma 7.39 and [**51**, Theorem 3.1] where we take $X = Y = \mathcal{X}_\infty$. □

We illustrate this approach with our chosen example, noting that computation in the present case is partially reused in section 7.4.4 when $p = 2$.

Example 7.41 (Brusselator – Condition on the generator for Fréchet differentiability of the semigroup)**.** Using the expression for f as given by (7.65), for $T > 0$ we have

$$\begin{aligned}&\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_\infty)} \\ &\qquad= \sup_{t\in[0,T]}\left\|\begin{pmatrix} c\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t) + b\widetilde{\boldsymbol{w}}_1^2(t) + 2c\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t) \\ -c\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t) - b\widetilde{\boldsymbol{w}}_1^2(t) - 2c\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t)\end{pmatrix}\right\|_{\mathcal{X}_\infty} \\ &\qquad\lesssim \sup_{t\in[0,T]}\left\{\|\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t)\|_\infty + \|\widetilde{\boldsymbol{w}}_1^2(t)\|_\infty + \|\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t)\|_\infty\right\} \\ &\qquad= \sup_{t\in[0,T]}\Big\{\|\widetilde{\boldsymbol{w}}_1(t)\|_\infty^2\|\widetilde{\boldsymbol{w}}_2(t)\|_\infty \\ &\qquad\qquad + \|\widetilde{\boldsymbol{w}}_1(t)\|_\infty^2 + \|\widetilde{\boldsymbol{w}}_1(t)\|_\infty\|\widetilde{\boldsymbol{w}}_2(t)\|_\infty\Big\} \\ &\qquad\leq \sup_{t\in[0,T]}\left\{\|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty}^3 + 2\|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty}^2\right\} \\ &\qquad\lesssim \|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty}^3 + 2\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty}^2,\end{aligned} \tag{7.77}$$

where the last inequality follows from (7.70) and the continuity of the orbits $t \mapsto \widetilde{\boldsymbol{w}}(t)$. It follows that

$$\lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\to 0} \frac{\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_\infty)}}{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}} \lesssim \lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\to 0} \|\widetilde{\boldsymbol{w}}(0)\|^2_{\mathcal{X}_\infty} + 2\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty} = 0, \tag{7.78}$$

showing that lemma 7.39 is applicable.

7.4.4. The case $p = 2$

The reason for considering the Hilbert space setting separately is that even with our simple working example, the nonlinear operator f fails to be Fréchet differentiable.

Example 7.42 (Brusselator - Nowhere Fréchet differentiability when $p = 2$)**.** Consider

$$\|f(z+h) - f(z) - df(z)h\|_{\mathcal{X}_2} = \left\| \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\|_{\mathcal{X}_2} \tag{7.79}$$

where δ_1, δ_2 are given by eq. (7.67). For $h \in \mathcal{X}_2$, a calculation yields

$$\frac{\left\| \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} \right\|^2_{\mathcal{X}_2}}{\|h\|^2_{\mathcal{X}_2}} \gtrsim \frac{\int_0^1 h_1^4(x) h_2(x)^2 dx}{\int_0^1 (h_1^2(x) + h_2^2(x)) dx}. \tag{7.80}$$

Consider the sequence $(g_n)_{n\in\mathbb{N}}$ where for every n, $g_n \in L^2[0,1]$ is defined by

$$g_n(x) = \begin{cases} n^{\frac{\alpha}{2}} & \text{if } x < \frac{1}{n}, \\ \left(\frac{1}{2}\right)^{\frac{n}{2}} & \text{if } x \geq \frac{1}{n}, \end{cases} \tag{7.81}$$

for some $0 \leq \alpha < 1$. We have

$$\frac{\int_0^1 g_n^6(x) dx}{\int_0^1 g_n^2(x) dx} = \frac{n^{3\alpha-1} + \frac{n-1}{n}(\frac{1}{2})^{3n}}{n^{\alpha-1} + \frac{n-1}{n}(\frac{1}{2})^{n}}. \tag{7.82}$$

Hence, if we choose $h_1 = h_2 = g_n$ and α such that $\frac{1}{3} \leq \alpha < 1$, then we see in the limit of $n \to \infty$, namely $\|h\|_{\mathcal{X}_2} \to 0$, that the left-hand side of (7.80) does not vanish, meaning f is not Fréchet differentiable in the arbitrarily chosen $z \in \mathcal{X}_2$.

In order to pursue the analysis, for the completeness we first prove existence and unicity of the solution the IVP on $\mathcal{X}_2$ in theorem 7.43, in a way similar to the case $p = \infty$. Secondly, we obtain an a priori estimate on the norm of the solution of the translated problem, as part of the same theorem. Thirdly, we give the definitions of (Y, X)-Fréchet differentiability and (Y, X)-exponential stability introduced by [**51**]. And finally, we show based on our example that with an appropriate choice of the spaces, namely with $Y = \mathcal{X}_\infty$ and $X = \mathcal{X}_2$, within this concept of stability the principle of linearized stability is preserved.

THEOREM 7.43 (Existence, uniqueness and a priori estimate on $\mathcal{X}_2$). *Let $\boldsymbol{w}$ denote the solution of (7.49) with initial condition $\boldsymbol{w}(0) = w_0 \in \mathcal{X}_2$. Assume $f : \mathcal{X}_2 \to \mathcal{X}_2$ is Lipschitz continuous, and that there exists a fixed point w_e such that $f(w_e) = 0$. Then $\boldsymbol{w} \in C(\mathbb{R}, \mathcal{X}_2)$ and for any $T > 0$, there exists $C > 0$ depending on T and f but not on W such that*

$$\|\boldsymbol{w} - w_e\|_{L^\infty([0,T],\mathcal{X}_2)} \leq C\|\boldsymbol{w}(0) - w_e\|_{\mathcal{X}_2}. \tag{7.83}$$

Again, the existence and uniqueness part of the proof is similar to the proof of [**63**, Theorem 3.1].

PROOF. For $\tau > 0$, consider $\mathbb{M}_p$ with $p = 2$ and the operator K defined on $\mathbb{M}_2$ by (7.71). We again need to show K is a contraction. First, notice that due to the boundedness of $\mathcal{L} : L^2[0,1] \to L^2[0,1]$, for all $x, y \in \mathcal{X}_2$ we have

$$\begin{aligned}
\|\mathcal{L}_\otimes(x-y)\|_{\mathcal{X}_2} &\leq (D_1 \vee D_2) \left\| \begin{pmatrix} \mathcal{L}(x_1 - y_1) \\ \mathcal{L}(x_2 - y_2) \end{pmatrix} \right\|_{\mathcal{X}_2} \\
&\leq (D_1 \vee D_2) \left(\|\mathcal{L}\|_{2,2}^2 \left(\|x_1 - y_1\|_2^2 + \|x_2 - y_2\|_2^2 \right) \right)^{\frac{1}{2}} \\
&\leq (D_1 \vee D_2) \|\mathcal{L}\|_{2,2} \|x - y\|_{\mathcal{X}_2}. \qquad (7.84)
\end{aligned}$$

Using (7.84) with $\|\mathcal{L}\|_{2,2} \leq 1$ and the Lipschitz continuity of f, one obtains

$$\begin{aligned}
\|K\boldsymbol{u} - K\boldsymbol{v}\|_{\mathbb{M}_2} &\leq \sup_{t\in[0,\tau]} \left\{ \int_0^t \|\mathcal{L}_\otimes (\boldsymbol{u}(s) - \boldsymbol{v}(s))\|_{\mathcal{X}_2} ds \right. \\
&\qquad \left. + \int_0^t \|f(\boldsymbol{u}(s)) - f(\boldsymbol{v}(s))\|_{\mathcal{X}_2} ds \right\} \\
&\leq \sup_{t\in[0,\tau]} \left\{ (D_1 \vee D_2) \|\mathcal{L}\|_{2,2} \int_0^t \|\boldsymbol{u}(s) - \boldsymbol{v}(s)\|_{\mathcal{X}_2} ds \right. \\
&\qquad \left. + L \int_0^t \|\boldsymbol{u}(s) - \boldsymbol{v}(s)\|_{\mathcal{X}_2} ds \right\} \\
&\leq ((D_1 \vee D_2) + L)\, \tau \|\boldsymbol{u} - \boldsymbol{v}\|_{\mathbb{M}_2},
\end{aligned}$$

and by choosing τ as in the proof of theorem 7.36 and by the same argument, existence and uniqueness of a solution in $C(\mathbb{R}, \mathcal{X}_2)$ follows. In order to show (7.83), let $\widetilde{\boldsymbol{w}}$ be the solution of the translated problem (7.73) with initial condition $\widetilde{w}_0 \in \mathcal{X}_2$. Notice that $\mathcal{L}_\otimes$ is dissipative, since

$$(\mathcal{L}_\otimes x, x)_{\mathcal{X}_2} = \sum_{m=1}^{2} D_m \left(\mathcal{L}x_m, x_m\right)_{L^2[0,1]} \leq 0 \tag{7.85}$$

for all $x \in \mathcal{X}_2$, and fix $T > 0$. Hence, dropping the explicit reference to the space in the notation of the inner products, for all $0 \le t \le T$ we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\widetilde{\boldsymbol{w}}(t)\|^2_{\mathcal{X}_2} &= \left(\dot{\widetilde{\boldsymbol{w}}}(t), \widetilde{\boldsymbol{w}}(t)\right) \\ &\le (\mathcal{L}_\otimes \widetilde{\boldsymbol{w}}(t), \widetilde{\boldsymbol{w}}(t)) + (f(\widetilde{\boldsymbol{w}}(t) + w_e), \widetilde{\boldsymbol{w}}(t)) \\ &\le \|f(\widetilde{\boldsymbol{w}}(t) + w_e)\|_{\mathcal{X}_2} \|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_2} \\ &\le L\|\widetilde{\boldsymbol{w}}(t)\|^2_{\mathcal{X}_2},\end{aligned}$$

where we used the Cauchy-Schwarz inequality and the Lipschitz continuity of f. By the Gronwall inequality

$$\|\widetilde{\boldsymbol{w}}(t)\|^2_{\mathcal{X}_2} \le e^{2Lt}\|\widetilde{\boldsymbol{w}}(0)\|^2_{\mathcal{X}_2} \le e^{2LT}\|\widetilde{\boldsymbol{w}}(0)\|^2_{\mathcal{X}_2}, \tag{7.86}$$

which implies

$$\|\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_2)} \le e^{LT}\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_2}.$$

□

Let us introduce weaker, local versions of Fréchet differentiability [**51**, Definition 2.1] and stability [**51**, Definitions 2.3 and 2.4].

Definition 7.44. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces and let $F : X \to X$ be a nonlinear operator, such that $\operatorname{dom} F \subset Y \subseteq X$. Then F is (Y, X)-Fréchet differentiable at z if there exists a bounded linear operator $df(z) : X \to X$ such that for all $h \in \operatorname{dom} f$,

$$\lim_{\|h\|_Y \to 0} \frac{\|F(z+h) - F(z) - df(z)h\|_X}{\|h\|_X} = 0 \tag{7.87}$$

Definition 7.45. The equilibrium z_e of the nonlinear system (5.1) is

(a) (Y, X)-locally stable if for all $\epsilon > 0$, there exists $\delta > 0$ such that if $\|z_0 - z_e\|_Y < \delta$, then $\|z(t) - z_e\|_X < \epsilon$ for all $t \ge 0$.
(b) (Y, X)-locally exponentially stable if there exists $\delta, \alpha, \beta > 0$ such that if $\|z_0 - z_e\|_Y < \delta$, then $\|z(t) - z_e\|_X \le \alpha e^{-\beta t}\|z_0 - z_e\|_X$ for all $t \ge 0$.

We are now in a position to formulate the Hilbert-space variants of lemma 7.39 and theorem 7.40, in their (Y, X)-local versions, resembling [**51**].

Lemma 7.46. *Let $\widetilde{\boldsymbol{w}}$ denote the solution of (7.73), and let $(\widetilde{S}(t))_{t\ge 0}$ be the nonlinear semigroup such that $\widetilde{\boldsymbol{w}}(t) = \widetilde{S}(t)\widetilde{w}_0$. Assume further there exists $w_e \in \mathcal{X}_2$ such that $f(w_e) = 0$. If for all $T > 0$*

$$\lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty} \to 0} \frac{\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_2)}}{\|\widetilde{w}_0\|_{\mathcal{X}_2}} = 0, \tag{7.88}$$

then the nonlinear semigroup $(\widetilde{S}(t))_{t\ge 0}$ is $(\mathcal{X}_\infty, \mathcal{X}_2)$-Fréchet differentiable in 0, and its Fréchet derivative is the strongly continuous semigroup $(\widetilde{T}(t))_{t\ge 0}$ generated by $\mathcal{L}_\otimes + df(w_e)$, where $df(w_e)$ is the Gâteaux derivative of $\mathcal{L}_\otimes + f(\cdot + w_e)$.

Proof. See the proof of [**51**, Lemma 3.1]. □

THEOREM 7.47. *Under our working assumptions, and assuming* (7.88) *applies, if w_e which satisfies $f(w_e) = 0$ is a globally exponentially stable equilibrium of the linear problem* (7.47) *on $\mathcal{X}_2$, then it is a $(\mathcal{X}_\infty, \mathcal{X}_2)$-locally-exponentially stable equilibrium of the nonlinear problem* (7.44)*. Conversely, if w_e is unstable in the linear system, it is also unstable in the nonlinear system.*

PROOF. See the proof of theorem 7.40. □

Example 7.48 (Brusselator – Local $(\mathcal{X}_\infty, \mathcal{X}_2)$ principle of linearized stability)**.** Starting from eq. (7.77) we have

(7.89)

$$
\begin{aligned}
&\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_2)} \\
&\qquad = \sup_{t\in[0,T]} \left\| \begin{pmatrix} c\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t) + b\widetilde{\boldsymbol{w}}_1^2(t) + 2c\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t) \\ -c\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t) - b\widetilde{\boldsymbol{w}}_1^2(t) - 2c\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t) \end{pmatrix} \right\|_{\mathcal{X}_2} \\
&\qquad = \sqrt{2} \sup_{t\in[0,T]} \left\| c\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t) + b\widetilde{\boldsymbol{w}}_1^2(t) + 2c\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t) \right\|_2 \\
&\qquad \lesssim \sup_{t\in[0,T]} \left\{ \|\widetilde{\boldsymbol{w}}_1^2(t)\widetilde{\boldsymbol{w}}_2(t)\|_2 + \|\widetilde{\boldsymbol{w}}_1^2(t)\|_2 + \|\widetilde{\boldsymbol{w}}_1(t)\widetilde{\boldsymbol{w}}_2(t)\|_2 \right\} \\
&\qquad \leq \sup_{t\in[0,T]} \left\{ \left(\|\widetilde{\boldsymbol{w}}_1(t)\|_\infty^2 + 2\|\widetilde{\boldsymbol{w}}_1(t)\|_\infty \right) \|\widetilde{\boldsymbol{w}}_2(t)\|_2 \right\} \\
&\qquad \leq \sup_{t\in[0,T]} \left\{ \|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty}^2 + 2\|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_\infty} \right\} \sup_{t\in[0,T]} \|\widetilde{\boldsymbol{w}}(t)\|_{\mathcal{X}_2} \\
&\qquad \lesssim \left(\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty}^2 + 2\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty} \right) \|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_2}^2,
\end{aligned}
$$

where the last inequality follows from eqs. (7.70) and (7.83) It follows that

$$
\begin{aligned}
\lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\to 0} \frac{\|f(\widetilde{\boldsymbol{w}} - w_e) - df(w_e)\widetilde{\boldsymbol{w}}\|_{L^\infty([0,T],\mathcal{X}_2)}}{\|\widetilde{w}_0\|_{\mathcal{X}_2}} \\
\lesssim \lim_{\|\widetilde{w}_0\|_{\mathcal{X}_\infty}\to 0} \|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty}^2 + 2\|\widetilde{\boldsymbol{w}}(0)\|_{\mathcal{X}_\infty} = 0,
\end{aligned}
$$

which guarantees $(\mathcal{X}_\infty, \mathcal{X}_2)$-Fréchet differentiability of the nonlinear semigroup. Observe that we have only verified the local condition on f, and (7.89) was not shown to exist when $\|\widetilde{w}_0\|_{\mathcal{X}_2} \to 0$.

7.5. Conclusion

The whole concept of diffusion-driven instability was revisited in this graphon-flavored chapter, as dictated by the third research objective. Going all the way back to the definition of the combinatorial Laplacian, we successively discussed our assumptions on connectedness and density, the heat equation, and finally scalar and vector semilinear equations on graphons. The aim was to transpose the usual spectral methods on graphs to the analysis of the continuum-limit version of the equations.

The spectral theorem proved well-suited to our needs within the convenient framework of dense symmetric graphs. This is both a limitation and an opportunity, in the sense that there is little doubt about the pertinence of generalizing the results. Extending the study could go along at least two directions. With the first one, we would seek to answer the type of questions about stability without the density or symmetry assumptions on the graphs. With the second one, we rather consider this chapter to be a starting point to address different types of problems, such as synchronization.

Overall, the shift from graphs to graphons goes with two difficulties. First, the inherent approximation procedure needs to be validated and second, the analysis of the continuum model bears new challenges, such as technicalities around the application of the principle of linear stability. We have applied recent works such as [**51**] to overcome that problem and draw conclusions in the context of our main example featuring polynomial nonlinearity. But overall, a graph-limit version of all main aspects of the theory of RD equations on graphs, including the topics discussed in the previous chapters, must still be developed.

Further research opportunities are numerous, and some suggestions are outlined in the following and final chapter.

CHAPTER 8

Conclusions and perspectives

8.1. A look back

Having reached the point of drawing conclusions, we will not abide by the linear-nonlinear dichotomy that implies a classification along the lines of chapters 2 to 4 *vs* chapters 5 to 7. We will rather present the contributions and perspectives according to the research questions listed in section 1.4 of the introduction. As fig. 8.1 shows, to answer these questions we went through the modeling and analysis of diffusion processes on temporal graphs; we discussed bifurcations and stability in reaction-diffusion systems; we established the validity and analyzed the continuum limit of two key evolution equations. The common ground between these topics is graphs, viewed as the support of a dynamical system. Graphs are everywhere, in this thesis but also in real life, anchoring our work with a sense of concreteness.

In what comes next, sections 8.1.1 to 8.1.3, we will summarize our contributions, raise questions and perspectives, and make concluding remarks about each of the research questions. Eventually in section 8.2, we shape a limited amount of possible follow-ups, this time with more freedom with respect to the contents of the individual chapters, before writing the final note.

8.1.1. On the first question

The rationale behind the first question was that dynamical graphs are well known to affect perceptibly the modeling, the analysis and the properties of diffusion processes. The same holds true for other variants of dynamical systems on networks, at least when diffusion is involved. Therefore, it was critical before attempting to address more complex problems to gain a firm understanding of the microscopic models of diffusion. This includes the scenario when multiple timescales coexist in the model, even if this case is largely neglected in the literature.

8.1.1.1. *Summary of the findings*

That models with multiple timescales are sparse in the literature is the reason why at the start of the thesis we have reviewed, and then complemented the taxonomy of random walk processes on graphs. We have added to the well-known active walks, appropriate for human or animal trajectories, and passive walks typically used for virus or information spreading, new combinations of

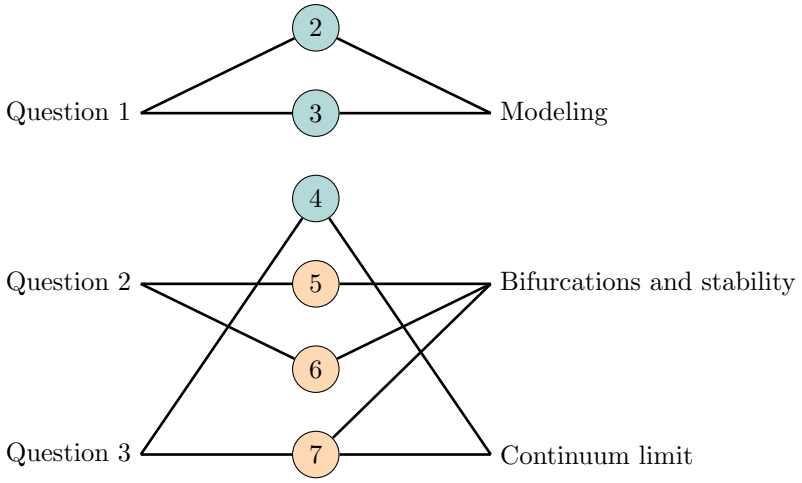


FIGURE 8.1. The thesis at a glance (bis). The nodes 2 to 7 on the central part are color-filled according to the type of equation (linear or nonlinear) of the corresponding chapters. The diagram relates the research questions to the topics that span the whole thesis.

active and passive processes on temporal networks. We have computed the mean resting time and the stationary state of the models, and have shown that in applications such as ranking, the outcome is determined not only by the rules of motion of the diffusing entity and those for the behavior of the edges, but also by a comparison between the associated characteristic times.

We have introduced a new model that emerged as the conjunction of the active node-centric and passive edge-centric walks, that is relevant in situations when an active agent is constrained by the dynamical properties of the underlying network. A typical example is the mobility of individuals on public transportation networks. This new model, despite its fairly simple microscopic rules and despite choosing memoryless distributions for the random processes involved, was shown to be stripped of its Markov properties both in time and in trajectory, when the underlying network has cycles. We followed an analytical approach to capture and quantify the causes that lie at the basis of this emergence of memory.

8.1.1.2. *Perspectives*

Despite its richness, even our most evolved model neglects certain aspects of real-life networks that could lead to interesting research directions. In particular, the assumption that the network can be described as a stationary process, or that the number of diffusion entities is conserved, calls for generalizations. Further, without even opening up the modeling framework to new types of processes, we believe it is of interest to characterize systematically, through metrics that go beyond the mean resting times, the random walk models we have introduced. The larger goal would be to, based on these metrics, design rules to decide

when certain timescales may be neglected and more basic models with only one timescale can be used.

In terms of applications, our preference would require us to first extend the modeling to simple epidemics models, to then try and confront our models that account for competing timescales due to finite edge activation times, with actual data.

8.1.1.3. *Concluding remark*

Overall the modeling and analysis of chapters 2 and 3 helped deepen our understanding of diffusion on temporal graphs, and provided us with a simple mechanism for emergence of non-markovianity in random walks. But importantly, it also laid the groundwork to study reaction-diffusion equations on switched networks in chapter 6.

8.1.2. On the second question

A significant portion of our study of reaction-diffusion models was centered on diffusion-driven instabilities. In this respect, at the light of the past chapters, we dare to say that we had some criticism about Turing's route to pattern formation. We have exposed the limits of the reaction-diffusion models that his idea relates to, and we have divulged any constraints on the emergence of diffusion-driven instabilities we were aware of. Yet we have also emphasized the simplicity and elegance of Turing's mechanism, and brought complements to it with our study of instabilities in RD on temporal networks.

8.1.2.1. *Summary of the findings*

Our first result about reaction-diffusion equations was to obtain all the solutions of the associated stationary problem, in a neighborhood of the bifurcation point from a simple eigenvalue. We have then discussed a generalization of the basic model, where delay is integrated in the diffusion. We have shown that instabilities driven by delayed diffusion may occur even in single species systems, and were able to determine the threshold value of the delay. The models with delay are varied in nature. Complementing our previous finding, we have given numerical evidence of the existence of a breathing pattern in a one species RD system with bistable reaction kinetics and delayed global feedback.

We have also shown that the temporal properties of the underlying graph could loosen the conditions for diffusion to have a destabilizing effect, when coupled with the local nonlinear reactions. For this to happen, the evolution of the graph needs to be sufficiently fast with respect to the local reaction. Based on a linear stability analysis, we were able to determine the threshold on the timescale separation. This adds to the effort to align the original, arguably too simplistic reaction-diffusion model with the complex reality.

8.1.2.2. *Perspectives*

The immediate perspectives are mostly theoretical. Our main concern would be to study more closely the properties of the solutions of a RD equation on a temporal graph, including the amplitude of their possibly periodic oscillations as a function of a parameter controlling the timescale of the evolution of the graph. On an equal footing, we deem necessary to formulate (in)stability results that would allow to draw conclusions on the nonlinear system, based on the linear stability analysis.

8.1.2.3. *Concluding remark*

All the criticism in the world about any given model, even if justified, should not hide another truth. Models represent some idealistic view, and there is no point in all too complex refinements. The authors of [**71**] make a clear case for Turing's idea, and embrace its simplicity. Besides that, while they acknowledge its relevance was debated at length, they review a number of compelling examples which have gradually reduced much of the skepticism surrounding the model, in the experimental biologists community. We hereby put to rest our discussion of Turing instabilities, but recommend the paper to the doubtful or curious reader.

8.1.3. On the third question

What we have achieved on this third theme can be viewed as the natural next steps with respect to the existing literature in the field, which broadly speaking addresses the need to study evolution equations on very large graphs as if they were continuous objects.

8.1.3.1. *Summary of the findings*

Regarding random walks, there was to best of our knowledge no previous work on the graph-limit version of the node-centric walk, even though it is an elementary model on static graphs. This is now covered by our work in the case of dense graphs, having adopted a Hilbert-space setting and having relied on semigroup methods. The same cannot be said about problems where diffusion is driven by the combinatorial Laplacian, in which case the continuum-limit was validated in fairly difficult problems involving fractional or higher order derivatives. But the central question of chapter 7, the stability analysis of a fixed point in a two-variable reaction-diffusion problem, was still open and it made sense to address it in the context of this work. Our investigation brought us back to using elements of spectral theory to properly characterize the combinatorial Laplacian, in a way that allowed to mimic the usual spectral methods on graphs, and apply them to the graph-limit setting. We have ultimately obtained a principle of linearized stability for bounded perturbations, in the case of models with polynomial nonlinearity.

8.1.3.2. *Perspectives*

In both chapters related to graph-limit theory, we have limited the discussion to dense graphs, and have devoted to sparse random graphs only short remarks. However, sparsity tends to be the rule in real-world networks, and the analysis we performed should be extended to sparse graphs. There already exists a substantial body of literature on the subject, including works in the field of dynamical systems. Repeatedly, it is the theory of L^p-graphons that is relied on in these works. Yet recently generalizations have emerged and offer more appropriate ways of dealing with sparse graphs [**23, 61, 132**]. Therefore we propose to reconsider existing results and pursue the work we did, exploiting the novel technical framework.

8.1.3.3. *Concluding remark*

Even if the results we have obtained seem to be merely the confirmation of our original intuition, some peculiarities due to the infinite-dimensional nature of the problems on the continuum have emerged, such as a technical condition of boundedness for the degree function or the type of nonlinearity allowed in the model. In our view, these are not simply details, and they are worth the effort. This being said, our methods may also serve practical purposes, such as numerical applications that require a discretization of a problem on a continuum based on a graph structure.

8.2. A look ahead

The research opportunities outlined above may be cataloged as think-deeper questions directly related to the content of the chapters. Nonetheless there would be merit in broadening our view to propose wider-looking continuation of our work. And this implies the need to fundamentally question our approach or basic assumptions.

8.2.1. Other directions

If we want to better understand dynamics on graphons and similar structures, it is advisable to take a step back and make sure definitions and concepts associated to dynamics on graphs have an equivalent on graphons. For instance, how do we characterize communicability in graphons, what is the graphon equivalent of shortest path, how do we account for long-range interactions, what is the Mellin or Laplace transform of a graphon [**39**]? Some pages of the book on the correspondence from graphs to graphons have been written already, but there is still some challenging way to go before it is complete. And there is interest to be expected from a vivid community producing works such as [**135**], where PageRank is studied in the continuum-limit of directed graphs.

If we want to follow through with the idea that graphs are not just mathematical representations, but that they may connect nodes that correspond to actual locations, as is the case in transportation networks for example, we may

have to give up the assumption that a graph is merely a discrete object. We have to start taking into account the length of the edges and associated metrics, turning each edge of the graph into a continuum. This is required in order to be able to bring into the modeling the time it takes for the diffusing entities to move from one node to the other. Our evolution equations would then take the form of boundary value problems on graphs, with imposed conditions on the ends of the edges. This would achieve a migration towards the world of dynamics on metric graphs in general, and the related topic of quantum random walks in particular, which by the way is a relatively unexplored and indubitably promising field of research.

If we want to gain more insight in the long-term behavior of reaction-diffusion systems and in the patterns they generate, we have to move beyond questions exclusively linked to stability. The theory of RD on a continuum is well-developed, more than it is for networks. On graphs, interesting results have been obtained using for instance weakly nonlinear approaches. But if we are willing to use less-standard operator-theoretic methods, we could exploit the framework of the Koopman operator, which would have us trade in a nonlinear finite-dimensional problem, for a linear though infinite-dimensional version [**79, 90**].

8.2.2. Final note

> "*My first feeling was that there was no way to continue. Writing isn't like math; in math, two plus two always equals four no matter what your mood is like. With writing, the way you feel changes everything.*"
> Stephenie Meyer, about her book *The Midnight Sun*

Overall, the questions we looked at originate in a few basic facts. Namely, some networks inherently possess temporal properties, and hence possibly affect in major ways the behavior of a hosted process. If there is a processing stage in a dynamical system, delays will presumably emerge and should be accounted for in the modeling. Diffusion, when interacting with nonlinear reactions, may have a destabilizing effect, instead of smoothing properties. Patterns and differentiation of nodes may be the result of a self-organized process. And graphs may be extremely large, with seemingly unbounded number of nodes, demanding new mathematical formulations.

A core idea throughout this dissertation is that the underlying graph is definitely not a mere discretization of the domain of the usual equation on a continuum. Dynamical systems on networks have numerous specificities and due to its complexity, the graph may have a dramatic impact on the methods for solving the equations and on the solutions themselves. This interplay helped define in advance what the first steps of this thesis would be, and was gradually uncovered with some new findings, questions and perspectives along the way.

APPENDIX A

Spectral measures

Linked to the spectral theorem in the main body (theorem 7.8) are some general concepts of spectral theory of operators. We introduce them in this appendix because they are required in a first encounter of the theorem, but also needed in practical applications. For details, we advise [**115**] as a reference volume on the matter from which the following material is taken.

A.1. Decomposition of a measure

Consider μ a measure on X. The set of pure points of μ is written $P_\mu = \{x \in X | \mu(\{x\}) > 0\}$. Define

$$\mu_{pp}(X) = \sum_{x \in P_\mu \cap X} \mu(\{x\}) \tag{A.1}$$

and $\mu_{cont} = \mu - \mu_{pp}$. Then we have $\mu_{cont}(\{p\}) = 0$ for all $p \in P_\mu$, and the unique decomposition

$$\mu = \mu_{pp} + \mu_{cont}. \tag{A.2}$$

The measure μ_{cont} can be further decomposed as $\mu_{cont} = \mu_{ac} + \mu_{sing}$, where μ_{ac} is absolutely continuous with respect the Lebesgue measure. That means $d\mu_{ac} = f dx$ for some locally L^1 function, where dx is the Lebesgue measure, and where μ_{sing} is singular relative to Lebesgue measure, i.e. μ_{sing} is continuous and such that $\mu_{sing}(S) = 0$ for some set S such that $\mathbb{R} \setminus S$ has Lebesgue measure zero. Ultimately, any measure μ on $\mathbb{R}$ can be decomposed uniquely as

$$\mu = \mu_{pp} + \mu_{ac} + \mu_{sing} \tag{A.3}$$

and the three are mutually singular. It follows that

$$L^2(\mathbb{R}, d\mu) = L^2(\mathbb{R}, d\mu_{pp}) \oplus L^2(\mathbb{R}, d\mu_{ac}) \oplus L^2(\mathbb{R}, d\mu_{sing}). \tag{A.4}$$

We are now ready to introduce spectral measures, which make the link with spectral theory, and play a key role in the spectral theorem.

A.2. Spectral measures

In order to be able to define spectral measures, one needs to give a meaning to $f(A)$, where f is a continuous function and A acts on a Hilbert space H. There comes the continuous functional calculus, which essentially defines a unique map

$$\phi : C(\sigma(A)) \to \mathbb{L}(H) \tag{A.5}$$

associated with a natural set of properties. For instance, if $P(x) = \sum_{k=1}^{N} a_k x^k$ is a polynomial, then $P(A) = \sum_{k=1}^{N} a_k A^k$, where powers of A are understood in the sense of composition of operators. Moreover, $\sigma(P(A)) = \{P(\lambda) | \lambda \in \sigma(A)\}$. We will simply write $\phi(f)$ as $f(A)$.

Now let us fix the operator A and $\psi \in H$. Then

$$f \to (\psi, f(A)\psi) \tag{A.6}$$

is a positive linear functional[1] on $C(\sigma(A))$. The Riesz-Markov theorem[2] guarantees the existence of μ_ψ, a unique **spectral measure associated with** ψ with support on $\sigma(A)$, such that

$$(\psi, f(A)\psi) = \int_{\sigma(A)} f(\lambda) d\mu_\psi. \tag{A.7}$$

Using spectral measures, one can extend the continuous functional calculus to bounded Borel functions on $\mathbb{R}$.

A.3. Towards the spectral theorem

Two lemmas combined, both of which relying directly on spectral measures, yield a direct proof of the spectral theorem, and importantly show what change of measures are needed to make it work. These lemmas require the notions of cyclic vectors and unitary operators.

Definition A.1 (Cyclic vector)**.** A vector $\psi \in H$ is a cyclic vector for A if finite linear combinations of the elements of $\left\{A^k \psi\right\}_{k=0}^{\infty}$ are dense in H.

Definition A.2 (Isomorphic spaces and unitary operator)**.** Two Hilbert spaces H_1 and H_2 are isomorphic is there exists a linear operator $U : H_1 \to H_2$ such that $(Ux, Uy)_{H_2} = (x, y)_{H_1}$, for all $x, y \in H_1$. The operator U is then called unitary.

As an important result of the definition, observe that unitary operators are norm preserving. Further note that cyclic vectors a neither unique, nor do they always exist, but when it is the case, they allow for the diagonalization of operators, as the following lemma shows.

Lemma A.3. *Let A be a bounded self-adjoint operator with cyclic vector ψ. Then, there exists a unique unitary operator $U : H \to L^2(\sigma(A), d\mu_\psi)$ such that*

$$UAU^{-1} f(\lambda) = \lambda f(\lambda). \tag{A.8}$$

The equality is in the sense of the elements of $L^2(\sigma(A), d\mu_\psi)$, meaning for instance equality between sequences, or between vectors.

---

[1] A positive linear functional on $C(X)$ is a linear functional ℓ such that $\ell(f) \geq 0$ for all f with $f \geq 0$ pointwise.

[2] This theorem essentially identifies a positive linear functional ℓ on $C(X)$ with a measure μ on X such that $\ell(f) = \int_X f d\mu$.

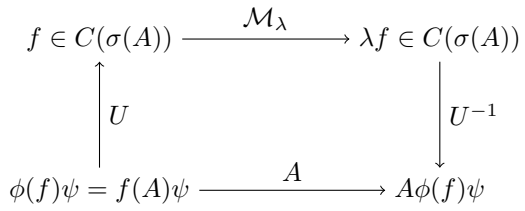


FIGURE A.1. Illustration of lemma A.3. The operator U of the lemma is essentially the inverse of ϕ, but it is dependent on the choice of ψ, which is not the case of ϕ. If ψ is cyclic, then $\text{clo}\{f(A)\psi | f \in C(\sigma(A))\}$ is the whole space. Here $\mathcal{M}_\lambda$ is the multiplication operator by the function $g(\lambda) = \lambda$.

The above lemma can only be applied when A possesses a cyclic vector, as figure A.1 illustrates. But actually, it will be useful in general, for when combined with the next lemma it gives the spectral theorem.

Lemma A.4. *Let A be a self-adjoint operator on a separable Hilbert space H. The there exists a direct sum decomposition $H = \oplus_{n=1}^{N} H_n$ with $N = 1, 2, \ldots$ or ∞ such that*

(a) *A leaves each H_n invariant, meaning that if $v \in H_n$, then $Av \in H_n$;*
(b) *for each n, there is a $\psi_n \in H_n$ that is cyclic for the restriction $A \restriction H_n$ of A to H_n, that is,*

$$H_n = \text{clo}\{f(A)\psi_n | f \in C(\sigma(A))\}. \tag{A.9}$$

The spectral theorem and a direct corollary are stated on page 146, and illustrated by examples.

APPENDIX B

Decompositions and bounds on the spectrum

This appendix details different decompositions of the spectrum used in chapters of the main text. They are also needed to complement the spectral measures of appendix A as a prerequisite to the application of the spectral theorem. We further introduce a straightforward generalization of the spectral abscissa of matrices to linear operators, and make the link with the growth bound of the generator of a strongly continuous semigroup. The main reference for this appendix is again [**115**], and we rely on [**38**] for the aspects related to semigroups.

B.1. Decomposition based on spectral measures

A convention chosen by some authors is to define the the discrete spectrum as the set of $\lambda \in \sigma(A)$ such that $\lambda\mathcal{I} - A$ is not injective; the residual spectrum as the set of $\lambda \in \sigma(A)$ such that $\lambda\mathcal{I} - A$ is injective but not surjective, and the range of $\lambda\mathcal{I} - A$ is not dense; the continuous spectrum as the set of $\lambda \in \sigma(A)$ such that $\lambda\mathcal{I} - A$ is injective but not surjective, and the range of $\lambda\mathcal{I} - A$ is dense. Then λ is in the continuous spectrum if and only if there exists v such that $||v|| = 1$, and $||Av - \lambda v||$ is arbitrarily small.

In this work we use a slightly different classification, that is based on spectral measures and that we summarize here, focusing on a Hilbert-space framework. Building on the decomposition of $L^2(\mathbb{R}, d\mu)$ given by (A.4), we first introduce a decomposition of the Hilbert space.

Definition B.1. Let A be a bounded self-adjoint operator on H. Let $H_{pp} = \{\psi | \mu_\psi \text{ is pure point}\}$, $H_{ac} = \{\psi | \mu_\psi \text{ is absolutely continuous}\}$, and let $H_{sing} = \{\psi | \mu_\psi \text{ is continuous singular}\}$.

The next theorem is an important result [**115**, Theorem VII.4].

THEOREM B.2. *$H = H_{pp} \oplus H_{ac} \oplus H_{sing}$, and each of these subspaces is invariant under A. Moreover, $A \restriction H_{pp}$ has a complete set of eigenvectors, $A \restriction H_{ac}$ has only absolutely continuous spectral measures and $A \restriction H_{sing}$ has only continuous singular spectral measures.*

This following sets offer a way of partitioning the spectrum as follows:

Definition B.3. The sets defined by

$$\begin{aligned}\sigma_{pp}(A) &= \{\lambda \text{ is an eigenvalue of } A\}\\ \sigma_{cont}(A) &= \sigma(A \restriction H_{cont} \equiv H_{sing} \oplus H_{ac})\\ \sigma_{ac}(A) &= \sigma(A \restriction H_{ac})\\ \sigma_{sing}(A) &= \sigma(A \restriction H_{sing})\end{aligned}$$

are the pure point, continuous, absolutely continuous, and (continuous) singular spectrum respectively.

Remark B.4. Since σ_{pp} was not defined as $\sigma(A \restriction H_{pp})$ but rather as the set of eigenvalues, it may happen that $\sigma \neq \sigma_{pp} \cup \sigma_{ac} \cup \sigma_{sing}$. But $\operatorname{clo} \sigma_{pp}(A) = \sigma(A \restriction H_{pp})$ and we have:

$$\sigma_{cont}(A) = \sigma_{ac}(A) \cup \sigma_{sing}(A), \tag{B.1}$$

$$\sigma(A) = \operatorname{clo} \sigma_{pp}(A) \cup \sigma_{cont}(A). \tag{B.2}$$

With the definitions used here, the singular spectrum can have nonzero Lebesgue measure. Moreover, the definition of the continuous spectrum does not agree with other author's choice to define the continuous spectrum as the the part of the spectrum which is disjoint from the point spectrum and the residual spectrum, as mentioned above.

B.2. Decomposition into the discrete and the essential spectrum

Another useful decomposition of the spectrum of an operator A into two disjoint subsets is given by the discrete spectrum $\sigma_{disc}(A)$ and the essential spectrum $\sigma_{ess}(A)$. These two subsets can be defined in terms of the dimensions of so-called spectral projections. We give some details, but we point to the fact that it is safe to go directly to theorems B.9 and B.10, which provide the needed characterizations of these subsets.

The functional calculus allows to define functions of an operator, $f \to f(A)$. Specializing to the case of indicator functions leads to spectral projections.

Definition B.5 (Spectral projection)**.** Let A be a bounded self-adjoint operator and Ω a Borel set of $\mathbb{R}$. $P_\Omega \equiv \chi_\Omega(A)$ ia called a spectral projection of A.

Remark B.6. The family of spectral projections P_Ω of an operator enjoy properties that are evocative of a measure. One can define based on such family a projection-valued measure (p.v.m.). If P_Ω is a p.v.m., then for any ϕ, $(\phi, P_\Omega \phi)$ is an ordinary measure. One can formulate a p.v.m. form of the spectral theorem [**115**, page 235].

The reason to introduce spectral projections resides in their connection with the spectrum.

Proposition B.7. *$\lambda \in \sigma(A)$ if and only if $P_{(\lambda-\epsilon,\lambda+\epsilon)}(A) \neq 0$ for any $\epsilon > 0$.*

Definition B.8 (Essential spectrum, discrete spectrum)**.** λ is in the essential spectrum of A, denoted $\lambda \in \sigma_{ess}(A)$, if and only if $P_{(\lambda-\epsilon,\lambda+\epsilon)}(A)$ is infinite dimensional for all $\epsilon > 0$.
If to the contrary $\lambda \in \sigma(A)$ but $P_{(\lambda-\epsilon,\lambda+\epsilon)}(A)$ is finite dimensional for some $\epsilon > 0$, then $\lambda \in \sigma_{disc}(A)$, the discrete spectrum of A.

THEOREM B.9 (Characterization of the discrete spectrum). *$\lambda \in \sigma_{disc}(A)$ if and only if both of the following are true:*

(a) *λ is an isolated point of $\sigma(A)$, meaning that for some $\epsilon > 0$, $(\lambda - \epsilon, \lambda + \epsilon) \cap \sigma(A) = \{\lambda\}$.*
(b) *λ is an eigenvalue of finite multiplicity, meaning that $\{\psi | A\psi = \lambda\psi\}$ is finite-dimensional.*

THEOREM B.10 (Characterization of the essential spectrum). *$\lambda \in \sigma_{ess}(A)$ if and only if at least one of the following holds:*

(a) *$\lambda \in \sigma_{cont} = \sigma_{ac}(A) \cup \sigma_{sing}(A)$*
(b) *λ is a limit point of $\sigma_{pp}(A)$*
(c) *λ is an eigenvalue of infinite multiplicity.*

B.3. Bounds

This section relates the growth bound of semigroup and the spectral bound of its generator, which corresponds with the spectral abscissa of matrices. For every strongly continuous semigroup $(T(t))_{t\geq 0}$, there exists $w \in \mathbb{R}$ an $M > 1$ such that $\|T(t)\| \leq Me^{wt}$ for all $t \geq 0$. Hence, we have the following definition.

Definition B.11 (Growth bound)**.** The growth bound of a strongly continuous semigroup $(T(t))_{t\geq 0}$ is defined by

$$\omega_0 := \inf \left\{\omega \in \mathbb{R} \,:\, \exists M_\omega > 1 \text{ such that } \|T(t)\| \leq M_\omega e^{\omega t},\, \forall t \geq 0.\right\}$$

Definition B.12 (Spectral bound)**.** The spectral bound of a linear operator A is defined by

$$\mathsf{s}(A) := \sup \{\operatorname{Re} \lambda \,:\, \lambda \in \sigma(A)\}.$$

THEOREM B.13. *For a strongly continuous semigroup $(T(t))_{t\geq 0}$ with generator A, it holds that $-\infty \leq \mathsf{s}(A) \leq \omega_0 \leq +\infty$.*

This result comes from [**38**, Corollary 1.13]. Note that the spectral bound and the growth bound of the generator of a uniformly continuous semigroup are equal, and one says that the spectral determined growth conditions hold. This will repeatedly be used in chapters 4 and 7.

Bibliography

[1] R. Al Jamal, A. Chow, and K. Morris, *Linearized stability analysis of nonlinear partial differential equations*, Mathematical Theory of Networks and Systems, (2014), pp. 847–852.

[2] R. Al Jamal and K. Morris, *Linearized stability of partial differential equations with application to stabilization of the Kuramoto–Sivashinsky equation*, SIAM Journal on Control and Optimization, 56 (2018), pp. 120–147.

[3] C. Angstmann, I. Donnelly, and B. Henry, *Continuous time random walks with reactions forcing and trapping*, Mathematical Modelling of Natural Phenomena, 8 (2013), pp. 17–27.

[4] C. Angstmann, I. Donnelly, and B. Henry, *Pattern formation on networks with reactions: A continuous-time random-walk approach*, Physical Review E, 87 (2013), p. 032804.

[5] C. Angstmann, I. Donnelly, B. Henry, and T. Langlands, *Continuous-time random walks on networks with vertex-and time-dependent forcing*, Physical Review E, 88 (2013), p. 022811.

[6] A. Anma, K. Sakamoto, and T. Yoneda, *Unstable subsystems cause Turing instability*, Kodai Mathematical Journal, 35 (2012), pp. 215–247.

[7] M. Asllani, T. Carletti, F. Di Patti, D. Fanelli, and F. Piazza, *Hopping in the crowd to unveil network topology*, Physical review letters, 120 (2018), p. 158301.

[8] F. Atai, *Complex time-delay systems : Theory and applications*, Springer-Verlag Berlin Heidelberg, 2010.

[9] R. Balescu, *Statistical dynamics: matter out of equilibrium*, Imperial College Press, 1997.

[10] A.-L. Barabasi, *The origin of bursts and heavy tails in human dynamics*, Nature, 435 (2005), p. 207.

[11] M. Belkin and P. Niyogi, *Towards a theoretical foundation for Laplacian-based manifold methods*, Journal of Computer and System Sciences, 74 (2008), pp. 1289 – 1308. Learning Theory 2005.

[12] I. V. Belykh, V. N. Belykh, and M. Hasler, *Blinking model and synchronization in small-world networks with a time-varying coupling*, Physica D: Nonlinear Phenomena, 195 (2004), pp. 188–206.

[13] D. ben Avraham and S. Havlin, *Diffusion and reactions in fractals and disordered systems*, Cambridge University Press, Cambridge, UK, 2000.

[14] S. Blanes, F. Casas, J. Oteo, and J. Ros, *The Magnus expansion and some of its applications*, Physics reports, 470 (2009), pp. 151–238.

[15] S. Blanes, F. Casas, and J. Ros, *Improved high order integrators based on the Magnus expansion*, BIT Numerical Mathematics, 40 (2000), pp. 434–450.

[16] C. Borgs, J. T. Chayes, H. Cohn, and N. Holden, *Sparse exchangeable graphs and their limits via graphon processes*, arXiv preprint arXiv:1601.07134, (2016).

[17] C. Borgs, J. T. Chayes, H. Cohn, and Y. Zhao, *An L^p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions*, arXiv preprint arXiv:1401.2906, (2014).

[18] ———, *An L^p theory of sparse graph convergence II: LD convergence, quotients and right convergence*, The Annals of Probability, 46 (2018), pp. 337–396.

[19] C. Borgs, J. T. Chayes, L. Lovász, V. Sós, and K. Vesztergombi, *Limits of randomly grown graph sequences*, European Journal of Combinatorics, 32 (2011), pp. 985 – 999.

[20] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Advances in Mathematics, 219 (2008), pp. 1801–1851.

[21] S. Brin and L. Page, *The anatomy of a large-scale hypertextual web search engine*, Computer networks and ISDN systems, 30 (1998), pp. 107–117.

[22] L. Cantini, C. Cianci, D. Fanelli, E. Massi, L. Barletti, and M. Asllani, *Stochastic amplification of spatial modes in a system with one diffusing species*, Journal of mathematical biology, 69 (2014), pp. 1585–1608.

[23] F. Caron and E. B. Fox, *Sparse graphs using exchangeable random measures*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 79 (2017), pp. 1295–1366.

[24] F. Casas, J. Oteo, and J. Ros, *Floquet theory: exponential perturbative treatment*, Journal of Physics A: Mathematical and General, 34 (2001), p. 3379.

[25] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, *Dynamics of person-to-person interactions from distributed RFID sensor networks*, PLOS ONE, 5 (2010), p. e11596.

[26] G. Cencetti, F. Battiston, T. Carletti, and D. Fanelli, *Turing-like patterns from purely reactive systems*, arXiv:1906.09048, (2019).

[27] C. Chicone, *Ordinary differential equations with applications*, vol. 34, Springer Science & Business Media, 2006.

[28] V. Colizza, R. Pastor-Satorras, and A. Vespignani, *Reaction–diffusion processes and metapopulation models in heterogeneous networks*, Nature Physics, 3 (2007), pp. 276–282.

[29] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, *On the Lambert W function*, Advances in Computational mathematics, 5 (1996), pp. 329–359.

[30] M. G. Crandall and P. H. Rabinowitz, *Bifurcation from simple eigenvalues*, Journal of Functional Analysis, 8 (1971), pp. 321–340.

[31] H. Crane, *Dynamic random networks and their graph limits*, The Annals of Applied Probability, 26 (2016), pp. 691–721.

[32] M. D'Elia, Q. Du, M. Gunzburger, and R. Lehoucq, *Nonlocal convection-diffusion problems on bounded domains and finite-range jump processes*, Computational Methods in Applied Mathematics, 17 (2017), pp. 707–722.

[33] J.-C. Delvenne, R. Lambiotte, and L. E. Rocha, *Diffusion on networked systems is a question of time or structure*, Nature communications, 6 (2015).

[34] J. C. Delvenne, S. N. Yaliraki, and M. Barahona, *Stability of graph communities across time scales*, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 12755–12760.

[35] L. M. Delves and J. Mohamed, *Computational methods for integral equations*, Cambridge University Press, 1985.

[36] P. Diaconis, S. Holmes, and S. Janson, *Threshold graph limits and random threshold graphs*, Internet Mathematics, 5 (2008), pp. 267–320.

[37] P. Diaconis and S. Janson, *Graph limits and exchangeable random graphs*, arXiv preprint arXiv:0712.2749, (2007).

[38] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, in Semigroup Forum, vol. 63, Springer, 2001, pp. 278–280.

[39] E. Estrada, L. Gambuzza, and M. Frasca, *Long-range interactions and network synchronization*, SIAM Journal on Applied Dynamical Systems, 17 (2018), pp. 672–693.

[40] S. Fedotov, *Non-markovian random walks and nonlinear reactions: subdiffusion and propagating fronts*, Physical Review E, 81 (2010), p. 011117.

[41] S. Fedotov and H. Stage, *Anomalous metapopulation dynamics on scale-free networks*, Physical review letters, 118 (2017), p. 098301.

[42] D. Figueiredo, P. Nain, B. Ribeiro, E. de Souza e Silva, and D. Towsley, *Characterizing continuous time random walks on time varying graphs*, in ACM SIGMETRICS Performance Evaluation Review, vol. 40, ACM, 2012, pp. 307–318.

[43] C. Gao, Y. Lu, and H. H. Zhou, *Rate-optimal graphon estimation*, The Annals of Statistics, 43 (2015), pp. 2624–2652.
[44] S. Gao and P. E. Caines, *The control of arbitrary size networks of linear systems via graphon limits: An initial investigation*, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), IEEE, 2017, pp. 1052–1057.
[45] L. Gauvin, A. Panisson, C. Cattuto, and A. Barrat, *Activity clocks: spreading dynamics on temporal networks of human contact*, Scientific reports, 3 (2013), p. 3099.
[46] E. Giné and V. Koltchinskii, *Empirical graph laplacian approximation of Laplace–Beltrami operators: Large sample results*, in High dimensional probability, Institute of Mathematical Statistics, 2006, pp. 238–259.
[47] D. S. Grebenkov and L. Tupikina, *Heterogeneous continuous-time random walks*, Physical Review E, 97 (2018), p. 012148.
[48] M. Gueuning, R. Lambiotte, and J.-C. Delvenne, *Backtracking and mixing rate of diffusion on uncorrelated temporal networks*, Entropy, 19 (2017), p. 542.
[49] D. Haim, G. Li, Q. Ouyang, W. D. McCormick, H. L. Swinney, A. Hagberg, and E. Meron, *Breathing spots in a reaction-diffusion system*, Physical review letters, 77 (1996), p. 190.
[50] J. Halatek and E. Frey, *Rethinking pattern formation in reaction–diffusion systems*, Nature Physics, 14 (2018), pp. 507–514.
[51] A. Hastir, J. Winkin, and D. Dochain, *Exponential stability of nonlinear infinite-dimensional systems: application to nonisothermal axial dispersion tubular reactors*. Submitted (2019).
[52] S. Hata, H. Nakao, and A. S. Mikhailov, *Sufficient conditions for wave instability in three-component reaction–diffusion systems*, Progress of Theoretical and Experimental Physics, 2014 (2014).
[53] M. Hein, J.-Y. Audibert, and U. v. Luxburg, *Graph Laplacians and their convergence on random neighborhood graphs*, Journal of Machine Learning Research, 8 (2007), pp. 1325–1368.
[54] T. Hoffmann, M. A. Porter, and R. Lambiotte, *Generalized master equations for non-Poisson dynamics on networks*, Physical Review E, 86 (2012), p. 046102.
[55] A. Holderrieth, *Matrix multiplication operators generating one parameter semigroups*, in Semigroup Forum, vol. 42, Springer, 1991, pp. 155–166.
[56] P. Holme, *Modern temporal network theory: a colloquium*, The European Physical Journal B, 88 (2015), pp. 1–30.
[57] P. Holme and J. Saramäki, *Temporal Networks*, Springer-Verlag, Berlin, Germany, 2013.
[58] B. D. Hughes, *Random walks and random environments: random walks*, vol. 1, Oxford University Press, 1995.
[59] L. I. Ignat and J. D. Rossi, *Decay estimates for nonlocal problems via energy methods*, Journal de mathématiques pures et appliquées, 92 (2009), pp. 163–187.
[60] S. Janson, *Connectedness in graph limits*, arXiv preprint arXiv:0802.3795, (2008).
[61] ———, *On edge exchangeable random graphs*, Journal of statistical physics, 173 (2018), pp. 448–484.
[62] H. Jasiulewicz and W. Kordecki, *Convolutions of Erlang and of Pascal distributions with applications to reliability*, Demonstratio Mathematica, 36 (2003), pp. 231–238.
[63] D. Kaliuzhnyi-Verbovetskyi and G. Medvedev, *The semilinear heat equation on sparse random graphs*, SIAM Journal on Mathematical Analysis, 49 (2017), pp. 1333–1355.
[64] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, *Small but slow world: How network topology and burstiness slow down spreading*, Physical Review E, 83 (2011), p. 025102.
[65] N. Kataoka and K. Kaneko, *Dynamical networks in function dynamics*, Physica D: Nonlinear Phenomena, 181 (2003), pp. 235–251.
[66] N. Kato, *A principle of linearized stability for nonlinear evolution equations*, Transactions of the American Mathematical society, 347 (1995), pp. 2851–2868.

[67] J. Kempe, *Quantum random walks: an introductory overview*, Contemporary Physics, 44 (2003), pp. 307–327.
[68] J. Klafter and I. M. Sokolov, *First steps in random walks: from tools to applications*, Oxford University Press, 2011.
[69] S. Klarsfeld and J. Oteo, *Recursive generation of higher-order terms in the Magnus expansion*, Physical Review A, 39 (1989), p. 3270.
[70] C. Koch, *Biophysics of computation: information processing in single neurons*, Oxford university press, 2004.
[71] S. Kondo and T. Miura, *Reaction-diffusion model as a framework for understanding biological pattern formation*, Science, 329 (2010), pp. 1616–1620.
[72] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná, *Hyperbolic geometry of complex networks*, Physical Review E, 82 (2010), p. 036106.
[73] C. Kuehn and S. Throm, *Power network dynamics on graphons*, SIAM Journal on Applied Mathematics, 79 (2019), pp. 1271–1292.
[74] R. Kutner and J. Masoliver., *The continuous time random walk, still trendy: fifty-year history, state of art and outlook*, The European Physical Journal B, 90 (2017), p. 50.
[75] K. Kuto, K. Osaki, T. Sakurai, and T. Tsujikawa, *Spatial pattern formation in a chemotaxis–diffusion–growth model*, Physica D: Nonlinear Phenomena, 241 (2012), pp. 1629–1639.
[76] R. Lambiotte, J. C. Delvenne, and M. Barahona, *Random walks, Markov processes and the multiscale modular organization of complex networks*, IEEE Trans. Netw. Sci. Eng., 1 (2014), pp. 76–90.
[77] R. Lambiotte and M. Rosvall, *Ranking and clustering of nodes in networks with smart teleportation*, Physical Review E, 85 (2012), p. 056107.
[78] R. Lambiotte, M. Rosvall, and I. Scholtes, *Understanding complex systems: From networks to optimal higher-order models*, arXiv preprint arXiv:1806.05977, (2018).
[79] Y. Lan and I. Mezić, *Linearization in the large of nonlinear systems and Koopman operator spectrum*, Physica D: Nonlinear Phenomena, 242 (2013), pp. 42–53.
[80] A. N. Langville and C. D. Meyer, *Deeper inside PageRank*, Internet Mathematics, 1 (2004), pp. 335–380.
[81] B. E. Lee, *Consensus and voting on large graphs: An application of graph limit theory*, Discrete & Continuous Dynamical Systems - A, 2018, 38 (4) : 1719-1744.
[82] S.-Y. Liu, A. Baronchelli, and N. Perra, *Contagion dynamics in time-varying metapopulation networks*, Physical Review E, 87 (2013), p. 032805.
[83] L. Lovász, *Random walks on graphs: A survey*, Combinatorics, Paul Erdos is eighty, 2 (1993), pp. 1–46.
[84] ———, *Large networks and graph limits*, vol. 60, American Mathematical Soc., 2012.
[85] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, Journal of Combinatorial Theory, Series B, 96 (2006), pp. 933–957.
[86] M. Lucas, D. Fanelli, T. Carletti, and J. Petit, *Desynchronization induced by time-varying network*, EPL (Europhysics Letters), 121 (2018), p. 50008.
[87] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Communications on pure and applied mathematics, 7 (1954), pp. 649–673.
[88] N. Masuda and R. Lambiotte, *A guide to temporal networks*, World Scientific, 1997.
[89] N. Masuda, M. A. Porter, and R. Lambiotte, *Random walks and diffusion on networks*, Physics Reports, 716-717 (2017), pp. 1–58.
[90] A. Mauroy, I. Mezić, and J. Moehlis, *Isostables, isochrons, and Koopman spectrum for the action–angle representation of stable fixed point dynamics*, Physica D: Nonlinear Phenomena, 261 (2013), pp. 19–30.
[91] G. S. Medvedev, *The nonlinear heat equation on dense graphs and graph limits*, SIAM Journal on Mathematical Analysis, 46 (2014), pp. 2743–2766.
[92] ———, *The nonlinear heat equation on W-random graphs*, Archive for Rational Mechanics and Analysis, 212 (2014), pp. 781–803.
[93] ———, *Small-world networks of Kuramoto oscillators*, Physica D: Nonlinear Phenomena, 266 (2014), pp. 13–22.

[94] G. S. Medvedev and X. Tang, *The Kuramoto model on power law graphs: Synchronization and contrast states*, Journal of Nonlinear Science, (2018), pp. 1–23.
[95] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson, *The unreasonable effectiveness of tree-based theory for networks with clustering*, Physical Review E, 83 (2011), p. 036112.
[96] V. Mendez, S. Fedotov, and W. Horsthemke, *Reaction-transport systems: mesoscopic foundations, fronts, and spatial instabilities*, Springer Science & Business Media, 2010.
[97] W. Michiels and S.-I. Niculescu, *Stability and stabilization of time-delay systems: an eigenvalue-based approach*, SIAM, 2007.
[98] P. C. Moan and J. Niesen, *Convergence of the Magnus series*, Foundations of Computational Mathematics, 8 (2008), pp. 291–301.
[99] R. Molontay and M. Nagy, *Two decades of network science: as seen through the co-authorship network of network scientists*, in Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2019, pp. 578–583.
[100] J. D. Murray, *Mathematical Biology* 1 *: An introduction*, Springer New York, 2002.
[101] ———, *Mathematical Biology 2 : Spatial Models and Biomedical Applications*, Springer New York, 2003.
[102] H. Nakao and A. S. Mikhailov, *Turing patterns in network-organized activator–inhibitor systems*, Nature Physics, 6 (2010), pp. 544–550.
[103] H. G. Othmer and L. Scriven, *Instability and dynamic pattern in cellular networks*, Journal of theoretical biology, 32 (1971), pp. 507–537.
[104] A. Otto, J. Wang, and G. Radons, *Delay-induced wave instabilities in single-species reaction-diffusion systems*, Physical Review E, 96 (2017), p. 052202.
[105] K. J. Palmer, *A generalization of Hartman's linearization theorem*, Journal of Mathematical Analysis and Applications, 41 (1973), pp. 753–758.
[106] F. Parise and A. Ozdaglar, *Graphon games*, in Proceedings of the 2019 ACM Conference on Economics and Computation, ACM, 2019, pp. 457–458.
[107] N. Perra, A. Baronchelli, D. Mocanu, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, *Random walks and search in time-varying networks*, Physical review letters, 109 (2012), p. 238701.
[108] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, *Activity driven modeling of time varying networks*, Scientific Reports, 2 (2012), p. 469.
[109] J. Petit, M. Asllani, D. Fanelli, B. Lauwens, and T. Carletti, *Pattern formation in a two-component reaction–diffusion system with delayed processes on a network*, Physica A: Statistical Mechanics and its Applications, 462 (2016), pp. 230–249.
[110] J. Petit, T. Carletti, M. Asllani, and D. Fanelli, *Delay-induced Turing-like waves for one-species reaction-diffusion model on a network*, Europhysics Letters, 111 (2015), p. 58002.
[111] J. Petit, B. Lauwens, D. Fanelli, and T. Carletti, *Theory of Turing patterns on time varying networks*, Physical Review Letters, 119 (2017), p. 148301.
[112] M. Porter and J. Gleeson, *Dynamical systems on networks*, Frontiers in Applied Dynamical Systems: Reviews and Tutorials, (2016).
[113] I. Prigogine and R. Lefever, *Symmetry breaking instabilities in dissipative systems II*, The Journal of Chemical Physics, 48 (1968), pp. 1695–1700.
[114] M. Reed and B. Simon, *Vol 4: Analysis of Operators*, Methods of modern mathematical physics, Academic Press, San Diego, 1978.
[115] ———, *Vol 1: Functional Analysis*, Methods of modern mathematical physics, Academic Press, San Diego, 1980.
[116] E. Renshaw and R. Henderson, *The correlated random walk*, Journal of Applied Probability, 18 (1981), pp. 403–414.
[117] L. Rosasco, M. Belkin, and E. D. Vito, *On learning with integral operators*, Journal of Machine Learning Research, 11 (2010), pp. 905–934.
[118] M. Rosvall, D. Axelsson, and C. T. Bergstrom, *The map equation*, The European Physical Journal Special Topics, 178 (2009), pp. 13–23.

[119] M. Rosvall and C. T. Bergstrom, *Maps of random walks on complex networks reveal community structure*, PNAS, 105 (2008), pp. 1118–1123.

[120] M. Salathé, M. Kazandjieva, J. W. Lee, P. Levis, M. W. Feldman, and J. H. Jones, *A high-resolution human contact network for infectious disease transmission*, Proceedings of the National Academy of Sciences, 107 (2010), pp. 22020–22025.

[121] A. Scherrer, P. Borgnat, E. Fleury, J.-L. Guillaume, and C. Robardet, *Description and simulation of dynamic mobility networks*, Computer Networks, 52 (2008), pp. 2842–2858.

[122] I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C. J. Tessone, and F. Schweitzer, *Causality-driven slow-down and speed-up of diffusion in non-markovian temporal networks*, Nature communications, 5 (2014), p. 5024.

[123] G. M. Schütz and S. Trimper, *Elephants can always remember: Exact long-range memory effects in a non-markovian random walk*, Physical Review E, 70 (2004), p. 045101.

[124] V. Sekara, A. Stopczynski, and S. Lehmann, *Fundamental structures of dynamic social networks*, PNAS, 113 (2016), pp. 9977–9982.

[125] S. Sen, P. Ghosh, S. S. Riaz, and D. S. Ray, *Time-delay-induced instabilities in reaction-diffusion systems*, Physical Review E, 80 (2009), p. 046212.

[126] H. Shinozaki and T. Mori, *Robust stability analysis of linear time-delay systems by lambert w function: Some extreme point results*, Automatica, 42 (2006), pp. 1791–1799.

[127] L. Speidel, R. Lambiotte, K. Aihara, and N. Masuda, *Steady state and mean recurrence time for random walks on stochastic temporal networks.*, Physical Review E, 91 (2015), p. 012806.

[128] M. Starnini, A. Baronchelli, A. Barrat, and R. Pastor-Satorras, *Random walks on temporal networks*, Physical Review E, 85 (2012), p. 056115.

[129] E. Steur, W. Michiels, H. Huijberts, and H. Nijmeijer, *Networks of diffusively time-delay coupled systems: Conditions for synchronization and its relation to the network topology*, Physica D: Nonlinear Phenomena, 277 (2014), pp. 22–39.

[130] D. J. Stilwell, E. M. Bollt, and D. G. Roberson, *Sufficient conditions for fast switching synchronization in time-varying network topologies*, SIAM Journal on Applied Dynamical Systems, 5 (2006), pp. 140–156.

[131] A. Turing, *The chemical basis of morphogenesis*, Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, (1952), pp. 37–72.

[132] V. Veitch and D. M. Roy, *The class of random graphs arising from exchangeable random measures*, arXiv preprint arXiv:1512.03099, (2015).

[133] F. Verhulst, *Nonlinear differential equations and dynamical systems*, Springer-Verlag New York, 1990.

[134] W. Woess, *Random walks on infinite graphs and groups*, vol. 138, Cambridge university press, 2000.

[135] A. Yuan, J. Calder, and B. Osting, *A continuum limit for the PageRank algorithm*, arXiv preprint arXiv:2001.08973, (2020).

[136] D. H. Zanette and A. S. Mikhailov, *Dynamical systems with time-dependent coupling: clustering and critical behaviour*, Physica D: Nonlinear Phenomena, 194 (2004), pp. 203–218.

[137] T. Zhang and H. Zang, *Delay-induced Turing instability in reaction-diffusion equations*, Physical Review E, 90 (2014), p. 052908.

[138] K. Zhao, M. Karsai, and G. Bianconi, *Entropy of dynamical social networks*, PLOS ONE, 6 (2011), p. e28116.