Lectures on
Network Systems

Francesco Bullo

With contributions by
Jorge Cortés
Florian Dörfler
Sonia Martínez

Lectures on Network Systems
Francesco Bullo
edition 1.6 – Jan 1, 2022
With contributions from J. Cortés, F. Dörfler, and S. Martínez

Textbook in landscape format
for presentation purposes

Lectures on Network Systems
Slide Version
Francesco Bullo
edition 1.6 – Jan 1, 2022
With contributions from J. Cortés, F. Dörfler, and S. Martínez

II Topics in Averaging Systems 495

III Nonlinear Systems 567

Part I

Linear Systems

Chapter 1

Motivating Problems and Systems

In this chapter, we introduce some example problems and systems from multiple disciplines to motivate our treatment of linear network systems in the following chapters. We look at the following examples:

(i) In the context of social influence networks, we discuss a classic model on how opinions evolve and possibly reach a consensus opinion in groups of individuals.

(ii) In the context of wireless sensor networks, we discuss a simple distributed averaging algorithms and, in the appendix, two advanced design problems for parameter estimation and hypothesis testing.

(iii) In the context of animal behavior, we present a flocking model involving a simple alignment rule.

(iv) In the context of dynamical flow systems, we discuss flows of commodities among compartments in both discrete-time and continuous-time, with classic examples including Markov chains and affine dynamics for water flows in desert ecosystems.

In all cases we are interested in presenting the basic models and motivating interest in understanding their dynamic behaviors, such as the existence and attractivity of equilibria.

1.1   Opinion dynamics in social influence networks

We consider a group of n individuals who must act together as a team. Each individual has his own subjective probability density function (pdf) p_i for the unknown value of some parameter (or more simply an estimate of the parameter). We assume now that individual i is appraised of the pdf p_j of each other member $j \neq i$ of the group. Then the model by (French Jr., 1956; Harary, 1959), see also the later (DeGroot, 1974), predicts that the individual will revise its pdf to be:

$$p_i^+ = \sum_{j=1}^{n} a_{ij} p_j, \tag{1.1}$$

Figure 1.1: Interactions in a social influence network

where a_{ij} denotes the weight that individual i assigns to the pdf of individual j when carrying out this revision. More precisely, the coefficient a_{ii} describes the attachment of individual i to its own opinion and a_{ij}, $j \neq i$, is an interpersonal influence weight that individual i accords to individual j.

We will refer to equation (1.1) as to the *French-Harary-DeGroot model* of opinion dynamics. In this model, the coefficients a_{ij} satisfy the following constraints: they are non-negative, that is, $a_{ij} \geq 0$, and, for each individual, the sum of self-weight and accorded weights equals 1, that is, $\sum_{j=1}^{n} a_{ij} = 1$ for all i. In mathematical terms, the matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

has non-negative entries and each of its rows has unit sum. Such matrices are said to be *row-stochastic*.

Scientific questions of interest include:

(i) Is this model of human opinion dynamics believable? Is there empirical evidence in its support?

(ii) How does one measure the coefficients a_{ij}?

(iii) Under what conditions do the pdfs converge to the same pdf? In other words, when do the agents achieve *consensus*? And to what final pdf?

(iv) What are more realistic, empirically-motivated models, possibly including stubborn individuals or antagonistic interactions?

1.2   Averaging algorithms in wireless sensor networks
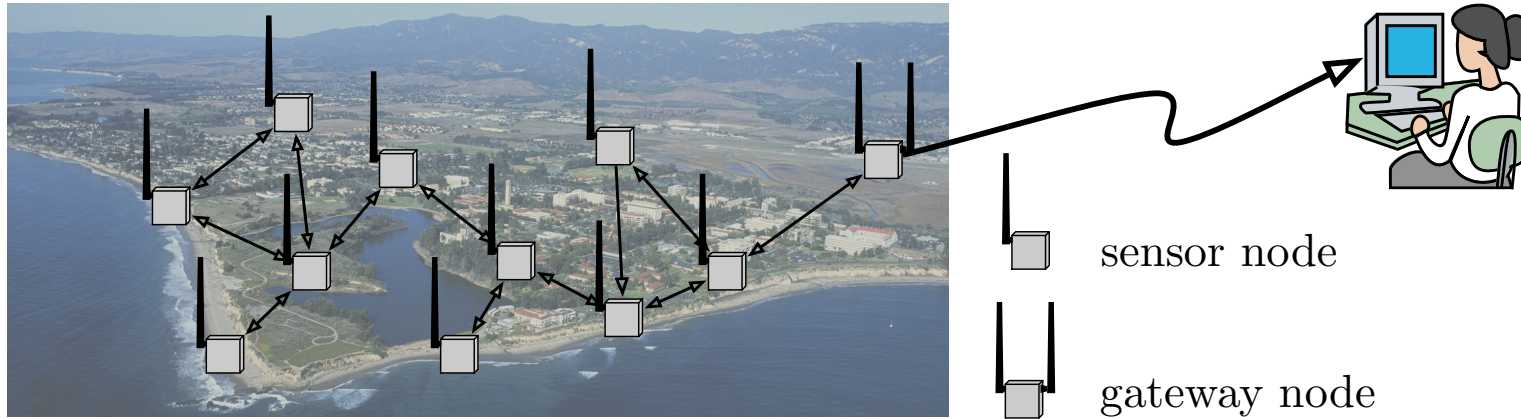
Figure 1.2: A wireless sensor network composed of a collection of spatially-distributed sensors in a field and a gateway node to carry information to an operator. The nodes are meant to measure environmental variables, such as temperature, sound, pressure, and cooperatively filter and transmit the information to an operator.

A wireless sensor network is a collection of spatially-distributed devices capable of measuring physical and environmental variables (e.g., temperature, vibrations, sound, light, etc), performing local computations, and transmitting information to neighboring devices and, in turn, throughout the network (including, possibly, an external operator).

Suppose that each node in a wireless sensor network has measured a scalar environmental quantity, say x_i. Consider the following simple distributed algorithm, based on the concepts of linear averaging: each node repeatedly executes

$$x_i^+ := \text{average}\left(x_i, \{x_j, \text{ for all neighbor nodes } j\}\right), \tag{1.2}$$

where x_i^+ denotes the new value of x_i. For example, for the graph in Figure 1.3, one can easily write $x_1^+ := (x_1 + x_2)/2$, $x_2^+ := (x_1 + x_2 + x_3 + x_4)/4$, and so forth. In summary, the algorithm's behavior is described by

$$x^+ = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} x = A_{\text{wsn}}x,$$

where the matrix A_{wsn} in equation is again row-stochastic.
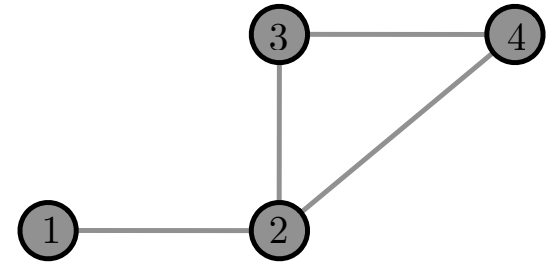
Figure 1.3: Example graph

Motivated by these examples from social influence networks and wireless sensor networks, we define the *averaging system* to be the dynamical system

$$x(k + 1) = Ax(k), \tag{1.3}$$

where A has non-negative entries and unit row sums. Here, k is the discrete-time variable.

Scientific questions of interest for the averaging model include:

(i) Does each node converge to a value? Is this value the same for all nodes?

(ii) Is this value equal to the average of the initial conditions? In other words, when do the agents achieve *average consensus*?

(iii) What properties do the graph and the corresponding matrix need to have in order for the algorithm to converge?

(iv) How quick is the convergence?

1.3   Flocking dynamics in animal behavior

Next, we draw inspiration from biology and we consider swarming and flocking behavior that many animal species exhibit, e.g., see Figure 1.4.

(a) A flock of snow geese (*Chen caerulescens*). Public domain image from the U.S. Fish and Wildlife Service.
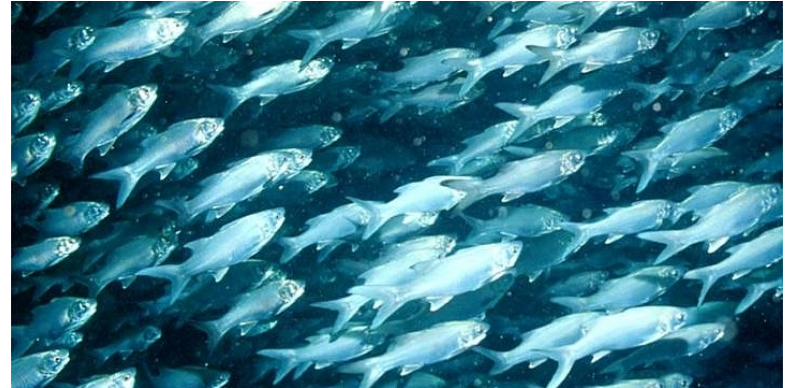
(b) A swarm of pacific threadfins (*Polydactylus sexfilis*). Public domain image from the U.S. National Oceanic and Atmospheric Administration.

Figure 1.4: Examples of animal flocking behaviors

To model this behavior as arising from decentralized interactions, we consider a simple "alignment rule" for each animal to steer towards the average heading of its neighbors; see Figure 1.5. This alignment rule amounts to a "spring-like" attractive force,
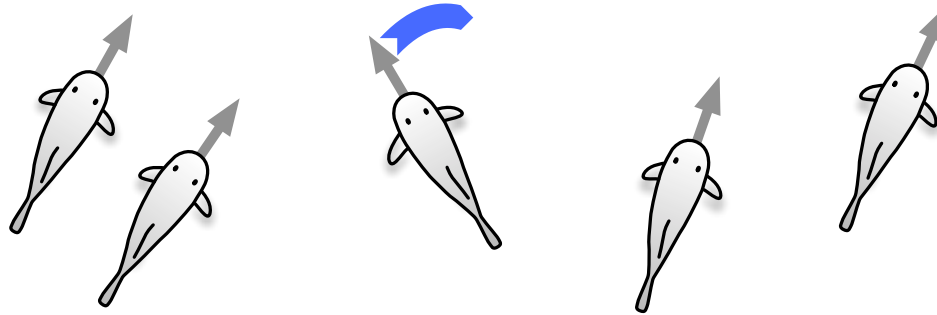
Figure 1.5: Alignment rule: the center fish rotates clockwise to align itself with the average heading of its neighbors.

described as follows:

$$
\dot{\theta}_i = \begin{cases} (\theta_j - \theta_i), & \text{if } i\text{th animal has one neighbor} \\ \frac{1}{2}(\theta_{j_1} - \theta_i) + \frac{1}{2}(\theta_{j_2} - \theta_i), & \text{if } i\text{th animal has two neighbors} \\ \frac{1}{m}(\theta_{j_1} - \theta_i) + \cdots + \frac{1}{m}(\theta_{j_m} - \theta_i), & \text{if } i\text{th animal has } m \text{ neighbors} \end{cases}
$$
$$
= \operatorname{average}\big(\{\theta_j, \text{ for all neighbors } j\}\big) - \theta_i, \tag{1.4}
$$

where we are assuming that each animal is a node with edges sensing the heading of other animals.

We can now proceed as before and define an averaging matrix A exactly as in the wireless sensor network example. Before proceeding, it is now customary to define a new matrix, called the *Laplacian matrix*, by

$$L = \mathrm{diag}(A\mathbb{1}_n) - A. \tag{1.5}$$

The interaction law (1.4) can be written the continuous-time averaging system:

$$\dot{\theta} = (A - I_n)\theta = -L\theta. \tag{1.6}$$

This dynamical system is usually referred to as *Laplacian flow*.

Note: this incomplete model does not concern itself with positions. In other words, we do not discuss collision avoidance and formation/cohesion maintenance. Moreover, the interaction pairs should be really state dependent. For example, we may assume that two animals see each other and interact if and only if their pairwise Euclidean distance is below a certain threshold. Finally, note that it is mathematically ill-posed to compute averages on a circle. For now, we will not worry about these matters.

Scientific questions of interest for this continuous-time averaging system are similar to those in the last two sections:

(i) how valid is this model in understanding and reproducing animal behavior?

(ii) what are equilibrium headings and when are they attractive?

(iii) what properties does the graph need to have to ensure a proper flocking behavior?

1.4 Dynamical flow systems in ecosystems

Dynamical flow systems, also called *compartmental systems*, model dynamical processes characterized by conservation laws and by the flow of material and commodities between units known as compartments. Dynamical flow systems are widespread in engineering applications; commodity flows among compartments include power, energy and water/gas networks, data routing and communication networks, traffic networks, and logistic networks.

In this section we introduce dynamical flow systems evolving in discrete and continuous time. We consider a low-dimensional schematic example from the study of ecosystems, where living and non-living components interact through nutrient cycles (water, nitrates, phosphates, etc) and energy flows. Specifically, we consider the widely-cited *water flow model for a desert ecosystem* (Noy-Meir, 1973), depicted in Figure 1.6. As illustrated in figure, each node of a dynamical flow network is a called a compartment and functions as a storage element for the commodity (i.e., water in this case). Each edge describes a flow of the commodity, including edges describing inflows from the environment and edges describing outflows into the environment.
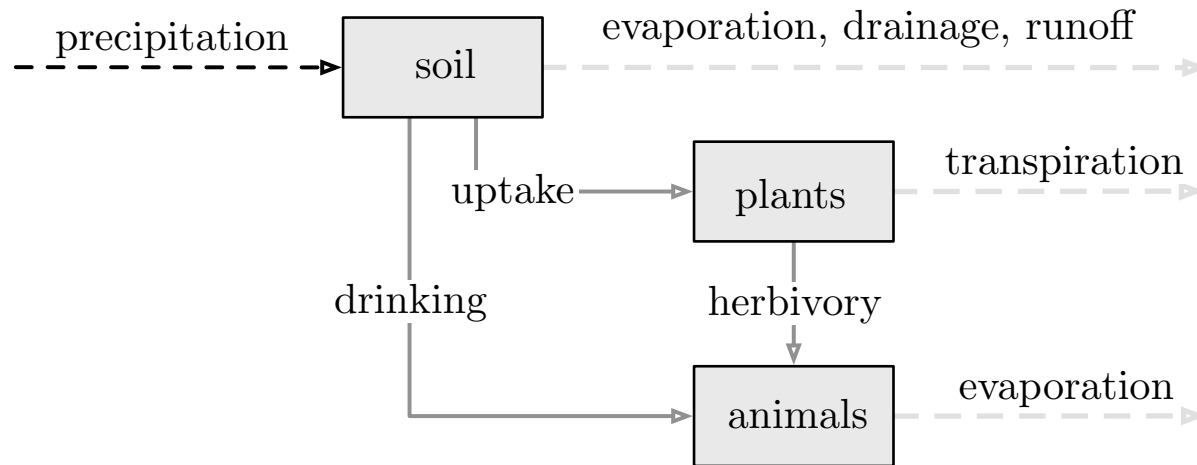
Figure 1.6: The Noy-Meir water flow model for a desert ecosystem. The black dashed line denotes an inflow from the outside environment. Each compartment functions as storage unit. The light-gray dashed lines denote outflows into the outside environment.

1.4.1   Discrete-time model

Given n interconnected compartments, e.g., as depicted in Figure 1.6, let

- $q_i(k)$ denote the quantity of commodity at compartment i at the discrete time $k \in \mathbb{N}$,

- a_{ij}, called *routing fractions* (or *split ratios* in traffic networks), denote the fraction of commodity at compartment i flowing to compartment j during one time-step, and

- $u_i \geq 0$ denote a non-negative supply to compartment i.

It is convenient to collect the routing fractions $a_{ij} \geq 0$ into a so-called *routing matrix* $A \in \mathbb{R}_{\geq 0}^{n \times n}$. With these definitions, we note the total commodity flowing at time k from i to j is $a_{ij}q_i(k)$. Writing the discrete-time mass balance equation for each compartment, we obtain the *discrete-time dynamical flow system*:

$$q_i(k+1) = \sum_{j=1}^{n} a_{ji}q_j(k) + u_i \qquad \Longleftrightarrow \qquad q(k+1) = A^\mathsf{T}q(k) + u. \tag{1.7}$$

Because the mass $q_i(k)$ either remains in the system or outflows into the environment at time $k+1$, we know that $q_i(k)$ is equal to $\sum_{j=1}^{n} a_{ij}q_i(k)$ plus a zero or positive outflow. Therefore, the ith row-sum of A satisfies:

$$\sum_{j=1}^{n} a_{ij} < 1 \quad \text{if and only if} \quad \text{compartment } i \text{ has an outflow into the environment,}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \quad \text{if and only if} \quad \text{compartment } i \text{ has no outflows into the environment.}$$

In summary, in general *open systems* with outflows, the routing matrix A is *row-substochastic*, that is, A is non-negative and at least one of its rows has sum strictly less than 1. In *closed systems* without outflows, the routing matrix A is row-stochastic, as in averaging models we reviewed earlier in the chapter.

For example, we report the routing matrix for the Noy-Meir water flow model in Figure 1.6. First, we let q_1, q_2, q_3 denote the quantity of water in the soil compartment, the plants compartment and the animals compartment, respectively. Second, we let $a_{\text{e-d-r}}$, $a_{\text{trnsp}}, a_{\text{evap}}, a_{\text{drnk}}, a_{\text{uptk}}, a_{\text{herb}}$, denote the routing fractions for, respectively, evaporation-drainage-runoff, transpiration, evaporation, drinking, uptake, and herbivory. This notation means that, for example, the water $q_1(k)$ in the soil at time k is split four-ways at time $k+1$: $a_{\text{e-d-r}}q_1(k)$ evaporates, $a_{\text{uptk}}q_1(k)$ is uptaken by plants, $a_{\text{drnk}}q_1(k)$ is drank by animals, and the remainder remains in the soil. Simple book-keeping leads to the routing matrix:

$$
A_{\text{Noy-Meir}} = \begin{bmatrix} 1 - a_{\text{e-d-r}} - a_{\text{uptk}} - a_{\text{drnk}} & a_{\text{uptk}} & a_{\text{drnk}} \\ 0 & 1 - a_{\text{trnsp}} - a_{\text{herb}} & a_{\text{herb}} \\ 0 & 0 & 1 - a_{\text{evap}} \end{bmatrix}. \tag{1.8}
$$

Note that each diagonal term $(A_{\text{Noy-Meir}})_{ii}$ is the fraction of water in compartment i that remains in compartment i after one time-step. As predicted, $A_{\text{Noy-Meir}}$ is row-substochastic in general and row-stochastic precisely when $a_{\text{e-d-r}} = a_{\text{trnsp}} = a_{\text{evap}} = 0$.

1.4.2 Continuous-time models

We now present the continuous-time version of the notions in the previous section. Given n interconnected compartments, e.g., as depicted in Figure 1.6, let

- $q_i(t)$ denote the quantity of commodity at compartment i at the continuous time $t \in \mathbb{R}_{\geq 0}$,

- f_{ij} denote the *flow rates* of commodity at compartment i flowing to compartment j, and

- $u_i \geq 0$ denote a non-negative supply to compartment i.

In other words, we assume that

$$\text{the flow of commodity from } i \text{ to } j \text{ at time } t \in \mathbb{R}_{\geq 0} = f_{ij}q_i(t). \tag{1.9}$$

for a positive constant *flow rate* f_{ij}. As before, by writing the mass balance equation at each compartment, the *continuous-time dynamical flow system* is

$$\dot{q}_i(t) = \sum_{j=1, j \neq i}^{n} (f_{ji}q_j(t) - f_{ij}q_i(t)) - f_{0,i}q_i(t) + u_i, \tag{1.10}$$

where we let $f_{0,i}$ denote the outflow rate at compartment i into the environment.

It is convenient to collect the flow rates f_{ij} into a so-called *flow rate matrix* $F \in \mathbb{R}_{\geq 0}^{n \times n}$, with zero diagonal entries by convention. Then, as in the flocking example in Section 1.3, we define the Laplacian matrix $L = \text{diag}(F\mathbb{1}_n) - F$ and we claim that $\sum_{j=1, j \neq i}^{n} (f_{ji}q_j - f_{ij}q_i) = (-L^\mathsf{T}q)_i$; see Exercise E1.2. In turn, the continuous-time dynamical flow system (1.10) can be written as

$$\dot{q}(t) = Cq(t) + u. \tag{1.11}$$

where the *compartmental matrix* C is defined by $C = -L^\mathsf{T} - \text{diag}(f_0)$ and where $\text{diag}(f_0)$ is a diagonal matrix with diagonal entries equal to the outflow rates. The matrix C has non-negative off-diagonal entries and it is therefore a so-called *Metzler matrix*. Moreover, C has non-positive column sums, which will play an important role in understanding its properties.

Finally, we report the flow rate matrix F and the compartmental matrix C for the water flow model in Figure 1.6. Corresponding to each edge in figure, we let $f_{\text{e-d-r}}$, f_{trnsp}, f_{evap}, f_{drnk}, f_{uptk}, f_{herb}, denote the flow rates for, respectively, evaporation-drainage-runoff, transpiration, evaporation, drinking, uptake, and herbivory. With this notation, we can write

$$
F_{\text{Noy-Meir}} = \begin{bmatrix} 0 & f_{\text{uptk}} & f_{\text{drnk}} \\ 0 & 0 & f_{\text{herb}} \\ 0 & 0 & 0 \end{bmatrix}, \quad
L_{\text{Noy-Meir}} = \begin{bmatrix} f_{\text{uptk}} + f_{\text{drnk}} & -f_{\text{uptk}} & -f_{\text{drnk}} \\ 0 & f_{\text{herb}} & -f_{\text{herb}} \\ 0 & 0 & 0 \end{bmatrix}, \text{ and}
$$

$$
C_{\text{Noy-Meir}} = -L_{\text{Noy-Meir}}^{\mathsf{T}} - \operatorname{diag}\begin{bmatrix} f_{\text{e-d-r}} \\ f_{\text{trnsp}} \\ f_{\text{evap}} \end{bmatrix} = \begin{bmatrix} -f_{\text{e-d-r}} - f_{\text{uptk}} - f_{\text{drnk}} & 0 & 0 \\ f_{\text{uptk}} & -f_{\text{trnsp}} - f_{\text{herb}} & 0 \\ f_{\text{drnk}} & f_{\text{herb}} & -f_{\text{evap}} \end{bmatrix}.
$$

1.4.3 Summary

We conclude by summarizing the model presented. First, the discrete-time dynamical flow model is

$$q(k+1) = A^{\mathsf{T}}q(k) + u(k), \tag{1.12}$$

where the routing matrix A is row-substochastic (or row-stochastic) and the supply vector u is typically non-negative. Second, the continuous-time dynamical flow model is

$$\dot{q}(t) = \big(-L^{\mathsf{T}} - \mathrm{diag}(f_0)\big)q(t) + u(t), \tag{1.13}$$

where L is a Laplacian matrix, the outflow vector f_0 is non-negative, and the supply vector u is non-negative. Note: this section has focused on linear models. We remark that various nonlinearities arise in important engineering applications; their modeling and analysis is postponed.

For both discrete and continuous-time flow systems, scientific questions of interest include:

(i) for constant inflows u, does the total mass in the system remain bounded?

(ii) for constant inflows u, is there a single (or multiple) final mass distribution among the nodes? In other words, does an equilibrium for the dynamics exist?

(iii) if an equilibrium exists, do all solutions converge to it?

(iv) does the mass at some nodes vanish asymptotically?

1.5   Appendix: Markov chains

In this section we provide a very basic review of (finite-dimensional) Markov chains and random walks over graphs. We consider both discrete-time and continuous-time models and show that the dynamics of the location probabilities for the random walk are special cases of dynamical flow systems.

Discrete-time Markov chains   Consider a discrete-time random walk on a graph, i.e., a sequence of locations on the graph selected in the following random fashion. At time k, we let

$$x_i(k) = \mathbb{P}[\text{location}(k) = i]. \tag{1.14}$$

We assume the *Markovian property* for this stochastic process: transition probabilities are independent of history and of time.
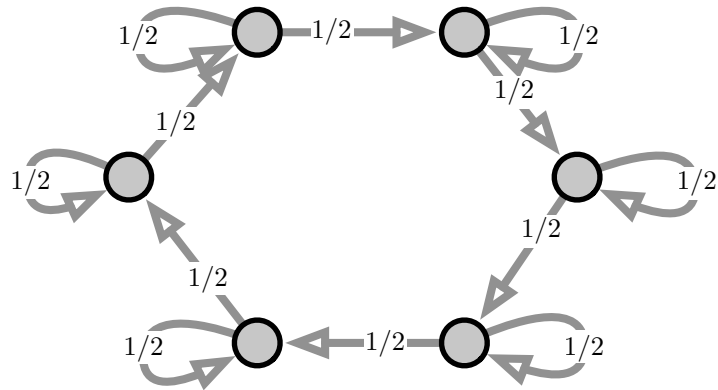
Figure 1.7: An example random walk on a graph: a random walker moves across the graph obeying a simple Markovian rule: the walker remains at its current node with probability 50% and moves clockwise with probability 50%.

Therefore, at each time k, transitions from node i to other nodes are described by a constant probability vector $\{a_{ij}\}_j$. Specifically, we let $a_{ij} = \mathbb{P}[\text{transition event } i \to j]$ denote the probability of a transition from note i to note j. Using Bayes' Theorem we obtain

$$x_i(k+1) = \mathbb{P}[\text{location}(k+1) = i] \tag{1.15}$$

$$= \sum_{j=1}^{n} \underbrace{\mathbb{P}[\text{location}(k+1) = i \mid \text{location}(k) = j]}_{\text{transition event } j \to i} \times \mathbb{P}[\text{location}(k) = j] \tag{1.16}$$

$$= \sum_{j=1}^{n} a_{ji} x_j(k) \qquad \Longleftrightarrow \qquad x(k+1) = A^\mathsf{T} x(k). \tag{1.17}$$

Accordingly, the matrix A is referred to as the *transition matrix* of the Markov chain. As for discrete-time flow systems, A has row sums equal to 1, that is, $\mathbb{1}_n^\mathsf{T} A^\mathsf{T} = \mathbb{1}_n^\mathsf{T}$ and this property immediately implies that the total mass $\mathbb{1}_n^\mathsf{T} x(k)$ is constant.

Continuous-time Markov chains   We next consider a continuous-time Markov chain model, where $x_i(t) = \mathbb{P}[\text{location}(t) = i]$, for $t \in \mathbb{R}$, instead of $k \in \mathbb{N}$. Here, the non-negative edge-weight a_{ji} denote the *rate of transition* from node j to node i. Clearly, rates are non-negative, but the sum of transition rates out of a given node does not necessarily sum to 1. In other words, if we let A denote the matrix containing transition rates (and set for simplicity $a_{ii} = 0$), then A is non-negative, but does not need to have unit row sums.

   Next, we derive the dynamics of a continuous-time Markov chain. First, we explain the probability of a transition over a short interval based on the transition rates:

$$\underbrace{\mathbb{P}[\text{location}(t+\tau) = i \mid \text{location}(t) = j]}_{\text{transition event } j \to i \text{ over duration } \tau} = \begin{cases} a_{ji}\tau + O(\tau^2), & \text{for } i \neq j, \\ 1 - \sum_{k \neq i} a_{ik}\tau + O(\tau^2), & \text{for } i = j. \end{cases} \tag{1.18}$$

Using again the Markovian assumption, we compute

$$x_i(t+\tau) = \sum_{j=1}^{n} \underbrace{\mathbb{P}[\text{location}(t+\tau) = i \mid \text{location}(t) = j]}_{\text{transition event } j \to i \text{ over duration } \tau} \times \mathbb{P}[\text{location}(t) = j]$$

$$= \left(1 - \sum_{k=1, k \neq i}^{n} a_{ik}\tau + O(\tau^2)\right)x_i(t) + \sum_{j=1, j \neq i}^{n} a_{ji}\tau x_j(t) + O(\tau^2),$$

and, in turn,

$$\frac{x_i(t+\tau) - x_i(t)}{\tau} + O(\tau) = -\left(\sum_{k=1}^{n} a_{ik}\right)x_i(t) + \sum_{j=1}^{n} a_{ji}x_j(t). \tag{1.19}$$

Finally, in the limit as the duration vanishes $\tau \to 0^+$, we obtain:

$$\dot{x}(t) = \left(-\operatorname{diag}(A^\mathsf{T}\mathbb{1}_n) + A^\mathsf{T}\right)x(t) := -L^\mathsf{T}x(t), \qquad \text{where } L = \operatorname{diag}(A\mathbb{1}_n) - A. \tag{1.20}$$

Here, the Laplacian matrix L is referred to as the *transition rate matrix* of the continuous-time Markov chain. As for continuous-time dynamical flow systems, L has zero row sums, that is, $\mathbb{1}^\mathsf{T}L^\mathsf{T} = \mathbb{0}_n^\mathsf{T}$ and this property implies that the total mass $\mathbb{1}_n^\mathsf{T}x(t)$ is constant.

1.6   Appendix: Robotic networks in cyclic pursuit and balancing

In this section we consider two simple examples of coordination motion in robotic networks. The standing assumption is that n robots, amicably referred to as "bugs," are placed and restricted to move on a circle of unit radius. Because of this bio-inspiration and because this language is common in the literature (Klamkin and Newman, 1971; Bruckstein et al., 1991), we refer to the following two problems as n-bugs systems.

On this unit circle the bugs' positions are angles measured counterclockwise from the positive horizontal axis. We let angles take value in $[0, 2\pi)$, that is, an arbitrary position θ satisfies $0 \leq \theta < 2\pi$. The bugs are numbered counterclockwise with identities $i \in \{1, \ldots, n\}$ and are at positions $\theta_1, \ldots, \theta_n$. It is convenient to identify $n+1$ with 1. We assume the bugs move in discrete times k in a counterclockwise direction by a controllable amount u_i (i.e., a control signal), that is:

$$\theta_i(k+1) = \mathrm{mod}(\theta_i(k) + u_i(k),\ 2\pi),$$

where $\mathrm{mod}(\vartheta,\ 2\pi)$ is the remainder of the division of ϑ by 2π and its introduction is required to ensure that $\theta_i(k+1)$ remains inside $[0, 2\pi)$.

1.6.1 Objective: optimal patrolling of a perimeter. Approach: Cyclic pursuit

We now suppose that each bug feels an attraction and moves towards the closest counterclockwise neighbor, as illustrated in Figure 1.8. Recall that the *counterclockwise distance from θ_i and θ_{i+1}* is the length of the counterclockwise arc from θ_i and θ_{i+1} and satisfies:

$$\text{dist}_{\text{cc}}(\theta_i, \theta_{i+1}) = \text{mod}(\theta_{i+1} - \theta_i,\ 2\pi),$$

In short, given a *control gain* $\kappa \in [0, 1]$, we assume that the ith bug sets adopts the *cyclic pursuit control law*

$$u_{\text{pursuit},i}(k) = \kappa \, \text{dist}_{\text{cc}}(\theta_i(k), \theta_{i+1}(k)).$$

(a) Cyclic pursuit control law   (b) Cyclic balancing control law
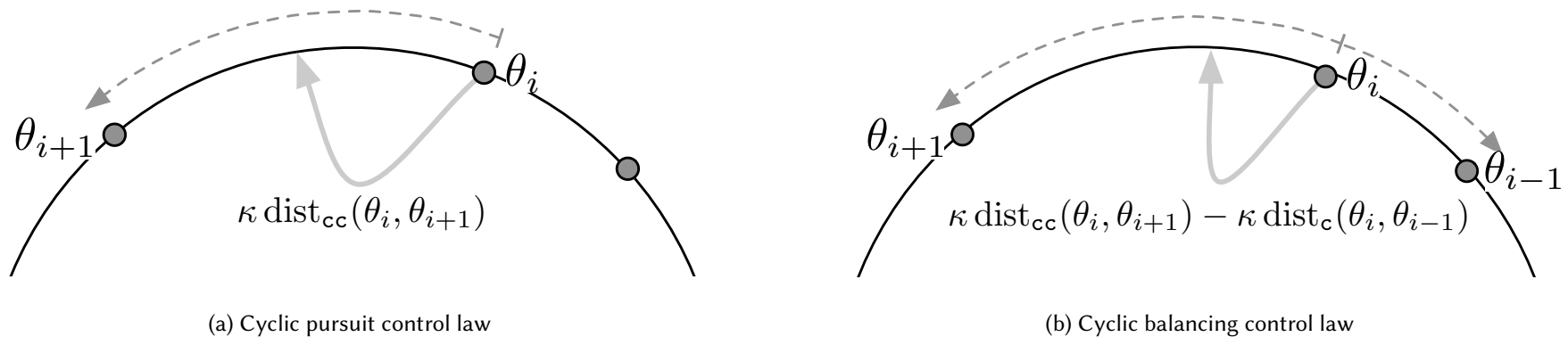
Figure 1.8: Cyclic pursuit and balancing are prototypical n-bug problems.

Scientific questions of interest include:

(i) Does this system have any equilibrium?

(ii) Is a rotating equally-spaced configuration a solution? An equally-spaced angle configuration is one for which $\text{mod}(\theta_{i+1} - \theta_i,\ 2\pi) = \text{mod}(\theta_i - \theta_{i-1},\ 2\pi)$ for all $i \in \{1, \ldots, n\}$. Such configurations are sometimes called *splay states*.

(iii) For which values of κ do the bugs converge to an equally-spaced configuration and with what pairwise distance?

1.6.2   Objective: optimal sensor placement. Approach: Cyclic balancing

Next, we suppose that each bug feels an attraction towards both the closest counterclockwise and the closest clockwise neighbor, as illustrated in Figure 1.8. Given a "control gain" $\kappa \in [0, 1/2]$ and the natural notion of clockwise distance, the ith bug adopts the *cyclic balancing control law*

$$u_{\mathrm{balancing},i}(k) = \kappa \operatorname{dist_{cc}}(\theta_i(k), \theta_{i+1}(k)) - \kappa \operatorname{dist_c}(\theta_i(k), \theta_{i-1}(k)),$$

where $\operatorname{dist_c}(\theta_i(k), \theta_{i-1}(k)) = \operatorname{dist_{cc}}(\theta_{i-1}(k), \theta_i(k))$.

Questions of interest are:

(i) Is a static equally-spaced configuration a solution?

(ii) For which values of κ do the bugs converge to a static equally-spaced configuration?

(iii) Is it true that the bugs will approach an equally-spaced configuration and that each of them will converge to a stationary position on the circle?

1.6.3 Robotic networks: cyclic balancing (continued)

It is unrealistic (among other aspects of this setup) to assume that the bugs know the absolute position of themselves and of their neighbors. Therefore, it is interesting to rewrite the dynamical system in terms of pairwise distances between nearby bugs.

For $i \in \{1, \ldots, n\}$, we define the relative angular distances (the lengths of the counterclockwise arcs) $d_i = \mathrm{dist}_{cc}(\theta_i, \theta_{i+1}) \geq 0$. (We also adopt the usual convention that $d_{n+1} = d_1$ and that $d_0 = d_n$). The change of coordinates from $(\theta_1, \ldots, \theta_n)$ to $(d_1, \ldots, d_n)$ leads us to rewrite the cyclic pursuit and the cyclic balancing laws as:

$$u_{\mathrm{pursuit},i}(k) = \kappa d_i,$$
$$u_{\mathrm{balancing},i}(k) = \kappa d_i - \kappa d_{i-1}.$$

In this new set of coordinates, one can show that the cyclic pursuit and cyclic balancing systems are, respectively,

$$d_i(k+1) = (1-\kappa)d_i(k) + \kappa d_{i+1}(k), \tag{1.21}$$
$$d_i(k+1) = \kappa d_{i+1}(k) + (1-2\kappa)d_i(k) + \kappa d_{i-1}(k). \tag{1.22}$$

These are two linear time-invariant dynamical systems with state $d = (d_1, \ldots, d_n)$ and governing equation described by the two $n \times n$ matrices:

$$A_{\mathrm{pursuit}} = \begin{bmatrix} 1-\kappa & \kappa & \ldots & 0 & 0 \\ 0 & 1-\kappa & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 1-\kappa & \kappa \\ \kappa & 0 & \ldots & 0 & 1-\kappa \end{bmatrix}, \tag{1.23}$$

and

$$A_{\mathrm{balancing}} = \begin{bmatrix} 1-2\kappa & \kappa & \ldots & 0 & \kappa \\ \kappa & 1-2\kappa & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 1-2\kappa & \kappa \\ \kappa & 0 & \ldots & \kappa & 1-2\kappa \end{bmatrix}.$$

We conclude with the following remarks.

(i) Equations (1.21) and (1.22) are correct if the counterclockwise order of the bugs is never violated. One can show that this is true for $\kappa < 1$ in the pursuit case and $\kappa < 1/2$ in the balancing case; we leave this proof to the reader in Exercise E1.6.

(ii) The matrices A_{pursuit} and $A_{\text{balancing}}$, for varying n and κ, are called Toeplitz and circulant based on the nonzero/zero patterns of their entries; we study the properties of such matrices in later chapters. Moreover, they have non-negative entries for the stated ranges of κ and are row-stochastic.

(iii) If one defines the agreement space, i.e., $\{(\alpha, \alpha, \ldots, \alpha) \in \mathbb{R}^n \mid \alpha \in \mathbb{R}\}$, then each point in this set is an equilibrium for both systems.

(iv) It must be true for all times that $(d_1, \ldots, d_n) \in \{x \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 2\pi\}$. This property is indeed the consequence of the non-negative matrices A_{pursuit} and $A_{\text{balancing}}$ being *doubly-stochastic*, i.e., each row-sum and each column-sum is equal to 1.

(v) We will later study for which values of κ the system converges to the agreement space.

1.7 Appendix: Design problems in wireless sensor networks

In this appendix we show how averaging algorithms are relevant in wireless sensor network problems and can be used to tackle more sophisticated than what shown in Section 1.2.

1.7.1 Wireless sensor networks: distributed parameter estimation

The next two examples are also drawn from the field of wireless sensor network, but they feature a more advanced setup and require a basic background in estimation and detection theory, respectively. The key lessons to be learned from these examples is that it is useful to have algorithms that compute the average of distributed quantities.

Following ideas from (Xiao et al., 2005; Garin and Schenato, 2010), we aim to estimate an unknown parameter $\theta \in \mathbb{R}^m$ via the measurements taken by a sensor network. Each node $i \in \{1, \ldots, n\}$ measures

$$y_i = B_i \theta + v_i,$$

where $y_i \in \mathbb{R}^{m_i}$, B_i is a known matrix and v_i is random measurement noise. We assume that

(A1) the noise vectors $v_1, \ldots, v_n$ are independent jointly-Gaussian variables with zero-mean $\mathbb{E}[v_i] = \mathbb{0}_{m_i}$ and positive-definite covariance $\mathbb{E}[v_i v_i^\mathsf{T}] = \Sigma_i = \Sigma_i^\mathsf{T}$, for $i \in \{1, \ldots, n\}$; and

(A2) the measurement parameters satisfy: $\sum_i m_i \geq m$ and $\begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}$ is full rank.

Given the measurements $y_1, \ldots, y_n$, it is of interest to compute a least-square estimate of θ, that is, an estimate of θ that minimizes a least-square error. Specifically, we aim to minimize the following *weighted least-square error*:

$$\min_{\widehat{\theta}} \sum_{i=1}^n \big\| y_i - B_i \widehat{\theta} \big\|_{\Sigma_i^{-1}}^2 = \sum_{i=1}^n \big(y_i - B_i \widehat{\theta}\big)^\mathsf{T} \Sigma_i^{-1} \big(y_i - B_i \widehat{\theta}\big).$$

In this weighted least-square error, individual errors are weighted by their corresponding inverse covariance matrices so that an accurate (respectively, inaccurate) measurement corresponds to a high (respectively, low) error weight. With this particular choice of weights, the least-square estimate coincides with the so-called maximum-likelihood estimate; see (Poor, 1998) for more details. Under assumptions (A1) and (A2), the optimal solution is

$$\widehat{\theta}^* = \Big(\sum_{i=1}^n B_i^\mathsf{T} \Sigma_i^{-1} B_i \Big)^{-1} \sum_{i=1}^n B_i^\mathsf{T} \Sigma_i^{-1} y_i.$$

This formula is easy to implement by a single processor with all the information about the problem, i.e., the parameters and the measurements.

To compute $\widehat{\theta}^*$ in the sensor (and processor) network, we perform two steps:

[Step 1:] we run two distributed algorithms in parallel to compute the average of the quantities $B_i^\mathsf{T}\Sigma_i^{-1}B_i$ and $B_i^\mathsf{T}\Sigma_i^{-1}y_i$.

[Step 2:] we compute the optimal estimate via

$$\widehat{\theta}^* = \operatorname{average}\left(B_1^\mathsf{T}\Sigma_1^{-1}B_1, \ldots, B_n^\mathsf{T}\Sigma_n^{-1}B_n\right)^{-1} \operatorname{average}\left(B_1^\mathsf{T}\Sigma_1^{-1}y_1, \ldots, B_n^\mathsf{T}\Sigma_n^{-1}y_n\right).$$

Questions of interest are:

(i) How do we design algorithms to compute the average of distributed quantities?

(ii) What properties does the graph need to have in order for such an algorithm to exist?

(iii) How do we design an algorithm with fastest convergence?

1.7.2   Wireless sensor networks: distributed hypothesis testing

We consider a distributed hypothesis testing problem; inspired by (Rao and Durrant-Whyte, 1993; Olfati-Saber et al., 2006). Let h_γ, for $\gamma \in \Gamma$ in a finite set Γ, be a set of two or more hypotheses about an uncertain event. For example, given an area of interest, we could have:

$h_0 =$ "no target is present",
$h_1 =$ "one target is present", and
$h_2 =$ "two or more targets are present".

Suppose that we know the *a priori probabilities* $p(h_\gamma)$ of the hypotheses and that n nodes of a sensor network take measurements y_i, for $i \in \{1, \ldots, n\}$, related to the event. Independently of the type of measurements, assume you can compute

$$p(y_i|h_\gamma) = \text{probability of measuring } y_i \text{ given that } h_\gamma \text{ is the true hypothesis.}$$

Also, assume that each observation is conditionally independent of all other observations, given any hypothesis.

(i) We wish to compute the *maximum a posteriori estimate*, that is, we want to identify which one is the most likely hypothesis, given the measurements. Note that, under the independence assumption, Bayes' Theorem implies that the *a posteriori probabilities* satisfy

$$p(h_\gamma|y_1, \ldots, y_n) = \frac{p(h_\gamma)}{p(y_1, \ldots, y_n)} \prod_{i=1}^{n} p(y_i|h_\gamma).$$

(ii) Observe that $p(h_\gamma)$ is known, and $p(y_1, \ldots, y_n)$ is a constant normalization factor scaling all posteriori probabilities equally. Therefore, for each hypothesis $\gamma \in \Gamma$, we need to compute

$$\prod_{i=1}^{n} p(y_i|h_\gamma),$$

or equivalently, we aim to exchange data among the sensors in order to compute:

$$\exp\left(\sum_{i=1}^{n} \log(p(y_i|h_\gamma))\right) = \exp\left(n \text{ average}\left(\log p(y_1|h_\gamma), \ldots, \log p(y_n|h_\gamma)\right)\right).$$

(iii) In summary, even in this hypothesis testing problem, we need algorithms to compute the average of the n numbers $\log p(y_1|h_\gamma), \dots,$ for each hypothesis γ.

Questions of interest here are the same as in the previous section.

1.8   Appendix: List of examples and applications

These lecture notes focus on a rigorous understanding of dynamics phenomena over networks, drawing examples from numerous application domains. As a guide for an instructor or reader with a specific interest, a list of examples and application is included here.

Analysis of physical and natural systems:

Network System	Sections, Examples, and Exercises
Electric networks	static and dynamics models of resistive circuits in Sections 6.1.3 and 7.1.2, equilibrium analysis in Section 6.3.2, Thomson's principle and energy routing in Exercise E6.14, Kirchhoff's and Ohm's laws in Section 9.3, synchronization of inductors/capacitors circuits in Exercise E8.5, and resistive circuits as compartmental systems in Exercise E10.13
Mechanical networks	spring networks in Section 6.1.2, equilibrium analysis in Section 6.3.2, grounded spring networks in Exercise E6.15, spring networks on a ring in Section 14.2, and symmetric flow systems and hydraulic flow systems in Section 10.5.1
Social influence systems and network science	French-Harary-DeGroot model in Section 1.1 and it analysis in Chapter 5, Friedkin-Johnsen model in Exercise E5.26, Abelson model in Section 7.1.1, centrality measures in Section 5.5, community detection in Section 6.4, and hubs and authorities in Exercise E5.15
Animal behavior, population dynamics, and ecosystems	flocking behavior in Section 1.3, Noy-Meir model in Section 1.4 and their analysis in Chapter 10, Leslie population dynamics in Exercise E4.19, and Lotka-Volterra models and analysis in Chapter 16

Design of engineering systems:

Network System	Sections, Examples, and Exercises
Networked control systems	control design for clock synchronization in Section 6.5, control design for synchronization of diffusively-coupled linear systems in Section 8.4, second-order Laplacian flows in Section 8.1.1, distributed estimation from relative measurements in Section 9.5, averaging-based integral control in Exercises E6.18 and E9.6
Robotic networks	design of robotic coordination control laws for cyclic pursuit on the circle in Section 1.6 and on the plane in Exercise E1.7, for rendezvous in Example E2.18, for deployment and centering in Exercises E5.14 and E7.10
Power networks	the analysis of coupled oscillators in Chapter 17 and of second-order Laplacian flows in Section 8.1.1
Parallel and scientific computation	Jacobi relaxation in Exercises E2.15 and E2.16, parallel averaging in Exercise E5.11, discretization of PDEs in Section 7.1.3, discretization of the Laplace PDE in Example E10.14, accelerated averaging in Section 11.5

1.9   Historical notes and further reading

Numerous other examples of multi-agent and large-scale interconnected systems can be found in the texts (Michel and Miller, 1977; Šiljak, 1978; Vidyasagar, 1981; Šiljak, 1991; Lakshmikantham et al., 1991; Wu, 2007; Ren and Beard, 2008; Bullo et al., 2009; Mesbahi and Egerstedt, 2010; Bai et al., 2011; Cristiani et al., 2014; Li and Duan, 2014; Fuhrmann and Helmke, 2015; Chen et al., 2015; Francis and Maggiore, 2016; Arcak et al., 2016; Porter and Gleeson, 2016; Fagnani and Frasca, 2017). Other, related, and instructive examples are presented in surveys such as (Martínez et al., 2007; Ren et al., 2007; Murray, 2007; Garin and Schenato, 2010). Textbooks, monographs and surveys on the broader and different theme of network science include (Newman, 2003; Boccaletti et al., 2006; Castellano et al., 2009; Easley and Kleinberg, 2010; Jackson, 2010; Newman, 2010; Spielman, 2017).

The opinion dynamics example in Section 1.1 is an illustration of the rich literature on social influence networks, starting with the early works by French Jr. (1956), Harary (1959), Abelson (1964), and DeGroot (1974). While the linear averaging model is by now known as the DeGroot model, the key ideas were already present in French Jr. (1956) and the main results (e.g., average consensus for doubly stochastic matrices) were already obtained by (Harary, 1959). Empirical evidence in support of the averaging model (including its variations) is described in (Friedkin and Johnsen, 2011; Friedkin et al., 2016; Chandrasekhar et al., 2020). An outstanding tutorial and survey on dynamic social networks is (Proskurnikov and Tempo, 2017). We postpone to Chapter 10 the literature review on compartmental systems.

The n-bugs problem is related to the study of pursuit curves and inquires about what the paths of n bugs are when they chase one another. We refer to (Klamkin and Newman, 1971; Watton and Kydon, 1969; Bruckstein et al., 1991; Marshall et al., 2004; Smith et al., 2005) for some classic works, surveys, and recent results.

1.10   Exercises

▶ E1.1   **Basic properties of averaging systems.** Given a row-stochastic matrix $A \in \mathbb{R}^{n \times n}$, consider the *averaging system* (1.3)

$$x(k+1) = Ax(k).$$

Show that

(i) for all initial conditions $x(0) \in \mathbb{R}^n$, all times $k \in \mathbb{N}$, and all indices $i \in \{1, \ldots, n\}$,

$$\min_i x_i(0) \leq \min_i x_i(k) \leq \min_i x_i(k+1) \leq \max_i x_i(k+1) \leq \max_i x_i(k) \leq \max_i x_i(0);$$

(ii) for all $x \in \mathbb{R}^n$, compute $y = Ax$ and show

$$\max_i y_i - \min_i y_i \leq \max_i x_i - \min_i x_i; \qquad\qquad\qquad \text{(E1.1)}$$

(iii) any consensus configuration, i.e., any point in $\{\beta \mathbb{1}_n \in \mathbb{R}^n \mid \beta \in \mathbb{R}\}$, is an equilibrium point of the averaging system; and

(iv) there exist row-stochastic matrices A such that $Ax = x$ and x is not a consensus configuration.

▶ E1.2   **Basic equivalences for dynamical flow networks.** In this exercise we review a few basic equivalences for a nonlinear dynamical flow network; for simplicity of notation, we assume the system is closed (no inflow or outflows).

Given n compartments, let $F_{i \to j}(q)$ denote the flow from compartment i to compartment j as function of the state q. Let $F_{\text{flows}}(q)$ denote the matrix with (i, j) entries $F_{i \to j}(q)$. Let $F \in \mathbb{R}_{\geq 0}^{n \times n}$ denote a zero-diagonal flow rate matrix, as in Section 1.4.2. Show

(i) the nonlinear model in components and in vector form are equivalent:

$$\dot{q}_i = \sum_{j=1, j \neq i}^{n} (F_{j \to i}(q) - F_{i \to j}(q)) \qquad \Longleftrightarrow \qquad \dot{q} = F_{\text{flows}}(q)^{\mathsf{T}} \mathbb{1}_n - F_{\text{flows}}(q) \mathbb{1}_n,$$

(ii) the flow linearity assumption in components and in vector form are equivalent:

$$F_{i \to j}(q) = f_{ij} q_i \qquad \Longleftrightarrow \qquad F_{\text{flows}}(q) = \operatorname{diag}(q) F,$$

(iii) linear dynamical flow networks are characterized by a negative transpose Laplacian matrix, as follows:

$$\dot{q} = F_{\text{flows}}(q)^{\mathsf{T}} \mathbb{1}_n - F_{\text{flows}}(q) \mathbb{1}_n \qquad \Longleftrightarrow \qquad \dot{q} = (F^{\mathsf{T}} - \operatorname{diag}(F \mathbb{1}_n)) q = -L^{\mathsf{T}} q.$$

E1.3   **Basic properties of dynamical flow networks.** Given a flow rate matrix F and outflow rate vector f_0, consider the linear model (1.11), namely, $\dot{q} = Cq + u$, where $C = -L^\mathsf{T} - \mathrm{diag}(f_0)$ and $L = \mathrm{diag}(F\mathbb{1}_n) - F$. Perform the following tasks:

    (i)   show that, if there are no inflows, i.e., if $u_i = 0$ for all i, then the total mass in the system does not increase with time,

   (ii)   write a formula for the diagonal and off-diagonal entries of the compartmental matrix C as a function of the flow rate constants, and

  (iii)   show that the column sums of C are non-positive.

E1.4   **Constants of motion.** In the study of mechanics, energy and momentum are two constants of motion, that is, these quantities are constant along each evolution of the mechanical system. Show that

(i) if A is a row stochastic matrix with $w^{\mathsf{T}}A = w^{\mathsf{T}}$, then $w^{\mathsf{T}}x(k) = w^{\mathsf{T}}x(0)$ for all times $k \in \mathbb{Z}_{\geq 0}$ where $x(k+1) = Ax(k)$; and

(ii) if L is a Laplacian matrix with with $w^{\mathsf{T}}L = \mathbb{0}_n^{\mathsf{T}}$, then $w^{\mathsf{T}}x(t) = w^{\mathsf{T}}x(0)$ for all times $t \in \mathbb{R}_{\geq 0}$ where $\dot{x}(t) = -Lx(t)$.

E1.5   **Simulating the averaging dynamics.** Simulate in your favorite programming language and software package the linear averaging algorithm in equation (1.2). Set $n = 5$, select the initial state equal to $(-2, -1, 0, +1, +2)$, and use the following undirected unweighted graphs, depicted in Figure E1.1:

  (i) the complete graph,

  (ii) the cycle graph, and

  (iii) the star graph with node 1 as center.

Which value do all nodes converge to? Is it equal to the average of the initial values? Verify that the evolution of the averaging dynamics for the star graph is as in Figure E1.2. Turn in your code, a few printouts (as few as possible), and your written responses.

Complete graph   Cycle graph   Star graph
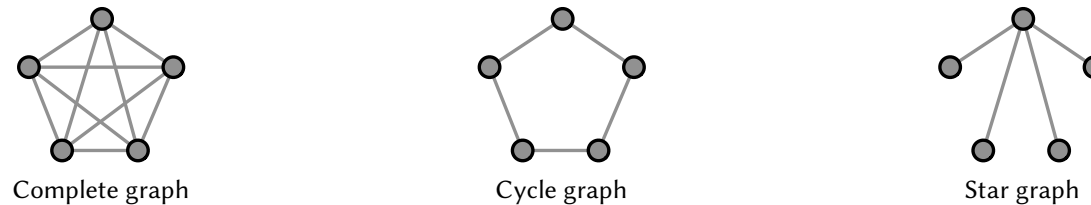
Figure E1.1: Three simple graphs
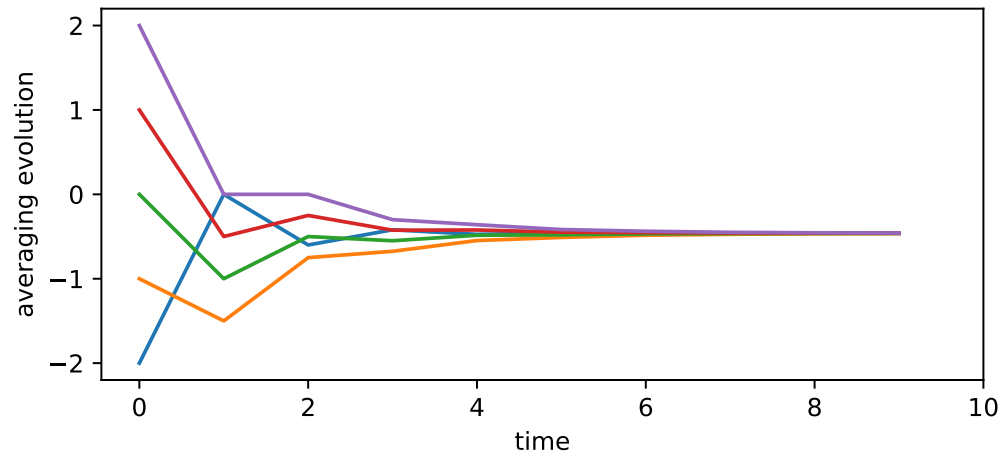
Figure E1.2: Linear averaging over a star graph: five distinct initial values converge to consensus.

E1.6   **Computing the bugs' dynamics.** Consider the cyclic pursuit and balancing dynamics described in Section 1.6. Verify

    (i) the cyclic pursuit closed-loop equation (1.21),

    (ii) the cyclic balancing closed-loop equation (1.22), and

    (iii) the counterclockwise order of the bugs is never violated.

Hint: Recall the distributive property of modular addition: $\mod(a \pm b, n) = \mod(\mod(a, n) \pm \mod(b, n), n)$.

E1.7   **Robotic coordination: continuous-time cyclic pursuit on the plane.** Consider four mobile robots on a plane with positions $p_i \in \mathbb{C}$, $i \in \{1, \ldots, 4\}$, and moving according to $\dot{p}_i = u_i$, where $u_i \in \mathbb{C}$ are the velocity commands. The task of the robots is rendezvous at a common point (while using only onboard sensors). A simple strategy to achieve rendezvous is *cyclic pursuit*: each robot i picks another robot, say $i + 1$, and pursues it. (Here $4 + 1 = 1$.) In other words, we set $u_i = p_{i+1} - p_i$ and obtain the closed-loop system (see also corresponding simulation below):

$$\begin{bmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \dot{p}_3 \\ \dot{p}_4 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}.$$
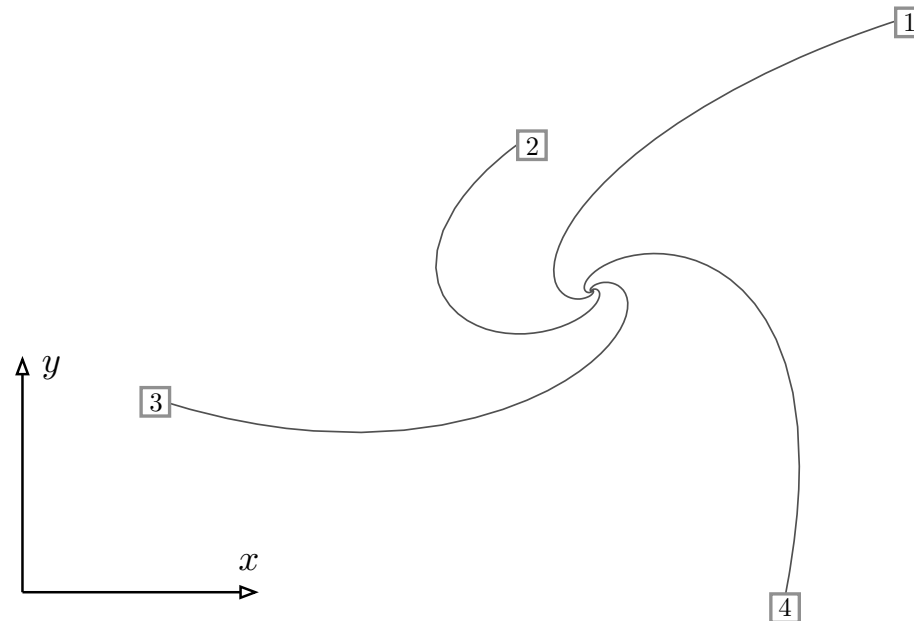
Prove that:

(i) the average robot position $\mathrm{average}(p(t)) = \sum_{i=1}^{4} p_i(t)/4$ remains constant for all $t \geq 0$;

(ii) the robots asymptotically rendezvous at the initial average robot position mass, that is,

$$\lim_{t \to \infty} p_i(t) = \mathrm{average}(p(0)) \quad \text{for } i \in \{1, \ldots, 4\};$$

(iii) if the robots are initially arranged in a square formation, then they remain in a square formation.

Hint: Given a matrix A with semisimple eigenvalues, the solution to $\dot{x} = Ax$ is given by the modal expansion $x(t) = \sum_{i=1}^{n} e^{\lambda_i t} v_i w_i^\mathsf{T} x(0)$, where v_i and w_i are the right and left eigenvectors associated to the eigenvalue λ_i and normalized to $w_i^\mathsf{T} v_i = 1$. The modal decomposition will be reviewed in Sections 2.1 and 11.1.

Chapter 2

Elements of Matrix Theory

In this chapter we review basic concepts from matrix theory with a special emphasis on the so-called Perron–Frobenius theory. These concepts will be useful when analyzing the convergence of the linear dynamical systems discussed in Chapter 1.

Notation

It is useful to start with some basic notations from matrix theory and linear algebra. We let $f \colon X \to Y$ denote a function from set X to set Y. We let $\mathbb{R}$, $\mathbb{N}$ and $\mathbb{Z}$ denote respectively the set of real, natural and integer numbers; also $\mathbb{R}_{\geq 0}$ and $\mathbb{Z}_{\geq 0}$ are the set of non-negative real numbers and non-negative integer numbers. For real numbers $a < b$, we let

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}, \qquad]a, b] = \{x \in \mathbb{R} \mid a < x \leq b\},$$
$$[a, b[= \{x \in \mathbb{R} \mid a \leq x < b\}, \qquad]a, b[= \{x \in \mathbb{R} \mid a < x < b\}.$$

Given a complex number $z \in \mathbb{C}$, its absolute value (sometimes referred to as modulus or magnitude) is denoted by $|z|$, its real part by $\Re(z)$ and its imaginary part by $\Im(z)$. We let i denote the imaginary unit $\sqrt{-1}$.

We let $\mathbb{1}_n \in \mathbb{R}^n$ (respectively $\mathbb{0}_n \in \mathbb{R}^n$) be the column vector with all entries equal to $+1$ (respectively 0). Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$, that is, e_i has all entries equal to zero except for the ith entry equal to 1. The 1-norm, 2-norm, and ∞-norm of a vector $x \in \mathbb{R}^n$ are defined by, respectively,

$$\|x\|_1 = |x_1| + \cdots + |x_n|, \qquad \|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}, \qquad \|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}.$$

We let I_n denote the n-dimensional identity matrix and $A \in \mathbb{R}^{n \times n}$ denote a square $n \times n$ matrix with real entries $\{a_{ij}\}$, $i, j \in \{1, \dots, n\}$. The matrix A is *symmetric* if $A^\mathsf{T} = A$.

For a matrix $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{C}$ is an *eigenvalue* and $v \in \mathbb{C}^n$ is a *right eigenvector*, or simply an *eigenvector*, if they together satisfy the eigenvalue equation $Av = \lambda v$. Sometimes it will be convenient to refer to (λ, v) as an *eigenpair*. A *left eigenvector* of the eigenvalue λ is a vector $w \in \mathbb{C}^n$ satisfying $w^\mathsf{T} A = \lambda w^\mathsf{T}$.

A symmetric matrix A is *positive definite*, denoted by $A \succ 0$ (resp. positive semidefinite, denoted by $A \succeq 0$) if all its eigenvalues are positive (resp. non-negative). We also let $A \prec 0$ and $A \preceq 0$ denote negative definite and negative semidefinite matrices. The *kernel* of A is the subspace $\mathrm{kernel}(A) = \{x \in \mathbb{R}^n \mid Ax = \mathbb{0}_n\}$, the *image* of A is $\mathrm{image}(A) = \{y \in \mathbb{R}^n \mid Ax = y, \text{ for some } x \in \mathbb{R}^n\}$, and the *rank* of A is the dimension of its image. Given vectors $v_1, \ldots, v_j \in \mathbb{R}^n$, their span is $\mathrm{span}(v_1, \ldots, v_j) = \{a_1 v_1 + \cdots + a_j v_j \mid a_1, \ldots, a_j \in \mathbb{R}\} \subset \mathbb{R}^n$.

2.1   Linear systems and the Jordan normal form

In this section we introduce a prototypical model for dynamical systems and study its stability properties via the so-called Jordan normal form, that is a key tool from matrix theory. We will later apply these results to the averaging model (1.3).

2.1.1 Discrete-time linear systems

We start with a basic definition.

Definition 2.1 (Discrete-time linear system). *A square matrix A defines a discrete-time linear system by*

$$x(k+1) = Ax(k), \quad x(0) = x_0, \tag{2.1}$$

or, equivalently by $x(k) = A^k x_0$, where the sequence $\{x(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ is called the solution, trajectory or evolution of the system.

Sometimes it is convenient to write $x^+ = f(x)$ to denote the system $x(k+1) = f(x(k))$.

We are interested in understanding when a solution from an arbitrary initial condition has an asymptotic limit as time diverges and to what value the solution converges. We formally define this property as follows.

Definition 2.2 (Semi-convergent and convergent matrices). *A matrix $A \in \mathbb{R}^{n \times n}$ is*

(i) *semi-convergent if $\lim_{k \to +\infty} A^k$ exists, and*

(ii) *convergent (also called Schur stable) if it is semi-convergent and $\lim_{k \to +\infty} A^k = \mathbb{0}_{n \times n}$.*

It is clear that, if A is semi-convergent with limiting matrix $A_\infty = \lim_{k \to +\infty} A^k$, then

$$\lim_{k \to +\infty} x(k) = A_\infty x_0.$$

In what follows we characterize the sets of semi-convergent and convergent matrices.

Remark 2.3 (Modal decomposition for symmetric matrices). *Before treating the general analysis method, we present the self-contained and instructive case of symmetric matrices. Recall that a symmetric matrix A has real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and corresponding orthonormal (i.e., orthogonal and unit-length) eigenvectors $v_1, \ldots, v_n$. Because the eigenvectors are an orthonormal basis for $\mathbb{R}^n$, we can write the* modal decomposition

$$x(k) = y_1(k)v_1 + \cdots + y_n(k)v_n,$$

where the ith normal mode *is defined by $y_i(k) = v_i^{\mathsf{T}} x(k)$. We then left-multiply the two equalities* (2.1) *by v_i^{T} and exploit $Av_i = \lambda_i v_i$ to obtain*

$$y_i(k+1) = \lambda_i y_i(k), \quad y_i(0) = v_i^{\mathsf{T}} x_0, \quad \implies \quad y_i(k) = \lambda_i^k (v_i^{\mathsf{T}} x_0).$$

In short, the evolution of the linear system (2.1) *is*

$$x(k) = \lambda_1^k (v_1^{\mathsf{T}} x_0) v_1 + \ldots + \lambda_n^k (v_n^{\mathsf{T}} x_0) v_n.$$

Therefore, each evolution starting from an arbitrary initial condition satisfies

(i) $\lim_{k \to \infty} x(k) = \mathbb{0}_n$ *if and only if $|\lambda_i| < 1$ for all $i \in \{1, \ldots, n\}$, and*

(ii) $\lim_{k \to \infty} x(k) = (v_1^{\mathsf{T}} x_0) v_1 + \cdots + (v_m^{\mathsf{T}} x_0) v_m$ *if and only if $\lambda_1 = \cdots = \lambda_m = 1$ and $|\lambda_i| < 1$ for all $i \in \{m+1, \ldots, n\}$.*

●

2.1.2   The Jordan normal form

In this section we review a very useful canonical decomposition of a square matrix. Recall that two $n \times n$ matrices A and B are *similar* if $B = TAT^{-1}$ for some invertible matrix T. Also recall that a similarity transform does not change the eigenvalues of a matrix.

Theorem 2.4 (Jordan normal form). *Each matrix $A \in \mathbb{C}^{n \times n}$ is similar to a block diagonal matrix $J \in \mathbb{C}^{n \times n}$, called the Jordan normal form of A, given by*

$$
J = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_m \end{bmatrix} \in \mathbb{C}^{n \times n},
$$

where each block J_i, called a Jordan block, is a square matrix of size j_i and of the form

$$
J_i = \begin{bmatrix} \lambda_i & 1 & \dots & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{j_i \times j_i}. \tag{2.2}
$$

Clearly, $m \leq n$ and $j_1 + \cdots + j_m = n$.

We refer to (Horn and Johnson, 1985) for a standard proof of this theorem. Note that the matrix J is unique, modulo a re-ordering of the Jordan blocks.

Regarding the eigenvalues of A, we note the following. The eigenvalues of J, and therefore also of A, are the (not necessarily distinct) complex numbers $\lambda_1, \dots, \lambda_m$. Given an eigenvalue λ,

(i) the *algebraic multiplicity* of λ is the sum of the sizes of all Jordan blocks with eigenvalue λ (or, equivalently, the multiplicity of λ as a root of the characteristic polynomial of A), and

(ii) the *geometric multiplicity* of λ is the number of Jordan blocks with eigenvalue λ (or, equivalently, the number of linearly-independent eigenvectors associated to λ).

An eigenvalue is

(i) *simple* if it has algebraic and geometric multiplicity equal precisely to 1, that is, a single Jordan block of size 1, and

(ii) *semisimple* if all its Jordan blocks have size 1, so that its algebraic and geometric multiplicity are equal.

Here is an example matrix in Jordan form and the multiplicities of its eigenvalues:

$$
\begin{bmatrix}
7 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 7 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 7 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 7 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 8 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 8 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 9
\end{bmatrix},
\quad
\begin{cases}
7 \text{ has algebraic mult. 4 and geometric mult. 2,} \\
\quad \text{so that } 7 \text{ is neither simple nor semisimple,} \\
8 \text{ has algebraic and geometric mult. 2, so it is semisimple,} \\
9 \text{ has algebraic and geometric mult. 1, so it is simple.}
\end{cases}
$$

Regarding the eigenvectors of A, Theorem 2.4 implies there exists an invertible matrix T such that

$$A = TJT^{-1} \tag{2.3}$$
$$\iff AT = TJ \tag{2.4}$$
$$\iff T^{-1}A = JT^{-1}. \tag{2.5}$$

Let $t_1, \ldots, t_n$ and $r_1, \ldots, r_n$ denote the columns and rows of T and T^{-1} respectively. If all eigenvalues of A are semisimple, then the equations (2.4) and (2.5) imply, for all $i \in \{1, \ldots, n\}$,

$$At_i = \lambda_i t_i \quad \text{and} \quad r_i A = \lambda_i r_i.$$

In other words, the ith column of T is the right eigenvector (or simply eigenvector) of A corresponding to the eigenvalue λ_i, and the ith row of T^{-1} is the corresponding left eigenvector of A.

If an eigenvalue is not semisimple, then it has larger algebraic than geometric multiplicity. For such eigenvalues, the columns of the matrix T are the right eigenvectors and the *generalized right eigenvectors* of A, whereas the rows of T^{-1} are the left eigenvectors and the *generalized left eigenvector* of A. For more details about generalized eigenvectors, we refer to reader to (Horn and Johnson, 1985).

Example 2.5 (Revisiting the wireless sensor network example). *As a numerical example, let us reconsider the wireless sensor network discussed in Section 1.2 and the 4-dimensional row-stochastic matrix A_{wsn}, which we report here for convenience:*

$$A_{\text{wsn}} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

With the aid of a symbolic mathematics program, we compute $A_{\text{wsn}} = TJT^{-1}$ where

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{24}(5 - \sqrt{73}) & 0 \\ 0 & 0 & 0 & \frac{1}{24}(5 + \sqrt{73}) \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 0 & -2 + 2\sqrt{73} & -2 - 2\sqrt{73} \\ 1 & 0 & -11 - \sqrt{73} & -11 + \sqrt{73} \\ 1 & -1 & 8 & 8 \\ 1 & 1 & 8 & 8 \end{bmatrix},$$

$$and \quad T^{-1} = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{96} + \frac{19}{96\sqrt{73}} & -\frac{1}{48} - \frac{5}{48\sqrt{73}} & \frac{1}{64} - \frac{3}{64\sqrt{73}} & \frac{1}{64} - \frac{3}{64\sqrt{73}} \\ -\frac{1}{96} - \frac{19}{96\sqrt{73}} & -\frac{1}{48} + \frac{5}{48\sqrt{73}} & \frac{1}{64} + \frac{3}{64\sqrt{73}} & \frac{1}{64} + \frac{3}{64\sqrt{73}} \end{bmatrix}.$$

Therefore, the eigenvalues of A are $1, 0, \frac{1}{24}(5 - \sqrt{73}) \approx -0.14$, and $\frac{1}{24}(5 + \sqrt{73}) \approx 0.56$. Corresponding to the eigenvalue 1, the right and left eigenvector equations are:

$$A_{\mathrm{wsn}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad and \quad \begin{bmatrix} 1/6 \\ 1/3 \\ 1/4 \\ 1/4 \end{bmatrix}^{\mathsf{T}} A_{\mathrm{wsn}} = \begin{bmatrix} 1/6 \\ 1/3 \\ 1/4 \\ 1/4 \end{bmatrix}^{\mathsf{T}} .$$

•

2.1.3   Semi-convergence and convergence for discrete-time linear systems

We can now use the Jordan normal form to study the powers of the matrix A. We start by computing

$$A^k = \underbrace{TJT^{-1} \cdot TJT^{-1} \cdot \ldots \cdot TJT^{-1}}_{k \text{ times}} = TJ^kT^{-1} = T \begin{bmatrix} J_1^k & 0 & \ldots & 0 \\ 0 & J_2^k & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & J_m^k \end{bmatrix} T^{-1},$$

so that, for a square matrix A with Jordan blocks J_i, $i \in \{1, \ldots, m\}$, the following statements are equivalent:

(i) A is semi-convergent (resp. convergent),

(ii) J is semi-convergent (resp. convergent), and

(iii) each block J_i is semi-convergent (resp. convergent).

Next, we compute the kth power of the generic Jordan block J_i with eigenvalue λ_i as a function of block size $1, 2, 3, \ldots, j_i$; they are, respectively,

$$
\begin{bmatrix} \lambda_i^k \end{bmatrix}, \quad
\begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} \\ 0 & \lambda_i^k \end{bmatrix}, \quad
\begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} & \binom{k}{2}\lambda_i^{k-2} \\ 0 & \lambda_i^k & k\lambda_i^{k-1} \\ 0 & 0 & \lambda_i^k \end{bmatrix},
$$

$$
\ldots, \quad
\begin{bmatrix}
\lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \binom{k}{2}\lambda_i^{k-2} & \cdots & \binom{k}{j_i-1}\lambda_i^{k-j_i+1} \\
0 & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \binom{k}{2}\lambda_i^{k-2} \\
0 & \cdots & 0 & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} \\
0 & \cdots & \cdots & 0 & \lambda_i^k
\end{bmatrix}, \quad (2.6)
$$

where the binomial coefficient $\binom{k}{m} = k!/(m!(k-m)!)$ satisfies $\binom{k}{m} \leq k^m/m!$. Note that, independently of the size of J_i, each entry of the kth power of J_i is upper bounded by a constant times $k^h \lambda_i^k$ for some non-negative integer h.

To study the limit as $k \to \infty$ of the generic block J_i^k, we study the limit as $k \to \infty$ of each term of the form $k^h \lambda_i^k$. Because exponentially-decaying factors dominate polynomially-growing terms, we know

$$\lim_{k \to \infty} k^h \lambda^k = \begin{cases} 0, & \text{if } |\lambda| < 1, \\ 1, & \text{if } \lambda = 1 \text{ and } h = 0, \\ \text{non-existent or unbounded}, & \text{if } (|\lambda| = 1 \text{ with } \lambda \neq 1) \text{ or } (|\lambda| > 1) \\ & \text{or } (\lambda = 1 \text{ and } h = 1, 2, \dots). \end{cases} \tag{2.7}$$

In summary, for each block J_i with eigenvalues λ_i, we can infer that:

  (i) a block J_i of size 1 is convergent if and only if $|\lambda_i| < 1$,

 (ii) a block J_i of size 1 is semi-convergent if and only if $|\lambda_i| < 1$ or $\lambda_i = 1$, and

(iii) a block J_i of size larger than 1 is semi-convergent and convergent if and only if $|\lambda_i| < 1$.
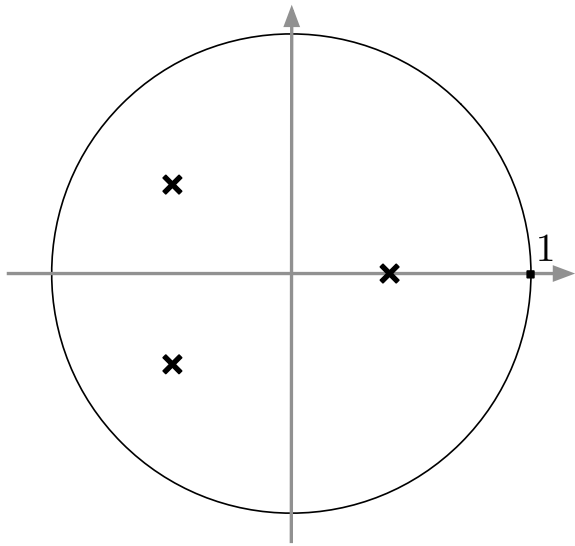
(a) The spectrum of a convergent matrix
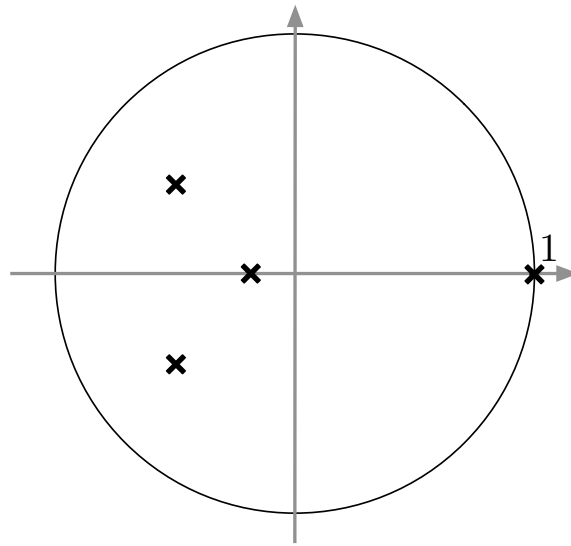
(b) The spectrum of a semi-convergent matrix, provided the eigenvalue 1 is semisimple.
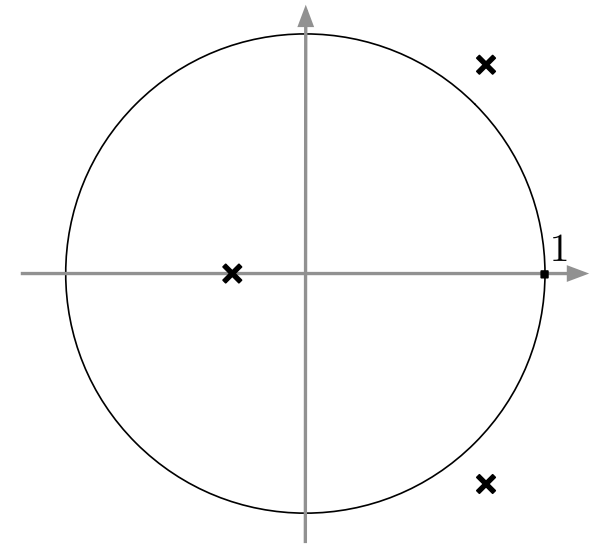
(c) The spectrum of a matrix that is not semi-convergent.

Figure 2.1: Eigenvalues and convergence properties of discrete-time linear systems

Based on this discussion, we are now ready to present necessary and sufficient conditions for semi-convergence and convergence of an arbitrary square matrix. We state these conditions using two useful definitions.

Definition 2.6 (Spectrum and spectral radius of a matrix). *Given a square matrix A,*

(i) *the spectrum of A, denoted $\mathrm{spec}(A)$, is the set of eigenvalues of A; and*

(ii) *the spectral radius of A is the maximum norm of the eigenvalues of A, that is,*

$$\rho(A) = \max\{|\lambda| \mid \lambda \in \mathrm{spec}(A)\},$$

or, equivalently, the radius of the smallest disk in $\mathbb{C}$ centered at the origin and containing the spectrum of A.

Theorem 2.7 (Convergence and spectral radius). *For a square matrix A, the following statements hold:*

(i) *A is convergent (i.e., $\lim_{k \to +\infty} A^k = \mathbb{0}_{n \times n}$) if and only if $\rho(A) < 1$,*

(ii) *A is semi-convergent and not convergent (i.e., $\lim_{k \to +\infty} A^k$ exists different from $\mathbb{0}_{n \times n}$) if and only if*

    (a) *1 is an eigenvalue,*

    (b) *1 is a semisimple eigenvalue, and*

    (c) *all other eigenvalues have magnitude less than 1.*

2.2   Row-stochastic matrices and their spectral radius

Motivated by the averaging model introduced in Chapter 1, we now consider in discrete-time linear systems defined by matrices with special properties. Specifically, we are interested in matrices with non-negative entries and whose row-sums are all equal to 1.
   The square matrix $A \in \mathbb{R}^{n \times n}$ is

  (i) *non-negative* (respectively *positive*) if $a_{ij} \geq 0$ (respectively $a_{ij} > 0$) for all i and j in $\{1, \dots, n\}$;

 (ii) *row-stochastic* if non-negative and $A\mathbb{1}_n = \mathbb{1}_n$;

(iii) *column-stochastic* if non-negative and $A^\mathsf{T}\mathbb{1}_n = \mathbb{1}_n$; and

(iv) *doubly-stochastic* if it is row- and column-stochastic.

   In the following, we write $A > 0$ and $v > 0$ (respectively $A \geq 0$ and $v \geq 0$) for a positive (respectively non-negative) matrix A and vector v.

Given a finite number of points $p_1, \ldots, p_n$ in $\mathbb{R}^n$, a *convex combination of $p_1, \ldots, p_n$* is a point of the form

$$\eta_1 p_1 + \eta_2 p_2 + \ldots + \eta_n p_n,$$

where the real numbers $\eta_1, \ldots, \eta_n$ satisfy $\eta_1 + \cdots + \eta_n = 1$ and $\eta_i \geq 0$ for all $i \in \{1, \ldots, n\}$. (For example, on the plane $\mathbb{R}^2$, the set of convex combinations of two distinct points is the segment connecting them and the set of convex combinations of three distinct points is the triangle (including its interior) defined by them.) The numbers $\eta_1, \ldots, \eta_n$ are called *convex combination coefficients* and each row of a row-stochastic matrix consists of convex combination coefficients.

2.2.1 The spectral radius for row-stochastic matrices

We now introduce a useful general method to localize the spectrum of a matrix and then use it to characterize the spectral radius of a row-stochastic matrix.

Theorem 2.8 (Geršgorin Disks Theorem). *For any square matrix $A \in \mathbb{R}^{n \times n}$,*

$$\mathrm{spec}(A) \subset \bigcup_{i \in \{1,\ldots,n\}} \underbrace{\left\{ z \in \mathbb{C} \;\middle|\; |z - a_{ii}| \leq \sum_{j=1, j \neq i}^{n} |a_{ij}| \right\}}_{\text{disk in the complex plane centered at } a_{ii} \text{ with radius } \sum_{j=1, j \neq i}^{n} |a_{ij}|} .$$

Proof. Consider the eigenvalue equation $Ax = \lambda x$ for the eigenpair (λ, x), where λ and $x \neq \mathbb{0}_n$ are in general complex. Choose the index $i \in \{1, \ldots, n\}$ so that

$$|x_i| = \max_{j \in \{1, \ldots, n\}} |x_j| > 0.$$

The ith component of the eigenvalue equation can be rewritten as

$$\lambda - a_{ii} = \sum_{j=1, j \neq i}^{n} a_{ij} x_j / x_i.$$

Now, take the complex magnitude of this equality and upper-bound its right-hand side:

$$|\lambda - a_{ii}| = \left| \sum_{j=1, j \neq i}^{n} a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j=1, j \neq i}^{n} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j=1, j \neq i}^{n} |a_{ij}|.$$

This inequality defines a set of the possible locations for the arbitrary eigenvalue λ of A. The statement follows by taking the union of such sets for each eigenvalue of A. ∎

Each disk in the theorem statement is referred to as a *Geršgorin disk*, or more accurately, as a *Geršgorin row disk*; an analogous disk theorem can be stated for Geršgorin column disks. Exercise E2.15 showcases an instructive application to distributed computing of numerous topics covered so far, including convergence notions and the Geršgorin Disks Theorem.

Lemma 2.9 (Spectral properties of a row-stochastic matrix). *For a row-stochastic matrix A,*

(i) 1 is an eigenvalue, and

(ii) $\mathrm{spec}(A)$ is a subset of the unit disk and $\rho(A) = 1$.

Proof. First, recall that A being row-stochastic is equivalent to two facts: $a_{ij} \geq 0$, $i, j \in \{1, \ldots, n\}$, and $A \mathbb{1}_n = \mathbb{1}_n$. The second fact implies that $\mathbb{1}_n$ is an eigenvector with eigenvalue 1. Therefore, by definition of spectral radius, $\rho(A) \geq 1$. Next, we prove that $\rho(A) \leq 1$ by invoking the Geršgorin Disks Theorem 2.8 to show that $\mathrm{spec}(A)$ is contained in the unit disk centered at the origin. The Geršgorin disks of a row-stochastic matrix as illustrated in Figure 2.2. Note that A being row-stochastic implies $a_{ii} \in [0, 1]$ and
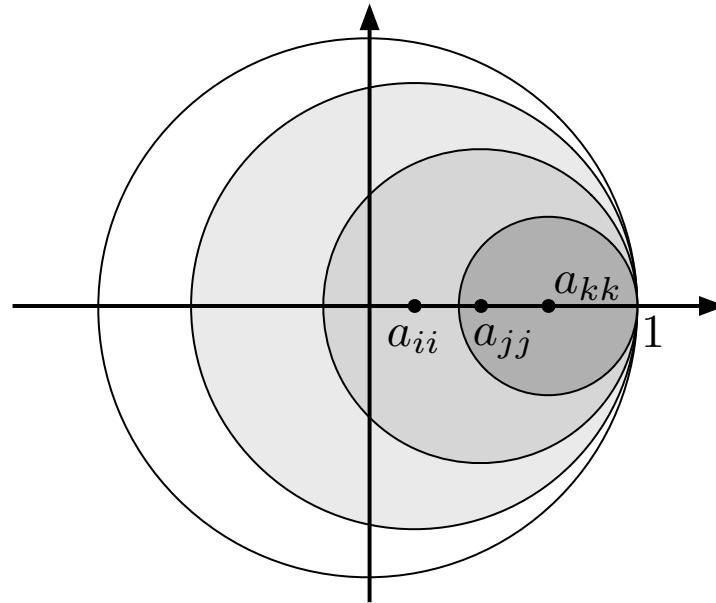
Figure 2.2: All Geršgorin disks of a row-stochastic matrix are contained in the unit disk.

$a_{ii} + \sum_{j \neq i} a_{ij} = 1$. Hence, the center of the ith Geršgorin disk belongs to the positive real axis between 0 and 1, and the right-most point in the disk is at 1.   ∎

Note: because 1 is an eigenvalue of each row-stochastic matrix A, clearly A is not convergent. But it is possible for A to be semi-convergent.

2.3   Perron–Frobenius theory

We have seen how row-stochastic matrices are not convergent; we now focus on characterizing those that are semi-convergent. To establish whether a row-stochastic matrix is semi-convergent, we introduce the widely-established Perron–Frobenius theory for non-negative matrices.

2.3.1 Classification of non-negative matrices

In the previous section we already defined non-negative and positive matrices. In this section we are interested in classifying non-negative matrices in terms of their zero/nonzero pattern and of the asymptotic behavior of their powers. We start by introducing simple example non-negative matrices and related comments in Table 2.1.

$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\quad$: $\operatorname{spec}(A_1) = \{1, 1\}$, the zero/nonzero pattern in A_1^k is constant, and $\lim_{k \to \infty} A_1^k = I_2$,

$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\quad$: $\operatorname{spec}(A_2) = \{1, -1\}$, the zero/nonzero pattern in A_2^k is periodic, and $\lim_{k \to \infty} A_2^k$ does not exist,

$A_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\quad$: $\operatorname{spec}(A_3) = \{0, 0\}$, the zero/nonzero pattern is $A_3^k = 0$ for all $k \geq 2$, and $\lim_{k \to \infty} A_3^k = 0$,

$A_4 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$ $\quad$: $\operatorname{spec}(A_4) = \{1, -1/2\}$, the zero/nonzero pattern is $A_4^k > 0$ for all $k \geq 2$, and $\lim_{k \to \infty} A_4^k = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$, and

$A_5 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ $\quad$: $\operatorname{spec}(A_5) = \{1, 1\}$, the zero/nonzero pattern in A_5^k is constant and $\lim_{k \to \infty} A_5^k$ is unbounded.

Table 2.1: Example 2-dimensional non-negative matrices and their properties

Based on these examples, we now introduce two sets of non-negative matrices with certain characteristic properties.

Definition 2.10 (Irreducible and primitive matrices). *For $n \geq 2$, an $n \times n$ non-negative matrix A is*

(i) *irreducible if $\sum_{k=0}^{n-1} A^k$ is positive,*

(ii) *primitive if there exists $k \in \mathbb{N}$ such that A^k positive.*

A matrix that is not irreducible is said to be reducible.

Note that A_1, A_3 and A_5 are reducible whereas A_2 and A_4 are irreducible. Moreover, note that A_2 is not primitive whereas A_4 is. Additionally, note that a positive matrix is clearly primitive. Finally, if there is $k \in \mathbb{N}$ such that A^k is positive, then (one can show that) all subsequent powers A^{k+1}, A^{k+2}, ... are necessarily positive as well; see Exercise E2.5.

Note: In other words, A is irreducible if, for any $(i, j) \in \{1, \ldots, n\}^2$ there is a $k = k(i, j) \leq (n-1)$ such that $(A^k)_{ij} > 0$. There are multiple equivalent ways to define irreducible matrices. We discuss four equivalent characterizations later in Theorem 4.3.

The following result is immediate consequences of the well-known Cayley-Hamilton Theorem.

Lemma 2.11 (A primitive matrix is irreducible). *If the $n \times n$ non-negative matrix A is primitive, then it is also irreducible.*

Proof. Assume by contradition that A is reducible so that $I_n + A + \cdots + A^{n-1}$ has at least one zero entry, say the entry i, j. Since $A \geq 0$, this implies that $(A^k)_{ij} = 0$ for each $k \in \{0, \ldots, n-1\}$. By the Caley-Hamilton Theorem, each matrix power A^h, $h \in \mathbb{N}$, is a linear combination of $I_n, A, \ldots, A^{n-1}$. But then also $(A^h)_{ij} = 0$, for each power $h \geq n$, which means that A is not primitive. ∎

As a consequence of this lemma we can draw the set diagram in Figure 2.3 describing the set of non-negative square matrices and its subsets of irreducible, primitive and positive matrices. Note that the inclusions in the diagram are strict in the sense that:

(i) A_3 is non-negative but not irreducible;

(ii) A_2 is irreducible but not primitive; and

(iii) A_4 is primitive but not positive.
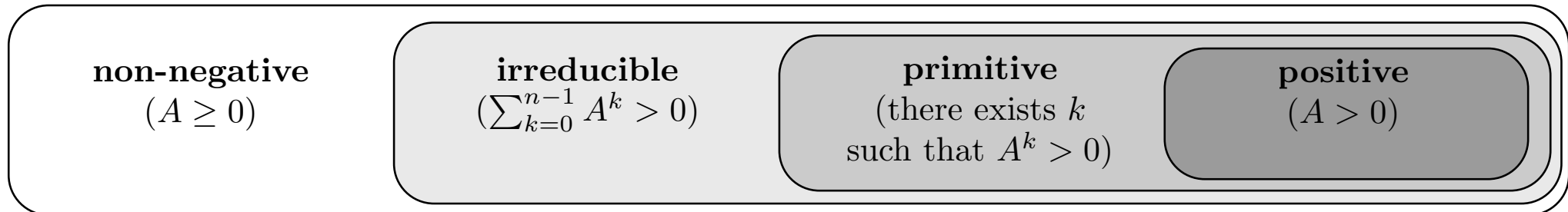
Figure 2.3: The set of non-negative square matrices and its subsets of irreducible, primitive and positive matrices.

2.3.2 Main results

We are now ready to state the main results in Perron–Frobenius theory and characterize the properties of the spectral radius of a non-negative matrix as a function of the matrix properties.

Theorem 2.12 (Perron–Frobenius Theorem). *Let $A \in \mathbb{R}^{n \times n}$, $n \geq 2$. If A is non-negative, then*

  (i) *there exists a real eigenvalue $\lambda \geq |\mu| \geq 0$ for all other eigenvalues μ,*

 (ii) *the right and left eigenvectors v and w of λ can be selected non-negative.*

If additionally A is irreducible, then

 (iii) *the eigenvalue λ is strictly positive and simple,*

 (iv) *the right and left eigenvectors v and w of λ are unique and positive, up to rescaling.*

If additionally A is primitive, then

 (v) *the eigenvalue λ satisfies $\lambda > |\mu|$ for all other eigenvalues μ.*

Some remarks and some additional statements are in order. For non-negative matrices, the real non-negative eigenvalue λ is the spectral radius $\rho(A)$ of A. We refer to λ as the *dominant eigenvalue* of A; it is also referred to as the *Perron eigenvalue*. We illustrate the results in the theorem in Figure 2.4.
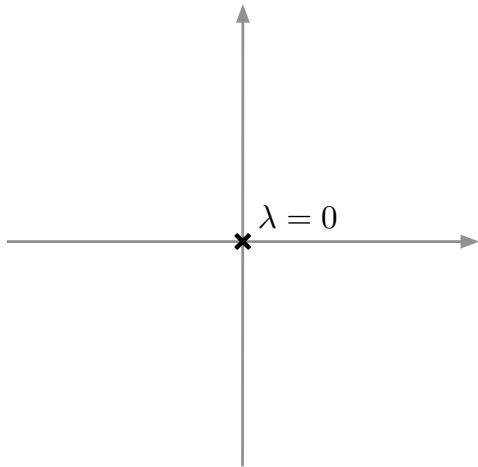
(a) The matrix A_3 is reducible: its dominant eigenvalue is 0 and so is its other eigenvalue.
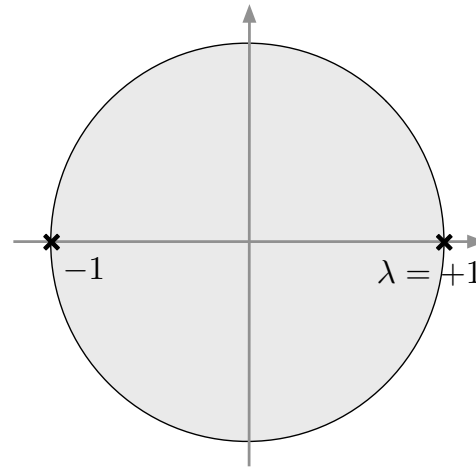
(b) The matrix A_2 is irreducible but not primitive: its dominant eigenvalues $+1$ is not strictly larger, in magnitude, than the other eigenvalue -1.
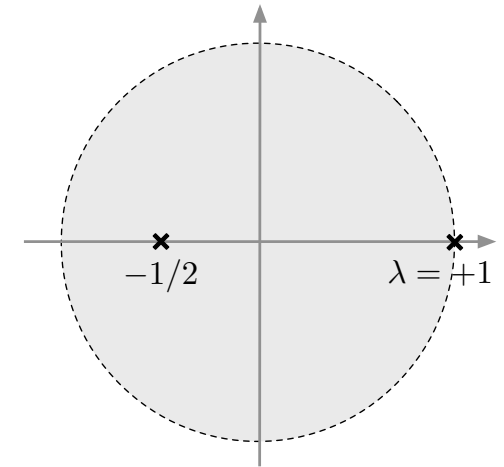
(c) The matrix A_4 is primitive: its dominant eigenvalue $+1$ is strictly larger than the other eigenvalue $-1/2$.

Figure 2.4: Spectra of non-negative matrices consistent with the Perron–Frobenius Theorem

For irreducible matrices, the right and left eigenvectors v and w (unique up to rescaling and selected non-negative) of the dominant eigenvalue λ are called the *right and left dominant eigenvector*, respectively. One can show that, up to rescaling, the right dominant eigenvector is the only positive right eigenvector of a primitive matrix A (a similar statement holds for the left dominant eigenvector); see also Exercise E2.4.

We refer to Theorem 4.11 in Section 4.5 for some useful bounds on the dominant eigenvalue and to Theorem 5.1 in Section 5.1 for a version of the Perron–Frobenius Theorem for reducible matrices.

2.3.3 Applications to matrix powers and averaging systems

Given a primitive non-negative matrix A, the Perron–Frobenius Theorem for primitive matrices has immediate consequences for the behavior of A^k as $k \to \infty$. We start with a semi-convergence result that applies to primitive matrices. We postpone the proof to Section 2.3.4.

Theorem 2.13 (Powers of non-negative matrices with a simple and strictly dominant eigenvalue). *Let A be a non-negative matrix. Assume the dominant eigenvalue λ is simple and strictly larger, in magnitude, than all other eigenvalues. Then A/λ is semi-convergent and*

$$\lim_{k\to\infty} \frac{A^k}{\lambda^k} = vw^\mathsf{T},$$

where v and w are the right and left dominant eigenvectors of A normalized so that $v \geq 0$, $w \geq 0$ and $v^\mathsf{T} w = 1$.

Note: The matrix vw^T is a rank-one projection matrix with numerous properties, which we discuss in Exercise E5.9.

We apply this theorem to a row-stochastic matrix A as arising in the French-Harary-DeGroot averaging model. For such a matrix, the dominant eigenvalue is $\lambda = 1$ and the corresponding right eigenvector is naturally selected to be $\mathbb{1}_n$. Therefore, if 1 is simple and strictly dominant, then

$$\lim_{k\to\infty} A^k = \mathbb{1}_n w^\mathsf{T} = \begin{bmatrix} w^\mathsf{T} \\ \vdots \\ w^\mathsf{T} \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & \dots & w_n \\ \vdots & \vdots & \vdots & \vdots \\ w_1 & w_2 & \dots & w_n \end{bmatrix},$$

where w is the left dominant eigenvector of A satisfying $w_1 + \cdots + w_n = 1$.

Example 2.14 (Revisiting the wireless sensor network example). *Let us reconsider the wireless sensor network discussed in Section 1.2 and the 4-dimensional row-stochastic matrix A_{wsn}. Recall that this matrix arises in the context of the French-Harary-DeGroot model averaging model. First, note that A_{wsn} is primitive because A_{wsn}^2 is positive:*

$$
A_{\text{wsn}} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \implies A_{\text{wsn}}^2 = \begin{bmatrix} 3/8 & 3/8 & 1/8 & 1/8 \\ 3/16 & 17/48 & 11/48 & 11/48 \\ 1/12 & 11/36 & 11/36 & 11/36 \\ 1/12 & 11/36 & 11/36 & 11/36 \end{bmatrix}.
$$

Therefore, the Perron–Frobenius Theorem 2.12 for primitive matrices applies to A_{wsn} and implies that the dominant eigenvalue 1 is simple and strictly dominant. Indeed, the four pairs of eigenvalues and right eigenvectors of A_{wsn} (as computed in Example 2.5) are $(1, \mathbb{1}_4)$ and

$$
\left(\frac{1}{24}(5 + \sqrt{73}), \begin{bmatrix} -2 - 2\sqrt{73} \\ -11 + \sqrt{73} \\ 8 \\ 8 \end{bmatrix} \right), \quad \left(\frac{1}{24}(5 - \sqrt{73}), \begin{bmatrix} 2(-1 + \sqrt{73}) \\ -11 - \sqrt{73} \\ 8 \\ 8 \end{bmatrix} \right), \quad \left(0, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right).
$$

Moreover, we know that A_{wsn} is semi-convergent. To apply the convergence results in Theorem 2.13, we compute its left dominant eigenvector, normalized to have unit sum, to be $w = [1/6, 1/3, 1/4, 1/4]^{\mathsf{T}}$ so that we have:

$$
\lim_{k \to \infty} A_{\text{wsn}}^k = \mathbb{1}_4 w^{\mathsf{T}} = \begin{bmatrix} 1/6 & 1/3 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/4 & 1/4 \\ 1/6 & 1/3 & 1/4 & 1/4 \end{bmatrix}.
$$

•

To fully understand the behavior of averaging systems and the properties of row-stochastic matrices, we study graph theory in Chapters 3 and 4 and postpone a comprehensive analysis to Chapter 5.

2.3.4   Selected proofs

We conclude this section with the proof of some selected statements.

Proof of Perron–Frobenius Theorem 2.12. We start by establishing that a primitive A matrix satisfies $\rho(A) > 0$. By contradiction, if $\operatorname{spec}(A) = \{0, \ldots, 0\}$, then the Jordan normal form J of A is nilpotent, that is, there is a $k^* \in \mathbb{N}$ so that $J^k = A^k = 0$ for all $k \geq k^*$. But this is a contradiction because A being primitive implies that there is $k^* \in \mathbb{N}$ so that $A^k > 0$ for all $k \geq k^*$.

Next, we prove that $\rho(A)$ is a real positive eigenvalue with a positive right eigenvector $v > 0$. We first focus on the case that A is a positive matrix, and later show how to generalize the proof to the case of primitive matrices. Without loss of generality, assume $\rho(A) = 1$. If (λ, x) is an eigenpair for A such that $|\lambda| = \rho(A) = 1$, then

$$|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x| = A|x| \quad \Longrightarrow \quad |x| \leq A|x|. \tag{2.8}$$

Here, we use the notation $|x| = (|x_i|)_{i \in \{1,\ldots,n\}}$, $|A| = \{|a_{ij}|\}_{i,j \in \{1,\ldots,n\}}$, and vector inequalities are understood component-wise. In what follows, we show $|x| = A|x|$. Adopting the notation $z = A|x|$ and $y = z - |x|$, equation (2.8) reads $y \geq 0$ and we aim to show $y = 0$. By contradiction, assume y has a non-zero component. Therefore, $Ay > 0$. Independently, we also know $z = A|x| > 0$. Thus, there must exist $\varepsilon > 0$ such that $Ay > \varepsilon z$. Eliminating the variable y in the latter equation, we obtain $A_\varepsilon z > z$, where we define $A_\varepsilon = A/(1+\varepsilon)$. The inequality $A_\varepsilon z > z$ implies $A_\varepsilon^k z > z$ for all $k > 0$. Now, observe that $\rho(A_\varepsilon) < 1$ so that $\lim_{k \to \infty} A_\varepsilon^k = \mathbb{0}_{n \times n}$ and therefore $0 > z$. Since we also knew $z > 0$, we now have a contradiction. Therefore, we know $y = 0$.

So far, we have established that $|x| = A|x|$, so that $(1, |x|)$ is an eigenpair for A. Also note that $A > 0$ and $x \neq 0$ together imply $A|x| > 0$. Therefore we have established that 1 is an eigenvalue of A with eigenvector $|x| > 0$. Next, observe that the above reasoning is correct also for primitive matrices if one replaces the first equality (2.8) by $|x| = |\lambda^k||x|$ and carries the exponent k throughout the proof.

In summary, we have established that there exists a real eigenvalue $\lambda > 0$ such that $\lambda \geq |\mu|$ for all other eigenvalues μ, and that each right (and therefore also left) eigenvector of λ can be selected positive up to rescaling. It remains to prove that λ is simple and is strictly greater than the magnitude of all other eigenvalues. For the proof of these two points, we refer to (Meyer, 2001, Chapter 8). ∎

Proof of Theorem 2.13. The proof is organized in three steps. First, because λ is simple, we write the Jordan normal form of A as

$$A = T \begin{bmatrix} \lambda & \mathbb{0}_{1\times(n-1)} \\ \mathbb{0}_{(n-1)\times 1} & B \end{bmatrix} T^{-1}, \tag{2.9}$$

where the block-diagonal matrix $B \in \mathbb{R}^{(n-1)\times(n-1)}$ contains the Jordan blocks of all eigenvalues of A except for λ. Because λ is strictly dominant, we know that $\rho(B/\lambda) < 1$, which in turn implies

$$\lim_{k\to+\infty} B^k/\lambda^k = \mathbb{0}_{(n-1)\times(n-1)}.$$

Recall $A^k = T \begin{bmatrix} \lambda & 0 \\ 0 & B \end{bmatrix}^k T^{-1}$ so that

$$\lim_{k\to+\infty} \left(\frac{A}{\lambda}\right)^k = T \left(\lim_{k\to+\infty} \begin{bmatrix} 1^k & 0 \\ 0 & (B/\lambda)^k \end{bmatrix} \right) T^{-1} = T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} T^{-1}. \tag{2.10}$$

Second, we let $v_1, \ldots, v_n$ (respectively, $w_1, \ldots, w_n$) denote the columns of T (respectively the rows of T^{-1}), that is, $T = [v_1 \ \ldots \ v_n]$, and $(T^{-1})^\mathsf{T} = [w_1 \ \ldots \ w_n]$. Equation (2.9) is equivalently written as

$$A \underbrace{\begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}}_{=T} = \underbrace{\begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}}_{=T} \begin{bmatrix} \lambda & 0 \\ 0 & B \end{bmatrix}.$$

The first column of the above matrix equation is $Av_1 = \lambda v_1$, that is, v_1 is the right dominant eigenvector v of A, up to rescaling. Recall that λ is simple so that its right eigenvector is unique up to rescaling. By analogous arguments, we find that w_1 is the left dominant eigenvector w of A, up to rescaling. With this notation, equation (2.10) leads to

$$\lim_{k \to +\infty} \left(\frac{A}{\lambda} \right)^k = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} w_1^\mathsf{T} \\ w_2^\mathsf{T} \\ \vdots \\ w_n^\mathsf{T} \end{bmatrix} = v_1 w_1^\mathsf{T}.$$

As third and final step, the $(1,1)$ entry of the matrix equality $T^{-1}T = I_n$ gives precisely the normalization $w_1^\mathsf{T} v_1 = 1$. In summary, v_1 and w_1 are the right and left dominant eigenvectors, up to rescaling, and they are known to satisfy $v_1^\mathsf{T} w_1 = 1$. Hence, $vw^\mathsf{T} = v_1 w_1^\mathsf{T}$. This concludes the proof of Theorem 2.13. ∎

2.4   Historical notes and further reading

For comprehensive treatments on matrix theory we refer to the classic texts by Gantmacher (1959), Horn and Johnson (1985), and Meyer (2001).

Regarding the main Perron–Frobenius Theorem 2.12, historically, Perron (1907) established the original result for the case of positive matrices. Frobenius (1912) provided the substantial extension to the settings of primitive and irreducible matrices. More historical information is given in (Meyer, 2001, Chapter 8).

Theorem 2.13 is generalized as follows: an irreducible row-stochastic matrix A with left dominant eigenvector w satisfies $\lim_{k\to\infty} \frac{1}{k}(I_n + A + \cdots + A^{k-1}) = \mathbb{1}_n w^\mathsf{T}$. We refer to (Meyer, 2001, Section 8.4) for more details on this result and to (Breiman, 1992, Chapter 6) for the more general Ergodic Theorem. (Essentially, (Hartfiel, 1998, Theorem 1.6) refers to Theorem 2.13 already as the Ergodic Theorem for Markov chains.)

2.5   Exercises

E2.1   **Simple properties of stochastic matrices.** Let $A_1, \ldots, A_k$ be $n \times n$ matrices, let $A_1 \ldots A_k$ be their product and let $\eta_1 A_1 + \cdots + \eta_k A_k$ be their convex combination with arbitrary convex combination coefficients. Show that

(i) if $A_1, \ldots, A_k$ are non-negative, then their product and all their convex combinations are non-negative,

(ii) if $A_1, \ldots, A_k$ are row-stochastic, then their product and all their convex combinations are row-stochastic, and

(iii) if $A_1, \ldots, A_k$ are doubly-stochastic, then their product and all their convex combinations are doubly-stochastic.

▶ E2.2   **Semi-convergence and Jordan block decomposition.** Consider a matrix $A \in \mathbb{C}^{n \times n}$, $n \geq 2$, with $\rho(A) = 1$. Show that the following statements are equivalent:

   (i) A is semi-convergent,
   (ii) either $A = I_n$ or there exists a nonsingular matrix $T \in \mathbb{C}^{n \times n}$ and a number $m \in \{1, \ldots, n-1\}$ such that

$$A = T \begin{bmatrix} I_m & \mathbb{0}_{m \times (n-m)} \\ \mathbb{0}_{(n-m) \times m} & B \end{bmatrix} T^{-1},$$

   where $B \in \mathbb{C}^{(n-m) \times (n-m)}$ is convergent, that is, $\rho(B) < 1$.

   Note: If A is real, then it is possible to find real-valued matrices T and B in statement (ii) by using the notion of real Jordan normal form (Hogben, 2013).

E2.3   **Row-stochastic matrices after pairwise-difference similarity transform.** For $n \geq 2$, let $A \in \mathbb{R}^{n \times n}$ be row stochastic. Define $T \in \mathbb{R}^{n \times n}$ by

$$T = \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ 1/n & 1/n & \cdots & 1/n \end{bmatrix}.$$

Perform the following tasks:

(i) for $x = [x_1, \ldots, x_n]^{\mathsf{T}}$, write Tx in components and show T is invertible,

(ii) show $TAT^{-1} = \begin{bmatrix} A_{\text{stable}} & \mathbb{0}_{n-1} \\ c^{\mathsf{T}} & 1 \end{bmatrix}$ for some $A_{\text{stable}} \in \mathbb{R}^{(n-1) \times (n-1)}$ and $c \in \mathbb{R}^{n-1}$,

(iii) if A is doubly-stochastic, then $c = 0$,

(iv) show that A primitive implies $\rho(A_{\text{stable}}) < 1$, and

(v) compute TAT^{-1} for $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

E2.4   **Uniqueness of the non-negative eigenvector in irreducible non-negative matrices.** Given a square matrix $A \in \mathbb{R}^{n \times n}$, show that:

(i) if v_1 is a right eigenvector of A corresponding to the eigenvalue λ_1, w_2 is a left eigenvector of A relative to λ_2, and $\lambda_1 \neq \lambda_2$, then $v_1 \perp w_2$; and

(ii) if A is non-negative and irreducible and $u \in \mathbb{R}_{\geq 0}^n$ is a right non-negative eigenvector of A, then u is an eigenvector corresponding to the eigenvalue $\rho(A)$.

Note: Statement (i) is sometimes referred to as the principle of bi-orthogonality.

▶ E2.5   **Powers of primitive matrices.** For any non-negative $A \in \mathbb{R}^{n \times n}$ and any numbers k and $m \in \mathbb{N}$, show that:

(i)   if the ith row of A^k is positive, then the ith row of $A^{k+m} > 0$,

(ii)   if the jth column of A^k is positive, then the jth column of $A^{k+m} > 0$,

(iii)   if $A^k > 0$, then $A^{k+m} > 0$.

E2.6   **Sufficient conditions for primitivity.** Let $A \in \mathbb{R}^{n \times n}$ be non-negative.

(i) Is the following statement true? If yes, explain why; if not, provide a counterexample.

If A has a zero entry, then A is reducible, because the zero entry can be moved to position $A_{n,1}$ via a permutation similarity transformation.

(ii) Show that A is primitive if there exists $r \in \{1, \ldots, n\}$ such that $A_{rj} > 0$ and $A_{ir} > 0$ for all $i, j \in \{1, \ldots, n\}$.

E2.7   **Some properties of doubly-stochastic matrices.** For $A \in \mathbb{R}^{n \times n}$, show that:

(i) if A is doubly-stochastic, then the matrix $A^\mathsf{T} A$ is doubly-stochastic and its spectrum satisfies $\operatorname{spec}(A^\mathsf{T} A) \subset [0, 1]$,

(ii) if A is doubly-stochastic and irreducible, then $A^\mathsf{T} A$ has positive diagonal and does not need to be irreducible (give a counterexample), and

(iii) if A is doubly-stochastic, irreducible and with positive diagonal, then $A^\mathsf{T} A$ is doubly-stochastic, irreducible and with positive diagonal.

Hint: Show that, if $\min_i a_{ii} = a_{\min} > 0$, then $A^\mathsf{T} A \geq a_{\min} A$.

E2.8   **On some non-negative matrices.** How many 2×2 matrices do there exist that are simultaneously doubly stochastic, irreducible and not primitive?

▶ E2.9 **Discrete-time affine systems.** Given $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$, consider the discrete-time affine system

$$x(k+1) = Ax(k) + b.$$

Assume A is convergent and show that

(i) the matrix $(I_n - A)$ is invertible,

(ii) the only equilibrium point of the system is $(I_n - A)^{-1}b$, and

(iii) $\lim_{k \to \infty} x(k) = (I_n - A)^{-1}b$ for all initial conditions $x(0) \in \mathbb{R}^n$.

Hint: Define a new sequence $y(k)$, $k \in \mathbb{Z}_{\geq 0}$, by $y(k) = x(k) - x^*$ for an appropriate x^*.

E2.10   **An affine averaging system.** Given a primitive doubly-stochastic matrix A and a vector b satisfying $\mathbb{1}_n^{\mathsf{T}} b = 0$, consider the *affine averaging system*

$$x(k+1) = Ax(k) + b.$$

Show that

(i) the quantity $\mathbb{1}_n^{\mathsf{T}} x(k)$ is constant for all k,

(ii) for each $\alpha \in \mathbb{R}$, there exists a unique equilibrium point x_α^* satisfying $\mathbb{1}_n^{\mathsf{T}} x_\alpha^* = \alpha$ and satisfying generically $x_\alpha^* \notin \mathrm{span}\{\mathbb{1}_n\}$, and

(iii) all solutions $\{x(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ satisfying $\mathbb{1}_n^{\mathsf{T}} x(0) = \alpha$ converge to x_α^*.

Hint: First, use Exercise E2.2 and study the properties of the similarity transformation matrix T and its inverse T^{-1}. Second, define $y(k) = T^{-1} x(k)$, show the evolution of $y_1(k)$ is decoupled from that of the other entries and apply E2.9.

▶ E2.11   **The Neumann series.** For $A \in \mathbb{C}^{n \times n}$, show that the following statements are equivalent:

(i) $\rho(A) < 1$,

(ii) $\lim_{k \to \infty} A^k = \mathbb{0}_{n \times n}$, and

(iii) the *Neumann series* $\sum_{k=0}^{\infty} A^k$ converges.

Additionally show that, if any and hence all of these conditions hold, then

(iv) the matrix $(I_n - A)$ is invertible, and

(v) $\displaystyle\sum_{k=0}^{\infty} A^k = (I_n - A)^{-1}$.

Hint: This statement is an extension of Theorem 2.7 and a generalization of the classic geometric series $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$, convergent for all $|x| < 1$. For the proof, the hint is to use the Jordan normal form.

E2.12 **Permutation and orthogonal matrices.** Consider the following three notions. A set G with a binary operation mapping two elements of G into another element of G, denoted by $(a, b) \mapsto a \star b$, is a *group* if:

(G1) $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in G$ (associativity property);

(G2) there exists $e \in G$ such that $a \star e = e \star a = a$ for all $a \in G$ (existence of an identity element); and

(G3) there exists $a^{-1} \in G$ such that $a \star a^{-1} = a^{-1} \star a = e$ for all $a \in G$ (existence of inverse elements).

A matrix $P \in \{0, 1\}^{n \times n}$ is a *permutation matrix* if each of its rows and each of its columns contains precisely only one entry equal to 1. A permutation matrix acts on a vector by permuting its entries.

A matrix $R \in \mathbb{R}^{n \times n}$ is an *orthogonal matrix* if $RR^{\mathsf{T}} = R^{\mathsf{T}}R = I_n$. In other words, the columns and rows of R are orthonormal vectors.

Prove that

(i) the set of $n \times n$ permutation matrices with the operation of matrix multiplication is a group;

(ii) the set of $n \times n$ orthogonal matrices with the operation of matrix multiplication is a group; and

(iii) each permutation matrix is orthogonal.

E2.13   **On doubly-stochastic and permutation matrices.** The following result is known as the Birkhoff – Von Neumann Theorem. For a matrix $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

  (i) A is doubly-stochastic; and

  (ii) A is a convex combination of permutation matrices.

Do the following:

  - show that the set of doubly-stochastic matrices is convex (i.e., given any two doubly-stochastic matrices A_1 and A_2, any matrix of the form $\lambda A_1 + (1 - \lambda)A_2$, for $\lambda \in [0, 1]$, is again doubly-stochastic);
  - show that (ii) $\implies$ (i);
  - find in the literature a proof of (i) $\implies$ (ii) and sketch it in one or two paragraphs.

E2.14   **Determinants of block matrices (Silvester, 2000).** Given square matrices $A, B, C, D \in \mathbb{R}^{n \times n}$, $n \geq 1$, useful identities are

$$
\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{cases} \det(D)\det(A - BD^{-1}C), & \text{if } D \text{ is invertible,} & \text{(E2.1a)} \\ \det(AD - BC), & \text{if } CD = DC, & \text{(E2.1b)} \\ \det(DA - BC), & \text{if } BD = DB. & \text{(E2.1c)} \end{cases}
$$

(i) Prove equality (E2.1a).

(ii) Prove equality (E2.1b) and (E2.1c) assuming D is invertible.

Hint: Show $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I_n & \mathbb{0}_{n \times n} \\ -D^{-1}C & I_n \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & B \\ \mathbb{0}_{n \times n} & D \end{bmatrix}$. We refer to (Silvester, 2000) for the complete proofs and for the additional identities

$$
\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{cases} \det(AD - CB), & \text{if } AC = CA, & \text{(E2.2a)} \\ \det(DA - CB), & \text{if } AB = BA. & \text{(E2.2b)} \end{cases}
$$

E2.15   **The Jacobi relaxation in parallel computation.** Consider n distributed processors that aim to collectively solve the linear equation $Ax = b$, where $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is invertible and its diagonal elements a_{ii} are nonzero. Each processor stores a variable $x_i(k)$ as the discrete-time variable k evolves and applies the following iterative strategy termed *Jacobi relaxation*. At time step $k \in \mathbb{N}$ each processor performs the local computation

$$x_i(k+1) = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^{n} a_{ij} x_j(k) \right), \quad i \in \{1, \dots, n\}.$$

Next, each processor $i \in \{1, \dots, n\}$ sends its value $x_i(k+1)$ to all other processors $j \in \{1, \dots, n\}$ with $a_{ji} \neq 0$, and they iteratively repeat the previous computation. The initial values of the processors are arbitrary.

(i)   Assume the Jacobi relaxation converges, i.e., assume $\lim_{k \to \infty} x(k) = x^*$. Show that $Ax^* = b$.

(ii) Give a necessary and sufficient condition for the Jacobi relaxation to converge.

(iii) Use Geršgorin Disks Theorem 2.8 to show that the Jacobi relaxation converges if the matrix A is *strictly row diagonally dominant*, that is, if $|a_{ii}| > \sum_{j=1, j \neq i}^{n} |a_{ij}|$, for all rows $i \in \{1, \dots, n\}$.

Note: We refer to (Bertsekas and Tsitsiklis, 1997) for a standard treatment of the Jacobi relaxation and related methods.

E2.16   **The Jacobi over-relaxation in parallel computation.** We now consider a more sophisticated version of the Jacobi relaxation presented in Exercise E2.15. Consider again n distributed processors that aim to collectively solve the linear equation $Ax = b$, where $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is invertible and its diagonal elements a_{ii} are nonzero. Each processor stores a variable $x_i(k)$ as the discrete-time variable k evolves and applies the following iterative strategy termed *Jacobi over-relaxation*. At time step $k \in \mathbb{N}$ each processor performs the local computation

$$x_i(k+1) = (1-\omega)x_i(k) + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1, j\neq i}^{n} a_{ij}x_j(k)\right), \quad i \in \{1, \ldots, n\},$$

where $\omega \in \mathbb{R}$ is an adjustable parameter. Next, each processor $i \in \{1, \ldots, n\}$ sends its value $x_i(k+1)$ to all other processors $j \neq i$ with $a_{ji} \neq 0$, and they iteratively repeat the previous computation. The initial values of the processors are arbitrary.

  (i)   Assume the Jacobi over-relaxation converges to $x^\star$ and show that $Ax^\star = b$ if $\omega \neq 0$.

  (ii) Find the expression governing the dynamics of the error variable $e(k) := x(k) - x^\star$.

  (iii) Suppose that A is strictly row diagonally dominant, that is $|a_{ii}| > \sum_{j\neq i}|a_{ij}|$, for all rows $i \in \{1, \ldots, n\}$. Use the Geršgorin Disks Theorem 2.8 to discuss the convergence properties of the algorithm for all possible values of $\omega \in \mathbb{R}$.

        Hint: Consider different thresholds for ω.

E2.17   **Simulation (cont'd).** This is a followup to Exercise E1.5. Consider the linear averaging algorithm in equation (1.2): set $n = 5$, select the initial state equal to $(1, -1, 1, -1, 1)$, and use (a) the complete graph (b) a cycle graph, and (c) a star graph with node 1 as center.

   (i) To which value do all nodes converge to?

   (ii) Compute the dominant left eigenvector of the averaging matrix associated to each of the three graphs and verify that the asymptotic result in Theorem 2.13 (illustrated in Example 2.14) is correct.

E2.18 **Robotic coordination: continuous- and discrete-time rendezvous on the real line.** Consider n robots moving on the line with positions $z_1, z_2, \ldots z_n \in \mathbb{R}$. In order to gather at a common location (i.e., reach rendezvous), each robot heads for the centroid of its neighbors, that is,

$$\dot{z}_i = \frac{1}{n-1} \left(\sum_{j=1, j\neq i}^{n} z_j \right) - z_i.$$

(i) Will the robots asymptotically rendezvous at a common location?

(ii) Consider the Euler discretization of the above closed-loop dynamics with sampling period $T > 0$:

$$z_i(k+1) = z_i(k) + T\left(\frac{1}{n-1} \left(\sum_{j=1, j\neq i}^{n} z_j(k) \right) - z_i(k) \right).$$

For which values of the sampling period T will the robots rendezvous?

Hint: Use the modal decomposition in Remark 2.3.

E2.19   **Perron eigenvalue as solution to optimization problems.** Given a positive matrix $A > 0$, show that

(i) the Perron eigenvalue and eigenvector of A are the unique solution of

$$\max_{(\mu,x)\in\mathbb{R}\times\mathbb{R}^n} \quad \mu$$

$$\text{subject to} \quad Ax \geq \mu x$$

$$x > 0 \text{ and } \mathbb{1}_n^{\mathsf{T}} x = 1;$$

(ii) the Perron eigenvector is the unique solution, and the Perron eigenvalue the optimal value, of

$$\max_{x\in\mathbb{R}^n} \quad \min_{i\in\{1,\dots,n\}} (Ax)_i/x_i$$

$$\text{subject to} \quad x > 0 \text{ and } \mathbb{1}_n^{\mathsf{T}} x = 1.$$

Note: The following equalities are referred to as the Collatz–Wielandt formula for the Perron eigenvalue λ:

$$\lambda = \max_{x>0,\mathbb{1}_n^{\mathsf{T}}x=1} \min_{i\in\{1,\dots,n\}} \frac{(Ax)_i}{x_i} = \min_{x>0,\mathbb{1}_n^{\mathsf{T}}x=1} \max_{i\in\{1,\dots,n\}} \frac{(Ax)_i}{x_i}.$$

E2.20 **The pseudoinverse and the singular value decomposition.** Let $A \in \mathbb{R}^{m \times n}$ with rank $r \leq \min\{m, n\}$. The *pseudoinverse* (also referred to as the *Moore-Penrose pseudoinverse*) of A is the unique matrix $A^\dagger \in \mathbb{R}^{n \times m}$ satisfying:

$$AA^\dagger A = A, \qquad A^\dagger AA^\dagger = A^\dagger,$$

$$AA^\dagger \text{ is symmetric, and } A^\dagger A \text{ is symmetric.}$$

(E2.3)

Show

   (i) if $m = n = r$, then $A^\dagger = A^{-1}$,

  (ii) if $m > n = r$, then $A^\dagger = (A^\mathsf{T} A)^{-1} A^\mathsf{T}$,

 (iii) if $r = m < n$, then $A^\dagger = A^\mathsf{T} (AA^\mathsf{T})^{-1}$.

Next, a triple (U, Σ, V) is a *singular value decomposition (SVD)* of A if

$$A = U\Sigma V^\mathsf{T},$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal and, for a positive diagonal matrix $\tilde{\Sigma} \in \mathbb{R}^{r \times r}$,

$$\Sigma = \begin{bmatrix} \tilde{\Sigma} & \mathbb{0}_{r \times (n-r)} \\ \mathbb{0}_{(m-r) \times r} & \mathbb{0}_{(m-r) \times (n-r)} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Show

 (iv) $A^\dagger = V\Sigma^\dagger U^\mathsf{T}$, where $\Sigma^\dagger = \begin{bmatrix} \tilde{\Sigma}^{-1} & \mathbb{0}_{r \times (m-r)} \\ \mathbb{0}_{(n-r) \times r} & \mathbb{0}_{(n-r) \times (m-r)} \end{bmatrix} \in \mathbb{R}^{n \times m}.$

Chapter 3

Elements of Graph Theory

Graph theory provides key concepts to model, analyze and design network systems and distributed algorithms; the language of graphs pervades modern science and technology and is therefore essential.

3.1 Graphs and digraphs

[Graphs] An *undirected graph* (in short, a *graph*) consists of a set V of *nodes* and of a set E of unordered pairs of nodes, called *edges*. For $u, v \in V$ and $u \neq v$, the set $\{u, v\}$ denotes an unordered edge.

[Neighbors and degrees in graphs] Two nodes u and v of a given graph are *neighbors* if $\{u, v\}$ is an undirected edge. Given a graph G, we let $\mathcal{N}_G(v)$ denote the set of neighbors of v.

The *degree* of v is the number of neighbors of v. A graph is *regular* if all the nodes have the same degree; e.g., in Figure 3.1, the cycle graph is regular with degree 2 whereas the complete bipartite graph $K(3,3)$ and the Petersen graph are regular with degree 3.

Example 3.1 (Basic graphs). *We define various basic graphs of dimension n as follows: Path graph: nodes are ordered in a sequence and edges connect subsequent nodes in the sequence. Cycle (or ring) graph: all nodes and edges can be arranged as the vertices and edges of a regular polygon. Star graph: edges connect a specific node, called the center, to all other nodes. Complete graph: every pair of nodes is connected by an edge. Complete bipartite graph: nodes are divided into two sets and every node of the first set is connected with every node of the second set. Figure 3.1 illustrates these five definitions and two other examples.* •

(a) *Path graph* with 6 nodes, denoted by P_6

(b) *Cycle (ring) graph* with 6 nodes, denoted by C_6

(c) *Star graph* with 6 nodes, denoted by S_6

(d) *Complete graph* with 6 nodes, denoted by K_6

(e) *Complete bipartite graph* with $3 + 3$ nodes, denoted by $K_{3,3}$

(f) *Two-dimensional grid graph* with 4×7 nodes, denoted by $G_{4,7}$

(g) *Petersen graph*
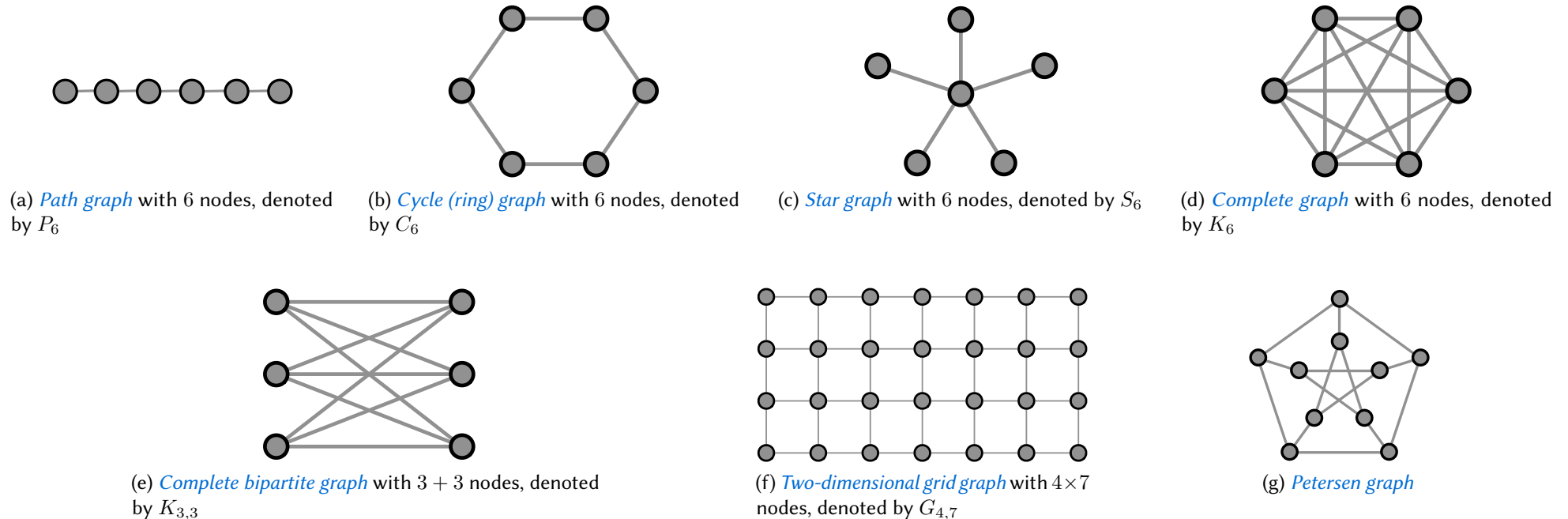
Figure 3.1: Example graphs.

3.1.1 Digraphs

[Digraphs and self-loops] A *directed graph* (in short, a *digraph*) of order n is a pair $G = (V, E)$, where V is a set with n elements called *nodes* and E is a set of ordered pairs of nodes called *edges*. In other words, $E \subseteq V \times V$. As for graphs, V and E are the *node set* and *edge set*, respectively. For $u, v \in V$, the ordered pair (u, v) denotes an edge *from u to v*. A digraph is *undirected* if $(v, u) \in E$ anytime $(u, v) \in E$. In a digraph, a *self-loop* is an edge from a node to itself. Consistently with a customary convention, self-loops are not allowed in graphs. We define and visualize some basic example digraphs in Figure 3.2.
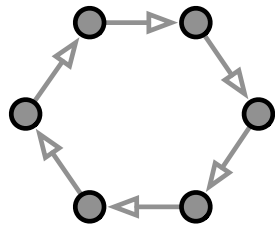
(a) *Cycle digraph* with 6 nodes   (b) *Complete digraph* with 6 nodes   (c) A digraph with no directed cycles
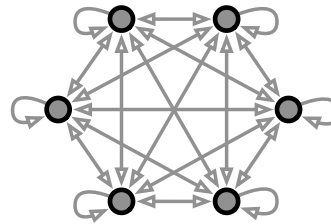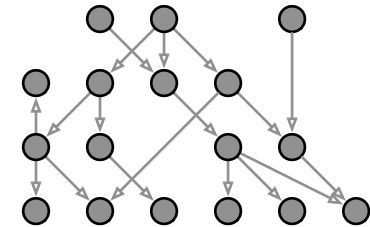
Figure 3.2: Example digraphs

[Subgraphs] A digraph (V', E') is a *subgraph* of a digraph (V, E) if $V' \subseteq V$ and $E' \subseteq E$. A digraph (V', E') is a *spanning subgraph* of (V, E) if it is a subgraph and $V' = V$. The subgraph of (V, E) *induced by $V' \subseteq V$* is the digraph (V', E'), where E' contains all edges in E between two nodes in V'.

[In- and out-neighbors] In a digraph G with an edge $(u, v) \in E$, u is called an *in-neighbor* of v, and v is called an *out-neighbor* of u. We let $\mathcal{N}^{\text{in}}(v)$ (resp., $\mathcal{N}^{\text{out}}(v)$) denote the set of in-neighbors, (resp. the set of out-neighbors) of v. Given a digraph $G = (V, E)$, an *in-neighbor* of a nonempty set of nodes U is a node $v \in V \setminus U$ for which there exists an edge $(v, u) \in E$ for some $u \in U$.

[In- and out-degree] The *in-degree $d_{\text{in}}(v)$* and *out-degree $d_{\text{out}}(v)$* of v are the number of in-neighbors and out-neighbors of v, respectively. Note that a self-loop at a node v makes v both an in-neighbor as well as an out-neighbor of itself. A digraph is *topologically balanced* if each node has the same in- and out-degrees (even if distinct nodes have distinct degrees).

3.2   Walks and connectivity in undirected graphs

[Walks] A *walk* (or *path*) in a graph is an ordered sequence of nodes such that any pair of consecutive nodes in the sequence is an edge of the graph. A walk is *simple* if no node appears more than once in it, except possibly for the case in which the initial node is the same as the final node. (Note: some authors adopt the term "path" to refer to a walk.)

[Connectivity and connected components] A graph is *connected* if there exists a walk between any two nodes. If a graph is not connected, then it is composed of multiple *connected components*, that is, multiple connected subgraphs.

[Cycles] A *cycle* is a simple walk that starts and ends at the same node and has at least three distinct nodes. A graph is *acyclic* if it contains no cycles. A connected acyclic graph is a *tree*.
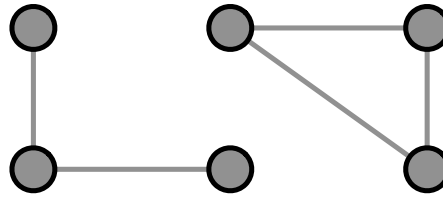
Figure 3.3: This graph has two connected components. The leftmost connected component is a tree, while the rightmost connected component is a cycle.

3.3   Walks and connectivity in digraphs

[Directed walks] A *directed walk* in a digraph is an ordered sequence of nodes such that any pair of consecutive nodes in the sequence is a directed edge of the digraph. A directed walk is *simple* if no node appears more than once in it, except possibly for the initial and final node.

[Cycles in digraphs] A *cycle* (or *directed cycle* in a digraph is a directed walk that starts and ends at the same node. It is customary to accept, as feasible cycles in digraphs, also cycles of length 1 (that is, a self-loop) and cycles of length 2 (that is, composed of just 2 nodes). A cycle is *simple* if no node appears more than once in it, except the initial and final nodes. The set of simple cycles of a directed graph is finite. A digraph is *acyclic* if it contains no cycles. In computer science, a directed acylic graph is sometimes referred to as a *DAG*.

[Sources and sinks] In a digraph, every node with in-degree 0 is called a *source*, and every node with out-degree 0 is called a *sink*. Every acyclic digraph has at least one source and at least one sink; see Figure 3.4 and Exercise E3.2.

(a) An acyclic digraph with one sink and two sources

(b) A directed cycle

Figure 3.4: Examples of sources and sinks

[Directed trees] A *directed tree* (sometimes called a *rooted tree*) is an acyclic digraph with the following property: there exists a node, called the *root*, such that any other node of the digraph can be reached by one and only one directed walk starting at the root. A *directed spanning tree* of a digraph is a spanning subgraph that is a directed tree.

3.3.1   Connectivity properties of digraphs

Next, we present four useful connectivity notions for a digraph G:

(i) G is *strongly connected* if there exists a directed walk from any node to any other node;

(ii) G is *weakly connected* if the undirected version of the digraph is connected;

(iii) G possesses a *globally reachable node* if one of its nodes can be reached from any other node by traversing a directed walk; and

(iv) G possesses a *directed spanning tree* if one of its nodes is the root of directed walks to every other node.

These notions are illustrated in Figure 3.5.
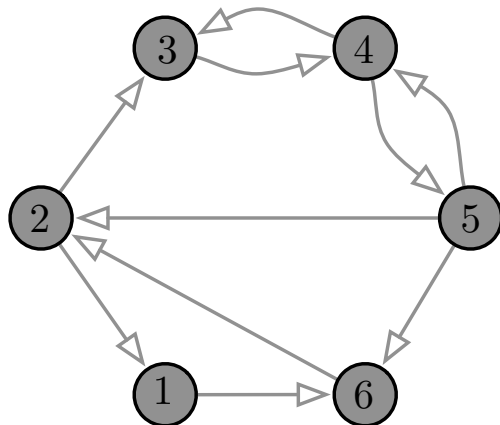
(a) A strongly connected digraph
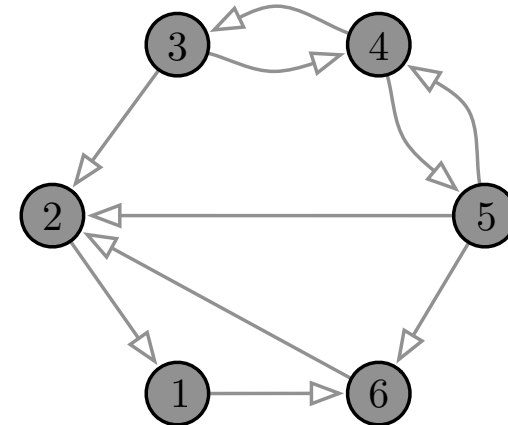
(b) A weakly connected digraph with a globally reachable node

Figure 3.5: Connectivity examples for digraphs

For a digraph $G = (V, E)$, the *reverse digraph* $G(\mathrm{rev})$ has node set V and edge set $E(\mathrm{rev})$ composed of all edges in E with reversed direction. Clearly, a digraph contains a directed spanning tree if and only if the reverse digraph contains a globally reachable node.

3.3.2 Periodicity of strongly-connected digraphs

[Periodic and aperiodic digraphs] A strongly-connected directed graph is *periodic* if there exists a $k > 1$, called the *period*, that divides the length of every simple cycle of the graph. In other words, a digraph is periodic if the greatest common divisor of the lengths of all its simple cycles is larger than one. A digraph is *aperiodic* if it is not periodic.

(a) A periodic digraph with period 2

(b) An aperiodic digraph with simple cycles of length 1 and 2.
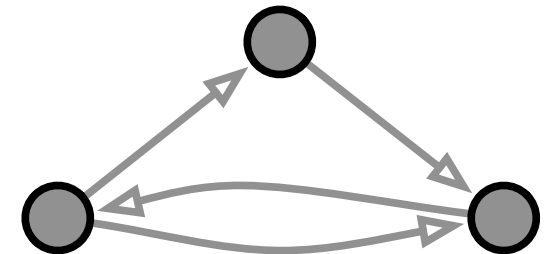
(c) An aperiodic digraph with simple cycles of length 2 and 3.

Figure 3.6: Example periodic and aperiodic digraphs.

Note: the definition of periodic digraph is well-posed because a digraph has only a finite number of simple cycles (because nodes are not repeated in simple walks). The notions of periodicity and aperiodicity only apply to digraphs and not to undirected graphs (where the notion of a cycle is defined differently). Any strongly-connected digraph with a self-loop is aperiodic.

3.3.3   Condensation digraphs

[Strongly connected components] A subgraph H is a *strongly connected component* of G if H is strongly connected and any other subgraph of G strictly containing H is not strongly connected.

[Condensation digraph] The *condensation digraph* of a digraph G, denoted by $C(G)$, is defined as follows: the nodes of $C(G)$ are the strongly connected components of G, and there exists a directed edge in $C(G)$ from node H_1 to node H_2 if and only if there exists a directed edge in G from a node of H_1 to a node of H_2. The condensation digraph has no self-loops. This construction is illustrated in Figure 3.7.

(a) An example digraph G (b) The strongly connected components of the digraph G (c) The condensation digraph $C(G)$
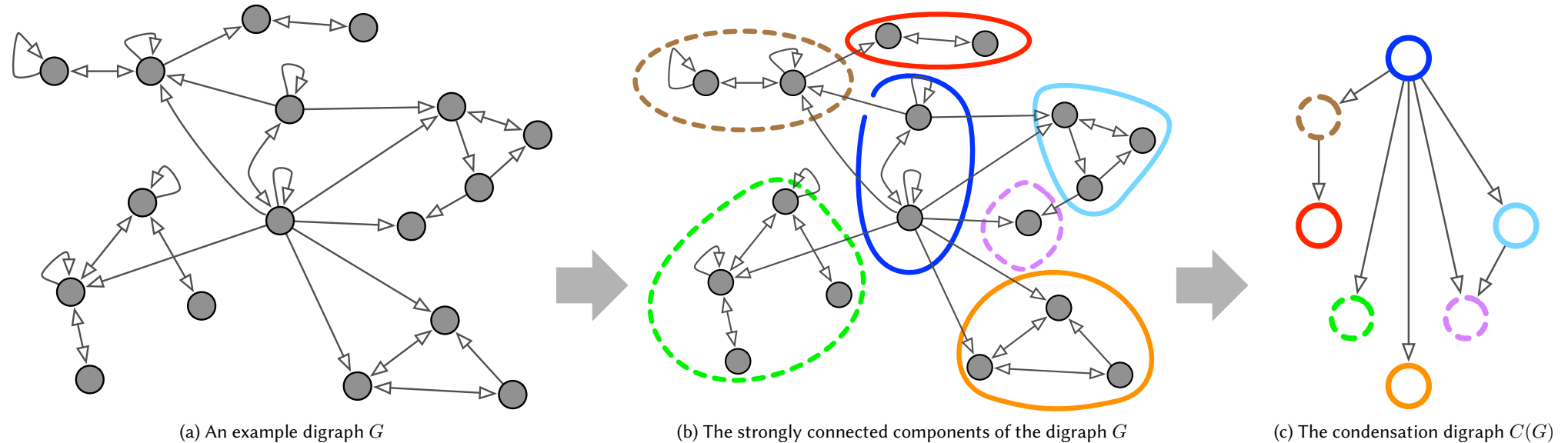
Figure 3.7: An example digraph, its strongly connected components and its condensation digraph.

Lemma 3.2 (Properties of the condensation digraph). *For a digraph G and its condensation digraph $C(G)$,*

   (i) *$C(G)$ is acyclic,*

  (ii) *G is weakly connected if and only if $C(G)$ is weakly connected, and*

 (iii) *the following statement are equivalent:*

     (a) *G contains a globally reachable node,*

     (b) *$C(G)$ contains a globally reachable node, and*

     (c) *$C(G)$ contains a unique sink.*

Proof. We prove statement (i) by contradiction. If there exists a cycle $(H_1, H_2, \ldots, H_m, H_1)$ in $C(G)$, then the set of nodes $H_1, \ldots, H_m$ are strongly connected in $C(G)$. But this implies that also the subgraph of G containing all node of $H_1, \ldots, H_m$ is strongly connected in G. But this is a contradiction with the fact that any subgraph of G strictly containing any of the $H_1, \ldots, H_m$ must be not strongly connected. Statement (ii) is intuitive and simple to prove; we leave this task to the reader.

Regarding statement (iii), we start by proving that (iii)a $\implies$ (iii)b. Let v be a a globally reachable node in G and let H denote the node in $C(G)$ containing v. Pick an arbitrary node $\bar{H}$ of $C(G)$ and let $\bar{v}$ be a node of G in $\bar{H}$. Since v is globally reachable, there exists a directed walk from $\bar{v}$ to v in G. This directed walk induces naturally a directed walk in $C(G)$ from $\bar{H}$ to H. This shows that H is a globally reachable node in $C(G)$.

Regarding (iii)b $\implies$ (iii)a, let H be a globally reachable node of $C(G)$ and pick a node v in H. We claim v is globally reachable in G. Indeed, pick any node $\bar{v}$ in G belonging to a strongly connected component $\bar{U}$ of G. Because H is globally reachable in $C(G)$, there exists a directed walk of the form $\bar{H} = H_0, H_1, \ldots, H_k, H_{k+1} = H$ in $C(G)$. One can now piece together a directed walk in G from $\bar{v}$ to v, by walking inside each of the strongly connected components H_i and moving to the subsequent strongly connected components H_{i+1}, for $i \in \{0, \ldots, k\}$.

The final equivalence between statement (iii)b and statement (iii)c is an immediate consequence of $C(G)$ being acyclic.   ■

Finally, we provide some example to compute and visualize condensation digraphs using the NetworkX library and Python; see Table 3.1.

```python
 1  # Python3 code to compute condensed digraphs via the NetworkX library
 2  # Import the NetworkX library and other tools
 3  import numpy as np; import matplotlib.pyplot as plt; import networkx as nx
 4  myfs= (30, 10); fig, axs343 = plt.subplots(1, 3, figsize=(myfs[0]*1.2, myfs[1]))
 5
 6  def plot_condensated_graph(G, axs3):
 7      """
 8      :param G: Digraph to be condensed
 9      :param axs3: 1 dim axs with at least 3 subplots locations
10      """
11
12      # Visualize initial digraph
13      Gpos = nx.spring_layout(G)   #setting the position with respect to G
14      nx.draw_networkx(G, Gpos, node_size=40, ax=axs3[0], connectionstyle='arc3, rad = 0.2', with_labels=False)
15
16      # Algorithm to compute the condensed digraph:
17      G_conden = nx.algorithms.components.condensation(G)
18
19      all_col = []
20      # Compute coloring scheme for the condensated graph
21      for u, node in G_conden.nodes(data=True):
22          sg = node['members']  # This contains a set of nodes from previous graph, that belongs to the condensated node
23          co = np.random.rand(1,3)
24          all_col.append(co)
25          nx.draw_networkx_nodes(G.subgraph(sg), Gpos, node_size=40, node_color=co, ax=axs3[1])
26          nx.draw_networkx_edges(G, Gpos, edgelist=G.edges(sg), edge_color=co, ax=axs3[1], connectionstyle='arc3, rad = 0.2')
27
28      nx.draw_networkx(G_conden, node_size=40, ax=axs3[2], node_color=all_col, connectionstyle='arc3, rad = 0.2', ...
            with_labels=False)
29      axs3[0].set_xlabel("Original digraph");  axs3[1].set_xlabel("Strongly connected components");
30      axs3[2].set_xlabel("Condensed digraph");
31
32  # Generate and plot random digraph:
33  G_random = nx.random_k_out_graph(40, 5, 1)
34  plot_condensated_graph(G_random, axs343)
```

Table 3.1: Python code for computation and visualization of condensed digraphs

3.4   Weighted digraphs

A *weighted digraph* is a triplet $G = (V, E, \{a_e\}_{e \in E})$, where the pair (V, E) is a digraph with nodes $V = \{v_1, \ldots, v_n\}$, and where $\{a_e\}_{e \in E}$ is a collection of strictly positive weights for the edges E.

   Note: for simplicity we let $V = \{1, \ldots, n\}$. It is therefore equivalent to write $\{a_e\}_{e \in E}$ or $\{a_{ij}\}_{(i,j) \in E}$.
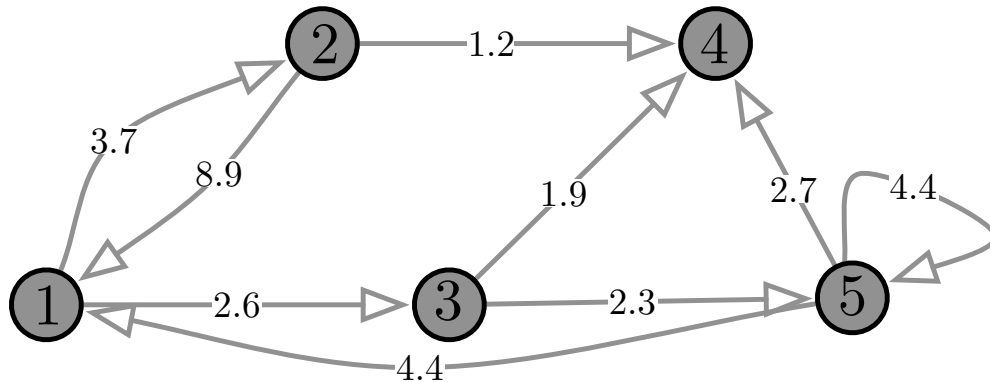
The set of weights for this weighted digraph is

$$a_{12} = 3.7, \ a_{13} = 2.6, \ a_{21} = 8.9,$$
$$a_{24} = 1.2, \ a_{34} = 1.9, \ a_{35} = 2.3,$$
$$a_{51} = 4.4, \ a_{54} = 2.7, \ a_{55} = 4.4.$$

   A digraph $G = (V = \{v_1, \ldots, v_n\}, E)$ can be regarded as a weighted digraph by defining its set of weights to be all equal to 1, that is, setting $a_e = 1$ for all $e \in E$. A weighted digraph is *undirected* if $a_{ij} = a_{ji}$ for all $i, j \in \{1, \ldots, n\}$.

The notions of connectivity and definitions of in- and out-neighbors, introduced for digraphs, remain equally valid for weighted digraphs. The notions of in- and out-degree are generalized to weighted digraphs as follows. In a weighted digraph with $V = \{v_1, \ldots, v_n\}$, the *weighted out-degree* and the *weighted in-degree* of node v_i are defined by, respectively,

$$d_{\text{out}}(v_i) = \sum_{j=1}^{n} a_{ij} \qquad \text{(i.e., } d_{\text{out}}(v_i) \text{ is the sum of the weights of all the out-edges of } v_i) ,$$

$$d_{\text{in}}(v_i) = \sum_{j=1}^{n} a_{ji} \qquad \text{(i.e., } d_{\text{in}}(v_i) \text{ is the sum of the weights of all the in-edges of } v_i) .$$

A weighted digraph is *weight-balanced* if each node has the same weighted in- and out-degrees (even if distinct nodes have distinct weighted degrees), that is, $d_{\text{out}}(v_i) = d_{\text{in}}(v_i)$ for all $v_i \in V$. For unweighted digraphs, weight-balance is the same property as topological balance.

3.5   Appendix: Database collections and software libraries

Useful collections of example networks are freely available online; here are some examples:

(i) A broad range of example networks is available online at the Stanford Large Network Dataset Collection, see http://snap.stanford.edu/data.

(ii) The SuiteSparse Matrix Collection (formerly known as the University of Florida Sparse Matrix Collection), available at http://suitesparse.com and described in (Davis and Hu, 2011), contains a large and growing set of sparse matrices and complex graphs arising in a broad range of applications; e.g., see Figure 3.8.

(iii) The UCI Network Data Repository, available at http://networkdata.ics.uci.edu, is an effort to facilitate the scientific study of networks; see also (DuBois, 2008).

(a) Western North America power grid

(b) Sampson monastery network

(c) EPANET 3 water supply network
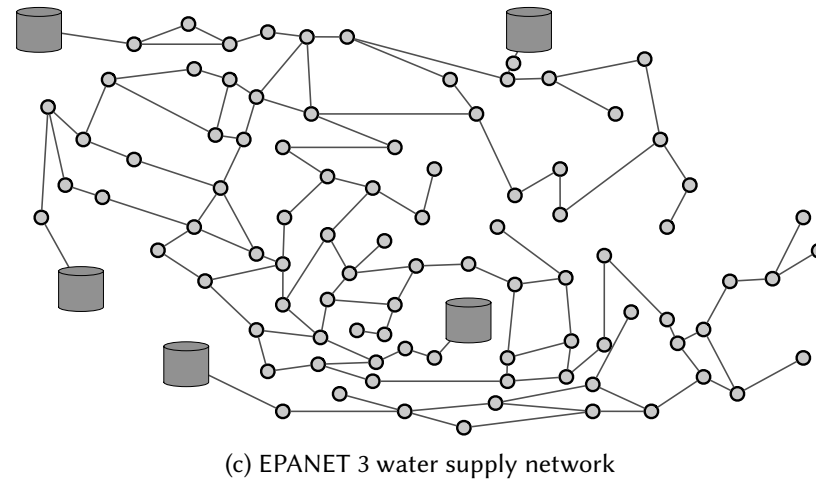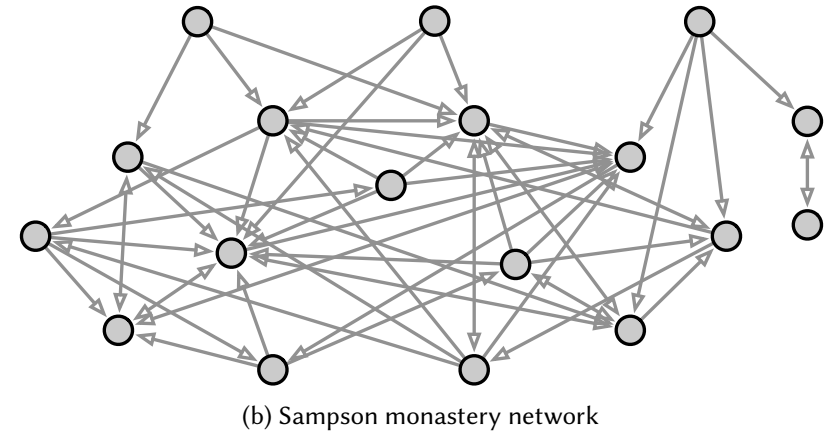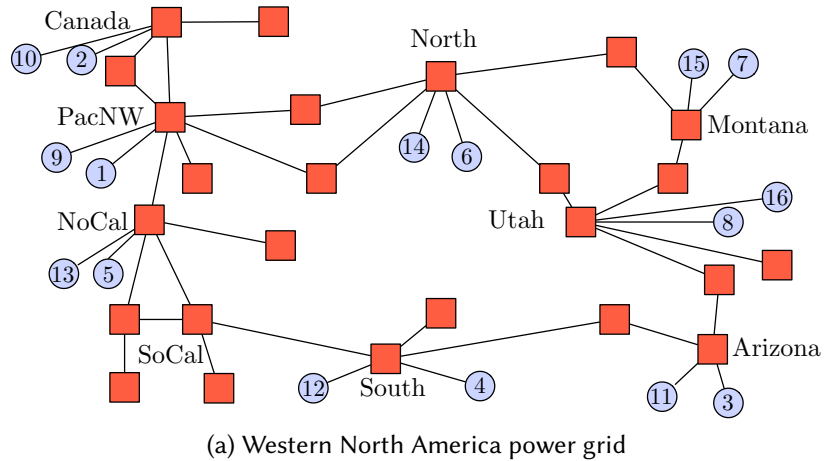
Figure 3.8: Example networks from distinct domains: Figure 3.8a illustrates a simplified aggregated model with 16 generators and 25 load busses of the Western North American power grid (Trudnowski et al., 1991); Figure 3.8b illustrates the Sampson monastery dataset (Sampson, 1969), describing social relations among a set of 18 monk-novitiates in an American monastery; Figure 3.8c illustrates the water supply network EPANET 3, described in (Rossman, 2000).

Useful software libraries for network analysis and visualization are freely available online; here are some examples:

(i) NetworkX, available at https://networkx.org, is a Python library for network analysis. For example, one feature is the ability to compute condensation digraphs.

(ii) Gephi, available at https://gephi.org, is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs.

(iii) Cytoscape, available at http://www.cytoscape.org, is an open-source software platform for visualizing complex networks and integrating them with attribute data.

(iv) Graphviz, available at http://www.graphviz.org/, is an open source graph visualization software.

3.6   Historical notes and further reading

Paraphrasing from Chapter 1 "Discovery!" in the classic work by Harary (1969),

> (Euler, 1741) became the father of graph theory as well as topology when he settled a famous unsolved problem of his day called the Königsberg Bridge Problem.

> Subsequent rediscoveries of graph theory by Kirchhoff (1847) and Cayley (1857) also had their roots in the physical world. Kirchhoff's investigations of electric networks led to his development of the basic concepts and theorems concerning trees in graphs, while Cayley considered trees arising from the enumeration of organic chemical isomers.

For modern comprehensive treatments we refer the reader to standard books in graph theory such as (Diestel, 2000; Bollobás, 1998).

An early reference about the condensation decomposition is (Harary, 1962).

A classic reference in graph drawing is (Fruchterman and Reingold, 1991). The layout of the three graphs in Figure 3.8 is obtained via the algorithm proposed by Hu (2005). Geometric representations of graphs are reviewed in the remarkable recent text (Lovász, 2019).

3.7   Exercises

E3.1   **Properties of undirected trees.** Consider an undirected graph G with n nodes and m edges (and without self-loops). Show that the following statements are equivalent:

    (i) G is a tree;

    (ii) G is connected and $m = n - 1$; and

    (iii) G is acyclic and $m = n - 1$.

▶ E3.2   **Topological sort of acyclic digraphs.** Let G be an acyclic digraph (DAG) with nodes $\{1, \ldots, n\}$. A *topological sort* of G is a re-numbering of the nodes of G with the property that, if (u, v) is an edge of G, then $u > v$.

   (i) Show that G contains at least one sink, i.e., a node without out-neighbors and at least one source, i.e., a node without in-neighbors.

   (ii) Provide an algorithm to perform a topological sort of G. Is the topological sort unique?

        Hint: Use high-level pseudo-code instructions such as "select a node satisfying property A" or "remove all edges satisfying property B."

   (iii) Show that, after topologically sorting the nodes of G, the adjacency matrix of G is lower-triangular, i.e., all its entries above the main diagonal are equal to zero.

E3.3 **Globally reachable nodes and disjoint closed subsets.** Consider a digraph $G = (V, E)$ with at least two nodes. Prove that the following statements are equivalent:

(i) G has a globally reachable node, and

(ii) for every pair S_1, S_2 of non-empty disjoint subsets of V, there exists a node that is an out-neighbor of S_1 or S_2.

E3.4   **Condensation digraphs.** Draw the condensation for each of the following digraphs.
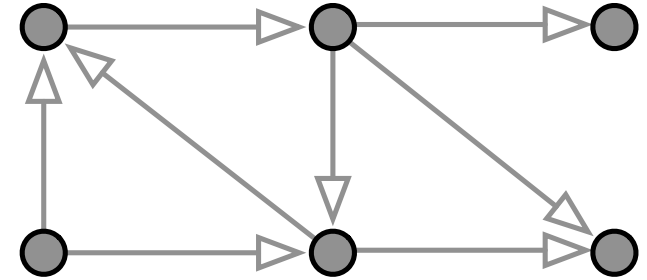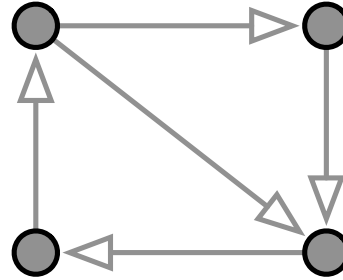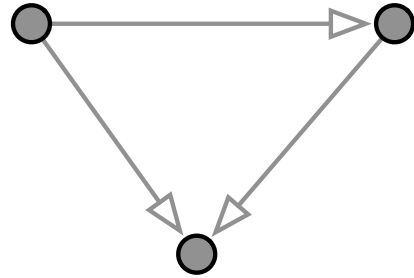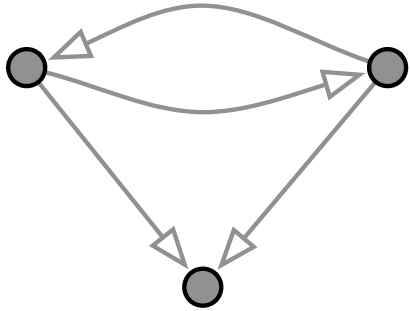
E3.5   **Directed spanning trees in the condensation digraph.** For a digraph G and its condensation digraph $C(G)$, show that the following statements are equivalent:

   (i) G contains a directed spanning tree, and

   (ii) $C(G)$ contains a directed spanning tree.

▶ E3.6   **Connectivity in weight-balanced digraphs.** Let G be a weight-balanced weighted digraph. Show that

(i) for every subset of nodes U, the total weight of edges from nodes outside U to nodes inside U equals the total weight of edges from nodes inside U to nodes outside U, and

(ii) if G is weakly connected, then it is strongly connected.

E3.7   **Line graph.** The *line graph* $\mathrm{Line}(G)$ of an undirected graph G is the undirected graph whose vertices are the edges of G and whose edge set is defined as follows: there is an edge between two vertices e_1, e_2 of $L(G)$ if and only if e_1 and e_2, regarded as edges of G, have a common node. Show that

    (i) if G is connected, then $\mathrm{Line}(G)$ is connected;

    (ii) if $d_1, \ldots, d_n$ are the degrees of the nodes of G and m denotes the number of edges of G, then

$$\text{number of edges in } \mathrm{Line}(G) = \frac{1}{2}\sum_{i=1}^{n} d_i^2 - m. \tag{E3.1}$$

E3.8 **Alternative definition of aperiodicity.** Given a directed graph G, a *recurrence time* of a node i is a natural number $k > 0$ such that there exists a directed walk from i to i of length k. The *period* of node i is the greatest common divisor of all recurrence times of i. Show that

    (i) all nodes belonging to the same strongly-connected component of G have the same period,

   (ii) if G is strongly connected, then G is aperiodic if and only if the period of each node is 1.

E3.9   **Euler's formula and the sparsity of planar graphs.** An undirected graph G is *planar* if, loosely speaking, it can be drawn on the plane in such a way that its edges intersect only at their endpoints.
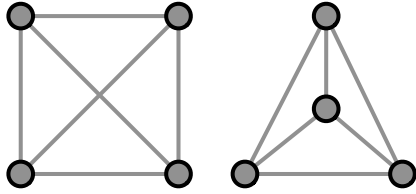
(a) Two drawings of K_4. The right figure demonstrates that K_4 is planar. Euler's formula is satisfied with $n = 4$, $m = 6$ and $f = 4$.
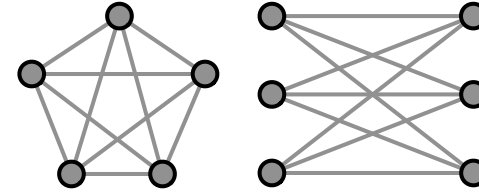
(b) K_5 and $K_{3,3}$ are not planar and play a fundamental role in Kuratowski's characterization of planar graphs; see (Diestel, 2000, Section 4.4).

Assuming a planar graph G is drawn on the plane without edge intersections, the edges of G divide the plane into regions called *faces*; all faces are bounded, except for the outer infinite face. For a connected undirected graph G with n nodes, m edges and f faces, *Euler's formula* states

$$n - m + f = 2.$$

Show that

(i)   each planar graph with $n \geq 3$ is sparse, in the sense that $m \leq 3n - 6$, and

(ii) the equality is strict when every face is a triangle.

Note: *Planar graphs are reviewed in (Diestel, 2000, Section 4.4) and (Godsil and Royle, 2001, Section 1.8).*

Chapter 4

Elements of Algebraic Graph Theory

In this chapter we present results on the adjacency matrices as part of the broader field of algebraic graph theory. The key results in this area relate, through necessary and sufficient conditions, matrix properties with graphical properties. For example, we will show how a matrix is primitive if and only if its associated digraph is strongly connected and aperiodic.

4.1 The adjacency matrix

Given a weighted digraph $G = (V, E, \{a_e\}_{e \in E})$, with $V = \{1, \ldots, n\}$, the *weighted adjacency matrix* of G is the $n \times n$ non-negative matrix A defined as follows: for each edge $(i, j) \in E$, the entry (i, j) of A is equal to the weight $a_{(i,j)}$ of the edge (i, j), and all other entries of A are equal to zero. In other words, $a_{ij} > 0$ if and only if (i, j) is an edge of G, and $a_{ij} = 0$ otherwise. Figure 4.1 shows an example of a weighted digraph.

$$A = \begin{bmatrix} 0 & 3.7 & 2.6 & 0 & 0 \\ 8.9 & 0 & 0 & 1.2 & 0 \\ 0 & 0 & 0 & 1.9 & 2.3 \\ 0 & 0 & 0 & 0 & 0 \\ 4.4 & 0 & 0 & 2.7 & 4.4 \end{bmatrix}.$$
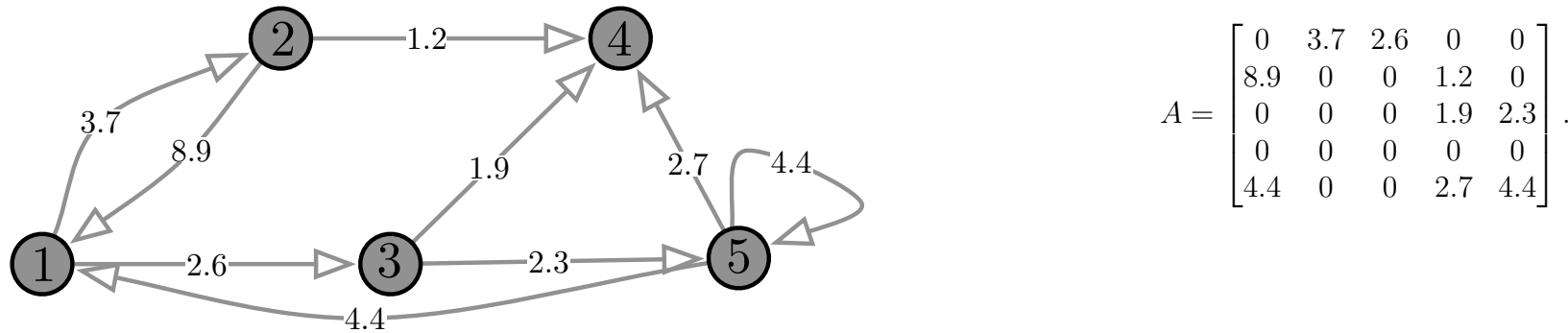
Figure 4.1: A weighted digraph and its adjacency matrix

The *binary adjacency matrix* $A \in \{0, 1\}^{n \times n}$ of a digraph $G = (V = \{1, \ldots, n\}, E)$ or of a weighted digraph is defined by

$$a_{ij} = \begin{cases} 1, & \text{if } (i, j) \in E, \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

Here, a binary matrix is any matrix with entries taking values in $0, 1$.

Finally, the *weighted out-degree matrix* D_{out} and the *weighted in-degree matrix* D_{in} of a weighted digraph are the diagonal matrices defined by

$$D_{\text{out}} = \text{diag}(A\mathbb{1}_n) = \begin{bmatrix} d_{\text{out}}(1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_{\text{out}}(n) \end{bmatrix}, \quad \text{and} \quad D_{\text{in}} = \text{diag}(A^{\mathsf{T}}\mathbb{1}_n),$$

where $\text{diag}(z_1, \ldots, z_n)$ is the diagonal matrix with diagonal entries equal to $z_1, \ldots, z_n$.

We conclude this section with some basic examples.

Example 4.1 (Basic graphs and their adjacency matrices). *Recall the definitions of walk, cycle, star, complete and complete bipartite graph from Example 3.1. Figure 4.2 illustrates their adjacency matrices.*
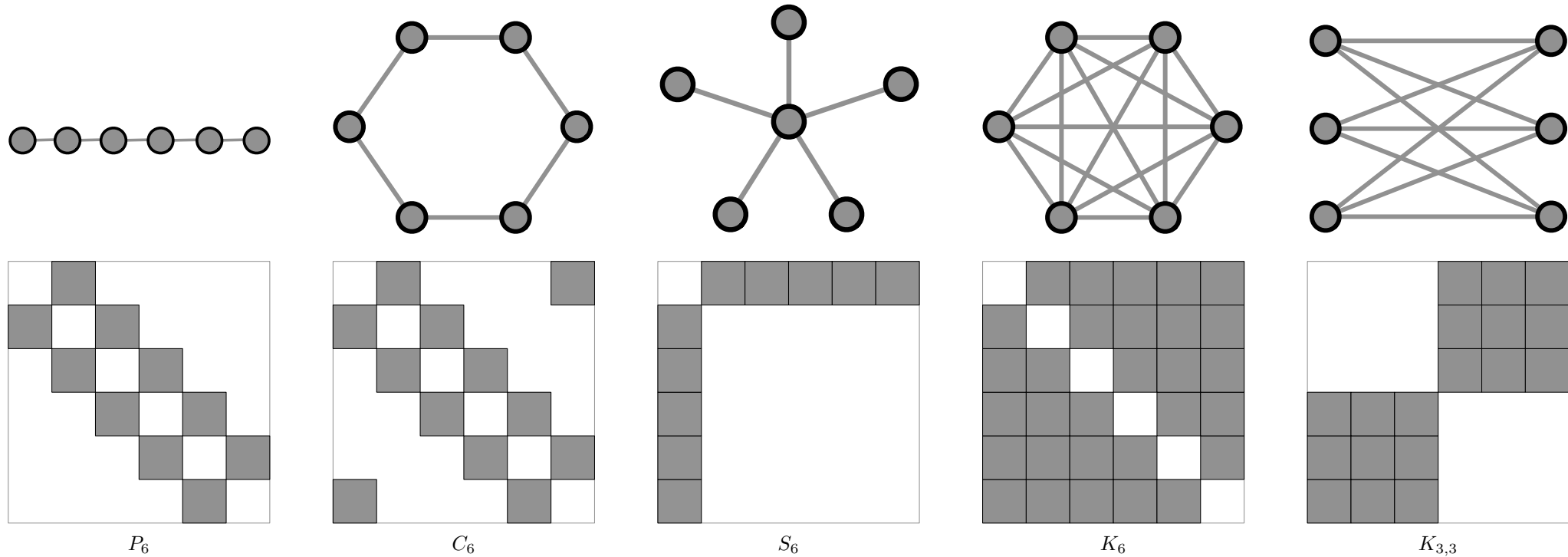
Figure 4.2: Walk, cycle, star, complete and complete bipartite graph (from Figure 3.1) and their binary adjacency matrices depicted in their respective *pixel picture* representation.

Note that the adjacency matrices of walk and cycle graphs have a particular structure. An $n \times n$ matrix T is Toeplitz (also called diagonal-constant) if there exist scalar numbers $a_{-(n-1)}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{n-1}$ such that

$$
T = \begin{bmatrix}
a_0 & a_1 & \cdots & \cdots & a_{n-1} \\
a_{-1} & a_0 & a_1 & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \cdots & a_{-1} & a_0 & a_1 \\
a_{-(n-1)} & \cdots & \cdots & a_{-1} & a_0
\end{bmatrix}.
$$

Two special cases are of interest, namely, those of tridiagonal Toeplitz and circulant matrices. For these two cases it is possible to compute eigenvalues and eigenvectors in closed form for arbitrary dimensions n; we refer to Exercises E4.2 and E4.3 for more information. We conclude with a table containing the adjacency spectrum of the basic graphs, i.e., the spectrum of their binary adjacency matrices.

Graph	Adjacency Matrix	Adjacency Spectrum
path graph P_n	Toeplitz tridiagonal	$\{2\cos(\pi i/(n+1)) \mid i \in \{1, \ldots, n\}\}$
cycle graph C_n	circulant	$\{2\cos(2\pi i/n)) \mid i \in \{1, \ldots, n\}\}$
star graph S_n	$e_1 e_{-1}^{\mathsf{T}} + e_{-1} e_1^{\mathsf{T}}$	$\{\sqrt{n-1}, 0, \ldots, 0, -\sqrt{n-1}\}$
complete graph K_n	$\mathbb{1}_n \mathbb{1}_n^{\mathsf{T}} - I_n$	$\{(n-1), -1, \ldots, -1\}$
complete bipartite $K_{n,m}$	$\begin{bmatrix} \mathbb{0}_{n\times n} & \mathbb{1}_{n\times m} \\ \mathbb{1}_{m\times n} & \mathbb{0}_{m\times m} \end{bmatrix}$	$\{\sqrt{nm}, 0, \ldots, 0, -\sqrt{nm}\}$

Table 4.1: Adjacency spectrum for basic graphs (we adopt the notation $e_{-i} = \mathbb{1}_n - e_i$)

We ask the reader to prove the statements in the table in Exercise E4.4.   ●

4.2 Algebraic graph theory: basic and prototypical results

In this section we review some basic and prototypical results that involve correspondences between graphs and adjacency matrices.

In what follows we let G denote a weighted digraph and A its weighted adjacency matrix or, equivalently, we let A be a non-negative matrix and G be its *associated weighted digraph* (i.e., the digraph with nodes $\{1, \dots, n\}$ and with weighted adjacency matrix A). We start with some straightforward statements, organized as a table of correspondences.

Digraph G	Non-negative matrix A (adjacency of G)
G is undirected	$A = A^{\mathsf{T}}$
G is weight-balanced	$A \mathbb{1}_n = A^{\mathsf{T}} \mathbb{1}_n$, that is, $D_{\text{out}} = D_{\text{in}}$
(no self-loops) node i is a sink	(zero diagonal) ith row-sum of A is zero
(no self-loops) node i is a source	(zero diagonal) ith column-sum of A is zero
each node has weighted out-degree equal to 1 ($D_{\text{out}} = I_n$)	A is row-stochastic
each node has weighted out- and in-degree equal to 1 ($D_{\text{out}} = D_{\text{in}} = I_n$)	A is doubly-stochastic

Next, we relate the powers of the adjacency matrix with the existence of directed walks in the digraph. We start with some simple observation.

- First, pick two nodes i and j and note that there exists a directed walk from i to j of length 1 (i.e., an edge) if and only if $a_{ij} = (A)_{ij} > 0$.

- Next, consider the formula for the matrix power:

$$(A^2)_{ij} = (i\text{th row of } A) \cdot (j\text{th column of } A) = \sum_{h=1}^{n} (A)_{ih}(A)_{hj}.$$

- This formula implies that:
  $(A^2)_{ij} > 0$ if and only if there exists a node h such that $(A)_{ih} > 0$ and $(A)_{hj} > 0$,
      if and only if (i, h) and (h, j) are edges of G,
      if and only if there exists a directed walk of length 2 from i to j.

- In short, we now know that a directed walk from i to j of length 2 exists if and only if $(A^2)_{ij} > 0$.

These observations lead to the following simple, but central result. We leave the detailed proof as Exercise E4.6.

Lemma 4.2 (Directed walks and powers of the adjacency matrix). *Let G be a weighted digraph with n nodes, with adjacency matrix A and binary adjacency matrix $A_{0,1} \in \{0,1\}^{n \times n}$. For all $i, j \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$*

(i) *the (i, j) entry of $A_{0,1}^k$ equals the number of walks of length k from node i to node j; and*

(ii) *the (i, j) entry of A^k is positive if and only if there exists a walk of length k from node i to node j.*

(Walks here are directed walks that possibly include self-loops.)

4.3   Graph theoretical characterization of irreducible matrices

In this section we provide three equivalent characterizations of the notion of irreducibility and we can now characterize certain connectivity properties of digraphs based on the powers of the adjacency matrix.

Before proceeding, we introduce a few useful concepts. First, $\{\mathcal{I}, \mathcal{J}\}$ is a *partition* of the index set $\{1, \ldots, n\}$ if $\mathcal{I} \cup \mathcal{J} = \{1, \ldots, n\}$, $\mathcal{I} \neq \emptyset$, $\mathcal{J} \neq \emptyset$, and $\mathcal{I} \cap \mathcal{J} = \emptyset$. Second, a *permutation matrix* is a square binary matrix with precisely one entry equal to 1 in every row and every column. (In other words, the columns of a permutation matrix are a reordering of the basis vectors $e_1, \ldots, e_n$; a permutation matrix acts on a vector by permuting its entries.) Finally, an $n \times n$ matrix A is *block triangular* if there exists $r \in \{1, \ldots, n-1\}$ such that

$$A = \left[\begin{array}{c|c} B & C \\ \hline \mathbb{0}_{(n-r) \times r} & D \end{array} \right],$$

where $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times (n-r)}$ and $D \in \mathbb{R}^{(n-r) \times (n-r)}$ are arbitrary.

We are now ready to state the main result of this section.

Theorem 4.3 (Strongly connected digraphs and irreducible adjacency matrices). *Let G be a weighted digraph with $n \geq 2$ nodes and with weighted adjacency matrix A. The following statements are equivalent:*

(i) *A is irreducible, that is, $\sum_{k=0}^{n-1} A^k > 0$;*

(ii) *there exists no permutation matrix P such that PAP^{T} is block triangular;*

(iii) *G is strongly connected;*

(iv) *for all partitions $\{\mathcal{I}, \mathcal{J}\}$ of the index set $\{1, \ldots, n\}$, there exists $i \in \mathcal{I}$ and $j \in \mathcal{J}$ such that (i, j) is a directed edge in G.*

Note: as the theorem establishes, there are four equivalent characterizations of irreducibility. In the literature, it is common to define irreducibility through property (ii) or (iv). We next see two simple examples.
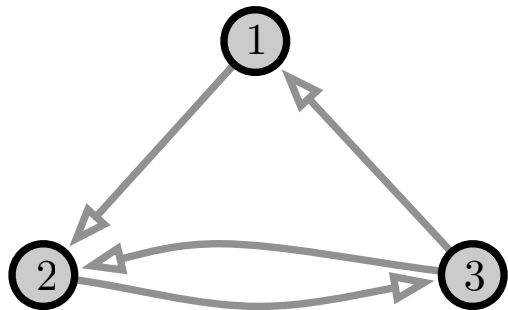
This digraph is strongly connected and, accordingly, its adjacency matrix is irreducible:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$
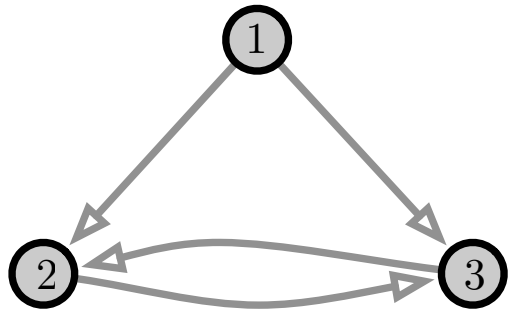
This digraph is not strongly connected (nodes 2 and 3 are globally reachable, but 1 is not) and, accordingly, its adjacency matrix is reducible:

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

4.3.1   Proof

Proof of Theorem 4.3. We start with the main equivalence. Regarding (i) $\implies$ (iii), pick two nodes i and j. Because $\sum_{k=0}^{n-1} A^k > 0$, there must exists k such that $(A^k)_{ij} > 0$. Lemma 4.2(ii) implies the existence of a walk of length k from i to j. Hence, G is strongly connected.

Conversely, regarding (iii) $\implies$ (i), because G is strongly connected, there exists a directed walk of length k' connecting node i to node j, for all i and j. By removing any closed sub-walk from such a walk (so that no intermediate node is repeated), one can compute a walk from i to j of length $k < n$. Hence, by Lemma 4.2(ii), the entry $(A^k)_{ij}$ is strictly positive and, in turn, so is the entire matrix sum $\sum_{k=0}^{n-1} A^k$.

Next, we establish the equivalence between the two graph-theoretical statements. Regarding (iii) $\implies$ (iv), pick a partition $\{\mathcal{I}, \mathcal{J}\}$ of the index set $\{1, \dots, n\}$ and two nodes $i_0 \in \mathcal{I}$ and $j_0 \in \mathcal{J}$. By assumptions there exists a directed walk from i_0 to j_0. Hence there must exist an edge from a node in $\mathcal{I}$ to a node in $\mathcal{J}$.

Regarding (iv) $\implies$ (iii), pick a node $i \in \{1, \dots, n\}$ and let $R_i \subset \{1, \dots, n\}$ be the set of nodes reachable from i, i.e., the set of nodes that belong to directed walks originating from node i. Denote the unreachable nodes by $U_i = \{1, \dots, n\} \setminus R_i$. Second, by contradiction, assume U_i is not empty. Then $R_i \cup U_i$ is a partition of the index set $\{1, \dots, n\}$ and statement (iv) implies the existence of a non-zero entry a_{jh} with $j \in R_i$ and $h \in U_i$. But then the node h is reachable. Therefore, $U_i = \emptyset$, and all nodes are reachable from i.

We establish the last two equivalences as follows. Regarding (ii) $\implies$ (iv), by contradiction, assume there exists a partition $(\mathcal{I}, \mathcal{J})$ of $\{1, \ldots, n\}$ such that $a_{ij} = 0$ for all $(i, j) \in \mathcal{I} \times \mathcal{J}$. Let $\pi \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be the permutation that maps all entries of $\mathcal{I}$ into the last $|\mathcal{I}|$ entries of $\{1, \ldots, n\}$. Here, $|\mathcal{I}|$ denotes the number of elements of $\mathcal{I}$. Let P be the corresponding permutation matrix. We now compute PAP^{T} and block partition it as:

$$PAP^{\mathsf{T}} = \begin{bmatrix} A_{\mathcal{J}\mathcal{J}} & A_{\mathcal{J}\mathcal{I}} \\ A_{\mathcal{I}\mathcal{J}} & A_{\mathcal{I}\mathcal{I}} \end{bmatrix},$$

where $A_{\mathcal{J}\mathcal{J}} \in \mathbb{R}^{|\mathcal{J}| \times |\mathcal{J}|}$, $A_{\mathcal{I}\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{I}|}$, $A_{\mathcal{I}\mathcal{J}} \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{J}|}$, and $A_{\mathcal{J}\mathcal{I}} \in \mathbb{R}^{|\mathcal{J}| \times |\mathcal{I}|}$. By construction, $A_{\mathcal{I}\mathcal{J}} = \mathbb{0}_{|\mathcal{I}| \times |\mathcal{J}|}$ so that PAP^{T} is block triangular, which is in contradiction with the assumed statement (ii).

Regarding (iv) $\implies$ (ii), by contradiction, assume there exists a permutation matrix P and a number $r < n$ such that

$$PAP^{\mathsf{T}} = \begin{bmatrix} B & C \\ \mathbb{0}_{(n-r) \times r} & D \end{bmatrix},$$

where the matrices $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times (n-r)}$, and $D \in \mathbb{R}^{(n-r) \times (n-r)}$ are arbitrary. The permutation matrix P defines a unique permutation $\pi \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ with the property that the columns of P are $\mathbb{e}_{\pi(1)}, \ldots, \mathbb{e}_{\pi(n)}$. Let $\mathcal{J} = \{\pi^{-1}(1), \ldots, \pi(r)^{-1}\}$ and $\mathcal{I} = \{1, \ldots, n\} \setminus \mathcal{J}$. Then, by construction, for any pair $(i, j) \in \mathcal{I} \times \mathcal{J}$, we know $a_{ij} = 0$, which is in contradiction with the assumed statement (iv). ∎

4.3.2   Connectivity and adjacency powers – with self-loops

Next we present two results, whose proof are analogous to those of the previous theorem and left to the reader as an exercise.

Lemma 4.4 (Global reachability and powers of the adjacency matrix). *Let G be a weighted digraph with $n \geq 2$ nodes and weighted adjacency matrix A. For any $j \in \{1, \ldots, n\}$, the following statements are equivalent:*

(i) *the jth node of G is globally reachable, and*

(ii) *the jth column of $\sum_{k=0}^{n-1} A^k$ is positive.*

Next, we notice that if node j is reachable from node i via a walk of length k and at least one node along that walk has a self-loop, then node j is reachable from node i via walks of length k, $k+1$, $k+2$, and so on. This observation and the last lemma lead to the following corollary.

Corollary 4.5 (Connectivity properties of the digraph and positive powers of the adjacency matrix: cont'd). *Let G be a weighted digraph with n nodes, weighted adjacency matrix A and a self-loop at each node. The following statements are equivalent:*

(i) G is strongly connected; and

(ii) A^{n-1} is positive, so that A is primitive.

For any $j \in \{1, \ldots, n\}$, the following two statements are equivalent:

(iii) the jth node of G is globally reachable; and

(iv) the jth column of A^{n-1} has positive entries.

Finally, we conclude this section with a clarification.

Remark 4.6 (Similarity transformations defined by permutation matrices). *Note that PAP^{T} is the similarity transformation of A defined by P because the permutation matrix P satisfies $P^{-1} = P^{\mathsf{T}}$; see Exercise E2.12. Moreover, note that PAP^{T} is simply a reordering of rows and columns. For example, consider* $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ *with* $P^{\mathsf{T}} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. *Note* $P \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ *and* $P^{\mathsf{T}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$

and compute

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \implies PAP^{\mathsf{T}} = \begin{bmatrix} a_{22} & a_{23} & a_{21} \\ a_{32} & a_{33} & a_{31} \\ a_{12} & a_{13} & a_{11} \end{bmatrix},$$

so that the entries of the 1st, 2nd and 3rd rows of A are mapped respectively to the 3rd, 1st and 2nd rows of PAP^{T} — and, at the same time, — the entries of the 1st, 2nd and 3rd columns of A are mapped respectively to the 3rd, 1st and 2nd columns of PAP^{T}. •

4.4   Graph theoretical characterization of primitive matrices

In this section we present the main result of this chapter, an immediate corollary and its proof; see also Figure 4.3.

Theorem 4.7 (Strongly connected and aperiodic digraphs and primitive adjacency matrices). *Let G be a weighted digraph with $n \geq 2$ nodes and with weighted adjacency matrix A. The following two statements are equivalent:*

(i) G is strongly connected and aperiodic; and

(ii) A is primitive, that is, there exists $k \in \mathbb{N}$ such that A^k is positive.

$$A \qquad\qquad A^2 \qquad\qquad A^3 \qquad\qquad A^4 \qquad\qquad A^5$$
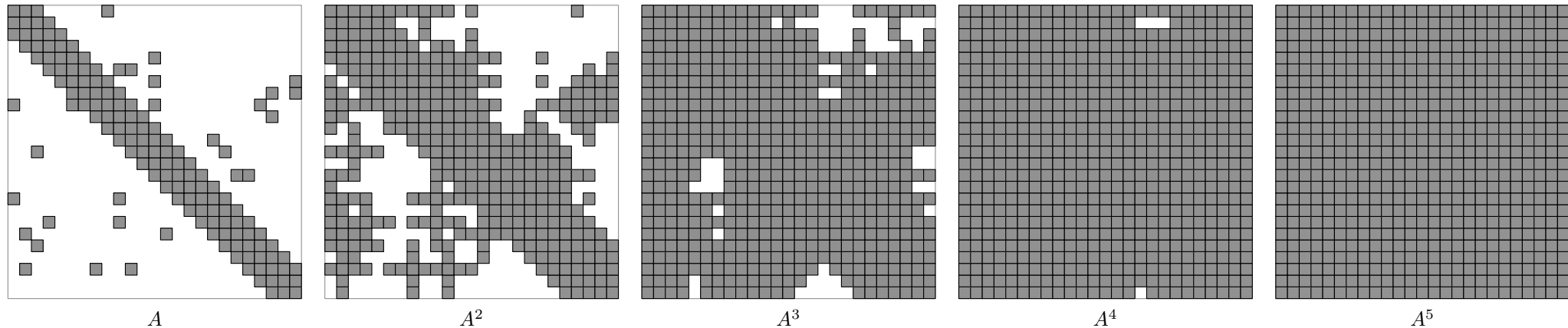
Figure 4.3: Pixel pictures of the increasing powers of a non-negative matrix $A \in \mathbb{R}^{25 \times 25}$. The digraph associated to A is strongly connected and has self-loops at each node; as predicted by Theorem 4.7, there exists $k = 5$ such that $A^k > 0$.

4.4.1 Proof

Before proving Theorem 4.7, we introduce a useful fact from number theory, whose proof we leave as Exercise E4.16. First, we recall a useful notion: a set of integers are *coprime* if its elements share no common positive factor except 1, that is, their greatest common divisor is 1. Loosely, the following lemma states that coprime numbers generate, via linear combinations with non-negative integer coefficients, all numbers larger than a given threshold.

Lemma 4.8 (Frobenius number). *Given a finite set $A = \{a_1, a_2, \ldots, a_n\}$ of positive integers, an integer M is said to be representable by A if there exist non-negative integers $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ such that $M = \alpha_1 a_1 + \ldots + \alpha_N a_N$. The following statements are equivalent:*

  (i) *there exists a (finite) largest unrepresentable integer, called the Frobenius number of A, and*

  (ii) *the greatest common divisor of A is 1.*

For example, one can show that, for the set of coprime numbers $A = \{3, 5\}$, the largest unrepresentable integer is 7. As another example, the set $A = \{5, 10\}$ (imagine nickels and dimes in US currency) is not coprime and, indeed, there is an infinite sequence of unrepresentable numbers (i.e., there is an infinite number of monetary amounts that cannot be obtained using only nickels and dimes).

Proof of Theorem 4.7. Regarding (i) $\implies$ (ii), pick any ordered pair (i, j). We claim that there exists a number $k(i, j)$ with the property that, for all $m > k(i, j)$, we have $(A^m)_{ij} > 0$, that is, there exists a directed walk from i to j of length m for all $m > k(i, j)$. If this claim is correct, then the statement (ii) is proved with $k = \max\{k(i, j) \mid i, j \in \{1, \ldots, n\}\}$. To show this claim, let $\{c_1, \ldots, c_N\}$ be the set of the simple cycles of G and let $\{\ell_1, \ldots, \ell_N\}$ be their lengths. Because G is aperiodic, the lengths $\{\ell_1, \ldots, \ell_N\}$ are coprime and Lemma 4.8 implies the existence of a number $h(\ell_1, \ldots, \ell_N)$ such that any number larger than $h(\ell_1, \ldots, \ell_N)$ is a linear combination of $\ell_1, \ldots, \ell_N$ with non-negative integer as coefficients. Because G is strongly connected, there exists a walk γ of arbitrary length $\Gamma(i, j)$ that starts at i, contains a node of each of the cycles $c_1, \ldots, c_N$, and terminates at j. Now, we claim that $k(i, j) = \Gamma(i, j) + h(\ell_1, \ldots, \ell_N)$ has the desired property. Indeed, pick any number $m > k(i, j)$ and write it as $m = \Gamma(i, j) + \beta_1 \ell_1 + \cdots + \beta_N \ell_N$ for appropriate numbers $\beta_1, \ldots, \beta_N \in \mathbb{N}$. As illustrated in Figure 4.4, a directed walk from i to j of length m is constructed by attaching to the walk γ the following cycles: β_1 times the cycle c_1, β_2 times the cycle c_2, $\ldots$, β_N times the cycle c_N.
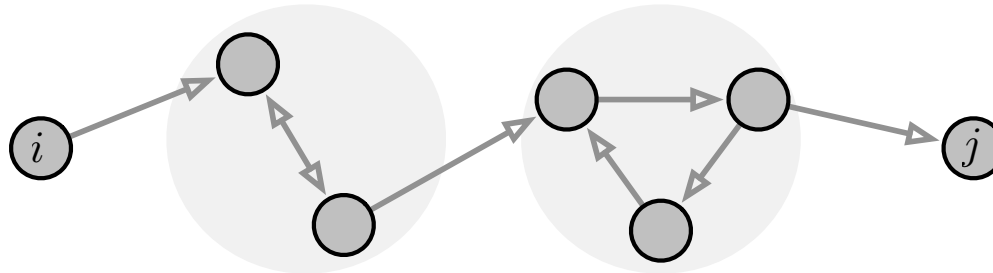
Figure 4.4: Illustration of the proof of Theorem 4.7: since 2 and 3 are coprime numbers, there exist only a finite set of numbers k for which there is no walk from i to j of length k.

Regarding (ii) $\implies$ (i), from Lemma 4.2 we know that $A^k > 0$ means that there are walks of length k from every node to every other node. Hence, the digraph G is strongly connected. Next, we prove aperiodicity. Because G is strongly connected, each node of G has at least one outgoing edge, that is, for all i, there exists at least one index j such that $a_{ij} > 0$. This fact implies that the matrix $A^{k+1} = A A^k$ is positive via the following simple calculation: $(A^{k+1})_{il} = \sum_{h=1}^{n} a_{ih}(A^k)_{hl} \geq a_{ij}(A^k)_{jl} > 0$. In summary, if A^k is positive for some k, then A^m is positive for all subsequent $m > k$ (see also Exercise E2.5). Therefore, there are closed walks in G of any sufficiently large length. By Exercise E3.8, this fact implies that G is aperiodic. ∎

4.5   Elements of spectral graph theory

In this section we switch topic and provide some elementary useful results on the spectral radius of a non-negative matrix A. These bounds play an important role, e.g., in understanding various monotonicity and convergence results in later chapters.

Lemma 4.9 (Bounds on the spectral radius of non-negative matrices, I). *For a non-negative matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$, vector $x \in \mathbb{R}_{\geq 0}^n$, $x \neq \mathbb{0}_n$, and scalars $r_1, r_2 > 0$, the following statements hold:*

  (i) *if $r_1 x \leq Ax$, then $r_1 \leq \rho(A)$,*

  (ii) *if $Ax \leq r_2 x$ and $x \in \mathbb{R}_{>0}^n$, then $\rho(A) \leq r_2$.*

Moreover, for an irreducible matrix A,

  (iii) *if $r_1 x \leq Ax \leq r_2 x$ and $r_1 x \neq Ax \neq r_2 x$, then $r_1 < \rho(A) < r_2$ and $x \in \mathbb{R}_{>0}^n$.*

Proof. Regarding statement (i), define $A_{r_1} = A/r_1$ and assume by absurd $\rho(A_{r_1}) < 1$. If $A_{r_1}x \geq x$, then multiplying the left and right hand side by A_{r_1} one can see that $A_{r_1}^k x \geq x$ for all $k \in \mathbb{N}$. But $\rho(A_{r_1}) < 1$ implies $\lim_{k\to\infty} A_{r_1}^k = \mathbb{0}_{n\times n}$. Therefore, for all nonzero non-negative x, we know there exists sufficiently large k such that $A_{r_1}^k x > x$. We found a contradiction and so we know $\rho(A/r_1) \geq 1$.

Regarding statement (ii), the Perron–Frobenius Theorem 2.12 for non-negative matrices implies the existence of $w \geq 0$, $w \neq \mathbb{0}_n$, such that $w^\mathsf{T} A = \rho(A)w^\mathsf{T}$. Left-multiplying $Ax \leq r_2 x$ by w^T leads to

$$\rho(A)(w^\mathsf{T}x) \leq r_2(w^\mathsf{T}x)$$

and the claims follows from noting that $x > 0$ implies $w^\mathsf{T}x > 0$.

Regarding statement (iii), we prove the lower bound. First, we note that the left dominant eigenvector w is now strictly positive because A is irreducible. This implies $w^\mathsf{T}x > 0$. Second, by assumption we know that $r_1 x_i \leq (Ax)_i$ for all i and that there exists at least one index j such that $r_1 x_j < (Ax)_j$. Therefore, since $w > 0$, we know $w^\mathsf{T}(r_1 x) < w^\mathsf{T}(Ax)$. This last inequality immediately implies $r_1 < \rho(A)$. The upper bound is obtained similarly.

The proof of positivity of x is left as an exercise to the reader. ∎

We next show how these bounds allow us to establish the monotonicity of the spectral radius as a function of matrix entries. The following lemma is the first result on this type of monotonicity property; we refer to Lemma 6.9, Example 6.13, Exercises E10.5 and E10.6 for similar results later in the book.

Lemma 4.10 (Monotonicity of spectral radius of non-negative matrices). *Let A and A' be non-negative matrices in $\mathbb{R}_{\geq 0}^{n \times n}$. Then*

(i) if $A \leq A'$, then $\rho(A) \leq \rho(A')$,

(ii) if additionally $A \neq A'$ and A' is irreducible, then $\rho(A) < \rho(A')$.

Proof. Regarding statement (i), from the Perron–Frobenius Theorem 2.12 for non-negative matrices let $(\rho(A), v)$ be a non-negative eigenpair for A. Compute

$$\rho(A)v = Av \le A'v. \tag{4.2}$$

Therefore Lemma 4.9(i) applied to the matrix A' states that $\rho(A) \le \rho(A')$. This proves statement (i). (An alternative proof relies upon Gelfand's formula; see Exercise E4.18.)

Next, we prove statement (ii) under the additional assumption that A is irreducible. From the Perron–Frobenius Theorem 2.12 for irreducible matrices let $(\rho(A), v)$ be a positive eigenpair for A. Equation (4.2) continues to hold, but $v > 0$ now implies also that $A'v \ne Av = \rho(A)v$. Therefore, Lemma 4.9(iii) implies that $\rho(A) < \rho(A')$.

Finally, we prove statement (ii) without the additional assumption that A is irreducible, i.e., we assume A is reducible and A' is irreducible. Define $B \in \mathbb{R}_{\ge 0}^{n \times n}$ by $b_{ij} = a_{ij}$ if $a_{ij} \ne 0$, and $b_{ij} = a'_{ij}/2$ if $a_{ij} = 0$. By design, B satisfies $A \le B \le A'$, and, therefore, statement (i) implies $\rho(A) \le \rho(B) \le \rho(A')$. Moreover, B is irreducible (since it has the same zero/positive pattern as A') and $B \ne A'$ so that the argument in the previous paragraph (where A is assumed irreducible) implies $\rho(B) < \rho(A')$. ■

Next, we specialize the results in Lemma 4.9 and provide a necessary and sufficient characterization for the spectral radius to be strictly less than the maximum out-degree. In what follows, recall that ith entry of the vector $A\mathbb{1}_n$ contains the ith row-sum of the matrix A and the out-degree of the ith node of the digraph associated to A. In other words, $d_{\text{out}}(i) = e_i^{\mathsf{T}} A\mathbb{1}_n$.

Theorem 4.11 (Bounds on the spectral radius of non-negative matrices, II). *For a non-negative matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$ with associated digraph G, the following statements hold:*

(i) $\min(A\mathbb{1}_n) \leq \rho(A) \leq \max(A\mathbb{1}_n)$; and

(ii) if $\min(A\mathbb{1}_n) < \max(A\mathbb{1}_n)$, then the following two statements are equivalent:

(a) for each node i with $e_i^{\mathsf{T}} A\mathbb{1}_n = \max(A\mathbb{1}_n)$, there exists a directed walk in G from node i to a node j with $e_j^{\mathsf{T}} A\mathbb{1}_n < \max(A\mathbb{1}_n)$; and

(b) $\rho(A) < \max(A\mathbb{1}_n)$.

An illustration of this result is given in Figure 4.5. Before providing the proof, we introduce a useful notion and establish a corollary.
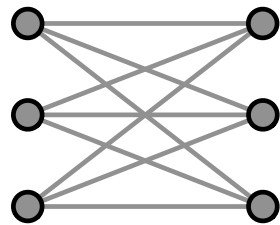
(a) Complete bipartite graph $K_{3,3}$ with binary adjacency matrix $A_{K_{3,3}}$
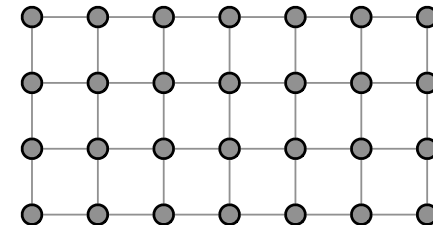
(b) Two-dimensional grid graph $G_{4,7}$ with binary adjacency matrix $A_{G_{4,7}}$

Figure 4.5: Illustration of Theorem 4.11 applied to the symmetric binary adjacency matrices of two undirected graphs. By counting the number of neighbors of each node (i.e., by computing the row sums of A) and observing that the grid graph is connected, we can establish that $\rho(A_{K_{3,3}}) = 3$ and $2 < \rho(A_{G_{4,7}}) < 4$.

We now apply this necessary and sufficient characterization to a useful class of non-negative matrices.

Definition 4.12 (Row-substochastic matrix). *A non-negative matrix $A \in \mathbb{R}^{n \times n}$ is row-substochastic if its row-sums are at most 1 and at least one row-sum is strictly less than 1, that is,*

$$A \mathbb{1}_n \leq \mathbb{1}_n, \text{ and there exists } i \in \{1, \ldots, n\} \text{ such that } \mathbb{e}_i^\top A \mathbb{1}_n < 1.$$

Note that any row-substochastic matrix A with at least one row-sum equal to 1 satisfies $\min(A \mathbb{1}_n) < \max(A \mathbb{1}_n)$ and that any irreducible row-substochastic matrix satisfies condition (ii)a in Theorem 4.11 because the associated digraph is strongly connected. These two observations allow us to characterize when row-substochastic matrices are convergent.

Corollary 4.13 (Convergent row-substochastic matrices). *Let A be row-substochastic with associated digraph G.*

(i) *A is convergent if and only if its G contains directed walks from each node with out-degree 1 to a node with out-degree less than 1, and*

(ii) *if A is irreducible, then A is convergent.*

We now present the proof of the main theorem in this section.

Proof of Theorem 4.11. We start by noting that statement (i) is an immediate consequence of Lemma 4.9(i) and (ii) with $x = \mathbb{1}_n$, $r_1 = \min(A\mathbb{1}_n)$ and $r_2 = \max(A\mathbb{1}_n)$.

Next, we establish that the condition (ii)a implies the bound (ii)b. It suffices to focus on row-substochastic matrices (if $\max(A\mathbb{1}_n) \neq 1$, we consider the row-substochastic matrix $A/\max(A\mathbb{1}_n)$). We now claim that:

(1) if $e_i^\mathsf{T} A^h \mathbb{1}_n < 1$ for some $h \in \mathbb{N}$, then $e_i^\mathsf{T} A^{h+1} \mathbb{1}_n < 1$,

(2) if i has an out-neighbor j (that is, $A_{ij} > 0$) with $e_j^\mathsf{T} A^h \mathbb{1}_n < 1$ for some $h \in \mathbb{N}$, then $e_i^\mathsf{T} A^{h+1} \mathbb{1}_n < 1$,

(3) there exists k such that $A^k \mathbb{1}_n < \mathbb{1}_n$, and

(4) $\rho(A) < 1$.

Regarding statement *(1)*, for a node i satisfying $e_i^\mathsf{T} A^h \mathbb{1}_n < 1$, we compute

$$e_i^\mathsf{T} A^{h+1} \mathbb{1}_n = e_i^\mathsf{T} A^h (A\mathbb{1}_n) \leq e_i^\mathsf{T} A^h \mathbb{1}_n < 1,$$

where we used the implication: if $\mathbb{0}_n \leq v \leq \mathbb{1}_n$ and $w \geq \mathbb{0}_n$, then $w^\mathsf{T} v \leq w^\mathsf{T} \mathbb{1}_n$. This proves statement *(1)*. Next, note that $A\mathbb{1}_n \leq \mathbb{1}_n$ implies $A^h \mathbb{1}_n \leq \mathbb{1}_n$. Moreover, $e_j^\mathsf{T} A^h \mathbb{1}_n < 1$ implies that the jth entry can be written as $(A^h \mathbb{1}_n)_j = e_j^\mathsf{T} A^h \mathbb{1}_n = 1 - (1 - e_j^\mathsf{T} A^h \mathbb{1}_n)$, where $(1 - e_j^\mathsf{T} A^h \mathbb{1}_n) > 0$. In summary,

$$A^h \mathbb{1}_n \leq \mathbb{1}_n - \left(1 - e_j^\mathsf{T} A^h \mathbb{1}_n\right) e_j.$$

Therefore, we compute

$$\begin{aligned}
e_i^\mathsf{T} A^{h+1} \mathbb{1}_n &= (e_i^\mathsf{T} A)(A^h \mathbb{1}_n) \\
&\leq (e_i^\mathsf{T} A)\left(\mathbb{1}_n - \left(1 - e_j^\mathsf{T} A^h \mathbb{1}_n\right) e_j\right) \\
&= e_i^\mathsf{T} A\mathbb{1}_n - \left(1 - e_j^\mathsf{T} A^h \mathbb{1}_n\right) e_i^\mathsf{T} A e_j \leq 1 - \left(1 - e_j^\mathsf{T} A^h \mathbb{1}_n\right) A_{ij} < 1.
\end{aligned}$$

This concludes the proof of statement *(2)*.

Regarding statement *(3)*, note that, if A is row-substochastic, then A^h is row-substochastic for any natural $h \geq 1$. Let S_h be the set of indices i such that the ith row-sum of A^h is strictly less than 1. Statement *(1)* implies $S_h \subseteq S_{h+1}$. Moreover, because of the existence of directed walks from every node to nodes with row-sum less than 1, statement *(2)* implies that there exists k such that $S_k = \{1, \ldots, n\}$. This proves statement *(3)*.

Next, define the maximum row-sum at time k by

$$\gamma = \max_{i \in \{1, \ldots, n\}} \sum_{j=1}^{n} (A^k)_{ij} < 1.$$

Given any natural number $k^* \geq k$, we can write $k^* = ak + b$ with a a positive integer and $b \in \{0, \ldots, k-1\}$. Note that

$$A^{k^*} \mathbb{1}_n \leq A^{ak} \mathbb{1}_n \leq \gamma^a \mathbb{1}_n.$$

The last inequality implies that, as $k^* \to \infty$ and therefore $a \to \infty$, the sequence A^{k^*} converges to 0. This fact proves statement *(4)* and, in turn, that the condition (ii)a implies the bound (ii)b.

Finally, we sketch the proof that the bound (ii)b implies the condition (ii)a. By contradiction, if condition (ii)a does not hold, then the condensation of G contains a sink whose corresponding row-sums in A are all equal to $\max(A\mathbb{1}_n)$. But to that sink corresponds an eigenvector of A whose eigenvalue is therefore $\max(A\mathbb{1}_n)$. We refer to Theorem 5.2 for a brief review of the properties of reducible non-negative matrix and leave to the reader the details of the proof. ∎

4.6   Appendix: Balancing and scaling nonnegative matrices

In this appendix we consider balancing and scaling problems. These optimization problems have a rich history, going back all the way to the 1960s. As reviewed by Idel (2016), these problems are related transportation planning in geography, contingency table analysis in statistics, pre-conditioning matrices in scientific computing, matching problems in combinatorial optimization.
   A nonnegative matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$ is

 (i) *balanced* if it has identical row and column sums, that is, $A \mathbb{1}_n = A^\mathsf{T} \mathbb{1}_n$, and

 (ii) (r, c)-*scaled*, for nonnegative vectors $r, c \in \mathbb{R}_{\geq 0}^n$, if it has row sums r and column sums c, that is, $A \mathbb{1}_n = r$ and $A^\mathsf{T} \mathbb{1}_n = c$.

   For example, it is known[1] that, for any tridiagonal matrix A, there exists a diagonal positive D such that DAD^{-1} is symmetric. Therefore any nonnegative tridiagonal matrix is balancable.
   Recall the notion of the *open simplex* $\mathring{\Delta}_n = \{x \in \mathbb{R}^n \mid x > 0, \mathbb{1}_n^\mathsf{T} x = 1\}$.

Theorem 4.14 (Balancing and block irreducibility). *Given a nonnegative $A \in \mathbb{R}_{\geq 0}^{n \times n}$, the following statements are equivalent:*

 (i) *there exists a positive diagonal matrix D such that DAD^{-1} is balanced (i.e., A is* balancable *and* balanced *by D);*

 (ii) *the digraph associated to A is the union of strongly connected components;*

 (iii) *the* balancing function *$f : \mathring{\Delta}_n \to \mathbb{R}$ defined by $f(x) = \sum_{i,j=1}^n a_{ij} x_i x_j^{-1}$ has a minimum point $x^* \in \mathring{\Delta}_n$ such that $\mathrm{diag}(x^*) A \, \mathrm{diag}(x^*)^{-1}$ is balanced;*

 (iv) *there exists a matrix $B \in \mathbb{R}_{\geq 0}^{n \times n}$ with the same pattern as A satisfying $B \mathbb{1}_n = B^\mathsf{T} \mathbb{1}_n$; and*

 (v) *the matrix $D_1 D_2 \dots D_k$ defined by the diagonal balancing algorithm, given below, approximately balances A.*

The scaling of A is unique and D is unique up to scalars for each irreducible block of A.

---

[1]For example see https://en.wikipedia.org/wiki/Tridiagonal_matrix

Diagonal balancing algorithm. At each iteration, the diagonal balancing algorithm identifies the index corresponding to the maximum violation of the "row sum is equal to colum sum" condition. The resulting row and columns are scaled in such a way as that their sums are equal.

Input: nonnegative matrix $A \in \mathbb{R}^{n \times n}$ and error threshold $\varepsilon > 0$
Output: a diagonal positive matrix D approximately balancing A, with a maximum "row sum minus column sum" error upper bounded by ε

1:    $A_1 := A, \quad k := 0$
2: **repeat**
3:      $k := k + 1$
4:      compute $r := A_k \mathbb{1}_n$ and $c := A_k^\mathsf{T} \mathbb{1}_n$
5:      find index i such that $|r_i - c_i| = \|r - c\|_\infty$
        // (this step identifies the index with maximum violation)
6:      define $D_k := \mathrm{diag}(1, \ldots, \underbrace{\sqrt{c_i / r_i}}_{i^{\text{th}} \text{ entry}}, \ldots, 1)$
7:      update $A_{k+1} := D_k A_k D_k^{-1}$
8: **until** $\|r - c\|_\infty \leq \varepsilon$
9: **return** $D_1 D_2 \ldots D_k$

Next, we consider scaling problems.

Theorem 4.15 (Scaling and matrix patterns). *Given a nonnegative $A \in \mathbb{R}_{\geq 0}^{n \times n}$ and nonnegative vectors vectors $r, c \in \mathbb{R}_{\geq 0}^n$ such that $\mathbb{1}^\mathsf{T} r = \mathbb{1}^\mathsf{T} c$, the following statements are equivalent:*

(i) there exists positive diagonal matrices D_1 and D_2 such that $D_1 A D_2$ is (r, c)-scaled;

(ii) there exists a matrix $B \in \mathbb{R}_{\geq 0}^{n \times n}$ with the same pattern as A satisfying $B \mathbb{1}_n = r$ and $B^\mathsf{T} \mathbb{1}_n = c$;

(iii) for every $\mathcal{I}, \mathcal{J} \subset \{1, \dots, n\}$ such that $A_{\mathcal{I}^c \mathcal{J}} = 0$, we have that $\sum_{i \in \mathcal{I}} r_i \geq \sum_{j \in \mathcal{J}} c_j$ and equality holds if and only if $A_{\mathcal{I}\mathcal{J}^c} = 0$; and

(iv) the matrices $L_1 L_2 \dots L_k$ and $R_1 R_2 \dots R_k$ defined by the SK scaling algorithm, given below, approximately (r, c)-scales A.

Sinkhorn-Knopp (SK) scaling algorithm. The SK algorithm alternates between scaling rows and columns of A. At each iteration, each row is multiplies by a positive constant to ensure the desired row sum; then the same operation is executed for each column. The symbol $\oslash$ mean entry-wise division.

Input: nonnegative matrix $A \in \mathbb{R}^{n \times n}$ and nonnegative vectors vectors $r, c \in \mathbb{R}_{\geq 0}^n$ such that $\mathbb{1}^\mathsf{T} r = \mathbb{1}^\mathsf{T} c$
Output: an approximately (r, c)-scaled matrix related to A by left and right multiplication by positive diagonal matrices

1: $A_{\text{tmp}} := A, \quad k := 1$
2: **repeat**
3: $\quad L_k := \operatorname{diag}(r \oslash (A_{\text{tmp}} \mathbb{1}_n)), A_{\text{tmp}} := L_k A_{\text{tmp}}$
$\quad$ // *(this step amounts to rescaling A_{tmp} so that its row sums are equal to r)*
4: $\quad R_k := \operatorname{diag}(c \oslash (A_{\text{tmp}}^\mathsf{T} \mathbb{1}_n)), A_{\text{tmp}} := A_{\text{tmp}} R_k$
$\quad$ // *(this step amounts to rescaling A_{tmp} so that its column sums are equal to c)*
5: $\quad k := k + 1$
6: **until** row sums of A_{tmp} are approximately equal to r
7: **return** $L_1 L_2 \dots L_k$ and $R_1 R_2 \dots R_k$

4.7   Historical notes and further reading

All results in this chapter are well documented in standard books on non-negative matrices such as (Seneta, 1981; Minc, 1988; Berman and Plemmons, 1994). For example, Lemma 4.9 is (Berman and Plemmons, 1994, Chapter 2, Theorem 1.11, page 28) and a slightly weaker version of Lemma 4.10 is (Berman and Plemmons, 1994, Chapter 2, Corollary 1.5, page 27). Also relevant are standard books on algebraic graph theory (Biggs, 1994; Godsil and Royle, 2001).

For more information on the Frobenius number we refer to (Owens, 2003) and, for an informal read, to Wikipedia:Coin_Problem.

More results on spectral graph theory and, specifically, a review and recent results on bounding the spectral radius of an adjacency matrix are given, for example, by Nikiforov (2002) and Das and Kumar (2004). Bounds on the eigenvalues of the Laplacian matrix are given in Section 6.2 and Exercise E6.6.

A comprehensive review of balancing and scaling problems is given by Idel (2016). Early references include (Osborne, 1960) on balancing and (Sinkhorn and Knopp, 1967) on scaling problems. Theorem 4.14 presents analysis results from (Idel, 2016, Theorem 5.1) and (Eaves et al., 1985, Theorem 3) and algorithmic results from (Schneider and Zenios, 1990). The results in Theorem 4.15 are taken from (Idel, 2016, Theorem 4.1). State-of-the-art algorithms are given by (Cohen et al., 2017).

4.8 Exercises

E4.1 **Example row-stochastic matrices and associated digraph.** Consider the row-stochastic matrices

$$A_1 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}, \quad A_2 = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad A_3 = \frac{1}{2} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

(i) Draw the digraphs G_1, G_2 and G_3 associated with these three matrices.

Using only the original definitions and without relying on the characterizations in Theorems 4.3 and 4.7, show that:

(ii) the matrices A_1, A_2 and A_3 are irreducible and primitive,

(iii) the digraphs G_1, G_2 and G_3 are strongly connected and aperiodic, and

(iv) the averaging algorithm defined by A_2 converges in a finite number of steps.

E4.2   **Tridiagonal Toeplitz matrices.** An $n \times n$ matrix A is *tridiagonal Toeplitz* if there exist numbers a, b, and c, with $a \neq 0$ and $c \neq 0$ such that

$$A = \begin{bmatrix} b & a & 0 & \dots & 0 \\ c & b & a & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & c & b & a \\ 0 & \dots & 0 & c & b \end{bmatrix}.$$

Show that the eigenvalues and right eigenvectors of a tridiagonal Toeplitz A are, for $j \in \{1, \dots, n\}$,

$$\lambda_j = b + 2a\sqrt{c/a}\cos\left(\frac{j\pi}{n+1}\right), \quad \text{and} \quad v_j = \begin{bmatrix} (c/a)^{1/2}\sin(1j\pi/(n+1)) \\ (c/a)^{2/2}\sin(2j\pi/(n+1)) \\ \vdots \\ (c/a)^{n/2}\sin(nj\pi/(n+1)) \end{bmatrix}.$$

E4.3   **Circulant matrices.** A matrix $C \in \mathbb{C}^{n \times n}$ is *circulant* if there exists numbers $c_0, \ldots, c_{n-1}$ such that

$$C = \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}.$$

In other words, a circulant matrix is fully specified by its first row; the remaining row of C are cyclic permutations of the first row. A circulant matrix is Toeplitz. Show that

(i) the eigenvalues and eigenvectors C are, for $j \in \{0, \ldots, n-1\}$,

$$\lambda_j = c_0 + c_1 \omega_j + c_2 \omega_j^2 + \cdots + c_{n-1} \omega_j^{n-1}, \quad \text{and} \quad v_j = \begin{bmatrix} 1 \\ \omega_j \\ \vdots \\ \omega_j^{n-1} \end{bmatrix},$$

where $\omega_j = \exp\left(\dfrac{2j\pi i}{n}\right), j \in \{0, \ldots, n-1\}$, are the nth complex roots of the number 1, and $i = \sqrt{-1}$.

(ii) for n even, $\kappa \in \mathbb{R}$, and $(c_0, c_1, \ldots, c_{n-1}) = (1 - 2\kappa, \kappa, 0, \ldots, 0, \kappa)$, the eigenvalues are

$$\lambda_j = 2\kappa \cos \frac{2\pi(j-1)}{n} + (1 - 2\kappa), \quad j \in \{1, \ldots, n\}.$$

Note: Circulant matrices enjoy numerous properties; e.g., if C_1 and C_2 are circulant, so are C_1^T, $C_1 + C_2$ and $C_1 C_2$. Additional properties are discussed for example by Davis (1979).

E4.4 **Adjacency spectrum of basic graphs.** Given the basic graphs in Examples 3.1 and 4.1 and the properties of tridiagonal Toeplitz and circulant matrices in Exercises E4.2 and E4.3, prove the statements in Table 4.1. In other words, show that, for $n \geq 2$,

(i) for the *path graph P_n*, the adjacency matrix is Toeplitz tridiagonal and the adjacency spectrum is $\{2\cos(\pi i/(n+1)) \mid i \in \{1, \ldots, n\}\}$;

(ii) for the *cycle graph C_n*, the adjacency matrix is circulant and the adjacency spectrum is $\{2\cos(2\pi i/n)) \mid i \in \{1, \ldots, n\}\}$;

(iii) for the *star graph S_n*, the adjacency matrix is $e_1 e_{-1} + e_{-1} e_1$, where $e_{-i} = \mathbb{1}_n - e_i$, and the adjacency spectrum is $\{\sqrt{n-1}, 0, \ldots, 0, -\sqrt{n-1}\}$;

(iv) for the *complete graph K_n*, the adjacency matrix is $\mathbb{1}_n \mathbb{1}_n^{\mathsf{T}} - I_n$, and the adjacency spectrum is $\{(n-1), -1, \ldots, -1\}$; and

(v) for the *complete bipartite graph $K_{n,m}$*, the adjacency matrix is $\begin{bmatrix} \mathbb{0}_{n \times n} & \mathbb{1}_{n \times m} \\ \mathbb{1}_{m \times n} & \mathbb{0}_{m \times m} \end{bmatrix}$ and the adjacency spectrum is $\{\sqrt{nm}, 0, \ldots, 0, -\sqrt{nm}\}$.

E4.5   **Edges and triangles in an undirected graph.** Let A be the binary adjacency matrix for an undirected graph G without self-loops. Recall that the trace of A is $\text{trace}(A) = \sum_{i=1}^{n} a_{ii}$.

   (i) Show $\text{trace}(A) = 0$.

   (ii) Show $\text{trace}(A^2) = 2|E|$, where $|E|$ is the u number of edges of G.

   (iii) Show $\text{trace}(A^3) = 6|T|$, where $|T|$ is the number of triangles of G. (A triangle is a complete subgraph with three nodes.)

   (iv) Verify results (i)–(iii) on the matrix $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

▶ E4.6   **Directed walks and powers of the adjacency matrix.** Prove Lemma 4.2.

E4.7 **An additional characterization of irreducibility.** In this exercise we provide an additional characterization of irreducibility to Theorem 4.3. Given a non-negative matrix A of dimension n, show that the following statements are equivalent:

(i) there exists no permutation matrix P such that PAP^T is block triangular, and

(ii) for any non-negative vector $y \in \mathbb{R}^n_{\geq 0}$ with $0 < k < n$ strictly positive components, the vector $(I_n + A)y$ has at least $k + 1$ strictly positive components.

E4.8   **An example reducible or irreducible matrix.** Consider the binary matrix:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Draw the digraph associated to A. Prove that A is irreducible or prove that A is reducible by providing a permutation matrix P that transforms A into an upper block-triangular matrix.

► E4.9   **Characterization of indecomposable matrices.** Following (Wolfowitz, 1963), we say a non-negative matrix A is *indecomposable* if its associated digraph contains a globally reachable node. Generalizing the proof of Theorem 4.7, show that the following statements are equivalent:

(i) A is indecomposable and the subgraph of globally reachable nodes is aperiodic, and

(ii) there exists an index $h \in \mathbb{N}$ such that A^h has a positive column.

E4.10   **A sufficient but not necessary condition for primitivity.** Assume the square matrix A is non-negative and irreducible. Show that

(i) if A has a positive diagonal element, then A is primitive,

(ii) if A is primitive, then it is false that A must have a positive diagonal element.

E4.11 **Primitive matrices are irreducible.** Prove Lemma 2.11 relying upon Theorem 4.3 instead of the Caley-Hamilton Theorem.

▶ E4.12   **The exponent of a primitive matrix.**

  (i)   Let G be the digraph with nodes $\{1, 2, 3\}$ and edges $\{(1, 2), (2, 3), (3, 1), (2, 1)\}$. Explain if and why G is strongly connected and aperiodic.

 (ii)   Recall a non-negative matrix A is primitive if there exists a number k such that $A^k > 0$; the smallest such number k is called the *exponent* of the primitive matrix A. Do one of the following:

   (a)   prove that the exponent of a primitive matrix $A \in \mathbb{R}^{n \times n}$ is less than or equal to n, or

   (b)   provide a counterexample.

Note: Wielandt (1950) proved that the exponent of a primitive matrix of dimension n is upper bounded by $(n-1)^2 + 1$ and that this bound is sharp in the sense that there exist primitive matrices for which $A^{(n-1)^2}$ is not positive. It is also known that the exponent of a primitive matrix with positive diagonal is at most $n - 1$. We refer to (Brualdi and Ryser, 1991, Section 3.5) for an elegant treatment, including a generalization of the example in this exercise.

E4.13 **Normalization of non-negative irreducible matrices.** Consider a weighted digraph G and with an irreducible adjacency matrix $A \in \mathbb{R}^{n \times n}$. The matrix A is not necessarily row-stochastic. Find a positive vector $w \in \mathbb{R}^n$ so that the normalized matrix

$$P = \frac{1}{\rho(A)} (\mathrm{diag}(w))^{-1} A \, \mathrm{diag}(w)$$

is non-negative, irreducible, and row-stochastic.

E4.14 **Eigenvalue shifting for stochastic matrices.** Let $A \in \mathbb{R}^{n \times n}$ be an irreducible row-stochastic matrix. Let E be a diagonal matrix with diagonal elements $E_{ii} \in \{0, 1\}$, with at least one diagonal element equal to zero. Show that AE and EA are convergent.

E4.15   **Decomposition of irreducible row-stochastic matrices.** Let $A \in \mathbb{R}^{n \times n}$ be row-stochastic and irreducible. Pick a dimension $1 < k < n$, and define the
block-matrix decomposition

$$A = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, \quad \text{where } W_{11} \in \mathbb{R}^{k \times k} \text{ and } W_{22} \in \mathbb{R}^{(n-k) \times (n-k)}.$$

Show that

(i) $I_k - W_{11}$ and $I_{n-k} - W_{22}$ are invertible, and

(ii) $(I_k - W_{11})^{-1} W_{12} \mathbb{1}_{n-k} = \mathbb{1}_k$ and $(I_{n-k} - W_{22})^{-1} W_{21} \mathbb{1}_k = \mathbb{1}_{n-k}$.

E4.16   **The Frobenius number.** Prove Lemma 4.8.
         Hint: Read up on the Frobenius number in (Owens, 2003).

E4.17   **Induced norms and Gelfand's formula.** In this exercise we review the notion of induced norm and some of its useful properties. Given a norm $\|\cdot\|$ on $\mathbb{C}^n$, the *induced norm* of a square matrix $A \in \mathbb{C}^{n\times n}$ is

$$\|A\| = \max \left\{ \|Ax\| \mid x \in \mathbb{C}^n \text{ and } \|x\| = 1 \right\} = \sup_{x \neq \mathbb{0}_n} \frac{\|Ax\|}{\|x\|}.$$

Specifically, in the context of p-norms, for $p \in \mathbb{N} \cup \{\infty\}$, it is well known that, for $x \in \mathbb{C}^n$ and $A \in \mathbb{C}^n$,

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \qquad\qquad \|A\|_1 = \max_{j \in \{1,\dots,n\}} \sum_{i=1}^{n} |a_{ij}|, \tag{E4.1}$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \qquad\qquad \|A\|_2 = \sqrt{\lambda_{\max}(A^*A)}, \tag{E4.2}$$

$$\|x\|_\infty = \max_{i \in \{1,\dots,n\}} |x_i| \qquad\qquad \|A\|_\infty = \max_{i \in \{1,\dots,n\}} \sum_{j=1}^{n} |a_{ij}|, \tag{E4.3}$$

where A^* is the complex conjugate of A. The induced 2-norm of A is also known as the maximum singular value of A and $\|A\|_2 = \rho(A)$, if $A = A^*$.
   For any $A \in \mathbb{C}^{n\times n}$ and any induced matrix norm $\|\cdot\|$, show that

(i) $\rho(A) \leq \|A\|$,

(ii) $\rho(A) \leq \|A^k\|^{1/k}$ for all $k \in \mathbb{Z}_{\geq 0}$, and

(iii) $\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}$, also known as *Gelfand's formula*.

E4.18   **Monotonicity properties of spectral radius of non-negative matrices.** Given $A \in \mathbb{C}^{n \times n}$, let $|A| \in \mathbb{R}_{\geq 0}^n$ denote the matrix with entries $|a_{ij}|$. Given $A \in \mathbb{C}^{n \times n}$ and $E \in \mathbb{R}_{\geq 0}^n$, use Gelfand's formula in Exercise E4.17 to show

$$\rho(A) \leq \rho(|A|) \leq \rho(|A| + E).$$

Note: This result implies that, for example, if A and A' be two non-negative $n \times n$ dimensional matrices, then $A \leq A'$ implies $\rho(A) \leq \rho(A')$.

E4.19 **Leslie population model.** The Leslie model is used in population ecology to model the changes in a population of organisms over a period of time; see the original reference (Leslie, 1945) and a comprehensive text (Caswell, 2006). In this model, the population is divided into n groups based on age classes; the indices i are ordered increasingly with the age, so that $i = 1$ is the class of the newborns. The variable $x_i(k)$, $i \in \{1, \ldots, n\}$, denotes the number of individuals in the age class i at time k; at every time step k the $x_i(k)$ individuals

- produce a number $\alpha_i x_i(k)$ of offsprings (i.e., individuals belonging to the first age class), where $\alpha_i \geq 0$ is a fecundity rate, and
- progress to the next age class with a survival rate $\beta_i \in [0, 1]$.

If $x(k)$ denotes the vector of individuals at time k, the Leslie population model reads

$$x(k+1) = Ax(k) = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & \alpha_n \\ \beta_1 & 0 & \cdots & 0 & 0 \\ 0 & \beta_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_{n-1} & 0 \end{bmatrix} x(k), \tag{E4.4}$$

where A is referred to as the *Leslie matrix*. Consider the following two independent sets of questions. First, assume $\alpha_i > 0$ for all $i \in \{1, \ldots, n\}$ and $0 < \beta_i \leq 1$ for all $i \in \{1, \ldots, n-1\}$.

(i) Prove that the matrix A is primitive.

(ii) Let $p_i(k) = \frac{x_i(k)}{\sum_{i=1}^{n} x_i(k)}$ denote the percentage of the total population in class i at time k. Call $p(k)$ the *population distribution* at time k. Compute $\lim_{k \to +\infty} p(k)$ as a function of the spectral radius $\rho(A)$ and the parameters (α_i, β_i), $i \in \{1, \ldots, n\}$.

Hint: Obtain a recursive expression for the components of the right dominant eigenvector of A

(iii) Assume $\beta_i = \beta > 0$ and $\alpha_i = \frac{\beta}{n}$ for $i \in \{1, \ldots, n\}$. What percentage of the total population belongs to the eldest class asymptotically, that is, what is $\lim_{k \to \infty} p_n(k)$?

(iv) Find a sufficient condition on the parameters (α_i, β_i), $i \in \{1, \ldots, n\}$, so that the population will eventually become extinct.

Second, assume $\alpha_i \geq 0$ for $i \in \{1, \ldots, n\}$ and $0 \leq \beta_i \leq 1$ for all $i \in \{1, \ldots, n-1\}$.

(v) Find a necessary and sufficient condition on $\alpha_1, \ldots, \alpha_n$, and $\beta_1, \ldots, \beta_{n-1}$, so that the Leslie matrix A is irreducible.

(vi) For an irreducible Leslie matrix (as in the previous point (v)), find a sufficient condition on the parameters (α_i, β_i), $i \in \{1, \ldots, n\}$, that ensures that the population will not go extinct.

E4.20 **Swiss railroads.** Consider the fictitious railroad map of Switzerland given in figure below.

   (i) Can a passenger go from any station to any other?

   (ii) Is the graph acyclic? Is it aperiodic? If not, what is its period?
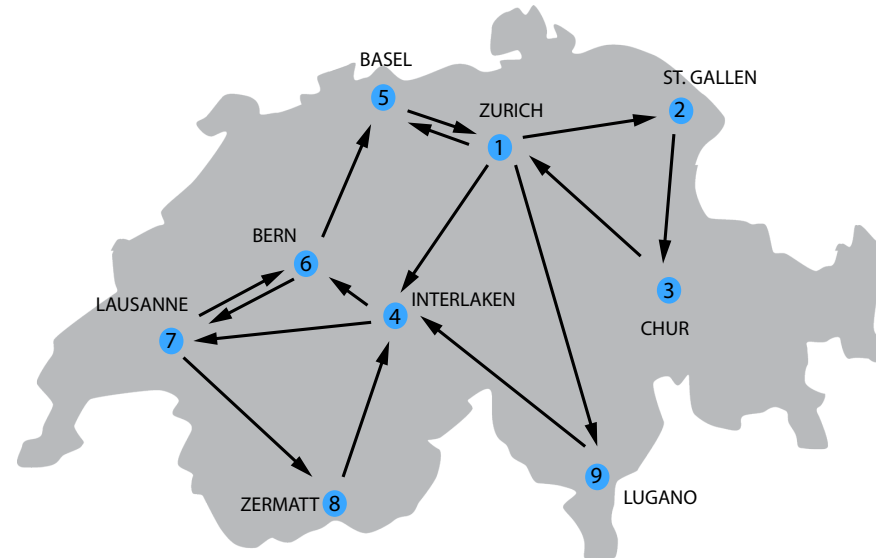
Next, write the unweighted adjacency matrix A of this transportation network and, relying upon A and its powers, answer the following questions:

   (iii) what is the number of links of the shortest walk connecting St. Gallen to Zermatt?

   (iv) is it possible to go from Bern to Chur using 4 links? And 5?

   (v) how many different routes, with strictly less then 9 links and possibly visiting the same station more than once, start from Zürich and end in Lausanne?

E4.21 **Normal, irreducible, row-stochastic matrices are doubly stochastic.** Assume $A \in \mathbb{R}^{n \times n}$ is row-stochastic, irreducible, and normal, i.e., $AA^{\mathsf{T}} = A^{\mathsf{T}}A$. Show that

(i) A is doubly-stochastic,

(ii) any row-stochastic circulant matrix (see E4.3) is doubly stochastic, and

(iii) any orthogonal row-stochastic matrix (see E2.12) is a permutation matrix.

Chapter 5

Discrete-time Averaging Systems

After discussing matrix theory and graph theory, we are ready to go back to the averaging model introduced in Chapter 1 with examples from sociology, wireless sensor networks and robotics. Recall that the discrete-time averaging systems, as given in equation (1.3), is

$$x(k+1) = Ax(k), \tag{5.1}$$

where the matrix $A = [a_{ij}]$ is row-stochastic.

This chapter presents comprehensive convergence results for this model, based on Perron–Frobenius theory and algebraic graph theory. We pay special attention to how the structure of the network determines its function, i.e., the asymptotic behavior of the averaging system. First, we discuss the emergence of consensus for primitive matrices and reducible matrices with a single sink and then we discuss the emerge of asymptotic disagreement for matrices with multiple sinks. Additionally, we provide alternative equivalent proofs of convergence to consensus via certain ergodicity coefficients. We then discuss the equal-neighbor and the Metropolis–Hastings models of row-stochastic matrices. Finally, we present some centrality notions from network science.
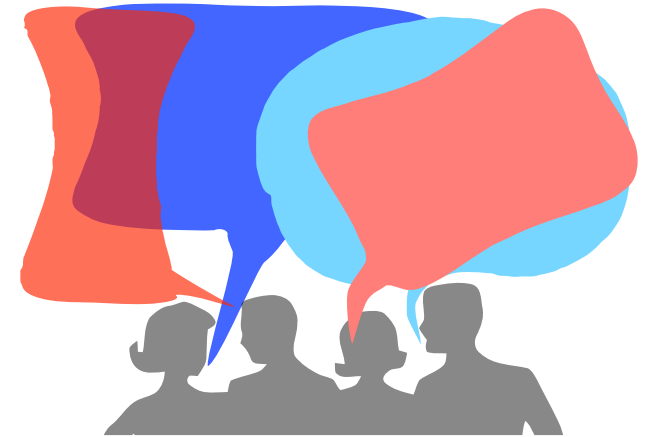

Figure 5.1: Opinion averaging is believed to be a key mechanism in social influence network.

5.1 Averaging systems achieving asymptotic consensus

We start by considering three simple averaging systems.

First example: Let us start the analysis where we left it off at the end of Chapter 2, i.e., with the wireless sensor network example illustrated in Figure 5.2. From the figure, we note that the weighted digraph G_{wsn} is strongly connected and aperiodic and that
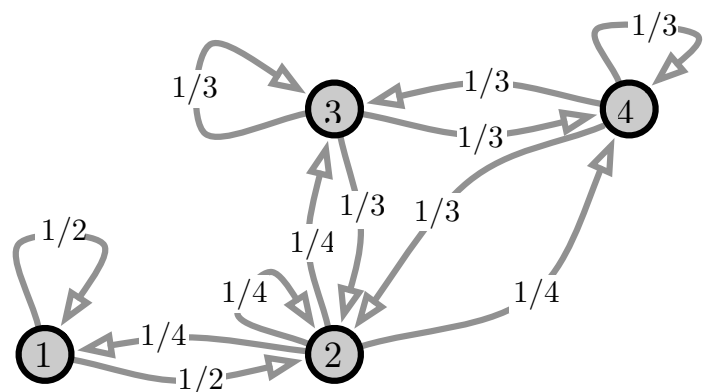

Figure 5.2: Wireless sensor network example introduced in Section 1.2 and studied in Example 2.14. This weighted directed graph G_{wsn} corresponds to the row-stochastic matrix $A_{\text{wsn}} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix}$, as defined in equation (1.2).

the row stochastic matrix A_{wsn} is not column stochastic. We now reason as follows. First, Theorem 4.7 (on the graph theoretical characterization of primitive matrices) states that the strongly connected and aperiodic digraph G_{wsn} has a primitive adjacency matrix A_{wsn}. Second, the Perron–Frobenius Theorem 2.12 states that the eigenvalue 1 is simple and strictly dominant for the primitive row-stochastic matrix A_{wsn}. Therefore, Theorem 2.13 (on the powers of non-negative matrices with a simple and strictly dominant eigenvalue) states that

$$\lim_{k \to \infty} A_{\text{wsn}}^k = \mathbb{1}_4 w^{\mathsf{T}},$$

where $w = [1/6, 1/3, 1/4, 1/4]^{\mathsf{T}}$ is the left dominant eigenvector of A_{wsn}. Next, each solution $x(k) = A_{\text{wsn}}^k x(0)$ to the averaging system $x(k+1) = A_{\text{wsn}} x(k)$ satisfies

$$\lim_{k \to \infty} x(k) = \lim_{k \to \infty} A_{\text{wsn}}^k x(0) = (\mathbb{1}_4 w^{\mathsf{T}}) x(0) = (w^{\mathsf{T}} x(0)) \mathbb{1}_4 = \begin{bmatrix} w^{\mathsf{T}} x(0) \\ \vdots \\ w^{\mathsf{T}} x(0) \end{bmatrix}.$$

In other words, the value at each node of the wireless sensor network converges to a *consensus* value $w^\mathsf{T}x(0) = (1/6)x_1(0) + (1/3)x_2(0) + (1/4)x_3(0) + (1/4)x_4(0)$. Note that the averaging algorithm A_{wsn} does not achieve *average consensus*, since the final value is not equal to the exact average of the initial conditions. Indeed, A_{wsn} is not column stochastic and node 2 has more influence than the other nodes.

Second example: As second example, Figure 5.3 illustrates the robotic pursuit digraph, denoted by G_{pursuit}, introduced in Section 1.6.
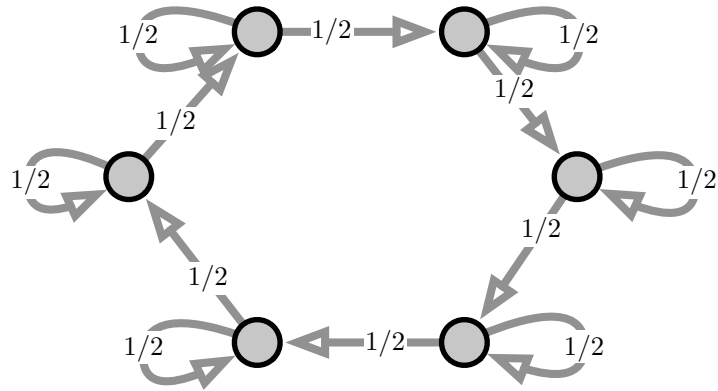

Figure 5.3: Robotic cyclic pursuit example introduced in Section 1.6, see equation (1.23) with $\kappa = 1/2$. This weighted digraph G_{pursuit} corresponds to the row-stochastic matrix $A_{\text{pursuit}} =$

$$
\begin{bmatrix}
1/2 & 1/2 & \dots & 0 & 0 \\
0 & 1/2 & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \ddots & \ddots & 1/2 & 1/2 \\
1/2 & 0 & \dots & 0 & 1/2
\end{bmatrix}.
$$

We note that G_{pursuit} is strongly connected, aperiodic, and weight-balanced. Therefore, the row-stochastic matrix A_{pursuit} is primitive and column stochastic and, in turn, the averaging system achieves average consensus:

$$
\lim_{k \to \infty} x(k) = \lim_{k \to \infty} A_{\text{pursuit}}^k x(0) = \operatorname{average}\big(x(0)\big) \mathbb{1}_n.
$$

Third example: As third example, we consider a reducible row-stochastic matrix whose associated digraph is not strongly connected. Such a matrix with its associated digraph and spectrum is illustrated in Figure 5.4. We note that this digraph has an aperiodic

(a) Pixel picture of a 10×10 row-stochastic matrix; we assume each row contains equal entries summing to 1.

(b) The corresponding digraph has an aperiodic subgraph of globally reachable nodes (all nodes in dark gray).

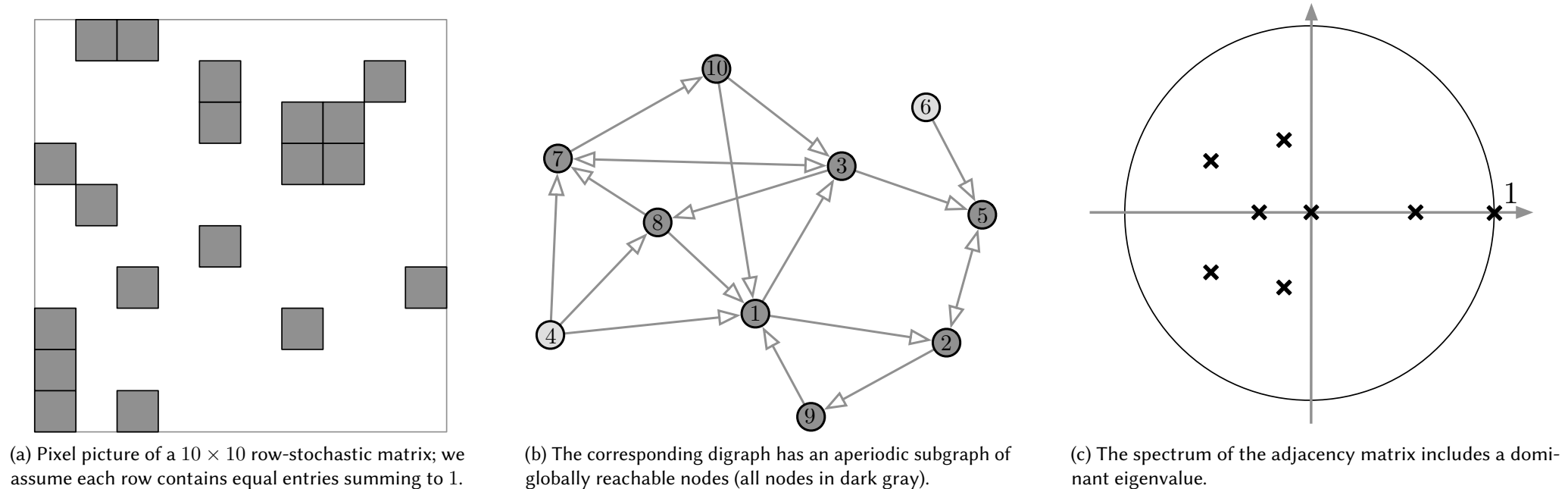(c) The spectrum of the adjacency matrix includes a dominant eigenvalue.

Figure 5.4: An example indecomposable row-stochastic matrix, its associated digraph (consistent with Theorem 5.1(A2)), and its spectrum (consistent with Theorem 5.1(A1))

subgraph of globally reachable nodes and that the eigenvalue 1 is still simple and strictly dominant. We call such row-stochastic matrices *indecomposable*; we refer to Exercise E4.9 for additional properties of such matrices. We will show that the associated averaging algorithm is akin to an averaging system of the form

$$x_1(k+1) = x_1(k),$$
$$x_2(k+1) = (1/2)x_1(k) + (1/2)x_2(k),$$

and still achieves consensus.

We are finally ready to state and prove the main result of this section.

Theorem 5.1 (Consensus for row-stochastic matrices with a globally-reachable aperiodic strongly-connected component).
Let A be a row-stochastic matrix and let G be its associated digraph. The following statements are equivalent:

(A1) *the eigenvalue 1 is simple and all other eigenvalues μ satisfy $|\mu| < 1$;*

(A2) *A is semi-convergent and $\lim_{k \to \infty} A^k = \mathbb{1}_n w^\mathsf{T}$, where $w \in \mathbb{R}^n$ satisfies $w \geq 0$, $\mathbb{1}_n^\mathsf{T} w = 1$, and $w^\mathsf{T} A = w^\mathsf{T}$; and*

(A3) *G contains a globally reachable node and the subgraph of globally reachable nodes is aperiodic.*

If any, and therefore all, of the previous conditions are satisfied, then the matrix A is said to be indecomposable and

(i) *$w \geq 0$ is the left dominant eigenvector of A and $w_i > 0$ if and only if node i is globally reachable;*

(ii) *the solution to the averaging model (5.1) $x(k+1) = Ax(k)$ satisfies*

$$\lim_{k \to \infty} x(k) = \left(w^\mathsf{T} x(0) \right) \mathbb{1}_n;$$

(iii) *if additionally A is doubly-stochastic, then $w = \frac{1}{n} \mathbb{1}_n$ (since $A^\mathsf{T} \mathbb{1}_n = \mathbb{1}_n$ and $\frac{1}{n} \mathbb{1}_n^\mathsf{T} \mathbb{1}_n = 1$) so that*

$$\lim_{k \to \infty} x(k) = \frac{\mathbb{1}_n^\mathsf{T} x(0)}{n} \mathbb{1}_n = \text{average}\left(x(0) \right) \mathbb{1}_n.$$

Note: statement (ii) implies that the limiting value is a weighted average of the initial conditions with relative weights given by the convex combination coefficients $w_1, \ldots, w_n$. The eigenvector is positive $w > 0$ if and only if the digraph associated to A is strongly connected. In digraphs that are not strongly connected, the initial values $x_i(0)$ of all nodes i which are not globally reachable have no effect on the final convergence value. In a social influence network, the coefficient w_i is regarded as the "social influence" of agent i. We illustrate the concept of social influence by introducing the Krackhardt's advice network in Figure 5.5.
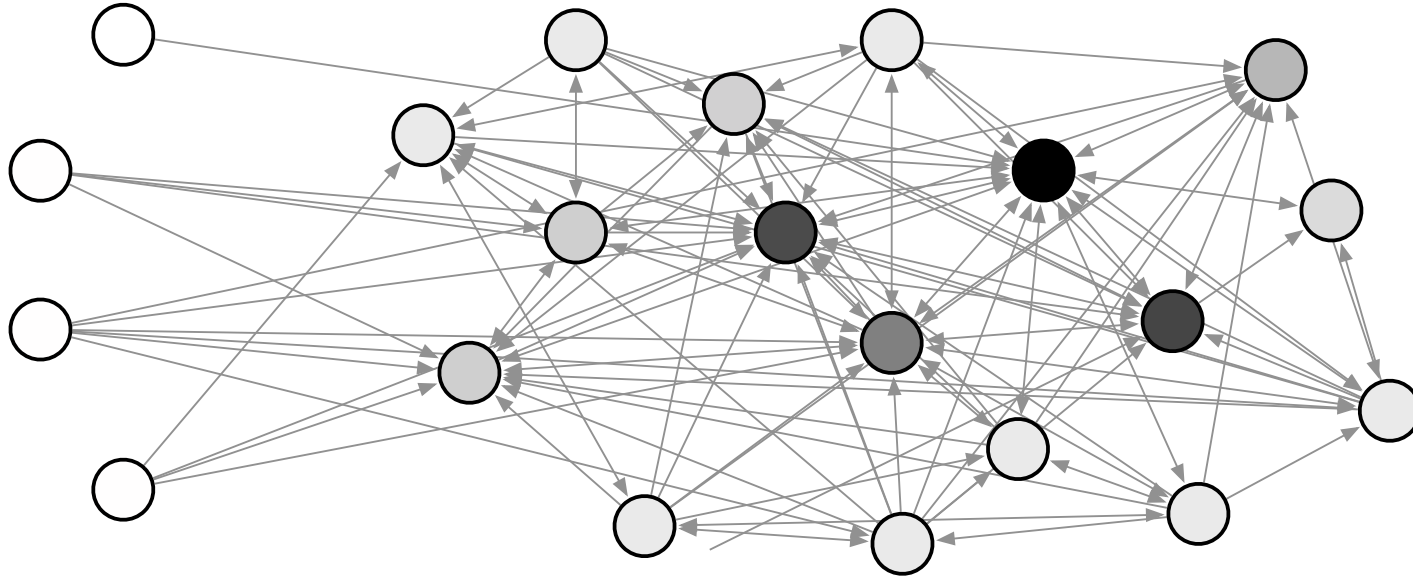

Figure 5.5: The empirically-observed *Krackhardt's advice network* (Krackhardt, 1987) describing the social influence among 21 managers in a manufacturing firm: each directed edge (i, j) means that manager i seeks advice from manager j. This weighted digraph contains an aperiodic subgraph of globally reachable nodes. Therefore, the corresponding row-stochastic matrix is reducible, but with simple and strictly dominant eigenvalue equal to 1. Moreover, the corresponding averaging system achieve consensus with social influence value illustrated by the the gray level of each node in figure.

Note: to clarify statement (A3) it is useful to review some properties of globally reachable nodes. We first recall a useful property from Lemma 3.2: G has a globally reachable node if and only if its condensation digraph has a globally reachable node (i.e., the condensation of G has a single sink). Second, it is easy to see that the set of globally reachable nodes induces a strongly connected component of G.

Note: The theorem is consistent with the following result stated in Exercise E3.6: a weight-balanced digraph with a globally reachable node is strongly connected.

Proof of Theorem 5.1. The statement (A1) $\implies$ (A2) is precisely Theorem 2.13 with $\lambda = 1$ (whose proof is given in Section 2.3.4).

Next, we prove that (A2) $\implies$ (A3). The assumption $\mathbb{1}_n^\top w = 1$ implies that at least one element, say the jth element, of w is positive. Because $\lim_{k \to \infty} A^k = \mathbb{1}_n w^\top$, we know that the jth column of $\lim_{k \to \infty} A^k$ has all-positive elements. Thus, for sufficiently large K, the jth column of A^K has all-positive elements, so there is a walk of length K from every node to the jth node. Thus, the jth node is globally reachable. Similarly, since the (j, j) entry of A^K converges to a positive number as $K \to \infty$, we know that there are walks of arbitrary length from j to j. By Exercise E3.8, the subgraph of globally reachable nodes is aperiodic. Hence, (A2) $\implies$ (A3).

Finally, we prove the implications (A3) $\Longrightarrow$ (A1) and (A2). By assumption the condensation digraph of A contains a sink that is globally reachable, hence it is unique. Assuming $0 < n_1 < n$ nodes are globally reachable, a permutation of rows and columns (see Exercise E3.2), brings the matrix A into the lower-triangular form

$$A = \begin{bmatrix} A_{11} & \mathbb{0}_{n_1 \times n_2} \\ A_{21} & A_{22} \end{bmatrix}, \tag{5.2}$$

where $A_{11} \in \mathbb{R}^{n_1 \times n_1}$, $A_{22} \in \mathbb{R}^{n_2 \times n_2}$, with $n_1 + n_2 = n$. The state vector x is correspondingly partitioned into $x_1 \in \mathbb{R}^{n_1}$ and $x_2 \in \mathbb{R}^{n_2}$ so that

$$x_1(k + 1) = A_{11} x_1(k), \tag{5.3}$$
$$x_2(k + 1) = A_{21} x_1(k) + A_{22} x_2(k). \tag{5.4}$$

In other words, x_1 and A_{11} are the variables and the matrix corresponding to the sink. Because the sink, as a subgraph of G, is strongly connected and aperiodic, A_{11} is primitive and row-stochastic and, by Theorem 2.13 on the powers of non-negative matrices, we compute

$$\lim_{k \to \infty} A_{11}^k = \mathbb{1}_{n_1} w_1^\mathsf{T},$$

where $w_1 > 0$ is the left eigenvector with eigenvalue 1 for A_{11} normalized so that $\mathbb{1}_{n_1}^\mathsf{T} w_1 = 1$.

We next analyze the matrix A_{22} as follows. Recall from Corollary 4.13 that an irreducible row-substochastic matrix has spectral radius less than 1. Now, because A_{21} cannot be zero (otherwise the sink would not be globally reachable), the matrix A_{22} is row-substochastic. Moreover, (after appropriately permuting rows and columns of A_{22}) it can be observed that A_{22} is a lower-triangular matrix such that each diagonal block is row-substochastic and irreducible (corresponding to each node in the condensation digraph). Therefore, we know $\rho(A_{22}) < 1$ and, in turn, $I_{n_2} - A_{22}$ is invertible. Because A_{11} is primitive and $\rho(A_{22}) < 1$, A is semi-convergent and $\lim_{k \to \infty} x_2(k)$ exists. This establishes that (A3) $\Longrightarrow$ (A1). Taking the limit as $k \to \infty$ in equation (5.4), some straightforward algebra shows that

$$\lim_{k \to \infty} x_2(k) = (I_{n_2} - A_{22})^{-1} A_{21} \left(\lim_{k \to \infty} x_1(k) \right) = (I_{n_2} - A_{22})^{-1} A_{21} \left(\mathbb{1}_{n_1} w_1^\mathsf{T} \right) x_1(0).$$

Since A is row-stochastic, we know $A_{21}\mathbb{1}_{n_1} + A_{22}\mathbb{1}_{n_2} = \mathbb{1}_{n_2}$ and hence $(I_{n_2} - A_{22})^{-1}A_{21}\mathbb{1}_{n_1} = \mathbb{1}_{n_2}$. Collecting these results, we write

$$\lim_{k \to \infty} \begin{bmatrix} A_{11} & \mathbb{0}_{n_1 \times n_2} \\ A_{21} & A_{22} \end{bmatrix}^k = \begin{bmatrix} \mathbb{1}_{n_1}w_1^\mathsf{T} & \mathbb{0}_{n_1 \times n_2} \\ \mathbb{1}_{n_2}w_1^\mathsf{T} & \mathbb{0}_{n_2 \times n_2} \end{bmatrix} = \mathbb{1}_n \begin{bmatrix} w_1 \\ \mathbb{0}_{n_2} \end{bmatrix}^\mathsf{T}.$$

This establishes that (A3) $\implies$ (A2) and (A1) $\implies$ (i). The implications (A2) $\implies$ (ii) and (A2) $\implies$ (iii) are straightforward. ■

5.2   Averaging systems reaching asymptotic disagreement

In this section we consider the general case of digraphs that do not contain globally reachable nodes, that is, digraphs whose condensation digraph has multiple sinks. Such an example digraph is the famous Sampson monastery network (Sampson, 1969); see Figure 5.6.
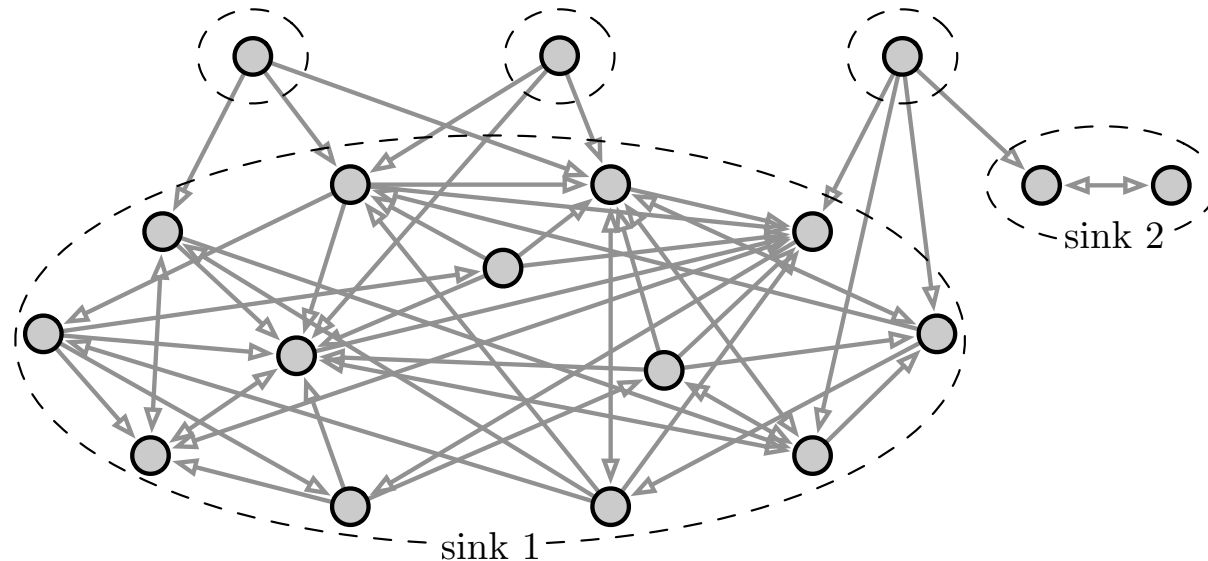

Figure 5.6: This image illustrates the Sampson monastery dataset (Sampson, 1969). This dataset describes the social relations among a set of 18 monk-novitiates in an isolated contemporary American monastery. This digraph contains two sinks in its condensation.

The main result of this section is a generalization of the consensus Theorem 5.1 in the previous section.

Theorem 5.2 (Convergence for row-stochastic matrices with multiple aperiodic sinks). *Let A be a row-stochastic matrix, G be its associated digraph, and $n_\mathrm{s} \geq 2$ be the number of sinks in the condensation digraph $C(G)$. The following statements are equivalent:*

(A1) *the eigenvalue 1 is semi-simple with multiplicity n_s and all other eigenvalues μ satisfy $|\mu| < 1$,*

(A2) *A is semi-convergent, and*

(A3) *each sink of $C(G)$, regarded as a subgraph of G, is aperiodic.*

If any, and therefore all, of the previous conditions are satisfied, then

  (i) *the left eigenvectors $w^p \in \mathbb{R}^n$, $p \in \{1, \ldots, n_\mathrm{s}\}$, of A corresponding to the eigenvalue 1 can be selected to satisfy: $w^p \geq 0$, $\mathbb{1}_n^\top w^p = 1$, and $w_i^p > 0$ if and only if node i belongs to sink p,*

 (ii) *the solution to the averaging model $x(k+1) = Ax(k)$ with initial condition $x(0)$ satisfies*

$$\lim_{k\to\infty} x_i(k) = \begin{cases} (w^p)^\top x(0), & \text{if node } i \text{ belongs to sink } p, \\ \displaystyle\sum_{p=1}^{n_\mathrm{s}} z_{i,p}\big((w^p)^\top x(0)\big), & \text{otherwise,} \end{cases}$$

    where $z_{i,p}$, $p \in \{1, \ldots, n_\mathrm{s}\}$, are convex combination coefficients and $z_{i,p} > 0$ if and only if there exists a directed walk from node i to the sink p.

Proof sketch. Rather than treating the general case, we work out a significant example with the key ideas of the general proof. We invite the reader to provide a proof in Exercise E5.25 and refer to (DeMarzo et al., 2003, Theorem 10) for additional details. Assume the condensation digraph of A is composed of three nodes, two of which are sinks, as in this figure.
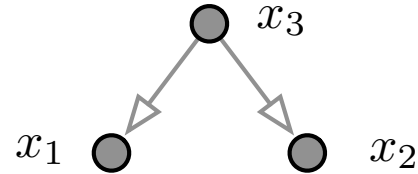
Therefore, after a permutation of rows and columns (see Exercise E3.2), A can be written as

$$A = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

and the state vector x is correspondingly partitioned into the vectors x_1, x_2 and x_3. The state equations are:

$$x_1(k+1) = A_{11}x_1(k), \tag{5.5}$$
$$x_2(k+1) = A_{22}x_2(k), \tag{5.6}$$
$$x_3(k+1) = A_{31}x_1(k) + A_{32}x_2(k) + A_{33}x_3(k). \tag{5.7}$$

By the properties of the condensation digraph and the assumption of aperiodicity of the sinks, the digraphs associated to the row-stochastic matrices A_{11} and A_{22} are strongly connected and aperiodic. Therefore, we immediately conclude that

$$\lim_{k\to\infty} x_1(k) = \left(w_1^\mathsf{T} x_1(0)\right)\mathbb{1}_{n_1} \quad \text{and} \quad \lim_{k\to\infty} x_2(k) = \left(w_2^\mathsf{T} x_2(0)\right)\mathbb{1}_{n_2},$$

where w_1 (resp. w_2) is the left eigenvector of the eigenvalue 1 for matrix A_{11} (resp. A_{22}) with the usual normalization $\mathbb{1}_{n_1}^\mathsf{T} w_1 = \mathbb{1}_{n_2}^\mathsf{T} w_2 = 1$.

Regarding the matrix A_{33}, the same discussion as in the proof of Theorem 5.1 ensures that $\rho(A_{33}) < 1$ and that, in turn, $I_{n_3} - A_{33}$ is nonsingular. We have now established that the eigenvalue 1 of A is semisimple with multiplicity $2 = n_s$ and that all other eigenvalues are strictly inside the unit disk. By taking the limit as $k \to \infty$ in equation (5.7), straightforward calculations show that

$$
\lim_{k \to \infty} x_3(k) = (I_{n_3} - A_{33})^{-1} \big(A_{31} \lim_{k \to \infty} x_1(k) + A_{32} \lim_{k \to \infty} x_2(k) \big)
$$
$$
= (w_1^\mathsf{T} x_1(0)) \big((I_{n_3} - A_{33})^{-1} A_{31} \mathbb{1}_{n_1} \big) + (w_2^\mathsf{T} x_2(0)) \big((I_{n_3} - A_{33})^{-1} A_{32} \mathbb{1}_{n_2} \big).
$$

Moreover, because A is row-stochastic, we know

$$
A_{31} \mathbb{1}_{n_1} + A_{32} \mathbb{1}_{n_2} + A_{33} \mathbb{1}_{n_3} = \mathbb{1}_{n_3},
$$

and, using again the fact that $I_{n_3} - A_{33}$ is nonsingular,

$$
\mathbb{1}_{n_3} = (I_{n_3} - A_{33})^{-1} A_{31} \mathbb{1}_{n_1} + (I_{n_3} - A_{33})^{-1} A_{32} \mathbb{1}_{n_2}.
$$

This concludes our incomplete proof of Theorem 5.2 for the simplified case $C(G)$ having three nodes and two sinks. ∎

Note that: convergence does not occur to consensus (not all components of the state are equal) and the final value of all nodes is independent of the initial values at nodes which are not in the sinks of the condensation digraph.

We conclude this section with a figure providing a summary of the asymptotic behavior of discrete-time averaging systems and its relationships with properties of matrices and graphs; see Figure 5.7.
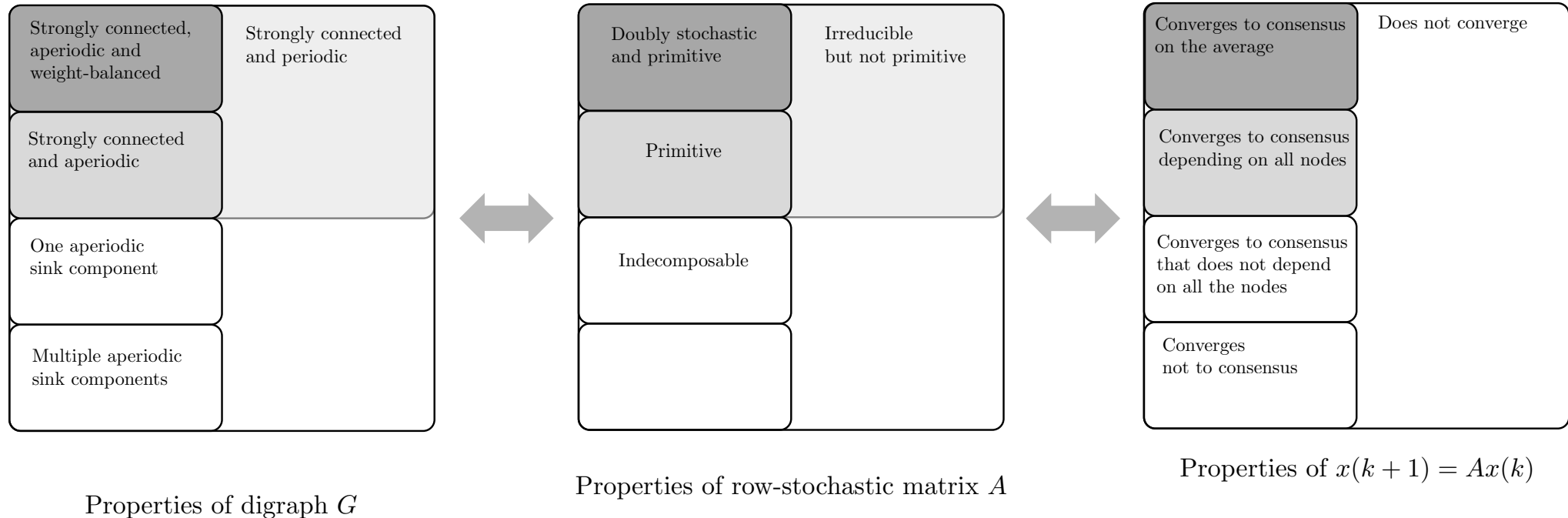

Properties of digraph G

Properties of row-stochastic matrix A

Properties of $x(k + 1) = Ax(k)$

Figure 5.7: Equivalent properties for a digraph G, a row-stochastic matrix A (being the adjacency matrix of G), and a discrete-time averaging dynamical system $x(k + 1) = Ax(k)$.

5.3 Consensus via disagreement and Lyapunov functions

In this section we present alternative analysis methods for averaging system. We postpone to Chapter 15 a treatment of Lyapunov stability theory.

5.3.1 Disagreement and deflated matrices

Consider a row-stochastic $A \in \mathbb{R}^{n \times n}$ with left dominant eigenvector $w \in \mathbb{R}_{\geq 0}^n$ normalized so that $\mathbb{1}_n^\top w = 1$. For the averaging system $x(k+1) = Ax(k)$, define

  (i) the *disagreement vector* $\delta(k) = x(k) - (w^\top x(0))\mathbb{1}_n \in \mathbb{R}^n$, and

  (ii) the *deflated matrix* $A - \mathbb{1}_n w^\top$.

In other words, subtract from the sequence $x(k)$ and from the matrix A their respective final values (here by final value of A is meant $\lim_k A^k$).

   Note: $w^\top x(k)$ is a weighted average of the entries of $x(k)$ and it is constant over time (e.g., see Exercise E1.4); therefore, $\delta(k) = x(k) - (w^\top x(k))\mathbb{1}_n$. Also note that $\mathbb{1}_n w^\top$ is a rank-1 projection matrix (for more properties of this matrix see Exercise E5.9). Finally, to motivate the nomenclature "deflated matrix," note that "deflating" the matrix A removes the influence of the dominant eigenvalue 1. Accordingly, we define the *essential spectral radius* of A by

$$\rho_{\text{ess}}(A) = \begin{cases} 0, & \text{if } \operatorname{spec}(A) = \{1, \ldots, 1\}, \\ \max\{|\lambda| \mid \lambda \in \operatorname{spec}(A) \setminus \{1\}\}, & \text{otherwise.} \end{cases}$$

Lemma 5.3 (Convergence of disagreement vector). *Given a row-stochastic matrix A with left dominant eigenvector w normalized so that $\mathbb{1}_n^{\mathsf{T}} w = 1$,*

(i) *the dynamics of the disagreement vector is*

$$\delta(k+1) = (A - \mathbb{1}_n w^{\mathsf{T}})\delta(k); \tag{5.8}$$

(ii) *if A is primitive, then the deflated matrix $A - \mathbb{1}_n w^{\mathsf{T}}$ has the same eigenvalues and eigenvectors of A, except the eigenvalue 1 which is replaced by 0 (with same right and left eigenvectors);*

(iii) *if A is primitive, then*

$$\rho(A - \mathbb{1}_n w^{\mathsf{T}}) = \rho_{\text{ess}}(A) < 1, \tag{5.9}$$

and, in turn, $\lim_{k \to \infty} \delta(k) = \mathbb{0}_n$ and $\lim_{k \to \infty} x(k) = (w^{\mathsf{T}} x(0))\mathbb{1}_n$.

Note: Lemma 5.3(iii) and the bounds established in Exercise E5.10 together imply that, for all $\varepsilon > 0$, there exists c_ε such that

$$\|\delta(k)\|_2 \le c_\varepsilon (\rho_{\text{ess}}(A) + \varepsilon)^k \|\delta(0)\|_2.$$

But the decrease is not monotonic, in the sense that it is not true in general that $\|\delta(k)\|_2 \le \|\delta(0)\|_2$.

Proof of Lemma 5.3. Regarding statement (i), note $\delta(k) = (I_n - \mathbb{1}_n w^{\mathsf{T}})x(k)$. Therefore

$$\delta(k+1) = (I_n - \mathbb{1}_n w^{\mathsf{T}})x(k+1) = (I_n - \mathbb{1}_n w^{\mathsf{T}})Ax(k) \tag{5.10}$$

$$= (A - \mathbb{1}_n w^{\mathsf{T}})x(k) = (A - \mathbb{1}_n w^{\mathsf{T}})(I_n - \mathbb{1}_n w^{\mathsf{T}})x(k), \tag{5.11}$$

where we used $\mathbb{1}_n w^{\mathsf{T}} A = \mathbb{1}_n w^{\mathsf{T}}$ and $(A - \mathbb{1}_n w^{\mathsf{T}})(I_n - \mathbb{1}_n w^{\mathsf{T}}) = A - \mathbb{1}_n w^{\mathsf{T}} - \mathbb{1}_n w^{\mathsf{T}} + \mathbb{1}_n w^{\mathsf{T}} \mathbb{1}_n w^{\mathsf{T}} = A - \mathbb{1}_n w^{\mathsf{T}}$.

Regarding statement (ii), the right eigenpair $(1, \mathbb{1}_n)$ of A corresponds to the eigenpair $(0, \mathbb{1}_n)$ of $A - \mathbb{1}_n w^{\mathsf{T}}$, since $(A - \mathbb{1}_n w^{\mathsf{T}})\mathbb{1}_n = \mathbb{0}_n$. Since A is primitive, all other eigenvalues have magnitude strictly less than 1. Next, pick an arbitrary right eigenpair (μ, z) of A with $|\mu| < 1$. Because w is a left eigenvector of 1 and z is a right eigenvector of eigenvalue $\mu \neq 1$, we know that $w^{\mathsf{T}} z = 0$, e.g., see Exercise E2.4(i). Therefore, we compute $(A - \mathbb{1}_n w^{\mathsf{T}})z = Az = \mu z$. This proves that (μ, z) is a right eigenpair of $A - \mathbb{1}_n w^{\mathsf{T}}$.

Statement (iii) is an immediate consequence of statement (ii). ∎

5.3.2 Quadratic disagreement

Define the *quadratic disagreement function* $V_{\mathsf{qd}} \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ by

$$V_{\mathsf{qd}}(x) = \frac{1}{n} \sum_{i,j=1}^{n} (x_i - x_j)^2. \tag{5.12}$$

Clearly, $V_{\mathsf{qd}}(x) \geq 0$ and $V_{\mathsf{qd}}(x) = 0$ if and only if x is a consensus vector.

Note: It is simple to check that $V_{\mathsf{qd}}(x) = x^{\mathsf{T}} \Pi_n x$, where $\Pi_n = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}} \in \mathbb{R}^{n \times n}$ is the orthogonal projection onto $\mathrm{span}\{\mathbb{1}\}^{\perp}$, is symmetric and positive-semidefinite. Also, it is simple to check that, when A is doubly-stochastic with $w = \frac{1}{n} \mathbb{1}_n$, the disagreement vector is $\delta(x) = x - \mathrm{average}(x) \mathbb{1}_n \in \mathbb{R}^n$ and $V_{\mathsf{qd}}(x) = \|\delta(x)\|_2^2$.

Given a symmetric row-stochastic matrix $A = A^{\mathsf{T}}$, define the *2-coefficient of ergodicity* of A by one of the following equal expressions:

$$\tau_2(A) = \left\| A - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}} \right\|_2 = \max_{\|y\|_2 = 1, y \perp \mathbb{1}_n} \|Ay\|_2. \tag{5.13}$$

where we recall the definition of induced matrix norm $\|A\|_2 = \max\{\|Ax\|_2 \mid \|x\|_2 = 1\}$ (e.g, for more details see Exercise E4.17).

Lemma 5.4 (Convergence of quadratic disagreement). *Given a symmetric row-stochastic matrix $A = A^\mathsf{T}$ with associated graph G,*

(i) for all $x \in \mathbb{R}^n$,

$$V_{\mathsf{qd}}(Ax) \leq \tau_2(A)^2 V_{\mathsf{qd}}(x);$$

(ii) if G is connected, then $\tau_2(A) = \rho_{\mathsf{ess}}(A) < 1$;

(iii) if G is connected, then any solution to $x(k+1) = Ax(k)$ satisfies

$$V_{\mathsf{qd}}(x(k)) \leq \underbrace{\tau_2(A)}_{<1}{}^{2k} V_{\mathsf{qd}}(x(0)), \quad \textit{for all } k \in \mathbb{N}.$$

Note: under the conditions in the theorem, the quadratic disagreement diminishes monotonically along each solution and we say that the function V_{qd} is a *quadratic Lyapunov function* for $x(k+1) = Ax(k)$.

Proof of Lemma 5.4. Regarding statement (i), for any $x \in \mathbb{R}^n$, we note that $\mathbb{1}_n^\mathsf{T} A x = \mathbb{1}_n^\mathsf{T} x$ so that

$$
\begin{aligned}
V_{\mathsf{qd}}(Ax) &= \|Ax - \operatorname{average}(Ax)\mathbb{1}_n\|_2^2 = \|A(x - \operatorname{average}(x)\mathbb{1}_n)\|_2^2 \\
&= \left\|A \tfrac{1}{\|x - \operatorname{average}(x)\mathbb{1}_n\|_2}(x - \operatorname{average}(x)\mathbb{1}_n)\right\|_2^2 V_{\mathsf{qd}}(x) \\
&\leq \max_{\|y\|_2 = 1, y \perp \mathbb{1}_n} \|Ay\|_2^2 \, V_{\mathsf{qd}}(x).
\end{aligned}
$$

Regarding statement (ii), recall the modal decomposition of symmetric matrices, e.g., see Remark 2.3. Since $A = A^\mathsf{T}$, we know A admits eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ with orthonormal eigenvectors $v_1, \ldots, v_n$. Since A is row stochastic and irreducible, we know $\lambda_1 = 1$, $\lambda_1 > \max_{j \in \{2, \ldots, n\}} |\lambda_j|$, $v_1 = \mathbb{1}_n/\sqrt{n}$. The modal decomposition of A is $A = \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T} + \sum_{i=2}^n \lambda_i v_i v_i^\mathsf{T}$. For any $y \in \operatorname{span}(\mathbb{1}_n)^\perp$,

$$
\|Ay\|_2^2 = \sum_{i=2}^n \lambda_i^2 (v_i^\mathsf{T} y)^2 \leq \left(\max_{j \in \{2, \ldots, n\}} \lambda_j^2 \right) \sum_{i=2}^n (v_i^\mathsf{T} y)^2 = \rho_{\mathsf{ess}}(A)^2 \|y\|_2^2,
$$

where we noted $\rho_{\mathsf{ess}}(A) = \max_{j \in \{2, \ldots, n\}} |\lambda_j| < 1$. This implies $\tau_2(A) \leq \rho_{\mathsf{ess}}(A)$. The converse follows from taking $y = v_j$ where $i = \operatorname{argmax}_{j \in \{2, \ldots, n\}} |\lambda_j|$. Statement (iii) follows from the previous statements. ∎

5.3.3 Max-min disagreement

Define the *max-min disagreement function* $V_{\text{max-min}} \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ by

$$V_{\text{max-min}}(x) = \max_{i \in \{1,\ldots,n\}} x_i - \min_{i \in \{1,\ldots,n\}} x_i = \max_{i,j \in \{1,\ldots,n\}} (x_i - x_j). \tag{5.14}$$

Clearly, $V_{\text{max-min}}(x) \geq 0$ and $V_{\text{max-min}}(x) = 0$ if and only if x is a consensus vector.

(Recall that Exercise E1.1 asks the reader to verify that any row-stochastic $A \in \mathbb{R}^{n \times n}$ and $x \in \mathbb{R}^n$ satisfy $V_{\text{max-min}}(Ax) \leq V_{\text{max-min}}(x)$. We will establish a stronger property below.)

Given a row-stochastic matrix A, define the 1-*coefficient of ergodicity* of A by one of the following equal expressions:

$$\begin{aligned}
\tau_1(A) &= \max_{\|y\|_1 = 1, y \perp \mathbb{1}_n} \|A^{\mathsf{T}} y\|_1 \\
&= \frac{1}{2} \max_{i,j \in \{1,\ldots,n\}} \sum_{h=1}^{n} |a_{ih} - a_{jh}| = 1 - \min_{i,j \in \{1,\ldots,n\}} \sum_{h=1}^{n} \min\{a_{ih}, a_{jh}\}.
\end{aligned} \tag{5.15}$$

(Exercise E5.22 asks the reader to prove that the three expressions for $\tau_1(A)$ are indeed equivalent, as claimed.)

Lemma 5.5 (Convergence of max-min disagreement). *Given a row-stochastic primitive matrix A with associated digraph G,*

(i) for all $x \in \mathbb{R}^n$, the max-min function satisfies

$$V_{\text{max-min}}(Ax) \leq \tau_1(A)V_{\text{max-min}}(x);$$

(ii) $\tau_1(A) < 1$ if and only if A is scrambling*, i.e., any two nodes have a common out-neighbor in G;*

(iii) if G contains a node that is globally reachable in h steps, then A^h is scrambling and any solution to $x(k+1) = Ax(k)$ satisfies

$$V_{\text{max-min}}(x(k)) \leq \underbrace{\tau_1(A^h)}_{<1}{}^{\lfloor k/h \rfloor} V_{\text{max-min}}(x(0)), \quad \text{for all } k \in \mathbb{N}.$$

Note: under the conditions in the theorem, the max-min disagreement diminishes monotonically along each solution and we say that the function $V_{\text{max-min}}$ is a *Lyapunov function* for $x(k+1) = Ax(k)$.

Note: G contains a node that is globally reachable in h steps, for some h (i.e., there exists h such that from each node there exists a directed path of length h to the specific node) if and only if G contains a globally reachable node and the strongly connected component of globally reachable nodes is aperiodic. This statement is a slight generalization of Theorem 4.7.

Note: if A is scrambling and each of its non-zero entries is lower bounded by $a_{\text{min}} > 0$, then statement (ii) can be strenghtened to state that $\tau_1(A) \leq 1 - a_{\text{min}}$.

We postpone to Chapter 15 a treatment of Lyapunov stability theory and to Chapter 11 a treatment of convergence factors and scalability.

Proof of Lemma 5.5. Regarding statement (i), for any two indices i and $j \in \{1, \ldots, n\}$, define $u_{(i,j),h} = a_{ih} - a_{jh} \in \mathbb{R}$, for $h \in \{1, \ldots, n\}$, so that

$$(Ax)_i - (Ax)_j = \sum_{h=1}^{n} u_{(i,j),h} x_h. \tag{5.16}$$

Note $\sum_{h=1}^{n} u_{(i,j),h} = \sum_{h=1}^{n} a_{ih} - \sum_{j=1}^{n} a_{jh} = 1 - 1 = 0$, because $A \mathbb{1}_n = \mathbb{1}_n$. Therefore, some entries of $u_{(i,j),h}$, for $h \in \{1, \ldots, n\}$, are positive and others negative. Define

$$\tau_{(i,j)} = \sum_{h \, : \, u_{(i,j),h} \geq 0} u_{(i,j),h} = - \sum_{h \, : \, u_{(i,j),h} < 0} u_{(i,j),h} \tag{5.17}$$

$$= \frac{1}{2} \sum_{h=1}^{n} |u_{(i,j),h}| = \frac{1}{2} \sum_{h=1}^{n} |a_{ih} - a_{jh}|, \tag{5.18}$$

where the third equality follows from realizing that the absolute sum of positive and negative entries is twice the sum of the positive entries alone. Next, we obtain a useful upper bound:

$$(Ax)_i - (Ax)_j = \sum_{h \, : \, u_{(i,j),h} \geq 0} u_{(i,j),h} x_h - \sum_{h \, : \, u_{(i,j),h} < 0} |u_{(i,j),h}| x_h$$

$$= \tau_{(i,j)} \left(\frac{\sum_{h \, : \, u_{(i,j),h} \geq 0} |u_{(i,j),h}| x_h}{\sum_{h \, : \, u_{(i,j),h} \geq 0} |u_{(i,j),h}|} - \frac{\sum_{h \, : \, u_{(i,j),h} < 0} |u_{(i,j),h}| x_h}{\sum_{h \, : \, u_{(i,j),h} < 0} |u_{(i,j),h}|} \right)$$

$$\leq \tau_{(i,j)} \left(\max_i x_i - \min_i x_i \right) \leq \tau_1(A) \left(\max_i x_i - \min_i x_i \right),$$

where we used the equality $\tau_1(A) = \frac{1}{2} \max_{i,j \in \{1, \ldots, n\}} \sum_{h=1}^{n} |a_{ih} - a_{jh}|$. This completes the proof of statement (i).

Regarding statement (ii), using the equality $\tau_1(A) = 1 - \min_{i,j \in \{1, \ldots, n\}} \sum_{h=1}^{n} \min\{a_{ih}, a_{jh}\}$, clearly $\tau_1(A) < 1$ if and only if $\sum_{h=1}^{n} \min\{a_{ih}, a_{jh}\} > 0$ for all i, j. But this condition holds if and only if for each pair of nodes i, j there exists a node h such that (i, h) and (j, h) are directed edges in G.

Regarding statement (iii), if the jth node in G that is globally reachable in h steps, then the jth column of A^h is strictly positive. Therefore A^h is scrambling, since each node has the node j as common out-neighbor in the digraph associated to A^h. The final bound follows from the previous statements. ∎

5.4 Appendix: Design of graphs weights

In this section we describe two widely-adopted algorithms to design weights for unweighted graphs.

5.4.1   The equal-neighbor model

Let G be a connected undirected graph, binary adjacency matrix A, and degree matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$, where $d_1, \ldots, d_n$ are the node degrees. Define the *equal-neighbor matrix*

$$A_{\text{equal-nghbr}} = D^{-1}A. \tag{5.19}$$

For example, consider the graph in Figure 5.8, for which we have:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \implies A_{\text{equal-nghbr}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \end{bmatrix}. \tag{5.20}$$
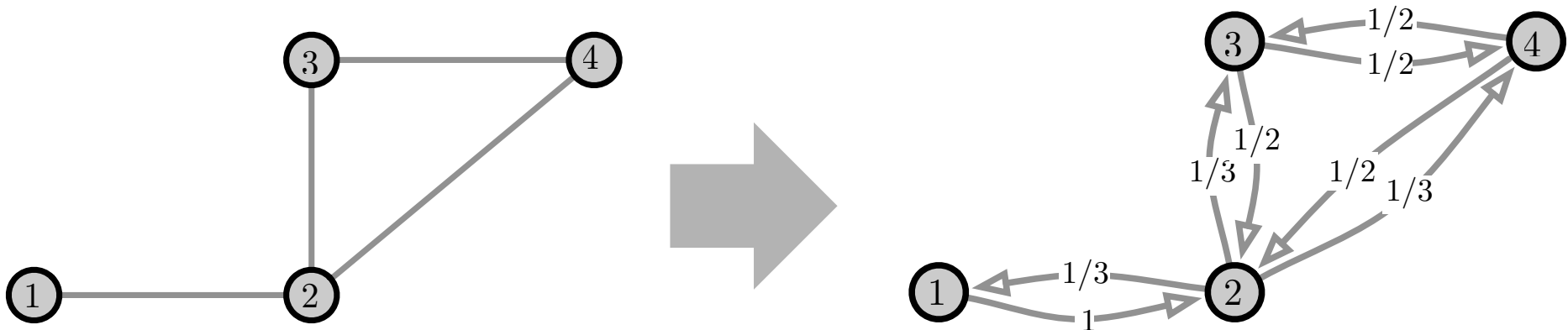
Figure 5.8: The equal-neighbor matrix

It is convenient now to introduce some simple generalizing notions. We say that an undirected graph G (possibly with self-loops) is aperiodic if G, regarded as a digraph, is aperiodic. To regard G as a digraph, we substitute each undirected edge of G with two directed edges (self-loops remain unchanged). Note that an acyclic undirected graph without self-loops is periodic with period 2.

Lemma 5.6 (The equal-neighbor row-stochastic matrix). *Let G be a connected weighted undirected graph (possibly with self-loops) with weighted adjacency matrix A and weighted degrees $d_1, \ldots, d_n$. For the equal-neighbor matrix $A_{\text{equal-nghbr}}$ defined as in (5.19),*

(i) *$A_{\text{equal-nghbr}}$ is well-defined, row-stochastic, and irreducible;*

(ii) *the left dominant eigenvector of $A_{\text{equal-nghbr}}$, normalized to have unit sum, is*

$$w_{\text{equal-nghbr}} = \frac{1}{\sum_{i=1}^{n} d_i} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix},$$

so that, assuming that G is aperiodic, the solution to the averaging model (5.1) $x(k+1) = Ax(k)$ satisfies

$$\lim_{k \to \infty} x_i(k) = \frac{1}{\sum_{i=1}^{n} d_i} \sum_{i=1}^{n} d_i x_i(0); \tag{5.21}$$

(iii) *$A_{\text{equal-nghbr}}$ is doubly-stochastic if and only if G is regular (i.e., all nodes have the same degree).*

For example, for the equal-neighbor matrix in equation (5.20) and Figure 5.8, one can easily verify that the dominant eigenvector is $\begin{bmatrix} 1 & 3 & 2 & 2 \end{bmatrix}^{\mathsf{T}}/8$.

Proof of Lemma 5.6. Because G is connected, each node degree is strictly positive, the degree matrix is invertible, and $A_{\text{equal-nghbr}}$ is well-defined. Because G is connected and because the zero/positive pattern of $A_{\text{equal-nghbr}}$ is the same as that of A, we know $A_{\text{equal-nghbr}}$ is irreducible. Next, we note a simple fact: any $v \in \mathbb{R}^n$ with non-zero entries satisfies $\text{diag}(v)^{-1}v = \mathbb{1}_n$. Let $d = A\mathbb{1}_n$ denote the vector of node degrees so that $D = \text{diag}(d)$. Statement (i) follows from

$$A_{\text{equal-nghbr}}\mathbb{1}_n = \text{diag}(d)^{-1}(A\mathbb{1}_n) = \text{diag}(d)^{-1}d = \mathbb{1}_n.$$

Statement (ii) follows from

$$A_{\text{equal-nghbr}}^{\mathsf{T}} w_{\text{equal-nghbr}} = A\,\text{diag}(d)^{-1}\left(\frac{1}{\mathbb{1}_n^{\mathsf{T}}d}d\right) = \frac{1}{\mathbb{1}_n^{\mathsf{T}}d}A\mathbb{1}_n = \frac{1}{\mathbb{1}_n^{\mathsf{T}}d}d = w_{\text{equal-nghbr}},$$

where we used the fact that A is symmetric. The convergence property follows because $A_{\text{equal-nghbr}}$ is irreducible and aperiodic. Statement (iii) is an immediate consequence of (ii). ∎

We conclude this section by reviewing the distributed averaging algorithm introduced in Section 1.2.

Example 5.7 (Averaging in wireless sensor networks). *As in equation* (1.2), *assume each node of a wireless sensor network contains a value x_i and repeatedly executes:*

$$x_i(k + 1) := \mathrm{average}\big(x_i(k), \{x_j(k), \text{ for all neighbor nodes } j\}\big), \tag{5.22}$$

or, more explicitly, $x_i(k + 1) = \frac{1}{1+d_i}\big(x_i(k) + \sum_{j\in\mathcal{N}(i)} x_j(k)\big)$. Algorithm (5.22) *can be written as:*

$$x(k + 1) = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} x(k) =: A_{\mathrm{wsn}}x(k),$$

where the matrix A_{wsn} is defined as in Section 1.2 *and where it is easy to verify that*

$$A_{\mathrm{wsn}} = (D + I_4)^{-1}(A + I_4).$$

Clearly, $A + I_4$ is the adjacency matrix of a graph that is equal to the graph in figure with the addition of a self-loop at each node; this new graph has degree matrix $D + I_4$. Therefore, the matrix A_{wsn} is an equal-neighbor matrix for the graph with added self-loops. We illustrate this observation in Figure 5.9. *From Lemma* 5.6 *we know that the left dominant eigenvector of A_{wsn} is*

$$w_{\mathrm{equal\text{-}neighbor+selfloops}} = \frac{1}{n + \sum_i d_i} \begin{bmatrix} d_1 + 1 \\ \vdots \\ d_n + 1 \end{bmatrix} = \begin{bmatrix} 1/6 \\ 1/3 \\ 1/4 \\ 1/4 \end{bmatrix},$$

because $(d_1, d_2, d_3, d_4) = (1, 3, 2, 2)$ and $n = 4$. This result is consistent with the eigenvector computed numerically in Example 2.5. ●

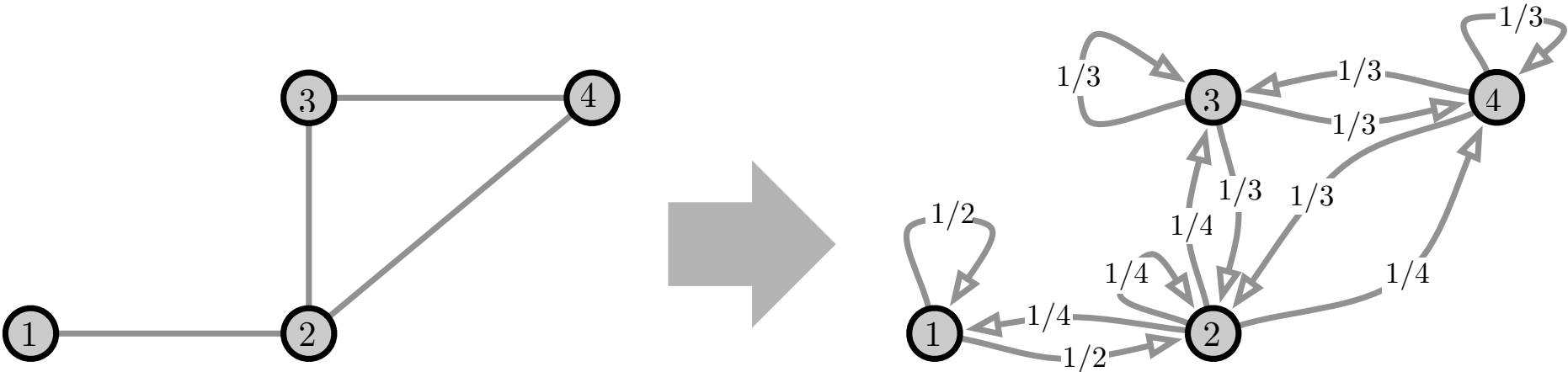Figure 5.9: The equal-neighbor matrix for an undirected graph with added self-loops

5.4.2 The Metropolis–Hastings model

Next, we suggest a second way of assigning weights to a graph for the purpose of designing an averaging algorithm (that achieves average consensus). Given an undirected unweighted graph G with edge set E and degrees $d_1, \ldots, d_n$, define the weighted adjacency matrix $A_{\text{Metr-Hast}}$, called othe *Metropolis–Hastings matrix*, by

$$(A_{\text{Metr-Hast}})_{ij} = \begin{cases} \dfrac{1}{1 + \max\{d_i, d_j\}}, & \text{if } \{i, j\} \in E \text{ and } i \neq j, \\[2ex] 1 - \displaystyle\sum_{\{i,h\}\in E, h\neq i} (A_{\text{Metr-Hast}})_{ih}, & \text{if } i = j, \\[2ex] 0, & \text{otherwise.} \end{cases}$$

In our example,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \implies A_{\text{Metr-Hast}} = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 1/4 & 5/12 & 1/3 \\ 0 & 1/4 & 1/3 & 5/12 \end{bmatrix}.$$

One can verify that the Metropolis–Hastings weights have the following properties:

(i) $(A_{\text{Metr-Hast}})_{ii} > 0$ for all nodes $i \in \{1, \ldots, n\}$, $(A_{\text{Metr-Hast}})_{ij} > 0$ for all pairs $\{i, j\} \in E$, and $(A_{\text{Metr-Hast}})_{ij} = 0$ otherwise;

(ii) $A_{\text{Metr-Hast}}$ is symmetric and doubly-stochastic;

(iii) $A_{\text{Metr-Hast}}$ is primitive if and only if G is connected; and

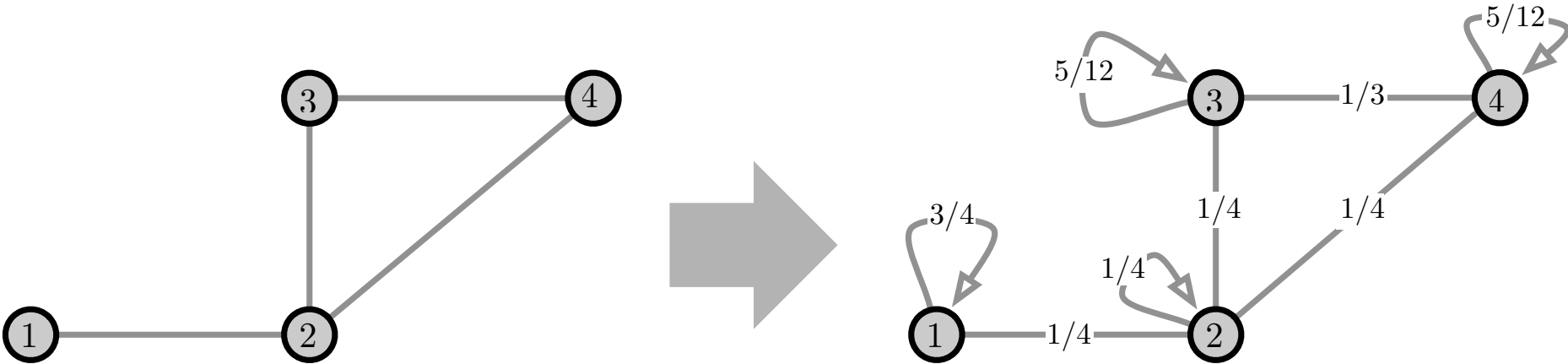(iv) the averaging model (5.1) $x(k+1) = Ax(k)$ achieves average consensus.

Figure 5.10: The Metropolis–Hastings model

5.5   Appendix: Design and computation of centrality measures

In network science it is of interest to determine the relative importance of a node in a network. There are many ways to do so and they are referred to as centrality measures or *centrality scores*. This section presents six centrality notions based on the adjacency matrix. We treat the general case of a weighted digraph G with weighted adjacency matrix A (warning: many articles in the literature deal with undirected graphs only.) The matrix A is non-negative, but not necessarily row stochastic. From the Perron–Frobenius theory, recall the following facts:

(i) if G is strongly connected, then the spectral radius $\rho(A)$ is an eigenvalue of maximum magnitude and its corresponding left eigenvector can be selected to be strictly positive and with unit sum (see Theorem 2.12); and

(ii) if G contains a globally reachable node, then the spectral radius $\rho(A)$ is an eigenvalue of maximum magnitude and its corresponding left eigenvector is non-negative and has positive entries corresponding to each globally reachable node (see Theorem 5.1).

Degree centrality   For an arbitrary weighted digraph G, the *degree centrality* $c_{\text{degree}}(i)$ of node i is its in-degree:

$$c_{\text{degree}}(i) = d_{\text{in}}(i) = \sum_{j=1}^{n} a_{ji}, \tag{5.23}$$

that is, the number of in-neighbors (if G is unweighted) or the sum of the weights of the incoming edges. Degree centrality is relevant, for example, in (typically unweighted) citation networks whereby articles are ranked on the basis of their citation records. (Warning: the notion that a high citation count is an indicator of quality is clearly a fallacy.)

Eigenvector centrality   One problem with degree centrality is that each in-edge has unit count, even if the in-neighbor has negligible importance. To remedy this potential drawback, one could define the importance of a node to be proportional to the weighted sum of the importance of its in-neighbors (see (Bonacich, 1972b) for an early reference). This line of reasoning leads to the following definition.

For a weighted digraph G with globally reachable nodes (or for an undirected graph that is connected), define the *eigenvector centrality* vector, denoted by c_{ev}, to be the left dominant eigenvector of the adjacency matrix A associated with the dominant eigenvalue and normalized to satisfy $\mathbb{1}_n^\mathsf{T} c_{ev} = 1$.

Note that the eigenvector centrality satisfies

$$A^\mathsf{T} c_{ev} = \frac{1}{\alpha} c_{ev} \qquad \Longleftrightarrow \qquad c_{ev}(i) = \alpha \sum_{j=1}^n a_{ji} c_{ev}(j). \tag{5.24}$$

where $\alpha = \frac{1}{\rho(A)}$ is the only possible choice of scalar coefficient in equation (5.24) ensuring that there exists a unique solution and that the solution, denoted c_{ev}, is strictly positive in a strongly connected digraph and non-negative in a digraph with globally reachable nodes. Note that this connectivity property may be restrictive in some cases.
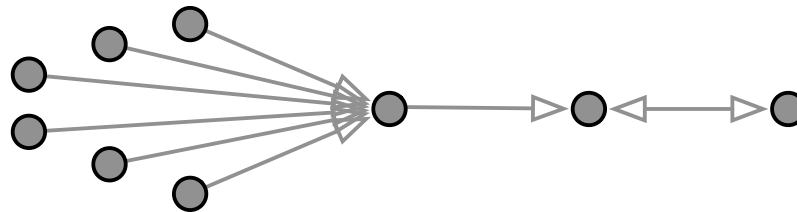
Figure 5.11: Comparing degree centrality versus eigenvector centrality: the node with maximum in-degree has zero eigenvector centrality in this graph

Katz centrality   For a weighted digraph G, pick an attenuation factor $\alpha < 1/\rho(A)$ and define the *Katz centrality* vector (see (Katz, 1953)), denoted by c_{K}, by the following equivalent formulations:

$$c_{\mathsf{K}}(i) = \alpha \sum_{j=1}^{n} a_{ji}(c_{\mathsf{K}}(j) + 1), \tag{5.25}$$

or

$$c_{\mathsf{K}}(i) = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \alpha^k (A^k)_{ji}. \tag{5.26}$$

Katz centrality has therefore two interpretations:

(i) the importance of a node is an attenuated sum of the importance and of the number of the in-neighbors – note indeed how equation (5.25) is a combination of equations (5.23) and (5.24), and

(ii) the importance of a node is α times number of length-1 walks into i (i.e., the in-degree) plus α^2 times the number of length-2 walks into i, etc. (From Lemma 4.2, recall that, for an unweighted digraph, $(A^k)_{ji}$ is equal to the number of directed walks of length k from j to i.)

Note how, for $\alpha < 1/\rho(A)$, equation (5.25) is well-posed and equivalent to

$$
\begin{aligned}
c_{\mathsf{K}} &= \alpha A^{\mathsf{T}}(c_{\mathsf{K}} + \mathbb{1}_n) \\
\iff c_{\mathsf{K}} + \mathbb{1}_n &= \alpha A^{\mathsf{T}}(c_{\mathsf{K}} + \mathbb{1}_n) + \mathbb{1}_n \\
\iff (I_n - \alpha A^{\mathsf{T}})(c_{\mathsf{K}} + \mathbb{1}_n) &= \mathbb{1}_n \\
\iff c_{\mathsf{K}} &= (I_n - \alpha A^{\mathsf{T}})^{-1}\mathbb{1}_n - \mathbb{1}_n \tag{5.27} \\
\iff c_{\mathsf{K}} &= \sum_{k=1}^{\infty} \alpha^k (A^{\mathsf{T}})^k \mathbb{1}_n,
\end{aligned}
$$

where we used the identity $(I_n - \mathcal{A})^{-1} = \sum_{k=0}^{\infty} \mathcal{A}^k$ valid for any matrix $\mathcal{A}$ with $\rho(\mathcal{A}) < 1$; see Exercise E2.11.

There are two simple ways to compute the Katz centrality. According to equation (5.27), for limited size problems, one can invert the matrix $(I_n - \alpha A^\mathsf{T})$. Alternatively, one can show (see Exercise E5.13) that the following iteration converges to the correct value:
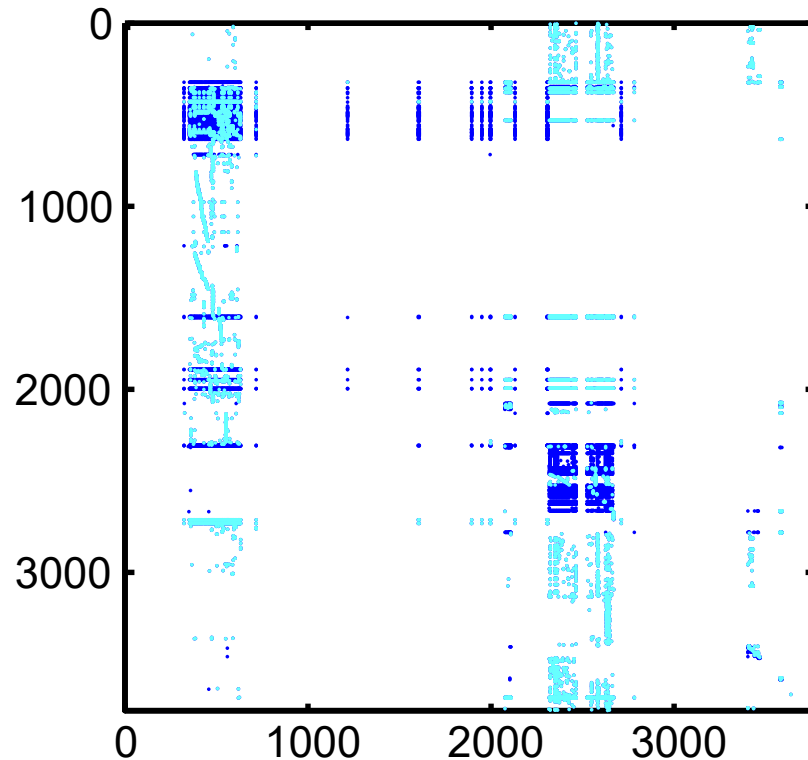$c_\mathsf{K}^+ := \alpha A^\mathsf{T}(c_\mathsf{K} + \mathbb{1}_n).$

Figure 5.12: The pattern in figure is the pixel picture of the so-called hyperlink matrix, i.e., the transpose of the adjacency matrix, for a collection of websites at the Lincoln University in New Zealand from the year 2006. Dark-colored points are nonzero entries of the adjacency matrix; light-colored points are outgoing links toward dangling nodes. Each empty column corresponds to a webpage without any outgoing link, that is, to a so-called dangling node. This network has 3756 nodes with 31,718 links. A fairly large portion of the nodes are dangling nodes: in this example, there are 3255 dangling nodes, which is over 85% of the total. Image courtesy of Hideaki Ishii and Roberto Tempo from data described in (Ishii and Tempo, 2014).

PageRank centrality   For a weighted digraph G with row-stochastic adjacency matrix (i.e., unit out-degree for each node), pick a convex combination coefficient $\alpha \in\ ]0, 1[$ and define the *PageRank centrality vector*, denoted by c_{pr}, as the unique positive solution to

$$c_{\mathrm{pr}}(i) = \alpha \sum_{j=1}^{n} a_{ji}c_{\mathrm{pr}}(j) + \frac{1-\alpha}{n}, \tag{5.28}$$

or, equivalently, to

$$c_{\mathrm{pr}} = Mc_{\mathrm{pr}}, \quad \mathbb{1}_n^{\mathsf{T}}c_{\mathrm{pr}} = 1, \quad \text{where } M = \alpha A^{\mathsf{T}} + \frac{1-\alpha}{n}\mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}. \tag{5.29}$$

(To establish the equivalence between these two definitions, the only non-trivial step is to notice that if c_{pr} solves equation (5.28), then it must satisfy $\mathbb{1}_n^{\mathsf{T}}c_{\mathrm{pr}} = 1$.)

Note that, for arbitrary unweighted digraphs and binary adjacency matrices $A_{0,1}$, it is natural to compute the PageRank vector with $A = D_{\mathrm{out}}^{-1}A_{0,1}$. We refer to (Ishii and Tempo, 2014; Gleich, 2015) for the important interpretation of the PageRank score as the stationary distribution of the so-called random surfer of an hyperlinked document network — it is under this disguise that the PageRank score was conceived by the Google co-founders and a corresponding algorithm led to the establishment of the Google search engine. In the Google problem it is customary to set $\alpha \approx .85$.

Closeness and betweenness centrality (based on shortest walks)   Degree, eigenvector, Katz and PageRank centrality are presented using the adjacency matrix. Next we present two centrality measures based on the notions of shortest walk and geodesic distance; these two notions belong to the class of *radial* and *medial* centrality measures (Borgatti and Everett, 2006).

   We start by introducing some additional graph theory. For a weighted digraph with n nodes, the *length* of a directed walk is the sum of the weights of edges in the directed walk. For $i, j \in \{1, \ldots, n\}$, a *shortest walk* from a node i to a node j is a directed walk of smallest length. Note: it is easy to construct examples with multiple shortest walks, so that the shortest walk is not unique. The *geodesic distance* $d_{i \to j}$ from node i to node j is the length of a shortest walk from node i to node j; we also stipulate that the geodesic distance $d_{i \to j}$ takes the value zero if $i = j$ and is infinite if there is no walk from i to j. Note: in general $d_{i \to j} \neq d_{j \to i}$. Finally, for $i, j, k \in \{1, \ldots, n\}$, we let $g_{i \to k \to j}$ denote the number of shortest walks from a node i to a node j that pass through node k.

   For a strongly-connected weighted digraph, the *closeness centrality score* of node $i \in \{1, \ldots, n\}$ is the inverse sum over the geodesic distances $d_{i \to j}$ from node i to all other nodes $j \in \{1, \ldots, n\}$, that is:

$$c_{\text{closeness}}(i) = \frac{1}{\sum_{j=1}^{n} d_{i \to j}}. \tag{5.30}$$

   For a strongly-connected weighted digraph, the *betweenness centrality score* of node $i \in \{1, \ldots, n\}$ is the fraction of all shortest walks $g_{k \to i \to j}$ from any node k to any other node j passing through node i, that is:

$$c_{\text{betweenness}}(i) = \frac{\sum_{j,k=1}^{n} g_{k \to i \to j}}{\sum_{h=1}^{n} \sum_{j,k=1}^{n} g_{k \to h \to j}}. \tag{5.31}$$

Summary   To conclude this section, in Table 5.1, we summarize the various centrality definitions for a weighted directed graph.

Measure	Definition	Assumptions
degree centrality	$c_{\text{degree}} = A^{\mathsf{T}} \mathbb{1}_n$	
eigenvector centrality	$c_{\text{ev}} = \alpha A^{\mathsf{T}} c_{\text{ev}}$	$\alpha = \dfrac{1}{\rho(A)}, G$ has a globally reachable node
PageRank centrality	$c_{\text{pr}} = \alpha A^{\mathsf{T}} c_{\text{pr}} + \dfrac{1-\alpha}{n} \mathbb{1}_n$	$\alpha < 1, A\mathbb{1}_n = \mathbb{1}_n$
Katz centrality	$c_{\text{K}} = \alpha A^{\mathsf{T}}(c_{\text{K}} + \mathbb{1}_n)$	$\alpha < \dfrac{1}{\rho(A)}$
closeness centrality	$c_{\text{closeness}}(i) = \dfrac{1}{\sum_{j=1}^{n} d_{i \to j}}$	G strongly connected
betweenness centrality	$c_{\text{betweenness}}(i) = \dfrac{\sum_{j,k=1}^{n} g_{k \to i \to j}}{\sum_{h=1}^{n} \sum_{j,k=1}^{n} g_{k \to h \to j}}$	G strongly connected

Table 5.1: Definitions of centrality measures for a weighted digraph G with adjacency matrix A

Figure 5.13 illustrates some centrality notions on an instructive example due to Brandes (2006). As it can be computed via the code in Table 5.2, a different node is the most central one in each metric; this variability is naturally expected and highlights the need to select a centrality notion relevant to the specific application of interest.

(a) degree centrality

(b) eigenvector centrality

(c) closeness centrality

(d) betweenness centrality
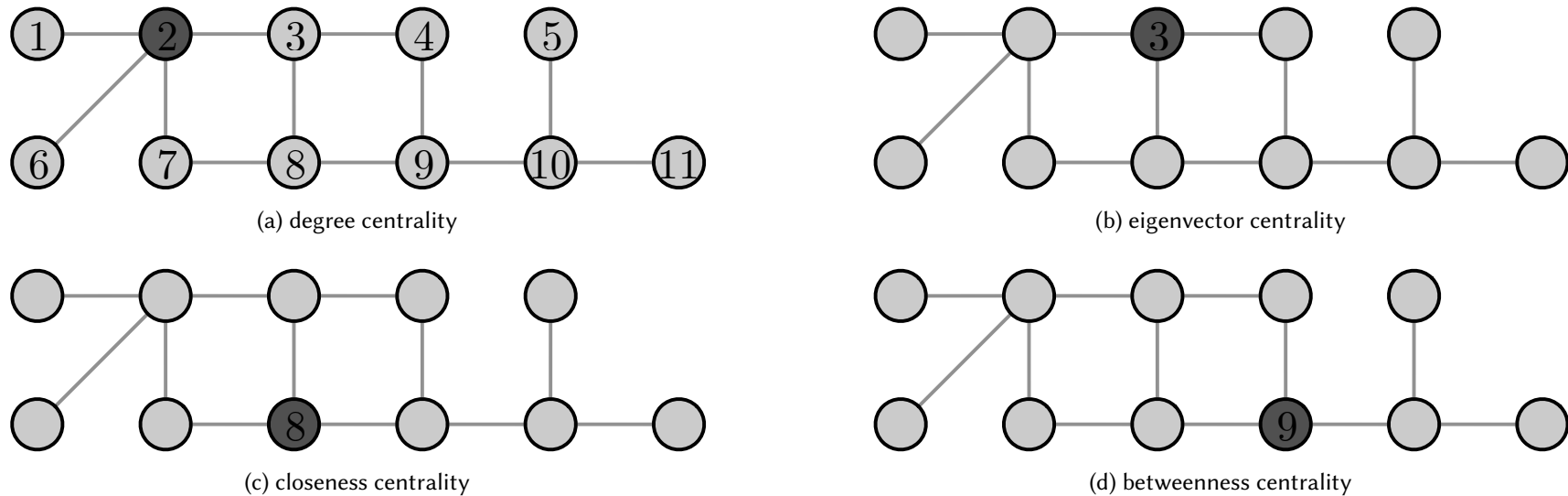
Figure 5.13: Degree, eigenvector, closeness, and betweenness centrality for an undirected unweighted graph. The dark node is the most central node in the respective metric; a different node is the most central one in each metric.

```python
1  # Python3 code for centrality computation via the NetworkX library
2  import networkx as nx
3  # define graph
4  G = nx.Graph(); G.add_nodes_from(range(1,11));  G.add_edges_from([(1,2), (2,6), (2,7), (2,3), (3,4), (3,8), (7,8), (8,9), ...
       (4,9), (9,10), (10,5), (10,11)])
5
6  # Node 2 has the highest degree centrality
7  degree_centrality = nx.degree_centrality(G)
8  # Node 3 has the highest eigenvector centrality
9  eigenvector_centrality = nx.eigenvector_centrality_numpy(G)
10 # Node 8 has highest closeness centrality
11 closeness_centrality = nx.closeness_centrality(G)
12 # Node 9 has highest closeness centrality
13 betweenness_centrality = nx.betweenness_centrality(G)
```

Table 5.2: Python code for centrality computation

5.6   Historical notes and further reading

The convergence of powers of row-stochastic matrices and of repeated averaging operations has been studied since the original work by Markov (1906), who proved an early version of Lemma 5.5. The τ_1 ergodicity coefficient is also referred to as the *Doeblin–Dobrushin* coefficient because of their early work Dobrushin (1956). Another early references on ergodicity coefficients and indecomposable stochastic matrices is (Wolfowitz, 1963); the interpretation of the word "ergodic" comes from the study of stochastic processes. An historic review is (Seneta, 1973), a notable review is (Ipsen and Selee, 2011), control theoretic approaches include (Liu et al., 2011), recent progress is detailed in (Marsli and Hall, 2020). In this literature, the max-min disagreement function $V_{\text{max-min}}$ is sometime referred to as the *diameter* or the *spread* of the vector.

For references on social influence networks and opinion dynamics we refer to Chapter 1. An early reference for Theorem 5.2 is (DeMarzo et al., 2003, Appendix C and, specifically, Theorem 10).

Opinion dynamics models are surveyed by Proskurnikov and Tempo (2017); for example, for nonlinear models based on bounded confidence we refer to (Hegselmann and Krause, 2002; MirTabatabaei and Bullo, 2012).

On the topic of computing optimal row-stochastic matrices, we postpone to Chapter 11 the study of related optimization problems.

A standard modern treatment of centrality notions is (Newman, 2010, Chapter 7); see also (Easley and Kleinberg, 2010, Chapter 14) for an introductory discussion. We also refer to (Brandes and Erlebach, 2005) for a comprehensive review of network analysis metrics and related computational algorithms, beyond centrality measures. Historically, centrality measures were originally studied in sociology, An incomplete list of early references and historical reviews in sociology includes (Bavelas, 1950) on closeness centrality, (Katz, 1953) on Katz centrality, (Freeman, 1977) on betweenness centrality, and (Bonacich, 1972a,b) on eigenvector centrality. Kleinberg (1999) generalizes centrality notions to networks with hubs and authorities; see Exercise E5.15.

PageRank is a centrality measure that has received tremendous recent attention due to the success of the Google search engines; this notion was popularized by (Brin and Page, 1998; Page, 2001), but see also the previous work (Friedkin, 1991) on total effective centrality and its relationship with PageRank (Friedkin and Johnsen, 2014). We refer to (Ishii and Tempo, 2014; Gleich, 2015; Nesterov, 2012) for recent works on PageRank and its multiple extensions and applications; we refer to (Ishii and Tempo, 2010; Zhao et al., 2013) for randomized distributed algorithms for PageRank computation.

5.7   Exercises

E5.1   **The final opinion of a French-Harary-DeGroot panel.** A conversation between 5 panelists is modeled according to the French-Harary-DeGroot model $x^+ = A_{\text{panel}}x$, where

$$
A_{\text{panel}} = \begin{bmatrix}
0.15 & 0.15 & 0.1 & 0.2 & 0.4 \\
0 & 0.55 & 0 & 0 & 0.45 \\
0.3 & 0.05 & 0.05 & 0 & 0.6 \\
0 & 0.4 & 0.1 & 0.5 & 0 \\
0 & 0.3 & 0 & 0 & 0.7
\end{bmatrix}.
$$

Assuming that the panel has sufficiently long deliberations, answer the following:

(i)   Draw the condensation of the associated digraph.

(ii)   Do the panelists finally agree on a common decision?

(iii)   In the event of agreement, does the initial opinion of any panelists get rejected? If so, which ones?

(iv)   Assume the panelists' initial opinions are their self-appraisals (that is, the self-weights $a_{11}, \ldots, a_{55}$) and compute the final opinion via elementary calculations.

E5.2 **Three averaging panels.** Consider the French-Harary-DeGroot opinion dynamics model $x(k+1) = Ax(k)$, where $x_i(k)$ denotes the opinion of individual i at time k and where A is row-stochastic. Recall that the coefficient $a_{ij} \in [0,1]$ is the influence of individual j on the update of the opinion of individual i, subject to the constraint $\sum_{j=1}^{n} a_{ij} = 1$. Consider the following three scenarios:

(i) Everybody gives equal weight to the opinion of everybody (including themselves).

(ii) Individual 1 gives equal weight to the opinions of herself and all others. Each individual $2, \ldots, n$ computes the average between her own opinion and that of individual 1.

(iii) Individual 1 does not change her opinion. Each individual $2, \ldots, n$ computes the average between her own opinion and that of individual 1.

For each scenario, derive the averaging matrix A, show that the opinions converge asymptotically to a final opinion vector, and characterize this final opinion vector.

▶ E5.3   **Designing averaging weights for an example topology.** Let G be the digraph in the following figure and let A be its adjacency matrix.
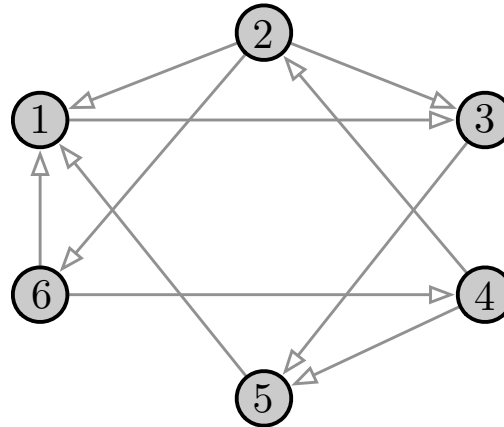
(i) Determine whether A is reducible or irreducible.

   • If A is reducible, compute a permutation matrix $P \in \{0, 1\}^{6 \times 6}$ such that

$$PAP^\mathsf{T} = \begin{bmatrix} B_{r \times r} & C_{r \times (n-r)} \\ \mathbb{0}_{(n-r) \times r} & D_{(n-r) \times (n-r)} \end{bmatrix} ,$$
                                                                                                                                                                (E5.1)

   for some matrices B, C, D and $r \geq 1$. Moreover, determine an edge to be added or removed from G, so that the accordingly modified adjacency matrix becomes irreducible.
   • If A is irreducible, determine an edge to be added or removed from G so that the accordingly modified adjacency matrix becomes reducible and compute the corresponding permutation matrix P as in (E5.1).

(ii) Design a new weighted digraph G' with weighted adjacency matrix A' by starting with G and performing only the following actions:

   • either add or remove (not both) a single edge from the digraph G, and
   • select the weight of each edge of G' to be any real value of your choice.

   Consider the iteration $x_{k+1} = A'x_k$, for $x_k \in \mathbb{R}^6$ and $k \in \mathbb{Z}_{\geq 0}$, and design a weighted digraph G' so that $\lim_{k \to \infty} x_k = \alpha \mathbb{1}_6$, for some $\alpha \in \mathbb{R}$.

▶ E5.4   **The equal-neighbor model over undirected topologies.** Given an irreducible symmetric non-negative matrix $A \in \mathbb{R}^{n \times n}$ and a vector $x \in \mathbb{R}^n_{>0}$, define

$$\mathcal{B}(x) = \mathrm{diag}(Ax)^{-1} A \, \mathrm{diag}(x). \tag{E5.2}$$

Show that

(i)   the matrix $\mathcal{B}(x)$ is well-defined and row-stochastic with same irreducible off-diagonal pattern as A,

(ii) has dominant left eigenvector $\pi(x) = \frac{1}{x^\mathsf{T} Ax} \, \mathrm{diag}(x) Ax$, and

(iii) is reversible, that is, $\mathrm{diag}(\pi(x))\mathcal{B}(x) = \mathcal{B}(x)^\mathsf{T} \, \mathrm{diag}(\pi(x))$.

Note: This classic result in Markov chain theory is related to Markov chains with maximal entropy (George et al., 2019).

E5.5   **The equal-neighbor row-stochastic matrix for weighted directed graphs.** Let G be a weighted digraph with n nodes, weighted adjacency matrix A and weighted out-degree matrix D_{out}. Define the *equal-neighbor-after-addition matrix*

$$A_{\text{equal-nghbr}} = (I_n + D_{\text{out}})^{-1}(I_n + A).$$

Show that

  (i) $A_{\text{equal-nghbr}}$ is row-stochastic;

  (ii) $A_{\text{equal-nghbr}}$ is primitive if and only if G is strongly connected; and

  (iii) $A_{\text{equal-nghbr}}$ is doubly-stochastic if G is weight-balanced and the weighted degree is constant for all nodes (i.e., $D_{\text{out}} = D_{\text{in}} = dI_n$ for some $d \in \mathbb{R}_{>0}$).

E5.6   **A stubborn individual.** Pick $\alpha \in \ ]0, 1[$, and consider the discrete-time averaging algorithm

$$x_1(k+1) = x_1(k),$$
$$x_2(k+1) = \alpha x_1(k) + (1 - \alpha)x_2(k).$$

Perform the following tasks:

   (i) compute the matrix A representing this algorithm and verify it is row-stochastic,
  (ii) compute the eigenvalues and left and right eigenvectors of A,
 (iii) draw the directed graph G representing this algorithm and discuss its connectivity properties,
  (iv) draw the condensation digraph of G,
   (v) compute the final value of this algorithm as a function of the initial values in two alternate ways:

      (a) invoking Exercise E2.9, and
      (b) invoking Theorem 5.1.

E5.7   **Individuals with self-confidence levels.** Consider 2 individuals, labeled $+1$ and -1, described by the self-confidence levels s_{+1} and s_{-1}. Assume $s_{+1} \geq 0$, $s_{-1} \geq 0$, and $s_{+1} + s_{-1} = 1$. For $i \in \{+1, -1\}$, define

$$x_i^+ := s_i x_i + (1 - s_i) x_{-i}.$$

Perform the following tasks:

(i) compute the matrix A representing this algorithm and verify it is row-stochastic,

(ii) compute A^2,

(iii) compute the eigenvalues, the right eigenvectors, and the left eigenvectors of A,

(iv) compute the final value of this algorithm as a function of the initial values and of the self-confidence levels. Is it true that an individual with higher self-confidence makes a larger contribution to the final value?

E5.8　　**Rescaling row stochastic matrices and left dominant eigenvectors.** Let $A \in \mathbb{R}^{n \times n}$ be row-stochastic and irreducible with dominant left eigenvector $v \in \mathbb{R}^n_{>0}$ (not necessarily normalized). Given $x \in \mathbb{R}^n$, define

$$\mathcal{A}(x) = \operatorname{diag}(x) + (I_n - \operatorname{diag}(x))A \in \mathbb{R}^{n \times n}.$$

Show that

(i)　$\mathcal{A}(x)$ is row-stochastic and has the same irreducible off-diagonal pattern as A if and only if $x \in [0, 1[^n$.

Next, assume $x \in [0, 1[^n$ and show

(ii)　if A is primitive, then so is $\mathcal{A}(x)$,

(iii)　a left dominant eigenvector of $\mathcal{A}(x)$ is $v(x) = (I_n - \operatorname{diag}(x))^{-1}v$,

(iv)　$w \in \mathbb{R}^n_{>0}$ (not necessarily normalized) is a left dominant eigenvector of $\mathcal{A}(x)$ if and only if

$$x = \mathbb{1}_n - \beta \operatorname{diag}(w)^{-1}v \in [0, 1[^n, \quad \text{for } 0 < \beta \leq \min_{i \in \{1, \ldots, n\}} w_i/v_i.$$

▶ E5.9 **The rank-one projection matrix defined by a primitive matrix.** From linear algebra recall:

- a square matrix B is a *projection matrix* if $B^2 = B$,
- a vector space V is the *direct sum* of two subspaces U and W, written $V = U \oplus W$, if each $v \in V$ defines unique $u \in U$ and $w \in W$ such that $v = u + w$, and
- a subspace U is *invariant under a linear map B* if $u \in U$ implies $Bu \in U$.

Let A be an n-dimensional primitive matrix with dominant eigenvalue λ, right dominant eigenvector $v > 0$ and left dominant eigenvector $w > 0$ with the normalization $v^\mathsf{T} w = 1$. Define the rank-one matrix $P_A := vw^\mathsf{T}$. Show that:

(i) $P_A = P_A^2$ is a projection matrix with image $\mathrm{span}\{v\}$ and it is an orthogonal projection if A is symmetric,

(ii) $I_n - P_A = (I_n - P_A)^2$ is a projection matrix whose image is $\mathrm{kernel}(P_A) = \{q \in \mathbb{R}^n \mid w^\mathsf{T} q = 0\} = \mathrm{span}\{w\}^\perp$,

(iii) $AP_A = P_A A = \lambda P_A$,

(iv) $\mathbb{R}^n = \mathrm{span}\{v\} \oplus \mathrm{span}\{w\}^\perp$ and both subspaces $\mathrm{span}\{v\}$ and $\mathrm{span}\{w\}^\perp$ are invariant under A,

(v) the restriction of A to the $\mathrm{span}\{v\}$ is multiplication by λ and the restriction of A to $\mathrm{span}\{w\}^\perp$ has eigenvalues equal to all eigenvalues of A except λ, and

(vi) $(A - \lambda vw^\mathsf{T})^k = A^k - \lambda^k vw^\mathsf{T}$ for all $k \in \mathbb{N}$.

Note: In some references, the matrix P_A is referred to as the spectral projector of A associated to the eigenvalue λ.

E5.10 **Bounds on the norm of a matrix power.** Given a matrix $B \in \mathbb{R}^{n \times n}$ and an index $k \in \mathbb{N}$, show that

(i) there exists $c > 0$ such that

$$\|B^k\|_2 \le c\, k^{n-1} \rho(B)^k,$$

(ii) for all $\varepsilon > 0$, there exists $c_\varepsilon > 0$ such that

$$\|B^k\|_2 \le c_\varepsilon (\rho(B) + \varepsilon)^k.$$

Hint: Use the Jordan normal form.

E5.11 **Average consensus via the parallel averaging algorithm.** Let G be a weighted graph with weighted adjacency matrix A and weighted degrees $d_1, \ldots, d_n$. Assume G is connected and aperiodic and consider the equal-neighbor matrix $A_{\text{en}} = \text{diag}(d_1, \ldots, d_n)^{-1} A$. Assign a value $x_i \in \mathbb{R}$ to each node i and consider the *parallel averaging algorithm:*
  1: each node i sets $y_i(0) = 1/d_i$ and $z_i(0) = x_i/d_i$
  2: the nodes run the averaging algorithms $y(k+1) = A_{\text{en}} y(k)$ and $z(k+1) = A_{\text{en}} z(k)$ for $k \in \mathbb{Z}_{\geq 0}$
  3: each node i sets $x_i(k) = z_i(k)/y_i(k)$ at each $k \in \mathbb{Z}_{\geq 0}$

Show that the parallel averaging algorithm

(i) is well posed, i.e., $y_i(k)$ does not vanish for any $i \in \{1, \ldots, n\}$ and $k \in \mathbb{Z}_{\geq 0}$, and

(ii) achieves average consensus, that is, $\lim_{k \to \infty} x(k) = \text{average}(x_1, \ldots, x_n) \mathbb{1}_n$.

Note: This algorithm is also referred to as the push sum iteration, because it may implemented over directional communication by "summing the pushed variables." This algorithm was originally introduced by Kempe et al. (2003) and later studied in (Olshevsky and Tsitsiklis, 2009; Benezit et al., 2010).

E5.12 **Computing centrality.** Write in your favorite programming language algorithms to compute degree, eigenvector, Katz and PageRank centralities. Compute these four centralities for the following undirected unweighted graphs (without self-loops):

  (i) the complete graph with 5 nodes;

  (ii) the cycle graph with 5 nodes;

  (iii) the star graph with 5 nodes; and

  (iv) the Zachary karate club network dataset. This dataset can be found for example at `https://en.wikipedia.org/wiki/Zachary%27s_karate_club`.

To compute the PageRank centrality, use $\alpha = .85$. To compute the Katz centrality of a matrix A, select for example $\alpha = 1/(2\rho(A))$.

E5.13 **Iterative computation of Katz centrality.** Given a graph with adjacency matrix A, show that the solution to the iteration $x(k+1) := \alpha A^{\mathsf{T}}(x(k) + \mathbb{1}_n)$ with $\alpha < 1/\rho(A)$ converges to the Katz centrality vector c_{K}, for all initial conditions $x(0)$.

E5.14   **Robotic coordination: deployment and centering as a discrete-time reducible averaging systems.** Consider $n \geq 3$ robots with positions $p_i \in \mathbb{R}$, $i \in \{1, \ldots, n\}$, dynamics $p_i(k+1) = u_i(k)$, where $u_i \in \mathbb{R}$ is a steering control input. Assume that the robots are indexed according to their initial position: $p_1(0) \leq p_2(0) \leq \cdots \leq p_n(0)$. Consider two walls at the positions $p_0 \leq p_1(0)$ and $p_{n+1} \geq p_n(0)$ so that all robots are contained between the walls. The walls are stationary, that is, $p_0(k+1) = p_0(k) = p_0$ and $p_{n+1}(k+1) = p_{n+1}(k) = p_{n+1}$ for all times k.

Consider the following coordination law: robots $i \in \{1, \ldots, n\}$ (each having two neighbors) move to the centroid of the local subset $\{p_{i-1}, p_i, p_{i+1}\}$ or, in other words,

$$p_i(k+1) = \frac{1}{3}(p_{i-1}(k) + p_i(k) + p_{i+1}(k)), \quad i \in \{1, \ldots, n\}.$$

Show that the robots become asymptotically uniformly spaced on the interval $[p_0, p_{n+1}]$.

E5.15 **Hubs and authorities (Kleinberg, 1999).** Let G be a digraph with node set $\{1, \ldots, n\}$ and edge set E. Assume G has a globally reachable node and the subgraph of globally reachable nodes is aperiodic.

We define two scores for each node $j \in \{1, \ldots, n\}$: the *hub score* $h_j \in \mathbb{R}$ and the *authority score* $a_j \in \mathbb{R}$. We initialize these scores with positive values and updated them simultaneously as follows: the hub score of node j is set equal to the sum of the authority scores of all nodex pointed to by j, and, similarly, the authority score of node j is set equal to the sum of the hub scores of all nodes pointing to j. In concise formulas, for $k \in \mathbb{N}$,

$$\begin{cases} h_j(k+1) = \sum_{i:\,(j,i)\in E} a_i\,, \\ a_j(k+1) = \sum_{i:\,(i,j)\in E} h_i\,. \end{cases} \tag{E5.3}$$

(i) Let $x(k) = \begin{bmatrix} h(k)^{\mathsf{T}} & a(k)^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$ denote the stacked vector of hub and authority scores. Provide an update equation for the hub and authority scores of the form

$$x(k+1) = Mx(k),$$

for some matrix $M \in \mathbb{R}^{2n \times 2n}$.

(ii) Will the sequence $x(k)$ converge as $k \to \infty$?

In what follows, we consider the modified iteration

$$y(k+1) = \frac{My(k)}{\|My(k)\|_2}\,,$$

where M is defined as in statement (i) above.

(iii) Will the sequence $y(k)$ converge as $k \to \infty$?

(iv) Show that the two subsequences of even and odd iterates $\{y(2k)\}_k$ and $\{y(2k+1)\}_k$ converge, that is,

$$\lim_{k \to \infty} y(2k) = y_{\mathrm{even}}(y_0), \quad \lim_{k \to \infty} y(2k+1) = y_{\mathrm{odd}}(y_0),$$

where $y_0 = x(0)$ is the stacked vector of initial hub and authority scores.

(v) Provide expressions for $y_{\mathrm{even}}(y_0)$ and $y_{\mathrm{odd}}(y_0)$.

E5.16   **Reversible primitive row-stochastic matrices.** Let A be a primitive row-stochastic $n \times n$ matrix and w be its left dominant eigenvector. The matrix A is *reversible* if

$$w_i A_{ij} = A_{ji} w_j, \quad \text{for all } i, j \in \{1, \ldots, n\},$$ (E5.4)

or, equivalently,

$$\operatorname{diag}(w) A = A^\mathsf{T} \operatorname{diag}(w).$$

Prove the following statements:

(i) if A is reversible, then its associated digraph is undirected, that is, if (i, j) is an edge, then so is (j, i),

(ii) if A is reversible, then $\operatorname{diag}(w)^{1/2} A \operatorname{diag}(w)^{-1/2}$ is symmetric and, hence, A has n real eigenvalues and n eigenvectors, and

(iii) if A is an equal-neighbor matrix for an unweighted undirected graph, then A is reversible.

Recall that, for $w = (w_1, \ldots, w_n) > 0$, the following definitions hold:

$$\operatorname{diag}(w)^{1/2} = \operatorname{diag}(\sqrt{w_1}, \ldots, \sqrt{w_n}), \quad \text{and} \quad \operatorname{diag}(w)^{-1/2} = \operatorname{diag}(1/\sqrt{w_1}, \ldots, 1/\sqrt{w_n}).$$

E5.17 **Maximum entropy random walk (Burda et al., 2009).** Let G be an unweighted connected graph with binary adjacency matrix $A \in \{0,1\}^n$. Let (λ, v) be the dominant eigenpair, i.e, $Av = \lambda v$ and $\mathbb{1}_n^\top v = 1$. Similarly to E4.13, define the square matrix P by

$$p_{ij} = \frac{1}{\lambda} \frac{v_j}{v_i} a_{ij}, \quad \text{for } i,j \in \{1,\dots,n\}.$$

Perform the following tasks:

(i) show that P is well defined, row stochastic, and irreducible,

(ii) pick $i,j \in \{1,\dots,n\}$ and $k \geq 1$. Assuming there exists a walk of length k from i to j, let $c_{ij}^{[k]}$ denote the product of the edge weights along the walk and show that

$$c_{ij}^{[k]} = \frac{1}{\lambda^k} \frac{v_j}{v_i},$$

(iii) let $w > 0$ be the left dominant eigenvector of P, normalized so that $\mathbb{1}_n^\top w = 1$, and show that

$$w_i = \frac{1}{\|v\|_2^2} v_i^2.$$

E5.18   **The role of the nodal degree in averaging systems.** Let G be an connected undirected graph without self-loops. Consider the averaging dynamics:

$$x(k+1) = Ax(k),$$

where $A = D^{-1}A_{01}$, D is the degree matrix, and A_{01} is the binary adjacency matrix of G.

(i) Under which conditions on G will the system converge to a final consensus state, i.e., an element of $\mathrm{span}\{\mathbb{1}_n\}$?

(ii) Assuming the state converges to a final consensus value, what is this steady state value?

(iii) Find a function $f(k, \lambda_2, \ldots, \lambda_n, d_1, \ldots, d_n)$ depending on the time step k, the eigenvalues $\lambda_2, \ldots, \lambda_n$ of A, and the degrees of the nodes $d_1, \ldots, d_n$ such that

$$\|e(k)\|_2 \le f(k, \lambda_2, \ldots, \lambda_n, d_1, \ldots, d_n)\|e(0)\|_2.$$

E5.19   **On the Metropolis-Hastings algorithm.** This exercise generalize the treatment in Section 5.4.2. Consider an irreducible $n \times n$ row-stochastic matrix A satisfying the symmetric structure condition: $a_{ij} \neq 0$ if and only if $a_{ji} \neq 0$, for all i, j. Given a positive unit-sum vector π (i.e., $\pi_i > 0$ and $\sum_{i=1}^{n} \pi_i = 1$), define the *Metropolis–Hastings matrix* M by

$$M_{ij} = \begin{cases} a_{ij} \min(1, R_{ij}), & \text{if } i \neq j, \\ a_{ii} + \sum_{k \neq i} a_{ik}(1 - \min(1, R_{ik})), & \text{if } i = j, \end{cases} \qquad (\text{E5.5})$$

where $R_{ij} = \dfrac{\pi_j a_{ji}}{\pi_i a_{ij}}$ for all i, j such $a_{ij} \neq 0$, and $R_{ij} = 1$ otherwise. Show

   (i) M is row-stochastic with the same topology as A,

   (ii) $\pi_i M_{ij} = \pi_j M_{ji}$, that is, M is reversible, and

   (iii) $\pi^\mathsf{T} M = \pi^\mathsf{T}$.

Note: The Metropolis algorithm was introduced by Metropolis et al. (1953) and extended to non-symmetric matrices by Hastings (1970); (Billera and Diaconis, 2001) contains a geometric interpretation and (Bierkens, 2016) a non-reversible extension. Regarding M as the transition matrix of a Markov chain, equation (E5.5) has the following interpretation: from node i, choose node j with probability a_{ij} and accept this choice with probability $\min(1, R_{ij})$, otherwise remain at node i.

E5.20 **Approximate optimization via random walks with Metropolis filter.** Given a finite set V, consider an optimization problem with objective function $f \colon V \to \mathbb{R}_{\geq 0}$. In order to amplify the differences in f, define a function $F \colon V \to \mathbb{R}_{\geq 0}$ by $F(v) = \exp(f(v)/T)$ for $0 < T \ll 1$. We now design an algorithm that allows us to draw a random sample from a distribution over V proportional to F; such an algorithm therefore allows us to pick a good solution to the optimization problem with high probability.

Given a regular graph with node set V, define the *Metropolis-filtered random walk* over V in the following two-step process. Let v_k denote the node visited at time k. First, select uniformly a random neighbor u of v_k. Second, if $F(u) \geq F(v_k)$, then move to u; otherwise, move to u with probability $F(u)/F(v_k)$ and stay at v_k with probability $1 - F(u)/F(v_k)$. Show that the Metropolis-filtered random walk is a reversible Markov chain with stationary distribution π satisfying

$$\pi(v) = \frac{F(v)}{\sum_{w \in V} F(w)}. \tag{E5.6}$$

Note: The Metropolis-filtered algorithm is due to Metropolis et al. (1953) and nicely reviewed by Lovász (1993). Note that an explicit representation of the graph G is not required.

E5.21   **Balanced vectors as positive sums of appropriate basis vectors.** Two vectors $v, w \in \mathbb{R}^n$ are *sign compatible* if $v_i w_i \geq 0$ for all $i \in \{1, \ldots, n\}$. If v and w are sign compatible, then

$$\|v + w\|_1 = \|v\|_1 + \|w\|_1. \tag{E5.7}$$

Being sign compatible is transitive, that is, if v and w are sign compatible and w and z are sign compatible, then also v and z are sign compatible.

Consider a vector $x \in \mathbb{R}^n$ such that $\mathbb{1}_n^\mathsf{T} x = 0$. Then there exist nonnegative coefficients $y_{ij} \geq 0$, $i, j \in \{1, \ldots, n\}$, $i \neq j$, such that

(i) $x = \frac{1}{2} \sum_{i,j=1}^n y_{ij}(\mathbb{e}_i - \mathbb{e}_j)$,

(ii) x and each $y_{ij}(\mathbb{e}_i - \mathbb{e}_j)$ are sign compatible, and

(iii) $\|x\|_1 = \sum_{i,j=1,i\neq j}^n y_{ij}$.

Note: This exercise is (Seneta, 1981, Lemma 2.4 at page 62).

E5.22   **Equivalent definitions of the 1-coefficient of ergodicity.** Given a row-stochastic matrix A, show that the expressions for $\tau_1(A)$ in equation (5.15) in Section 5.3 are equal, i.e., prove that

$$\max_{\|y\|_1=1, y\perp \mathbb{1}_n} \|A^\mathsf{T} y\|_1 = \frac{1}{2} \max_{i,j\in\{1,\ldots,n\}} \sum_{h=1}^{n} |a_{ih} - a_{jh}| = 1 - \min_{i,j\in\{1,\ldots,n\}} \sum_{h=1}^{n} \min\{a_{ih}, a_{jh}\}.$$

E5.23   **An alternative ergodicity coefficient.** Given a row-stochastic matrix $A \in \mathbb{R}^{n \times n}$, define its *column-maximum row-minimum entry*, denoted $\gamma(A)$, by

$$\gamma(A) = \max_{j \in \{1,\ldots,n\}} \min_{i \in \{1,\ldots,n\}} a_{ij} \in [0,1]. \tag{E5.8}$$

It is useful to clarify how to compute this quantity: for each column j the quantity $b_j = \min_i a_{ij}$ is the smallest entry over the n rows, and then $\gamma(A) = \max_j b_j$ is the largest of these entries over the n columns. Show that

(i)   for all $x \in \mathbb{R}^n$, the max-min function satisfies

$$V_{\text{max-min}}(Ax) \leq \big(1 - \gamma(A)\big) V_{\text{max-min}}(x);$$

(ii) $\gamma(A) > 0$ if and only if A has a strictly positive column;

(iii)   if there exists an index $h \in \mathbb{N}$ such that A^h has a positive column, then $k \mapsto V_{\text{max-min}}(x(k))$ converges exponentially fast to zero in the sense that, for all time $k \in \mathbb{N}$,

$$V_{\text{max-min}}(x(k)) \leq \underbrace{\big(1 - \gamma(A^h)\big)}_{<1}{}^{\lfloor k/h \rfloor} V_{\text{max-min}}(x(0)).$$

E5.24 **Perturbation bounds on the left dominant eigenvector.** Let A and $A + \Delta$ be row-stochastic and irreducible; let $v(A)$ and $v(A + \Delta)$ denote their left dominant eigenvectors, normalized to have unit sum. If A is scrambling so that $\tau_1(A) < 1$, then

$$\|v(A) - v(A + \Delta)\|_1 \leq \frac{\|\Delta\|_\infty}{1 - \tau_1(A)}. \tag{E5.9}$$

Note: This result is originally by Seneta (1988). A comparison of various perturbation bounds, also referred to sensitivity bounds or condition numbers, is given by Cho and Meyer (2001).

Exercises with solution

E5.25 **Necessary and sufficient conditions for semi-convergence.** With the same notation as in Theorem 5.2, prove that the following three properties are equivalent:

(A1) the eigenvalue 1 is semi-simple with multiplicity n_s and all other eigenvalues μ satisfy $|\mu| < 1$,

(A2) A is semi-convergent,

(A3) each sink of the condensation of G, regarded as a subgraph of G, is aperiodic.

Note: Gantmacher (1959) calls "regular" the semi-convergent row-stochastic matrices and "fully regular" the semi-convergent row-stochastic matrices whose limiting matrix has rank one, i.e., the indecomposable row-stochastic matrices.

Answer: The equivalence (A1) $\iff$ (A2), is a consequence of Theorem 2.7, where necessary and sufficient conditions for semi-convergence are given on the spectrum of A. In this case A is row-stochastic so that 1 is an eigenvalue. (Recall Theorem 2.7 states that A is semi-convergent if and only if 1 is semi-simple and all other eigenvalues have magnitude less than 1.) In the implication (A2) $\implies$ (A1), the fact that the multiplicity is n_s can be seen by writing the matrix in its Jordan normal form.

Regarding (A3) $\implies$ (A1): This statement amounts to one of the claims of Theorem 5.2, for which we provide a full proof here. Suppose each sink of the condensation of G, viewed as a subgraph of G, is aperiodic. By definition, each sink of the condensation of G, viewed as a subgraph of G, has no outgoing edges. Thus, if we reorder A by a permutation which groups each of the n_s sinks of the condensation of G, A has the following block structure:

$$
PAP^{\mathsf{T}} = \begin{bmatrix}
A_{11} & 0 & \dots & 0 & 0 \\
0 & A_{22} & \dots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \dots & A_{n_s n_s} & 0 \\
A_{U1} & A_{U2} & \dots & A_{U n_s} & A_{UU}
\end{bmatrix}.
$$

Since PAP^{T} is row-stochastic ($PAP^{\mathsf{T}}\mathbb{1}_n = PA\mathbb{1}_n = P\mathbb{1}_n = \mathbb{1}_n$), we also have that the constituent diagonal blocks $A_{11}, A_{22}, \dots, A_{n_s n_s}$ are row-stochastic. If we consider the collection of nodes U which do not belong to any of the n_s sinks, there must be an out-going edge from U or else one of those nodes would belong to a sink. Thus, $\begin{bmatrix} A_{U1} & A_{U2} & \dots & A_{U n_s} \end{bmatrix} \neq 0$ and the matrix A_{UU} is thus row substochastic. Every node i in U must have a walk to a node j belonging to a sink. Clearly, j is not in U. But then every i must have a walk to a node k with a directed edge to a node not in U. Thus, by Corollary 4.13, A_{UU} is convergent: $\rho(A_{UU}) < 1$.

Since each sink of the condensation of G is strongly connected (since every node in the condensation is a maximal strongly connected component) and aperiodic (by hypothesis) when viewed as a subgraph of G, each of the matrices $A_{11}, A_{22}, \dots, A_{n_s n_s}$ is primitive by Theorem 4.7 and thus have 1 as a strictly dominant eigenvalue by the Perron–Frobenius Theorem 2.12.

Since the eigenvalues of a block triangular matrix are the union of the eigenvalues of the blocks and the eigenvalues of a matrix are invariant under a symmetric permutation (or indeed any similarity transformation), A has eigenvalue 1 with multiplicity n_s and all other eigenvalues strictly smaller in magnitude. Moreover, since 1 has multiplicity one in each of the n_s blocks $A_{11}, \dots, A_{n_s n_s}$, 1 is semisimple.

Regarding (A1) $\implies$ (A3), suppose that 1 is a semisimple eigenvalue of multiplicity n_s and all other eigenvalues are strictly smaller. Consider the block structure of PAP^{T} shown above. Recall that the eigenvalues of A are the union of the eigenvalues of the diagonal blocks $A_{11}, A_{22}, \dots, A_{n_s n_s}, A_{UU}$. Since $A_{11}, A_{22}, \dots, A_{n_s n_s}$

are row stochastic and thus have eigenvalue 1, we conclude that all of the eigenvalues of A_{UU} are smaller than 1, $\rho(A_{UU}) < 1$, and that $A_{11}, A_{22}, \ldots, A_{n_s n_s}$ all have simple eigenvalue 1 with all eigenvalues strictly smaller than 1.

Now, assume by contradiction that the jth sink S of the condensation of G, viewed as a subgraph of G, is periodic with period k. Define $B := A_{jj}$. Pick any node k in S and assign each node ℓ in S a number between 0 and $k-1$ defined to be the length of any directed walk from k to ℓ modulo the period k. Since S is a node in the condensation of G, S is strongly connected so every node has a number. The number of a node is well-defined since if there are walks of length p and p' from k to ℓ, then concatenating with a walk of length L from ℓ to k (which exists because S is strongly connected) gives cycles with lengths $p + L$ and $p' + L$. Since $p + L$ and $p' + L$ are both multiples of k, $p - p'$ is as well so p and p' have the same remainder modulo k.

Notice that all edges from nodes with number m must be to nodes with number $(m+1) \mod k$. Thus, after permuting B such that the nodes with the same number are grouped together, we have

$$
PBP^{\mathsf{T}} = \begin{bmatrix} 0 & B_{01} & 0 & \ldots & 0 & 0 \\ 0 & 0 & B_{12} & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & B_{(k-2)(k-1)} \\ B_{(k-1)0} & 0 & 0 & \ldots & 0 & 0 \end{bmatrix}.
$$

Since B is row stochastic (and so is PBP^{T}) we have that $\mathbb{1}_{|S|}$ is an eigenvector of PBP^{T} with eigenvalue 1. This implies that $B_{i(i+1)}\mathbb{1}_{|S|/p} = \mathbb{1}_{|S|/p}$ for $0 \le i \le k-2$ and $B_{(k-1)0}\mathbb{1}_{|S|/p} = \mathbb{1}_{|S|/p}$. Define the primitive kth root of unity $\omega := \exp(2\pi \mathrm{i}/|S|)$ where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. Observe that $\omega^k = 1$. Then

$$
PBP^{\mathsf{T}} \begin{bmatrix} \mathbb{1}_{|S|/p} \\ \omega\mathbb{1}_{|S|/p} \\ \omega^2\mathbb{1}_{|S|/p} \\ \vdots \\ \omega^{k-2}\mathbb{1}_{|S|/p} \\ \omega^{k-1}\mathbb{1}_{|S|/p} \end{bmatrix} = \begin{bmatrix} 0 & B_{01} & 0 & \ldots & 0 & 0 \\ 0 & 0 & B_{12} & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & B_{(k-2)(k-1)} \\ B_{(k-1)0} & 0 & 0 & \ldots & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbb{1}_{|S|/p} \\ \omega\mathbb{1}_{|S|/p} \\ \omega^2\mathbb{1}_{|S|/p} \\ \vdots \\ \omega^{k-2}\mathbb{1}_{|S|/p} \\ \omega^{k-1}\mathbb{1}_{|S|/p} \end{bmatrix}
$$

$$
= \omega \begin{bmatrix} \mathbb{1}_{|S|/p} \\ \omega\mathbb{1}_{|S|/p} \\ \omega^2\mathbb{1}_{|S|/p} \\ \vdots \\ \omega^{k-2}\mathbb{1}_{|S|/p} \\ \omega^{k-1}\mathbb{1}_{|S|/p} \end{bmatrix}.
$$

Thus ω is an eigenvalue of B with $|\omega| = 1$, contradicting the fact that all eigenvalues of B other than 1 have magnitude smaller than 1. Thus, we conclude that each of the sinks of the condensation of G must be aperiodic.

E5.26 **Persistent disagreement in the opinion dynamics model by Friedkin and Johnsen (1999).** Let A be a row-stochastic matrix whose associated digraph describes an *interpersonal influence network*. Let each individual possess an *openness level* $\lambda_i \in [0,1]$, $i \in \{1, \dots, n\}$, describing how open is the individual to changing her initial opinion about a subject; set $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Consider the *Friedkin-Johnsen model* of opinion dynamics

$$x(k+1) = \Lambda A x(k) + (I_n - \Lambda) x(0). \tag{E5.1}$$

In other words, in this model, each individual i exhibits an attachment $(1 - \lambda_i)$ to its initial opinion $x_i(0)$, $x_i(k)$ represents the current opinion and $x_i(0)$ represents a prejudice by individual i. Consider the following two assumptions:

(A1) at least one individual has a strictly positive attachment to its initial opinion, that is, $\lambda_i < 1$ for at least one individual i; and

(A2) the interpersonal influence network contains directed walks from each individual with openness level equal to 1 to an individual with openness level less than 1.

Note that, if Assumption (A1) is not satisfied and therefore $\Lambda = I_n$, then we recover the French-Harary-DeGroot opinion dynamics model introduced in Section 1.1 and analyzed in this chapter. In what follows, let Assumption (A1) hold.

(i) Show that the matrix ΛA is convergent if and only if Assumption (A2) holds.

   Hint: Recall Corollary 4.13

Next, under Assumption (A2), perform the following tasks:

(ii) show that the so-called total influence matrix $V = (I_n - \Lambda A)^{-1}(I_n - \Lambda)$ is well-defined and row-stochastic,

   Hint: Review Exercises E2.9 and E2.11

(iii) show that the limiting opinions satisfy $\lim_{k \to +\infty} x(k) = V x(0)$,

(iv) show that A and V have the same left dominant eigenvector when $\Lambda = \lambda I_n$, for $0 < \lambda < 1$,

(v) compute the matrix V and state whether two individuals will achieve consensus or maintain persistent disagreement for the following pairs of matrices:

$$A_1 = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \text{ and } \Lambda_1 = \mathrm{diag}(1/2, 1),$$

$$A_2 = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \text{ and } \Lambda_2 = \mathrm{diag}(1/4, 3/4).$$

Note: Friedkin and Johnsen (1999, 2011) make the additional assumption that $\lambda_i = 1 - a_{ii}$, for $i \in \{1, \dots, n\}$; this assumption couples the openness level with the interpersonal influences and has the effect of enhancing stubborness of the individuals. This assumption is not needed here. The model (E5.1) is also referred to the averaging model with stubborn individuals. Other properties of this model are studied in (Bindel et al., 2015; Friedkin et al., 2016; Ravazzi et al., 2015).

Answer: We start by proving fact (i). Because at least one individual i has $\lambda_i < 1$ for some i, then at least one of the rows of ΛA has sum less than 1 and so ΛA is row-substochastic. Assumption (A2) states that the digraph associated to ΛA contains directed walks from each node with out-degree equal to 1 to a node with out-degree less than 1 in ΛA. But this property of the digraph is precisely a necessary a sufficient condition for ΛA to be convergent, as stated in Corollary 4.13

Next we prove statement (ii). Since ΛA is convergent, we know from Exercise E2.9 that $I_n - \Lambda A$ is invertible. From Exercise E2.11 on the Neumann series, we write the matrix V as

$$V = \sum_{k=0}^{\infty} (\Lambda A)^k (I_n - \Lambda). \tag{E5.2}$$

The entries of V are non-negative because V is the infinite convergent sum of non-negative matrices. Next, because A is stochastic, we reason:

$$(I_n - \Lambda)\mathbb{1}_n = (I_n - \Lambda)\mathbb{1}_n$$
$$\implies (I_n - \Lambda A)\mathbb{1}_n = (I_n - \Lambda)\mathbb{1}_n$$
$$\implies \mathbb{1}_n = (I_n - \Lambda A)^{-1}(I_n - \Lambda)\mathbb{1}_n,$$

so that $V = (I_n - \Lambda A)^{-1}(I_n - \Lambda)$ has row-sums equal to 1. Hence V is well-defined and row-stochastic.

Next, statement (iii) is an immediate consequence of Exercise E2.9: every solution to an affine system $x(k+1) = Ax(k) + b$, with A convergent, will converge to the unique equilibrium point $(I_n - A)^{-1}b$. With the notation here, the unique equilibrium point is $(I_n - \Lambda A)^{-1}(I_n - \Lambda)x(0) = Vx(0)$.

Regarding statement (iv), let w be the left dominant eigenvector of A, that is, $w^\mathsf{T} A = w^\mathsf{T}$ and left-multiply equation (E5.2) by w to obtain:

$$w^\mathsf{T} V = w^\mathsf{T} \sum_{k=0}^{\infty} (\lambda A)^k (I_n - \lambda I_n) = (1 - \lambda)w^\mathsf{T} \sum_{k=0}^{\infty} \lambda^k A^k$$

$$= (1 - \lambda)\Big(\sum_{k=0}^{\infty} \lambda^k\Big)(w^\mathsf{T} A^k) = (1 - \lambda)\Big(\frac{1}{1 - \lambda}\Big)w^\mathsf{T} = w^\mathsf{T}.$$

For statement (v), with the help of a computer, we easily compute V to be, respectively

$$V_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad \text{and } V_2 = \begin{bmatrix} 15/16 & 1/16 \\ 9/16 & 7/16 \end{bmatrix},$$

so that consensus is achieved in the first case, but not in the second case for generic initial opinions.

Chapter 6

The Laplacian Matrix

The previous chapters studied adjacency matrices and their application to discrete-time averaging dynamics. This chapter introduces and characterizes a second relevant matrix associated to a digraph, called the Laplacian matrix. Laplacian matrices appear in numerous applications and enjoy numerous useful properties.

6.1   The Laplacian matrix

Definition 6.1 (Laplacian matrix of a digraph). *Given a weighted digraph G with adjacency matrix A and out-degree matrix $D_{\text{out}} = \text{diag}(A\mathbb{1}_n)$, the Laplacian matrix of G is*

$$L = D_{\text{out}} - A.$$

In components $L = (\ell_{ij})_{i,j\in\{1,\ldots,n\}}$

$$\ell_{ij} = \begin{cases} -a_{ij}, & \text{if } i \neq j, \\ \displaystyle\sum_{h=1,h\neq i}^{n} a_{ih}, & \text{if } i = j, \end{cases}$$

or, for an unweighted undirected graph,

$$\ell_{ij} = \begin{cases} -1, & \text{if } \{i,j\} \text{ is an edge and not a self-loop,} \\ d(i), & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

An example is illustrated in Figure 6.1.
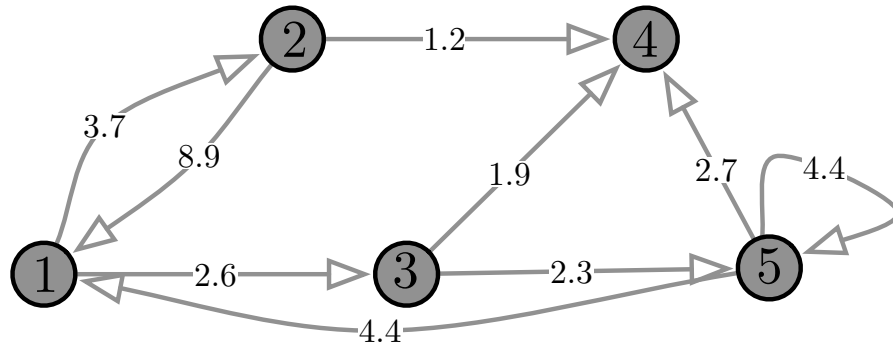
$$L = \begin{bmatrix} 6.3 & -3.7 & -2.6 & 0 & 0 \\ -8.9 & 10.1 & 0 & -1.2 & 0 \\ 0 & 0 & 4.2 & -1.9 & -2.3 \\ 0 & 0 & 0 & 0 & 0 \\ -4.4 & 0 & 0 & -2.7 & 7.1 \end{bmatrix}.$$

Figure 6.1: A weighted digraph and its Laplacian matrix

As a second example, we consider the complete undirected graph K_n. Recall from Table 4.1 that the adjacency matrix is $A_{K_n} = \mathbb{1}_n \mathbb{1}_n^\mathsf{T} - I_n$ and that each node has degree $n - 1$. Given the projection matrix $\Pi_n = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T}$ from Section 5.3, we compute the Laplacian of K_n to be

$$L_{K_n} = (n - 1)I_n - (\mathbb{1}_n \mathbb{1}_n^\mathsf{T} - I_n) = nI_n - \mathbb{1}_n \mathbb{1}_n^\mathsf{T} = n\,\Pi_n. \tag{6.1}$$

Note:

(i) the sign pattern of L is important $-$ diagonal elements are non-negative (zero or positive) and off-diagonal elements are non-positive (zero or negative);

(ii) the Laplacian matrix L of a digraph G does not depend upon the existence and values of self-loops in G;

(iii) the graph G is undirected (i.e., symmetric adjacency matrix) if and only if L is symmetric. In this case, $D_{\text{out}} = D_{\text{in}} = D$ and $A = A^{\mathsf{T}}$;

(iv) in a directed graph, $\ell_{ii} = 0$ (instead of $\ell_{ii} > 0$) if and only if node i has zero out-degree;

(v) L is said to be *irreducible* if G is strongly connected.

6.1.1 Useful equalities

We conclude this section with some useful equalities. We start with the obvious

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j. \tag{6.2}$$

First, for $x \in \mathbb{R}^n$,

$$(Lx)_i = \sum_{j=1}^{n} \ell_{ij} x_j = \ell_{ii} x_i + \sum_{j=1, j \neq i}^{n} \ell_{ij} x_j = \left(\sum_{j=1, j \neq i}^{n} a_{ij} \right) x_i + \sum_{j=1, j \neq i}^{n} (-a_{ij}) x_j$$

$$= \sum_{j=1, j \neq i}^{n} a_{ij}(x_i - x_j) = \sum_{j \in \mathcal{N}^{\text{out}}(i)} a_{ij}(x_i - x_j). \tag{6.3}$$

Additionally, if G has no self-loops and $d_{\text{out}}(i) = \sum_{j \in \mathcal{N}^{\text{out}}(i)} a_{ij} > 0$, then the values

$$\{ a_{ij}/d_{\text{out}}(i), \text{ for all out-neighbors } j \}$$

are convex combination coefficients defining a weighted average and

$$(Lx)_i = d_{\text{out}}(i)\big(x_i - \text{weighted-average}(\{x_j, \text{ for all out-neighbors } j\}) \big),$$

$$\stackrel{\text{for unit weights}}{=} d_{\text{out}}(i)\big(x_i - \text{average}(\{x_j, \text{ for all out-neighbors } j\}) \big).$$

Second, assume $L = L^\mathsf{T}$ (i.e., $a_{ij} = a_{ji}$) and compute:

$$
\begin{aligned}
x^\mathsf{T} L x &= \sum_{i=1}^{n} x_i (Lx)_i = \sum_{i=1}^{n} x_i \left(\sum_{j=1, j \neq i}^{n} a_{ij}(x_i - x_j) \right) \\
&= \sum_{i,j=1}^{n} a_{ij} x_i (x_i - x_j) = \left(\frac{1}{2} + \frac{1}{2} \right) \sum_{i,j=1}^{n} a_{ij} x_i^2 - \sum_{i,j=1}^{n} a_{ij} x_i x_j \\
&\overset{\text{by symmetry}}{=} \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} x_i^2 + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} x_j^2 - \sum_{i,j=1}^{n} a_{ij} x_i x_j \\
&= \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x_i - x_j)^2 = \sum_{\{i,j\} \in E} a_{ij}(x_i - x_j)^2.
\end{aligned}
\tag{6.4}
$$

The function $x \mapsto x^\mathsf{T} L x$ is sometimes referred to as the *Laplacian potential function*, because of the energy and power interpretation we present in the next two subsections. Note that the quadratic function $x \mapsto x^\mathsf{T} L x$ is a generalization of the quadratic disagreement function $V_{\mathsf{qd}}(x) = x^\mathsf{T} \Pi_n x$, defined in Section 5.3 when studying convergence to consensus. Specifically, for the complete undirected graph, equation (6.1) implies $x^\mathsf{T} L_{K_n} x = \frac{n}{2} V_{\mathsf{qd}}(x)$.

6.1.2 The Laplacian in mechanical networks of springs
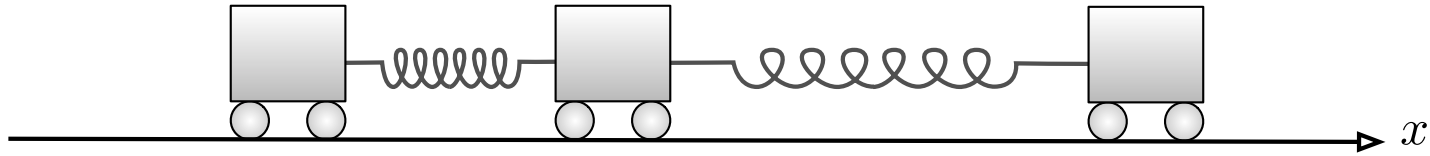
Figure 6.2: A spring network

A spring network is a collection of rigid bodies interconnected by springs. Let $x_i \in \mathbb{R}$ denote the displacement of the ith rigid body. Assume that each spring is ideal linear-elastic and let a_{ij} be the spring constant, i.e., the stiffness, of the spring connecting the ith and jth bodies.

   Define a graph as follows: the nodes are the rigid bodies $\{1, \ldots, n\}$ with locations $x_1, \ldots, x_n$, and the edges are the springs with weights a_{ij}. Each node i is subject to a force

$$\mathsf{F}_i = \sum_{j \neq i} a_{ij}(x_j - x_i) = -(L_{\text{stiffness}}x)_i,$$

where $L_{\text{stiffness}}$ is the Laplacian for the spring network (modeled as an undirected weighted graph). Moreover, recalling that the spring $\{i, j\}$ stores the quadratic energy $\frac{1}{2}a_{ij}(x_i - x_j)^2$, the total elastic energy is

$$\mathsf{E}_{\text{elastic}} = \frac{1}{2} \sum_{\{i,j\} \in E} a_{ij}(x_i - x_j)^2 = \frac{1}{2}x^{\mathsf{T}}L_{\text{stiffness}}x.$$

   In this role, the Laplacian matrix is referred to as the *stiffness matrix*. Stiffness matrices can be defined for spring networks in arbitrary dimensions (not only on the line) and with arbitrary topology (not only a chain graph, or line graph, as in figure). More complex spring networks can be found, for example, in finite-element discretization of flexible bodies and finite-difference discretization of diffusive media.

6.1.3   The Laplacian in electrical networks of resistors
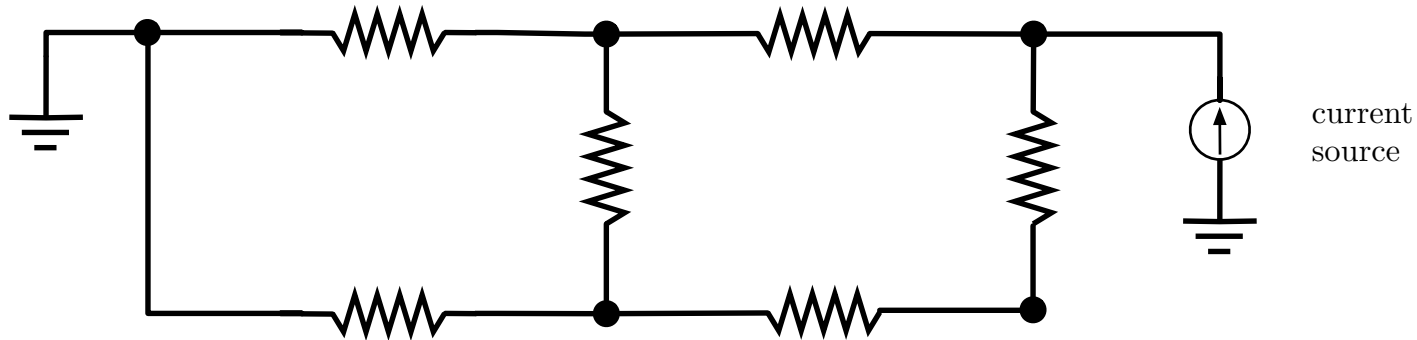
Figure 6.3: A resistive circuit

Suppose the graph is an *electrical network* with pure resistors and ideal voltage or current sources: (i) each graph node $i \in \{1, \ldots, n\}$ is possibly connected to an ideal voltage or current source or to ground, (ii) each edge is a resistor, say with resistance r_{ij} between nodes i and j. (This is an undirected weighted graph.)

Ohm's law along each edge $\{i, j\}$ gives the current flowing from i to j as

$$c_{i \to j} = (\mathsf{v}_i - \mathsf{v}_j)/r_{ij} = a_{ij}(\mathsf{v}_i - \mathsf{v}_j),$$

where v_i is the voltage at node i and a_{ij} is the inverse resistance, called conductance. We set $a_{ij} = 0$ whenever two nodes are not connected by a resistance and let $L_{\text{conductance}}$ denote the Laplacian matrix of conductances. Kirchhoff's current law says that at each node i:

$$c_{\text{injected at } i} = \sum_{j=1, j \neq i}^{n} c_{i \to j} = \sum_{j=1, j \neq i}^{n} a_{ij}(\mathsf{v}_i - \mathsf{v}_j).$$

Hence, the vector of injected currents c_{injected} and the vector of voltages at the nodes v satisfy

$$c_{\text{injected}} = L_{\text{conductance}} \, \mathsf{v}. \tag{6.5}$$

Moreover, the power dissipated on resistor $\{i, j\}$ is $c_{i \to j}(v_i - v_j)$, so that the total dissipated power is

$$P_{\text{dissipated}} = \sum_{\{i,j\} \in E} a_{ij}(v_i - v_j)^2 = v^\mathsf{T} L_{\text{conductance}} v.$$

6.2 Properties of the Laplacian matrix

6.2.1   Properties of the Laplacian matrix: L has zero row-sums

In this section we present various properties of Laplacian matrices.

Lemma 6.2 (Zero row-sums). *Let G be a weighted digraph with Laplacian L and n nodes. Then*

$$L\mathbb{1}_n = \mathbb{0}_n.$$

In equivalent words, 0 is an eigenvalue of L with right eigenvector $\mathbb{1}_n$. We refer to the eigenvalue 0 of L and to its left and right eigenvectors as *dominant*, since the eigenvalue 0 plays the same role for L as the eigenvalue 1 for a row-stochastic matrix.

Proof of Lemma 6.2. For all rows i, the ith row-sum is zero:

$$\sum_{j=1}^{n} \ell_{ij} = \ell_{ii} + \sum_{j=1,j\neq i}^{n} \ell_{ij} = \Big(\sum_{j=1,j\neq i}^{n} a_{ij} \Big) + \sum_{j=1,j\neq i}^{n} (-a_{ij}) = 0.$$

Equivalently, in vector format (remembering the weighted out-degree matrix D_{out} is diagonal and contains the row-sums of A):

$$L\mathbb{1}_n = D_{\text{out}}\mathbb{1}_n - A\mathbb{1}_n = \begin{bmatrix} d_{\text{out}}(1) \\ \vdots \\ d_{\text{out}}(n) \end{bmatrix} - \begin{bmatrix} d_{\text{out}}(1) \\ \vdots \\ d_{\text{out}}(n) \end{bmatrix} = \mathbb{0}_n.$$

∎

Based on this lemma, we now extend the notion of Laplacian matrix to a setting in which there is no digraph to start with.

Definition 6.3 (Laplacian matrix). *A matrix $L \in \mathbb{R}^{n \times n}$, $n \geq 2$, is Laplacian if*

(i) *its row-sums are zero,*

(ii) *its non-diagonal entries are non-positive, and*

(iii) *its diagonal entries are non-negative.*

Note: property (iii) is a consequence of (i) and (ii).

A Laplacian matrix L induces a weighted digraph G without self-loops in the natural way, that is, by letting (i, j) be an edge of G if and only if $\ell_{ij} < 0$. With this definition, L is the Laplacian matrix of G.

6.2.2 L has zero column-sums iff G is weight balanced

Next, we study when also the columns of a Laplacian matrix have vanishing sums.

Lemma 6.4 (Zero column-sums). *Let G be a weighted digraph with Laplacian L and n nodes. The following statements are equivalent:*

(i) G is weight-balanced, that is, $D_{\text{out}} = D_{\text{in}}$; and

(ii) $\mathbb{1}_n^\mathsf{T} L = \mathbb{0}_n^\mathsf{T}$.

Proof. Pick $j \in \{1, \dots, n\}$ and compute

$$(\mathbb{1}_n^\mathsf{T} L)_j = (L^\mathsf{T} \mathbb{1}_n)_j = \sum_{i=1}^n \ell_{ij} = \ell_{jj} + \sum_{i=1, j \neq i}^n \ell_{ij} = d_{\text{out}}(j) - d_{\text{in}}(j),$$

where the last equality follows from

$$\ell_{jj} = d_{\text{out}}(j) - a_{jj} \quad \text{and} \quad \sum_{i=1, j \neq i}^n \ell_{ij} = -(d_{\text{in}}(j) - a_{jj}).$$

In summary, we know that $\mathbb{1}_n^\mathsf{T} L = \mathbb{0}_n^\mathsf{T}$ if and only if $D_{\text{out}} = D_{\text{in}}$. ∎

6.2.3 Spectrum of L

Next, we study the eigenvalues of a Laplacian matrix.

Lemma 6.5 (Spectrum of the Laplacian matrix). *Given a weighted digraph G with Laplacian L, the eigenvalues of L different from 0 have strictly-positive real part.*

Proof. Recall $\ell_{ii} = \sum_{j=1, j\neq i}^{n} a_{ij} \geq 0$ and $\ell_{ij} = -a_{ij} \leq 0$ for $i \neq j$. By the Geršgorin Disks Theorem 2.8, we know that each eigenvalue of L belongs to at least one of the row disks

$$\left\{ z \in \mathbb{C} \mid |z - \ell_{ii}| \leq \sum_{j=1, j\neq i}^{n} |\ell_{ij}| \right\} = \left\{ z \in \mathbb{C} \mid |z - \ell_{ii}| \leq \ell_{ii} \right\}.$$
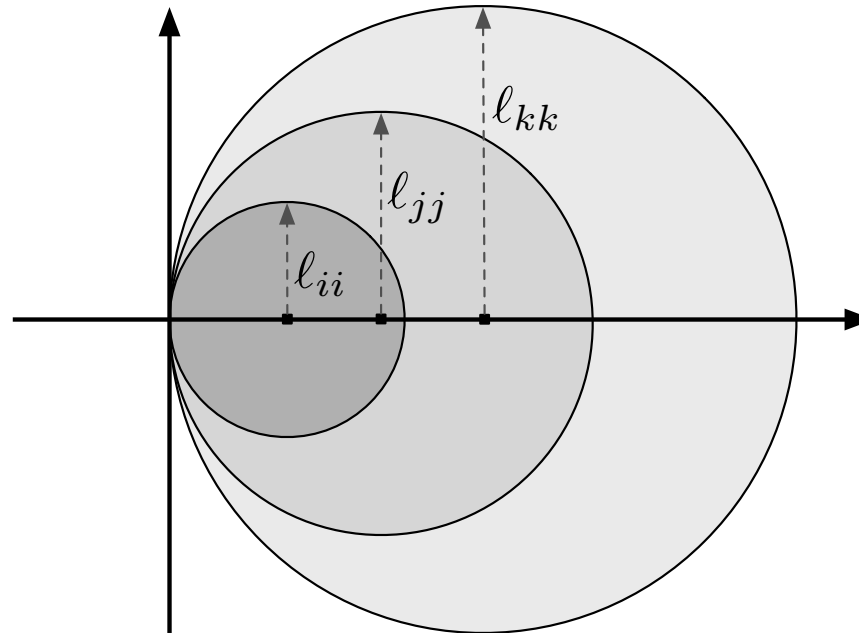
These disks, with radius equal to the center, contain the origin and complex numbers with positive real part. ∎

The following result is a consequence of Theorem 5.2.

Theorem 6.6 (Semisimplicity of the zero eigenvalue of a Laplacian matrix). *Let L be the Laplacian matrix of a weighted digraph G with n nodes. Let $n_s \geq 1$ denote the number of sinks in the condensation digraph of G. Then*

(i) the eigenvalue 0 is semisimple with multiplicity n_s,

(ii) the following statements are equivalent:

    (a) G contains a globally reachable node,

    (b) the eigenvalue 0 is simple, and

    (c) $\mathrm{rank}(L) = n - 1$.

Proof. First, we ask the reader to read and solve Exercise E6.1; with the notation in the exercise, we observe that if $C(G_L)$ has n_s sinks, so does $C(G_{\mathcal{A}_{L,\varepsilon}})$. Theorem 5.2 then implies that the eigenvalue 1 of $\mathcal{A}_{L,\varepsilon}$ is semisimple with multiplicity n_s and there exist left eigenvectors w^p, $p \in \{1, \ldots, n_s\}$, such that $(w^p)^\mathsf{T}\mathcal{A}_{L,\varepsilon} = (w^p)^\mathsf{T}$. But the same vectors therefore satisfy $(w^p)^\mathsf{T}L = \mathbb{0}_n^\mathsf{T}$, which implies that the eigenvalue 0 of L is semisimple with multiplicity n_s. $\blacksquare$

6.3   Symmetric Laplacian matrices and the algebraic connectivity

We now specialize the results in the previous section to the setting of a weighted directed graph G with a symmetric adjacency matrix $A = A^\mathsf{T} \in \mathbb{R}^{n \times n}$.

Note: a digraph G with symmetric Laplacian is an undirected graph possibly with self-loops. Therefore, the number of sinks of G is equal to the number of connected components of G.

6.3.1 Laplacian eigenvalues and algebraic connectivity

Assuming A and, therefore, L is symmetric, we know that all eigenvalues of L are real and that

(i) at least one is zero by Lemma 6.2, and

(ii) all eigenvalues are non-negative by Lemma 6.5.

Therefore, by convention, we write these eigenvalues as

$$0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

Definition 6.7 (Algebraic connectivity). *The second smallest eigenvalue λ_2 of a symmetric Laplacian L of a weighted digraph G is called the algebraic connectivity of G. The algebraic connectivity and its associated eigenvector are also referred to as the Fiedler eigenvalue and Fiedler eigenvector (in recognition of the early work by Fiedler (1973)).*

Theorem 6.6 directly implies the following simple results.

Corollary 6.8. *For a weighted undirected graph G with symmetric Laplacian L:*

(i) G is connected if and only if $\lambda_2 > 0$; and

(ii) the multiplicity of 0 as an eigenvalue of L is equal to the number of connected components of G.

Moreover, the algebraic connectivity has numerous properties and is related to numerous concepts in graph theory. We here present only a few selected results.

Lemma 6.9 (Properties of the algebraic connectivity). *Consider a weighted undirected graph with symmetric adjacency matrix A, symmetric Laplacian matrix L, and algebraic connectivity λ_2. The algebraic connectivity satisfies:*

(i) *the* variational description:

$$\lambda_2 = \min_{\|x\|_2=1, \, x\perp \mathbb{1}_n} x^\mathsf{T} L x, \tag{6.6}$$

(ii) *the* monotonicity property:

$$A \leq A' \quad \implies \quad \lambda_2 \leq \lambda_2',$$

where A' is a symmetric adjacency matrix with algebraic connectivity λ_2'.

The proof of these statements is postponed to Exercise E6.5. Statement (i) is a consequence of the Courant-Fisher Theorem (Meyer, 2001, Chapter 7), also called the Min-Max Theorem. Statement (ii) explains in what sense λ_2 is monotonic with respect to edge weights (a similar result for the spectral radius of an adjacency matrix is given in Lemma 4.10). Corollary 6.8(i) and Lemma 6.9(ii) together explain that λ_2 is a measure of how well-connected a graph is, thereby justifying the name algebraic connectivity.

We conclude this section with some basic examples.

Example 6.10 (Basic graphs and their algebraic connectivity). *Recall the definitions of path, cycle, star, complete and complete bipartite graph from Examples 3.1 and 4.1. We here report a table containing their algebraic connectivity and Laplacian spectrum and leave their proof to the reader in Exercise E6.9.*

Graph	Algebraic connectivity	Laplacian spectrum
path graph P_n	$2(1 - \cos(\pi/n)) \sim \pi^2/n^2$	$\{0\} \cup \{2(1 - \cos(\pi i/n)) \mid i \in \{1, \ldots, n-1\}\}$
cycle graph C_n	$2(1 - \cos(2\pi/n)) \sim 4\pi^2/n^2$	$\{0\} \cup \{2(1 - \cos(2\pi i/n)) \mid i \in \{1, \ldots, n-1\}\}$
star graph S_n	1	$\{0, 1, \ldots, 1, n\}$
complete graph K_n	n	$\{0, n, \ldots, n\}$
complete bipartite $K_{n,m}$	$\min(n, m)$	$\{0, m, \ldots, m, n, \ldots, n, m+n\}$, where m has multiplicity $n-1$ and n has multiplicity $m-1$

Table 6.1: The algebraic connectivity and Laplacian spectrum for basic graphs. Loosely speaking, the sparsely-connected graphs P_n and C_n have algebraic connectivity $\lambda_2 \sim 1/n^2$, the star graph S_n has constant algebraic connectivity, and the fully connected K_n has linearly growing algebraic connectivity.

•

6.3.2 Laplacian systems and Laplacian pseudoinverses

In this section we study Laplacian systems because of their rich structure and numerous applications. We start by revisiting the mechanical and electric example systems introduced in Sections 6.1.2 and 6.1.3. We are interested in understanding when equilibrium configurations exist and, if so, how to compute them.

(a) A spring network subject to a balanced force   (b) A resistive circuit subject to a balanced current injection
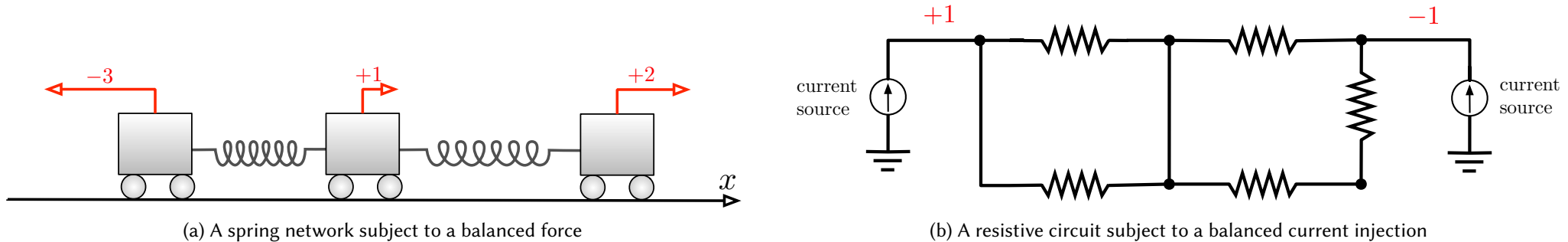
Figure 6.4: Laplacian systems

The spring network consists of n masses at positions $x \in \mathbb{R}^n$, described by a stiffness Laplacian matrix L, and subject to a load force f_{load} as in Figure 6.4a. Similarly, the resistive circuit consists of n nodes at voltage v and subject to a current injection c_{injected}. The force balance equation for the equilibrium displacement in the spring network and the flow balance equation for voltage equilibrium in the resistive circuit are

$$L_{\text{stiffness}} x = f_{\text{load}}, \quad \text{and} \quad L_{\text{conductance}} v = c_{\text{injected}}.$$

We formalize these two problems as follows.

Definition 6.11. *A Laplacian system is a linear system of equations in the variable $x \in \mathbb{R}^n$ of the form*

$$Lx = b, \tag{6.7}$$

where $L \in \mathbb{R}^{n \times n}$ is a Laplacian matrix and $b \in \mathbb{R}^n$.

To characterize the solutions to a Laplacian system we recall the notion of pseudoinverse matrix from Exercise E2.20 and present the following result.

Lemma 6.12 (The pseudoinverse Laplacian matrix). *Consider the symmetric Laplacian matrix L of a connected graph with decomposition $L = U \operatorname{diag}(0, \lambda_2, \ldots, \lambda_n)U^\mathsf{T}$, where $U \in \mathbb{R}^{n \times n}$ is orthonormal. Then*

(i) $\operatorname{image}(L) = \mathbb{1}_n^\perp$ so that the system $Lx = b$ admit solutions if and only if $b \perp \mathbb{1}_n$,

(ii) if $b \in \mathbb{R}^n$ is balanced, that is, $b \perp \mathbb{1}_n$, then the set of solutions to the Laplacian system is

$$\{L^\dagger b + \beta \mathbb{1}_n \mid \beta \in \mathbb{R}\},$$

(iii) the pseudoinverse of L is

$$L^\dagger = U \begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 1/\lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/\lambda_n \end{bmatrix} U^\mathsf{T} \tag{6.8}$$

and satisfies

$$L^\dagger = (L^\dagger)^\mathsf{T} \succeq 0, \quad L^\dagger \mathbb{1}_n = (L^\dagger)^\mathsf{T} \mathbb{1}_n = \mathbb{0}_n, \quad \text{and} \quad LL^\dagger = L^\dagger L = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T}.$$

In short, a Laplacian system is a static equilibrium problem with zero-mean (current / force) injections b that determine nodal (voltage / displacement) equilibrium variables $x^* = L^\dagger b$ uniquely up to a uniform displacement $\beta \mathbb{1}_n$.

We ask the reader to prove Lemma 6.12 in Exercise E6.10 and present here an example application.

Example 6.13 (Effective resistance). *Given a connected weighted undirected graph G with Laplacian L, regard G as a resistive circuit with conductances along the edges. The *effective resistance* between nodes i and j, denoted by r_{ij}^{eff}, is the potential difference induced between i and j when a unit of current is injected at i and extracted at j; see Figure 6.5*
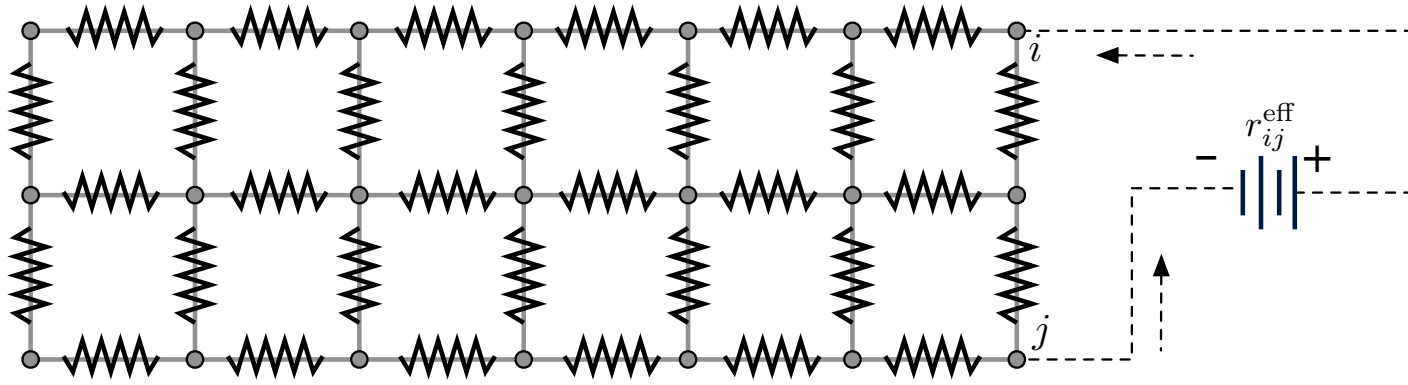
Figure 6.5: The effective resistance r_{ij}^{eff}

Simple calculations (starting with $c_{injected} = \mathbb{e}_i - \mathbb{e}_j$, where $\mathbb{e}_i$ is the ith base vector of $\mathbb{R}^n$) show that the effective resistance between i and j satisfies

$$r_{ij}^{\text{eff}} = (\mathbb{e}_i - \mathbb{e}_j)^{\mathsf{T}} L^{\dagger} (\mathbb{e}_i - \mathbb{e}_j). \tag{6.9}$$

Indeed, by definition, we have $r_{ij}^{\text{eff}} = v_i - v_j$. We combine $Lv = c_{injected}$ and $c_{injected} = \mathbb{e}_i - \mathbb{e}_j$ to obtain $v = L^{\dagger}(\mathbb{e}_i - \mathbb{e}_j)$. Therefore, $v_i - v_j = (\mathbb{e}_i - \mathbb{e}_j)^{\mathsf{T}} v = (\mathbb{e}_i - \mathbb{e}_j)^{\mathsf{T}} L^{\dagger}(\mathbb{e}_i - \mathbb{e}_j)$, which is equation (6.9).

The resistance distance is a distance function on the graph in the following sense. For all $i, j, k \in \{1, \ldots, n\}$,

(i) $r_{ij}^{\mathrm{eff}} \geq 0$ and $r_{ij}^{\mathrm{eff}} = 0$ if and only if $i = j$ (positive definiteness),

(ii) $r_{ij}^{\mathrm{eff}} = r_{ji}^{\mathrm{eff}}$ (symmetry), and

(iii) $r_{ij}^{\mathrm{eff}} \leq r_{ik}^{\mathrm{eff}} + r_{kj}^{\mathrm{eff}}$ (subadditivity).

Moreover, the resistance distance satisfies the so-called Rayleigh monotonicity property, i.e., with the same notation as in Lemma 6.9,

$$A \leq A' \quad \Longrightarrow \quad r_{ij}^{\mathrm{eff}} \geq (r_{ij}^{\mathrm{eff}})' \quad \text{for all } i, j.$$

We refer to (Klein and Randić, 1993) for additional properties and applications.   ●

6.4   Appendix: Community detection via algebraic connectivity

As just presented, the algebraic connectivity λ_2 of an undirected and weighted graph G is positive if and only if G is connected. We build on this insight and show that the algebraic connectivity does not only provide a binary connectivity measure, but it also quantifies the "bottleneck" of the graph. To develop this intuition, we study the problem of *community detection* in a large-scale undirected graph. This problem arises, for example, when identifying group of friends in a social network by means of the interaction graph.

Specifically, we consider the problem of partitioning the nodes V of an undirected connected graph G in two sets V_1 and V_2 so that

$$V_1 \cup V_2 = V, \ V_1 \cap V_2 = \emptyset, \ \text{ and } V_1, V_2 \neq \emptyset.$$

Of course, there are many such partitions. We measure the quality of a partition by the sum of the weights of all edges that need to be *cut* to separate the nodes V_1 and V_2 into two disconnected components. Formally, the *size of the cut* separating V_1 and V_2 is

$$J(V_1, V_2) = \sum_{i \in V_1, j \in V_2} a_{ij} \, .$$

We are interested in finding the cut with minimal size that identifies the two groups of nodes that are most loosely connected. The problem of minimizing the cut size J is combinatorial and computationally hard since we need to consider all possible partitions of the node set V. We present here a tractable approach based on a so-called relaxation step. First, define a vector $x \in \{-1, +1\}^n$ with entries $x_i = 1$ for $i \in V_1$ and $x_i = -1$ for $i \in V_2$. Then the cut size J is a function of x and can be rewritten via the Laplacian potential as

$$J(x) = \frac{1}{8} \sum_{i,j=1}^{n} a_{ij}(x_i - x_j)^2 = \frac{1}{4} x^{\mathsf{T}} L x$$

and the minimum cut size problem is:

$$\underset{x \in \{-1,1\}^n \setminus \{-\mathbb{1}_n, \mathbb{1}_n\}}{\text{minimize}} \quad x^{\mathsf{T}} L x.$$

(Here we exclude the cases $x \in \{-\mathbb{1}_n, \mathbb{1}_n\}$ because they correspond to one of the two sets being empty.) Second, since this problem is still computationally hard, we relax the problem from binary decision variables $x_i \in \{-1, +1\}$ to continuous decision variables

$y_i \in [-1, 1]$ (or $\|y\|_\infty \leq 1$), where we exclude $y \in \mathrm{span}(\mathbb{1}_n)$ (corresponding to one of the two groups being empty). Then the minimization problem becomes

$$\underset{y \in \mathbb{R}^n,\, y \perp \mathbb{1}_n,\, \|y\|_\infty = 1}{\text{minimize}} y^\top L y.$$

As a third and final step, we consider a 2-norm constraint $\|y\|_2 = 1$ instead of an ∞-norm constraint $\|y\|_\infty = 1$ (recall that $\|y\|_\infty \leq \|y\|_2 \leq \sqrt{n}\|y\|_\infty$) to obtain the following heuristic:

$$\underset{y \in \mathbb{R}^n,\, y \perp \mathbb{1}_n,\, \|y\|_2 = 1}{\text{minimize}} y^\top L y. \tag{6.10}$$

The variational description in Lemma 6.9(i) now states that the unique minimum of the relaxed optimization problem (6.10) is λ_2 and the minimizer is v_2, the eigenvector associated to λ_2 normalized to have unit 2-norm. We can then use as a heuristic $x = \mathrm{sign}(v_2)$ to find the desired partition $\{V_1, V_2\}$. Hence, the algebraic connectivity λ_2 (the Fiedler eigenvalue) is an estimate for the size of the minimum cut and the signs of the entries of the Fiedler eigenvector v_2 identify the associated partition in the graph. In this sense, Fiedler eigenvalue and eigenvector describe size and location of the "bottleneck" of a graph.

To illustrate these concepts, we borrow an example computational problem from (Gleich, 2006). We construct a randomly generated graph as follows. First, we partition $n = 1000$ nodes in two groups V_1 and V_2 of sizes 450 and 550 nodes, respectively. Second, we connect any pair of nodes in the set V_1 (respectively V_2) with probability 50% (respectively 40%). Third and finally, any two nodes in distinct groups, $i \in V_1$ and $j \in V_2$, are connected with a probability of 15%. The sparsity pattern of the associated adjacency matrix is shown in the left panel of Figure 6.6. No obvious partition is visible at first glance since the indices are not necessarily sorted, that is, V_1 is not necessarily $\{1, \ldots, 450\}$. The second panel displays the sorted entries of the eigenvector v_2 showing a sharp transition between positive and negative entries. Finally, the third panel displays the correspondingly sorted adjacency matrix $\tilde{A}$ clearly indicating the partition $V = V_1 \cup V_2$. The Python code to generate Figure 6.6 is in the Table 6.2 below.
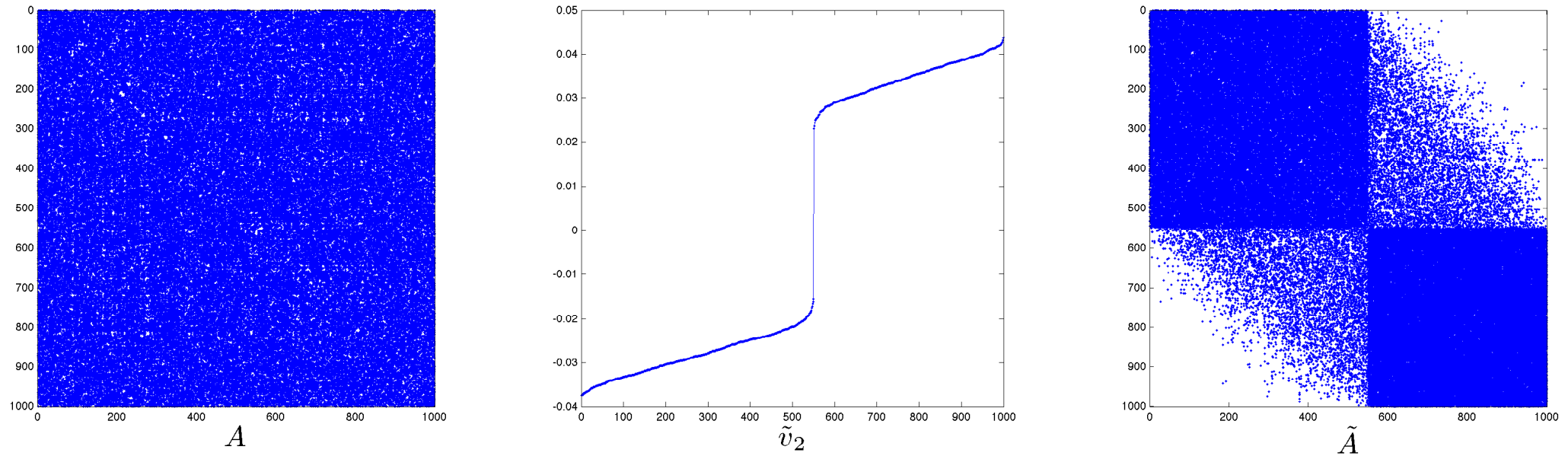
Figure 6.6: The left figure shows a randomly-generated sparse adjacency matrix A for a graph with 1000 nodes. The central figure displays the eigenvector $\tilde{v}_2$ which is identical to the normalized eigenvector v_2 after sorting the entries according to their magnitude, and the right figure displays the correspondingly sorted adjacency matrix $\tilde{A}$. For additional analysis of this problem, we refer the reader to (Gleich, 2006).

```python
 1  # Python3 code for community detection
 2  import numpy as np; import scipy.sparse.linalg as sla;
 3  import matplotlib.pyplot as plt; from matplotlib.colors import ListedColormap
 4  # Draw binary plot of adjacency matrix
 5  def plot_matrix_binary(M, ax, name=''):
 6      blue_map = ListedColormap(["blue", "white"]);     zeros = M == 0;
 7      im = ax.imshow(zeros, cmap=blue_map);     ax.set_xticks([]);     ax.set_yticks([])
 8
 9  # For a given graph size, randomly assign the nodes to two groups
10  n = 1000; group_size = 450; A = np.zeros([1000, 1000]);
11  x = np.random.permutation(n) - 1  # Random permutation of indices for groups
12  group1 = x[0:group_size]; group2 = x[group_size:];
13  # Assign probabilities of connecting nodes
14  p_group1 = 0.5; p_group2 = 0.4; p_between_groups = 0.15
15  # Construct adjacency matrix
16  A[np.ix_(group1,group1)] = (np.random.rand(group_size, group_size) < p_group1) * 1  # Ensure cast to integer
17  A[np.ix_(group2,group2)] = (np.random.rand(n-group_size,n-group_size) < p_group2) * 1
18  A[np.ix_(group1,group2)] = (np.random.rand(group_size, n-group_size) < p_between_groups) * 1
19  # Ensure symmetry by copying the just created upper triangle part
20  A = np.triu(A,1); A = A + A.T;
21
22  # Construct Laplacian, Fiedler eigenpair, and sorting indices
23  L = np.diag(np.sum(A, 1)) - A
24  D, V = sla.eigs(L, 2, which='SM')
25  V_sort_ind = np.argsort(V[:, 1])
26
27  # Init plot
28  custom_figsize= (6, 4); fig, axs64 = plt.subplots(3, 1, figsize=(custom_figsize[0]*1.2, custom_figsize[1]*3))
29  # Plot binary matrix
30  plot_matrix_binary(A, axs64[0]);  axs64[0].set_xlabel("$A$")
31  # Plot the eigenvector values sorted by magnitude
32  axs64[1].plot(np.sort(V[:, 1]))
33  axs64[1].set_aspect(1 / axs64[1].get_data_ratio())  # Workaround to make it square without equal axis ticks
34  axs64[1].set_xlabel(r"$\tilde{v}_2$")
35  # Plot the adjacency matrix sorted by the eigenvector
36  plot_matrix_binary(A[np.ix_(V_sort_ind,V_sort_ind)], axs64[2]); axs64[2].set_xlabel(r"$\widetilde{A}$");
```

Table 6.2: Python code for community detection

6.5 Appendix: Control design for clock synchronization

In this section we consider an idealized network of heterogeneous clocks and design a control strategy to ensure they achieve synchronization.

Consider n simplified clocks modeled as discrete-time integrators: $x_i(k+1) = x_i(k) + d_i$. The initial value $x_i(0)$ is called the *initial offset* and d_i is called the *clock speed* (or skew); see Figure 6.7. Assume that we can control each clock according to

$$x(k+1) = x(k) + d + u(k). \tag{6.11}$$

Define the *average clock speed* by $d_{\mathrm{ave}} = \mathrm{average}(d) = \mathbb{1}_n^\mathsf{T} d/n$ and the *average time* by $t_{\mathrm{ave}}(k) = d_{\mathrm{ave}}k + x_{\mathrm{ave}}(0)$.
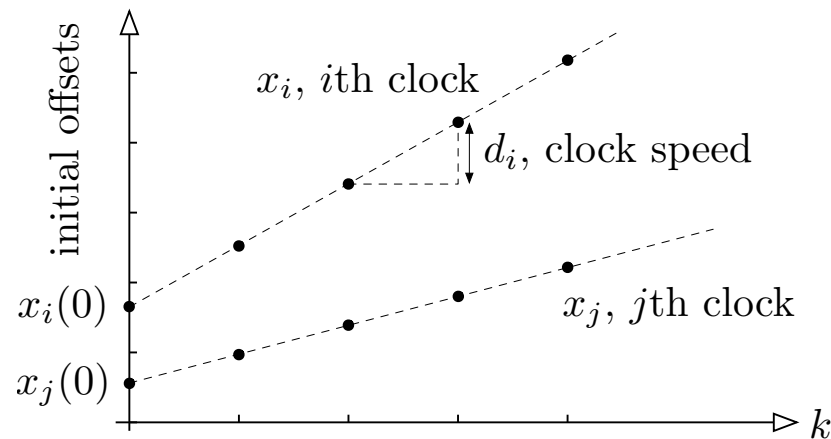
Figure 6.7: Two clocks with different initial offset $x_i(0) \neq x_j(0)$ and speeds $d_i \neq d_j$.

The *clock synchronization problem* is to design a control law u such that, for all clocks i and j,

$$\lim_{k \to \infty} x_i(k) - x_j(k) = 0.$$

Averaging-based proportional control   Suppose the clocks are interconnected by an connected undirected graph so that each node i can measure the errors $(x_j(k) - x_i(k))$ for some neighbors j. For each edge $\{i, j\}$, let $\kappa_{ij} = \kappa_{ji} > 0$ be a control gain (and set $\kappa_{pq} = 0$ whenever $\{p, q\}$ is not an edge), and select the *averaging-based proportional control law*

$$x_i(k+1) = x_i(k) + d_i + \sum_{j=1}^{n} \kappa_{ij}(x_j(k) - x_i(k)).$$

To analyze this control design, we proceed as follows. First, if $L = L^\mathsf{T}$ denotes the Laplacian matrix defined by these control gains, then the control is $u(k) = -Lx(k)$ and the closed-loop system is

$$x(k+1) = (I_n - L)x(k) + d.$$

For $\max_{i \in \{1,\ldots,n\}} \sum_{j=1, j \neq i}^{n} \kappa_{ij} < 1$, the matrix $I_n - L$ is non-negative and therefore row-stochastic.

Note: we now see that the closed-loop system is an averaging system with a forcing term; this is the reason we call this control action proportional/averaging.

Second, recall the average time $t_{\text{ave}}(k)$ and define the *error system* $y(k) = x(k) - t_{\text{ave}}(k)\mathbb{1}_n$. One can show that

$$y(k+1) = (I_n - L)y(k) + (d - d_{\text{ave}}\mathbb{1}_n),$$

and that this system is precisely an affine averaging system as studied in Exercise E2.10. According to Exercise E2.10(iii), we know that, generically, $y(k) \to y_{\text{final}} \notin \text{span}\{\mathbb{1}_n\}$ so that

$$\lim_{k \to \infty} x_i(k) - x_j(k) = \lim_{k \to \infty} y_i(k) - y_j(k) \neq 0.$$

In other words, proportional control keeps the errors bounded (they would naturally diverge without it), but does not achieve vanishing errors and therefore does not solve the clocks synchronization problem.

Proportional/averaging and integral control   We now introduce a so-called integrator state w_i at each node, pick an *integral control gain* γ, and design the *averaging-based proportional-integral control law* as

$$u(k) = -Lx(k) - w(k),$$
$$w(k+1) = w(k) + \gamma Lx(k),$$

so that the closed-loop system dynamics is

$$x(k+1) = (I_n - L)x(k) - w(k) + d,$$
$$w(k+1) = w(k) + \gamma Lx(k), \tag{6.12}$$

with corresponding system matrix in block form $\begin{bmatrix} I_n - L & -I_n \\ \gamma L & I_n \end{bmatrix}$.

The rationale for integral control is that, when in steady state with $w(k+1) = w(k)$, the integral equation in (6.12) enforces $\mathbb{0}_n = Lx(k)$. Hence, if the closed loop (6.12) admits a steady state, then necessarily all clocks must be synchronized. It is natural to assume a zero initial state for the initial integral state $w(0) = \mathbb{0}_n$.

Lemma 6.14 (Asymptotic clock synchronization). *Consider n clocks (6.11) with heterogeneous initial offsets $x_i(0)$, speeds d_i, average speed $d_{\text{ave}} = \text{average}(d)$, and average time $t_{\text{ave}}(k) = d_{\text{ave}}k + x_{\text{ave}}(0)$. Assume the undirected communication graph among them is connected. Select proportional/averaging gains κ_{ij} for all edges $\{i, j\}$ and an integral control gain γ satisfying*

$$\max_{i \in \{1,\dots,n\}} \sum_{j=1}^n \kappa_{ij} < 1, \quad and \quad 0 < \gamma < 1. \tag{6.13}$$

Then the proportional/averaging integral control ensures that, in the closed loop, the clocks synchronize and

$$\lim_{k \to \infty} \left(x(k) - t_{\text{ave}}(k)\mathbb{1}_n \right) = \mathbb{0}_n.$$

In other words, the clocks asymptotically synchronizes and their time grows linearly with a speed equal to the average clock speed.

Proof. We start by studying the evolution of the affine dynamical system (6.12) using the modal decomposition as illustrated in Section 2.1. Being a symmetric Laplacian matrix, L has real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ with corresponding orthonormal eigenvectors $v_1 = \mathbb{1}_n/\sqrt{n}, v_2, \ldots, v_n$. By left-multiplying the closed-loop system dynamics (6.12) by $v_\alpha^\mathsf{T}, \alpha \in \{1, \ldots, n\}$, we obtain the following n decoupled 2-dimensional systems:

$$\begin{bmatrix} x_\alpha(k+1) \\ w_\alpha(k+1) \end{bmatrix} = \begin{bmatrix} 1 - \lambda_\alpha & -1 \\ \gamma\lambda_\alpha & 1 \end{bmatrix} \begin{bmatrix} x_\alpha(k) \\ w_\alpha(k) \end{bmatrix} + \begin{bmatrix} d_\alpha \\ 0 \end{bmatrix}, \quad \alpha \in \{1, \ldots, n\}, \tag{6.14}$$

where $x_\alpha(k) = v_\alpha^\mathsf{T} x(k)$, $w_\alpha(k) = v_\alpha^\mathsf{T} w(k)$, and $d_\alpha = v_\alpha^\mathsf{T} d$. From this decomposition, the full state can be reconstructed by

$$x(k) = \sum_{\alpha=1}^n x_\alpha(k) v_\alpha = x_{\mathrm{ave}}(k)\mathbb{1}_n + \sum_{\alpha=2}^n x_\alpha(k) v_\alpha,$$

$$w(k) = \sum_{\alpha=1}^n w_\alpha(k) v_\alpha = w_{\mathrm{ave}}(k)\mathbb{1}_n + \sum_{\alpha=2}^n w_\alpha(k) v_\alpha.$$

where $x_{\mathrm{ave}}(k) = \mathrm{average}(x(k))$ and $w_{\mathrm{ave}}(k) = \mathrm{average}(w(k))$.

For $\alpha = 1$, after a simple rescaling, equation (6.14) reads

$$\begin{bmatrix} x_{\mathrm{ave}}(k+1) \\ w_{\mathrm{ave}}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\mathrm{ave}}(k) \\ w_{\mathrm{ave}}(k) \end{bmatrix} + \begin{bmatrix} d_{\mathrm{ave}} \\ 0 \end{bmatrix}.$$

Because $w(0) = \mathbb{0}_n$, we compute $w(k) = \mathbb{0}_n$ and $x_{\mathrm{ave}}(k) = t_{\mathrm{ave}}(k) = d_{\mathrm{ave}}k + x_{\mathrm{ave}}(0)$.

It now suffices to show that the solutions to the $n-1$ equations (6.14), for $\alpha \in \{2, \ldots, n\}$, satisfy $\lim_{k\to\infty} x_\alpha(k) = 0$. Simple calculations show that the only equilibrium solutions to the $n-1$ equations (6.14), for $\alpha \in \{2, \ldots, n\}$, are $x_\alpha^* = 0$ and $w_\alpha^* = -d_\alpha$. Hence, it suffices to show that all eigenvalues of the $n-1$ matrices of dimension 2×2 have magnitude strictly less than 1. For $\alpha \in \{2, \ldots, n\}$, the $n-1$ characteristic equations are

$$(z-1)^2 + \lambda_\alpha(z - 1 + \gamma) = 0.$$

We claim that these polynomials have both roots strictly inside the unit circle if and only if, for all $\alpha \in \{2, \ldots, n\}$,

$$0 < \gamma < 1, \quad \text{and} \quad 0 < \lambda_\alpha < 4/(2 - \gamma). \tag{6.15}$$

Recall from the proof of, and the discussion following, Lemma 6.5 that

$$\lambda_i \le \lambda_n < 2 \max_{i \in \{1, \ldots, n\}} \sum_{j=1}^{n} \kappa_{ij}.$$

But by the assumption (6.13) we know $\max_{i \in \{1, \ldots, n\}} \sum_{j=1}^{n} \kappa_{ij} < 1$, hence $\lambda_n < 2 \times 1 < 4/(2 - \gamma)$ for all $0 < \gamma < 1$. Hence, the inequalities (6.15) are satisfied.

To verify that the inequalities (6.15) imply that all roots have magnitude less than 1, we use the so-called bilinear transform method. This method is based on the equivalence between the following two properties: the original polynomial has roots strictly inside the unit disk and the transformed polynomial has roots with strictly negative real part. We proceed as follows: we take $z = (1 + s)/(1 - s)$ and substitute it into the polynomial $(z - 1)^2 + \lambda_\alpha(z - 1 + \gamma)$ so that, removing the denominator, we obtain the polynomial $(4 - 2\lambda_\alpha + \lambda_\alpha\gamma)s^2 - \lambda_\alpha(2\gamma - 2)s + \lambda_\alpha\gamma$. By the Routh-Hurwitz stability criterion, this polynomial has roots with negative real part if and only if all three coefficients are strictly positive or strictly negative. Some elementary calculations show that all three coefficients may never be negative and that all three coefficients are positive if and only if the inequalities (6.15) hold. ∎

6.6   Historical notes and further reading

Standard books on algebraic graph theory with extensive characterizations of adjacency and Laplacian matrices include (Biggs, 1994) and (Godsil and Royle, 2001). Laplacian matrices and their algebraic connectivity are surveyed by (Mohar, 1991; Merris, 1994; Maia de Abreu, 2007). Laplacian systems are discussed in (Vishnoi, 2013).

The rank of the Laplacian, as characterized in Theorem 6.6, was studied as early as in (Fife, 1972; Foster and Jacquez, 1975). A mathematical approach is given in (Agaev and Chebotarev, 2000) which features the first necessary and sufficient characterization. We also refer to the more recent (Lin et al., 2005; Ren and Beard, 2005) for the specific case of $\mathrm{rank}(L) = n - 1$.

The generalized inverse of the Laplacian matrix appears in some applications and is studied by Gutman and Xiao (2004). An informative overview is given by Dörfler et al. (2018).

The ground-breaking work in (Fiedler, 1973) established the use of the eigenvalues of the Laplacian matrix for example as a way to quantify graph connectivity and to perform clustering, as illustrated in Section 6.4. For surveys on community detection we refer to (Porter et al., 2009; Fortunato, 2010).

The example on clock synchronization via proportional/averaging and integral control in Section 6.5 is taken from (Carli et al., 2008a). More realistic settings are studied in (Schenato and Fiorentin, 2011; Carli and Zampieri, 2014; Mallada et al., 2015). Surveys include (Sundararaman et al., 2005; Sivrikaya and Yener, 2004; Simeone et al., 2008).

Complex-valued graphs, adjacency and Laplacian matrices are studied in (Reff, 2012); see also (Lin et al., 2013; Dong and Qiu, 2014) for some related applications.

6.7   Exercises

▶ E6.1   **Row-stochastic matrices associated to a Laplacian.** Let $L \in \mathbb{R}^{n \times n}$ be a Laplacian matrix and define $\ell_{\max} = \max_{i \in \{1,\ldots,n\}} \ell_{ii}$. Pick $\varepsilon < \frac{1}{\ell_{\max}}$ and define the ε-scaled matrix associated to L by

$$\mathcal{A}_{L,\varepsilon} = I_n - \varepsilon L. \tag{E6.1}$$

Let G_L be the weighted digraph without self-loops associated to L and $G_{\mathcal{A}_{L,\varepsilon}}$ be the weighted digraph associated to $\mathcal{A}_{L,\varepsilon}$. Show that:

(i) $\mathcal{A}_{L,\varepsilon}$ is row-stochastic with a strictly positive diagonal,
(ii) $G_{\mathcal{A}_{L,\varepsilon}}$ has all edges in G_L and self-loops at each node,
(iii) $\mathcal{A}_{L,\varepsilon}$ is doubly-stochastic if and only if G_L is weight-balanced,
(iv) $\mathcal{A}_{L,\varepsilon}$ is primitive if and only if G_L is strongly connected,
(v) (λ_A, v) is a right eigenpair for $\mathcal{A}_{L,\varepsilon}$ if and only if $\big((1 - \lambda_A)/\varepsilon, v\big)$ is a right eigenpair for L.

Note: The matrix $\mathcal{A}_{L,\varepsilon}$ corresponds to the Euler discretization with step-size ε of the continuous-time Laplacian flow $\dot{x} = -Lx$, introduced in Section 1.3 and studied in the next chapter.

▶ E6.2   **Example Laplacian spectra.** Let $G^\star$ be a graph with 8 nodes and with Laplacian matrix $L(G^\star) \in \mathbb{R}^{8 \times 8}$. For i $= \sqrt{-1}$, assume the spectrum of $L(G^\star)$ is

$$\operatorname{spec}(L(G^\star)) = \{0, 0, 0.5104, 1.6301, 2, 2.2045 - 1.0038\mathrm{i}, 2.2045 + 1.0038\mathrm{i}, 2.8646\}\,.$$

Consider the graphs G_1, G_2, and G_3 shown below. Argue why the following statements are true:

  (i) G_1 cannot be $G^\star$,
 (ii) G_2 cannot be $G^\star$, and
(iii) G_3 cannot be $G^\star$.
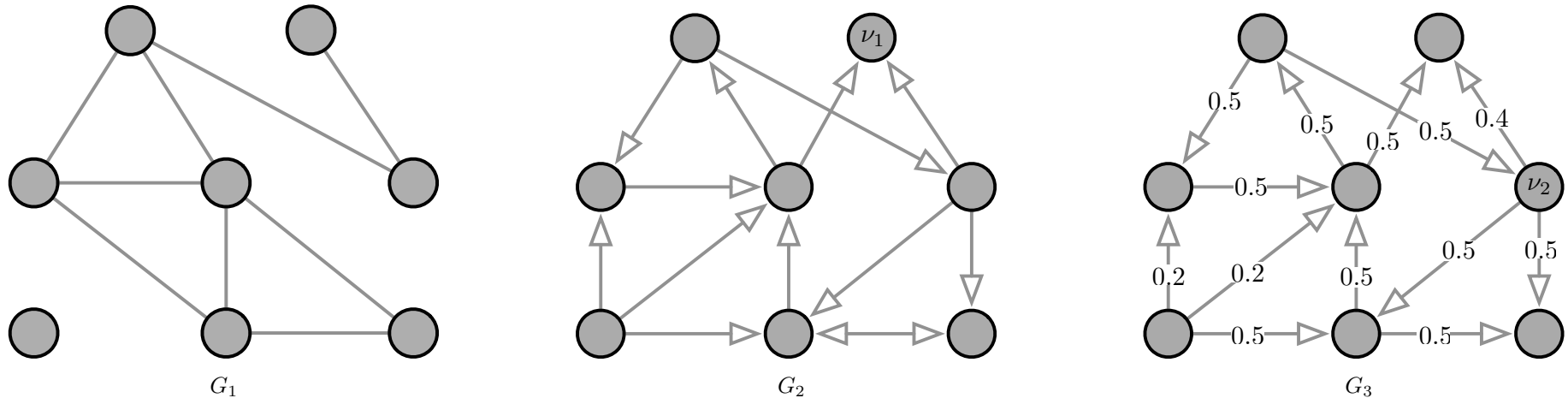
Figure E6.1: Example graphs and digraphs with 8 noses

E6.3   **A symmetric Laplacian matrix is positive semidefinite.** Let G be a weighted undirected graph with symmetric Laplacian matrix $L \in \mathbb{R}^{n \times n}$. Assume G is connected and let (λ_2, v) denote the Fiedler eigenpair. Let $\Pi_n = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T}$. Show that

(i) $L \succeq 0$ without relying on the Geršgorin Disks Theorem 2.8 and Lemma 6.5,

(ii) the Fiedler eigenvector satisfies $v \perp \mathbb{1}_n$ and $v^\mathsf{T} L v = \lambda_2 \|v\|_2^2$,

(iii) for any $x \in \mathbb{R}^n$ with $x_{\text{ave}} = \mathbb{1}_n^\mathsf{T} x / n$,

$$x^\mathsf{T} L x \geq \lambda_2 \|x - x_{\text{ave}} \mathbb{1}_n\|_2^2,$$

with equality if x is parallel to v.

(iv) $\Pi_n = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T}$ is the orthogonal projection onto $\mathbb{1}_n^\perp$ and $n\Pi_n$ is the Laplacian of the complete undirected graph, so that $\Pi_n = \Pi_n^\mathsf{T} \succeq 0$, and

(v) $\Pi_n L = L \Pi_n = L$ and $L \succeq \lambda_2 \Pi_n$.

Note: Statement (v) implies that L satisfies the so-called Lyapunov linear matrix inequality (LMI) $\Pi_n L + L^\mathsf{T} \Pi_n \succeq 2\lambda_2 \Pi_n$.

E6.4   **The Laplacian matrix of a weight-balanced digraph.** Let G be a weighted digraph with Laplacian matrix L. Prove the following statements are equivalent:

  (i) G is weight-balanced,

  (ii) $L + L^\mathsf{T}$ is the Laplacian matrix of the undirected digraph associated to the adjacency matrix $A + A^\mathsf{T}$.

  Moreover, with the notation in Exercise , show that

  (iii) if G is weight-balanced, then L satisfies the Lyapunov LMI $\Pi_n L + L^\mathsf{T}\Pi_n \succeq \lambda_2(L + L^\mathsf{T})\Pi_n$, and

  (iv) if additionally G is weakly connected, then $\lambda_2(L + L^\mathsf{T}) > 0$.

E6.5   **The algebraic connectivity is monotonic with respect to edge weights.** Show Lemma 6.9.

***Note:** To establish the variational characterization in equation (6.6), recall from (Meyer, 2001, Chapter 7) the Courant-Fisher minimax characterization of the eigenvalues of a symmetric matrix $P = P^\mathsf{T} \in \mathbb{R}^{n \times n}$:*

$$\lambda_k = \min_{S \in \mathcal{S}_k} \max_{x \in S, \|x\|=1} x^\mathsf{T} P x, \tag{E6.2}$$

where $\mathcal{S}_k$ is the set of k-dimensional vector subspaces of $\mathbb{R}^n$.

E6.6   **Upper and lower bound on largest Laplacian eigenvalue.** Let G be an undirected graph with symmetric Laplacian matrix $L = L^\mathsf{T} \in \mathbb{R}^{n \times n}$, Laplacian eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, and maximum degree $d_{\max} = \max_{i \in \{1,\ldots,n\}} d_i$. Show that the maximum eigenvalue λ_n satisfies:

$$d_{\max} \leq \lambda_n \leq 2d_{\max}.$$

Hint: Review the proof of Lemmas 6.5 and 6.9.

Note: Several other bounds are reviewed in (Maia de Abreu, 2007). For example, for an unweighted undirected graph with n nodes and minimum and maximum degree $d_{\min}, d_{\max}$, it is known that

$$2d_{\min} - n + 2 \leq \lambda_2 \leq \frac{n}{n-1} d_{\max}.$$

E6.7   **The Laplacian potential function in a directed graph (Gao et al., 2008).** Recall that the *quadratic form* associated with a symmetric matrix $B \in \mathbb{R}^{n \times n}$ is the function $x \mapsto x^{\mathsf{T}} B x$. Let G be a weighted digraph with n nodes and define the *Laplacian potential function* $\Phi_G \colon \mathbb{R}^n \to \mathbb{R}$ by

$$\Phi_G(x) = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x_j - x_i)^2.$$

Show that:

(i) Φ_G is the quadratic form associated with the symmetric positive-semidefinite matrix

$$P = \frac{1}{2}(D_{\text{out}} + D_{\text{in}} - A - A^{\mathsf{T}}),$$

(ii) $P = \frac{1}{2}\big(L + L(\mathrm{rev})\big)$, where the Laplacian of the reverse digraph is $L(\mathrm{rev}) = D_{\text{in}} - A^{\mathsf{T}}$.

E6.8   **Scaled Laplacian matrices.** Let $L = L^{\mathsf{T}} \in \mathbb{R}^{n \times n}$ be the Laplacian matrix of a connected, undirected, and symmetrically weighted graph. Given scalars $d_1, \ldots, d_n$, define the matrices A and B by

$$A := \mathrm{diag}\{d_1, \ldots, d_n\}L \quad \text{and} \quad B := L\,\mathrm{diag}\{d_1, \ldots, d_n\}.$$

(i) Give necessary and sufficient conditions on $\{d_1, \ldots, d_n\}$ for A to be a Laplacian matrix.
(ii) Give necessary and sufficient conditions on $\{d_1, \ldots, d_n\}$ for B to be a Laplacian matrix.
(iii) Give a sufficient condition on $\{d_1, \ldots, d_n\}$ for A and B to be symmetric.
(iv) Assuming $d_i \neq 0$, $i \in \{1, \ldots, n\}$, do A and B possess a zero eigenvalue? If so, what are the corresponding right and left eigenvectors for A and B?

E6.9   **Laplacian spectrum of basic graphs.** Given the basic graphs in Examples 3.1 and 4.1 and the properties of tridiagonal Toeplitz and circulant matrices in Exercises E4.2 and E4.3, compute the spectrum of the Laplacian matrix (and therefore also prove the statements in Table 6.1) for basic graphs. Specifically, show that, for $n \geq 2$,

    (i) for the *path graph* P_n, the Laplacian spectrum is $\{0\} \cup \{2(1 - \cos(\pi i/n)) \mid i \in \{1, \ldots, n-1\}\}$;

    (ii) for the *cycle graph* C_n, the Laplacian spectrum is $\{0\} \cup \{2(1 - \cos(2\pi i/n)) \mid i \in \{1, \ldots, n-1\}\}$;

    (iii) for the *star graph* S_n, the Laplacian spectrum is $\{0, 1, \ldots, 1, n\}$ where 1 has multiplicity $n-2$;

    (iv) for the *complete graph* K_n, the Laplacian spectrum is $\{0, n, \ldots, n\}$; and

    (v) for the *complete bipartite graph* $K_{n,m}$, the Laplacian spectrum is $\{0, m, \ldots, m, n, \ldots, n, m+n\}$, where m has multiplicity $n-1$ and n has multiplicity $m-1$.

E6.10   **The pseudoinverse Laplacian matrix.** Prove Lemma 6.12.

E6.11   **The regularized Laplacian matrix.** Let L be the Laplacian matrix of a weighted connected undirected graph with n nodes. Given a scalar $\beta \in \mathbb{R}$, define the *regularized Laplacian matrix* $L_{\mathrm{reg},\beta} = L + \frac{\beta}{n}\mathbb{1}_n\mathbb{1}_n^T$. Show that

  (i) $L_{\mathrm{reg},\beta}$ is nonsingular for $\beta \neq 0$,

  (ii) $L_{\mathrm{reg},\beta}$ is positive definite for $\beta > 0$, and

  (iii) the inverse of $L_{\mathrm{reg},\beta}$ satisfies

$$L_{\mathrm{reg},\beta}^{-1} = \left(L + \frac{\beta}{n}\mathbb{1}_n\mathbb{1}_n^T\right)^{-1} = L^\dagger + \frac{1}{\beta\,n}\mathbb{1}_n\mathbb{1}_n^T.$$

E6.12 **The Green matrix of a Laplacian matrix.** Assume L is the Laplacian matrix of a weighted connected undirected graph with n nodes. Show that

(i) the matrix $L + \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T}$ is positive definite,

(ii) the so-called *Green matrix*

$$X = \left(L + \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T}\right)^{-1} - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T} \tag{E6.3}$$

is the unique solution to the system of equations:

$$\begin{cases} LX = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T}, \\ \mathbb{1}_n^\mathsf{T} X = \mathbb{0}_n^\mathsf{T}, \end{cases}$$

(iii) $X = L^\dagger$. In other words, the Green matrix formula (E6.3) is an alternative definition of the Laplacian pseudoinverse.

E6.13   **Laplacian systems, Gaussian elimination and Kron reduction.** Consider an undirected and connected graph and its associated Laplacian matrix $L \in \mathbb{R}^{n \times n}$. Consider the associated Laplacian system $y = Lx$, where $x \in \mathbb{R}^n$ is unknown and $y \in \mathbb{R}^n$ is a given vector. Verify that an elimination of x_n from the last row of this equation yields the following reduced set of equations:

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} + \underbrace{\begin{bmatrix} -L_{1n}/L_{nn} \\ \vdots \\ -L_{n-1,n}/L_{nn} \end{bmatrix}}_{=A} y_n = \underbrace{\begin{bmatrix} \ddots & \vdots & \cdot^{\cdot^{\cdot}} \\ \cdots & L_{ij} - \frac{L_{in} \cdot L_{jn}}{L_{nn}} & \cdots \\ \cdot_{\cdot_{\cdot}} & \vdots & \ddots \end{bmatrix}}_{=L_{\mathrm{red}}} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \end{bmatrix},
$$

where the (i, j)-element of L_{red} is given by $L_{ij} - L_{in} \cdot L_{jn}/L_{nn}$. Show that the matrices $A \in \mathbb{R}^{n-1 \times 1}$ and $L \in \mathbb{R}^{(n-1) \times (n-1)}$ obtained after Gaussian elimination have the following properties:

   (i) A is non-negative and column-stochastic matrix with at least one strictly positive element; and

  (ii) L_{red} is a symmetric and irreducible Laplacian matrix.

Hint: To show the irreducibility of L_{red}, verify the following property regarding the *fill-in* of the matrix L_{red}: The graph associated to the Laplacian L_{red} has an edge between nodes i and j if and only if (i) either $\{i, j\}$ was an edge in the original graph associated to L, (ii) or $\{i, n\}$ and $\{j, n\}$ were edges in the original graph associated to L.

Note: *The matrix L_{red} is called the Kron reduction of L with respect to node n. The properties of this reduction process are discussed in (Dörfler and Bullo, 2013).*

E6.14   **Thomson's Principle and current flows.** Consider a connected and undirected resistive electrical network with n nodes, with external nodal current injections $c \in \mathbb{R}^n$ satisfying the balance condition $\mathbb{1}_n^{\mathsf{T}} c = 0$, and with resistances $r_{ij} > 0$ for every undirected edge $\{i, j\} \in E$. For simplicity, we set $r_{ij} = \infty$ if there is no edge connecting i and j. As shown earlier in this chapter, Kirchhoff's and Ohm's laws lead to the network equations

$$c_{\text{injected at } i} = \sum_{j \in \mathcal{N}(i)} c_{i \to j} = \sum_{j \in \mathcal{N}(i)}^{n} \frac{1}{r_{ij}} (v_i - v_j),$$

where v_i is the potential at node i and $c_{i \to j} = 1/r_{ij} \cdot (v_i - v_j)$ is the current flow from node i to node j. Consider now a more general set of current flows $f_{i \to j}$ (for all $i, j \in \{1, \dots, n\}$) "routing energy through the network" and compatible with the following basic assumptions:

  (i)   Skew-symmetry: $f_{i \to j} = -f_{j \to i}$ for all $i, j \in \{1, \dots, n\}$;
  (ii) Consistency: $f_{i \to j} = 0$ if $\{i, j\} \notin E$;
  (iii) Conservation: $c_{\text{injected at } i} = \sum_{j \in \mathcal{N}(i)} f_{i \to j}$ for all $i \in \{1, \dots, n\}$.

Show that among all possible current flows $f_{i \to j}$, the physical current flow $f_{i \to j} = c_{i \to j} = 1/r_{ij} \cdot (v_i - v_j)$ uniquely minimizes the energy dissipation:

$$\begin{aligned}
\underset{f_{i \to j}, \, i,j \in \{1, \dots, n\}}{\text{minimize}} \quad & J = \frac{1}{2} \sum_{i,j=1}^{n} r_{ij} f_{i \to j}^2 \\
\text{subject to} \quad & f_{i \to j} = -f_{j \to i} && \text{for all } i, j \in \{1, \dots, n\}, \\
& f_{i \to j} = 0 && \text{for all } \{i, j\} \notin E, \\
& c_{\text{injected at } i} = \sum_{j \in \mathcal{N}(i)} f_{i \to j} && \text{for all } i \in \{1, \dots, n\}.
\end{aligned}$$

This result is known as Thomson's Principle for electric circuits, e.g., see (Doyle and Snell, 1984).

Hint: The solution requires knowledge of the Karush-Kuhn-Tucker (KKT) conditions for optimality; this is a classic topic in nonlinear constrained optimization discussed in numerous textbooks, e.g., in (Luenberger and Ye, 2008).

E6.15 **Grounded spring networks subject to loads.** Consider a connected spring networks with n moving masses. Assume one of the masses is connected to a wall with a spring, as in figure. We refer to such a spring network as *grounded*.
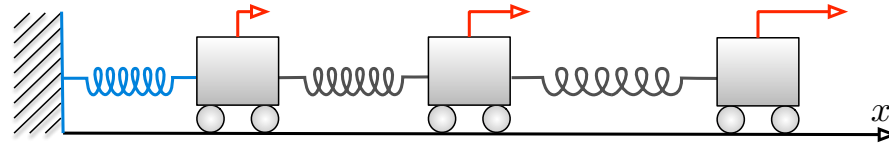
Let $L_{\text{free},n+1}$ be the $(n+1) \times (n+1)$ Laplacian matrix for the spring network of the n masses and the wall. Let L_{grounded} be the $n \times n$ *grounded Laplacian* constructed by removing the row and column of $L_{\text{free},n+1}$ corresponding to the wall. And let $L_{\text{free},n}$ be the $n \times n$ Laplacian matrix describing the spring network among the n moving masses without the spring connection to the wall; $L_{\text{free},n} = L_{\text{stiffness}}$, as defined in Section 6.1.2.

For such a grounded spring network,

(i) derive an expression relating L_{grounded} to $L_{\text{free},n}$,

(ii) show that L_{grounded} is positive definite,

(iii) compute the equilibrium displacement for an arbitrary load force f_{load} applied to the n moving masses.

Note: We refer to Chapter 10 and Exercise E10.11 for a comprehensive treatment of grounded Laplacian matrices as compartmental matrices.

E6.16   **Maximum power dissipation.** As in Subsection 6.1.3, consider an electrical network composed by three voltage sources (v_1, v_2, v_3) connected by three resistors (each with unit resistance in an undirected ring topology. Let L be the Laplacian matrix of conductances. Recall that the total power dissipated by the circuit is

$$\text{P}_{\text{dissipated}} = v^\mathsf{T} L v.$$

What is the maximum dissipated power if the voltages v satisfy $\|v\|_2 = 1$?
Hint: Recall the notion of induced 2-norm.

E6.17 **From algebraic to node connectivity.** Consider an unweighted undirected graph $G = (V, E)$ with second smallest eigenvalue $\lambda_2(G)$. Given a subset of nodes $S \subseteq V$, we define a graph $G' = (V', E')$ by *deleting* the nodes in S from G as follows: we let $V' = V \backslash S$ and E' contain all the edges in E except for those connected to a node in S. The *node connectivity* $\kappa(G)$ of G is defined by

$$\kappa(G) = \begin{cases} 0, & \text{if } G \text{ is disconnected,} \\ \text{minimum number of nodes whose deletion disconnects } G, & \text{otherwise.} \end{cases}$$

Show that

(i) $0 \le \lambda_2(G) \le \lambda_2(G') + |S|$, where $|S|$ is the cardinality of S,

(ii) $\lambda_2(G) \le \kappa(G)$.

Hint: Let $z \in \mathbb{R}^{|V'|}$, $\|z\|_2 = 1$, denote the Fiedler eigenvector of the Laplacian $L(G')$ associated with $\lambda_2(G')$. You may find it useful to define $q \in \mathbb{R}^{|V|}$ such that $q_i = z_i$ for every $i \in V'$ and $q_i = 0$ for every $i \in S$.

Exercises with solution

E6.18   **Averaging-based PID control.** Consider a set of n controllable agents governed by the second-order dynamics

$$\dot{x}_i(t) = y_i(t), \tag{E6.1a}$$
$$\dot{y}_i(t) = u_i(t) + \eta_i, \tag{E6.1b}$$

where $i \in \{1, \ldots, n\}$ is the index set, $t \mapsto u_i(t) \in \mathbb{R}$ is a control input to agent i, and $\eta_i \in \mathbb{R}$ is an unknown constant disturbance affecting agent i. Given an undirected, connected, and weighted graph G with adjacency matrix $A = A^\mathsf{T} \in \mathbb{R}^{n \times n}$, assume each agent i can measure its velocity $y_i \in \mathbb{R}$ and the relative position $x_i - x_j$ for each neighbor j. The agent then implements the *averaging-based proportional, integral, derivative (PID) controller* defined by

$$u_i(t) = -\sum_{j=1}^{n} a_{ij}\big(x_i(t) - x_j(t)\big) - y_i(t) - q_i(t), \tag{E6.2a}$$
$$\dot{q}_i(t) = y_i(t) - \sum_{j=1}^{n} a_{ij}\big(q_i(t) - q_j(t)\big), \tag{E6.2b}$$

where $q_i \in \mathbb{R}$ is a dynamic control state for each agent $i \in \{1, \ldots, n\}$. Show that

(i)   the average state $x_{\text{ave}}(t) = \frac{1}{n}\sum_{i=1}^{n} x_i(t)$ is bounded for all $t \geq 0$,
(ii) the set of equilibria $(x^\star, y^\star, q^\star)$ of the closed-loop system (E6.1)-(E6.2) satisfies $x^\star = \beta \mathbb{1}_n + L^\dagger \eta$ and $y^\star = \mathbb{0}_n$, where β is an appropriate constant and L is the graph Laplacian, and
(iii) all trajectories converge to these closed-loop equilibria.
      Hint: Recall the Routh-Hurwitz Criterion for third-degree polynomials: The polynomial $s^3 + a_2 s^2 + a_1 s + a_0$ has roots with strictly negative real part if and only if $a_0 > 0$, $a_2 > 0$, and $a_1 a_2 > a_0$.

Note: In other words the averaging-based PID control achieves zero velocity and position consensus with an error proportional to the disturbance.

Answer: In compact vector form, the closed-loop model reads as

$$\dot{x} = y, \tag{E6.3a}$$
$$\dot{y} = -Lx - y - q + \eta, \tag{E6.3b}$$
$$\dot{q} = y - Lq, \tag{E6.3c}$$

where $x, y, q \in \mathbb{R}^n$ are vectors with components x_i, y_i, q_i for $i \in \{1, \ldots, n\}$, and $L = \text{diag}(\sum_{j=1}^{n} A_{ij}) - A$ is the Laplacian matrix of G. Note the state matrix is

$$\begin{bmatrix} 0 & I_n & 0 \\ -L & -I_n & -I_n \\ 0 & I_n & -L \end{bmatrix}. \tag{E6.4}$$

Regarding statement (i), consider the 3-dimensional vector of aggregate states

$$
\begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{q} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} q_i \end{bmatrix},
$$

whose dynamics are given by

$$
\dot{\bar{x}} = \bar{y}, \tag{E6.5a}
$$
$$
\dot{\bar{y}} = -\bar{y} - \bar{q} + \bar{\eta}, \tag{E6.5b}
$$
$$
\dot{\bar{q}} = \bar{y}, \tag{E6.5c}
$$

where $\bar{\eta} = \sum_{i=1}^{n} \bar{\eta}_i$. The average dynamics (E6.5) admit a globally exponentially stable equilibrium which can be seen, e.g., by integrating the third equation (E6.5c) as $\bar{q}(t) = \bar{x}(t) - \bar{x}_0 + \bar{q}_0$ and inserting the solution into (E6.5) which yields the second-order system

$$
\ddot{\bar{x}} = -\dot{\bar{x}} - \bar{x} - \sigma, \tag{E6.6}
$$

where the $\sigma = \bar{\eta} - \bar{x}_0 + \bar{q}_0$. The system (E6.6) admits the globally exponentially stable equilibrium $\bar{x} = \sigma = \bar{\eta} - \bar{x}_0 + \bar{q}_0$. Hence, the aggregate state and so the average state is bounded.

Regarding statement (ii), the equilibria of the closed loop system (E6.3) are given by

$$
\mathbb{0}_n = y^\star, \tag{E6.7a}
$$
$$
\mathbb{0}_n = -Lx^\star - y^\star - q^\star + \eta, \tag{E6.7b}
$$
$$
\mathbb{0}_n = y^\star - Lq^\star. \tag{E6.7c}
$$

From (E6.7a) we have that $y^\star = \mathbb{0}_n$, which yields the remaining equations

$$
\mathbb{0}_n = -Lx^\star - q^\star + \eta, \tag{E6.8a}
$$
$$
\mathbb{0}_n = -Lq^\star. \tag{E6.8b}
$$

Equation (E6.8b) yields $q^\star = c_1 \cdot \mathbb{1}_n$ for some $c_1 \in \mathbb{R}$. By multiplying equations (E6.8a) from the left by $\mathbb{1}_n^\mathsf{T}$, we obtain $\mathbb{1}_n^\mathsf{T} q^\star = n \cdot c_1 = \mathbb{1}_n^\mathsf{T} \eta$ and, in turn,

$$
q^\star = c_1 \cdot \mathbb{1}_n = \frac{\mathbb{1}_n^\mathsf{T} \eta}{n} \mathbb{1}_n.
$$

Equation (E6.8a) now reduces to

$$
Lx^\star = \eta - q^\star = \left(I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T} \right) \eta.
$$

This is a Laplacian system, as studied in Lemma 6.12, and therefore the set of solutions is $\{L^\dagger \eta + c_2 \mathbb{1}_n \mid c_2 \in \mathbb{R}\}$, where we used the identity $L^\dagger(I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T})\eta = L^\dagger \eta$. Finally, equation (E6.6) implies that the constant c_2 must equal σ, that is, any equilibrium is of the form

$$x^\star = (\bar{\eta} - \bar{x}_0 + \bar{q}_0) \cdot \mathbb{1}_n + L^\dagger \eta.$$

In summary, the unique equilibrium is

$$(x^*, y^*, q^*) = \left((\bar{\eta} - \bar{x}_0 + \bar{q}_0) \cdot \mathbb{1}_n + L^\dagger \eta, \ \mathbb{0}_n, \ \frac{\mathbb{1}_n^\mathsf{T}\eta}{n}\mathbb{1}_n \right).$$

Regarding statement (iii), we start with a preliminary small result. For $\lambda \in \mathbb{R}$, define

$$C_\lambda = \begin{bmatrix} 0 & 1 & 0 \\ -\lambda & -1 & -1 \\ 0 & 1 & -\lambda \end{bmatrix},$$

whose characteristic polynomial is:

$$\det(sI_3 - C_\lambda) = s^3 + (1+\lambda)s^2 + (1+2\lambda)s + \lambda^2.$$

If $\lambda > 0$, then the Routh-Hurwitz Criterion for a polynomial of the form $s^3 + a_2s^2 + a_1s + a_0$ is satisfied because

$$a_2 = 1 + \lambda > 0, \qquad a_0 = \lambda^2 > 0, \qquad \text{and} \qquad a_1a_2 = (1+2\lambda)(1+\lambda) > \lambda^2 = a_0.$$

Therefore, the matrix C_λ is Hurwitz for any $\lambda > 0$.

Consider now the following change of coordinates shifting the desired equilibria to the origin

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{q} \end{bmatrix} = \begin{bmatrix} x - x^\star \\ y - y^\star \\ q - q^\star \end{bmatrix},$$

where $(x^\star, y^\star, q^\star)$ are an (more precisely, the unique) equilibrium of the closed loop (E6.3). Then the dynamics (E6.3) read as

$$\dot{\tilde{x}} = \tilde{y}, \tag{E6.9a}$$
$$\dot{\tilde{y}} = -L\tilde{x} - \tilde{y} - \tilde{q}, \tag{E6.9b}$$
$$\dot{\tilde{q}} = \tilde{y} - L\tilde{q}, \tag{E6.9c}$$

with corresponding state matrix as in equation (E6.4). Since L is a symmetric irreducible Laplacian matrix, it has eigenvalues $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n$ with corresponding orthonormal eigenvectors $v_1 = \mathbb{1}_n/\sqrt{n}, v_2, \ldots, v_n$. If we now adopt the modal decomposition as in Remark 2.3 (and in Section 6.5), we obtain

$$\frac{d}{dt}\begin{bmatrix} \tilde{x}_j \\ \tilde{y}_j \\ \tilde{q}_j \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\lambda_j & -1 & -1 \\ 0 & 1 & -\lambda_j \end{bmatrix} \begin{bmatrix} \tilde{x}_j \\ \tilde{y}_j \\ \tilde{q}_j \end{bmatrix} = C_{\lambda_j}\begin{bmatrix} \tilde{x}_j \\ \tilde{y}_j \\ \tilde{q}_j \end{bmatrix},$$

where $\tilde{x} = \sum_{j=1}^n \tilde{x}_j v_j$, $\tilde{y} = \sum_{j=1}^n \tilde{y}_j v_j$, and $\tilde{q} = \sum_{j=1}^n \tilde{q}_j v_j$. The proof of statement (iii) now follows because the evolution of the first mode (the average or aggregate state) was analyzed in statement (i) and the evolution of the all other modes is exponentially stable since $\lambda_j > 0$ implies C_{λ_j} is Hurwitz.

Chapter 7

Continuous-time Averaging Systems

In this chapter we consider averaging algorithms in which the variables evolve in continuous time, instead of discrete time. In other words, we consider a certain class of differential equations and show when their asymptotic behavior is the emergence of consensus.

7.1   Example systems

We present here some simple examples of continuous-time averaging systems, along the lines of the flocking dynamics example in Section 1.3.

7.1.1 Example #1: Continuous-time opinion dynamics

This first example is taken from (Abelson, 1964) and provides a continuous-time analog to the French-Harary-DeGroot discrete-time averaging model

$$x(k + 1) = Ax(k), \tag{7.1}$$

that we studied in detail in Chapters 1-5. Loosely speaking, we assume that

(i) there exists a time period $\tau \in \mathbb{R}$ satisfying $0 < \tau \ll 1$ such that the discrete-time indexes k and $k + 1$ correspond to real times $t = k\tau$ and $t + \tau = (k + 1)\tau$, respectively, and

(ii) the edge weights of the influence systems are of the form $a_{ij} = \overline{a_{ij}}\tau$, where the coefficients $\overline{a_{ij}}$ can be regarded as *contact rates* between the individuals.

We now compute the opinion change from time k to time $k + 1$:

$$x(k+1) - x(k) = (A - I_n)x(k) = -Lx(k),$$

where $L = I_n - A$ is the Laplacian of the matrix A. Note that the second assumption (ii) implies that the Laplacian matrix L satisfies $L = \overline{L}\tau$ for a Laplacian matrix $\overline{L}$ containing the contact rates. Therefore, we can write

$$\frac{x(t+\tau) - x(t)}{\tau} = \frac{x(k+1) - x(k)}{\tau} = -\overline{L}x(t),$$

and, taking the limit as $\tau \to 0^+$ in the left-hand side, we obtain the Abelson's continuous-time opinion dynamics model:

$$\dot{x}(t) = -\overline{L}x(t). \tag{7.2}$$

As we mentioned in Section 1.3, we refer to this equation as to the Laplacian flow. In summary, we have learnt that, if the edge weights of the influence systems are of the form $a_{ij} = \overline{a_{ij}}\tau$, then the solution to the French-Harary-DeGroot discrete-time averaging system (7.1) converges to the solution to the Abelson continuous-time averaging system (7.2) in the limit as $\tau \to 0^+$.

Note: Because of this correspondence, we expect to see consensus emerge along solutions to the systems (7.2), at least for certain classes of digraphs.

7.1.2   Example #2: A simple RC circuit
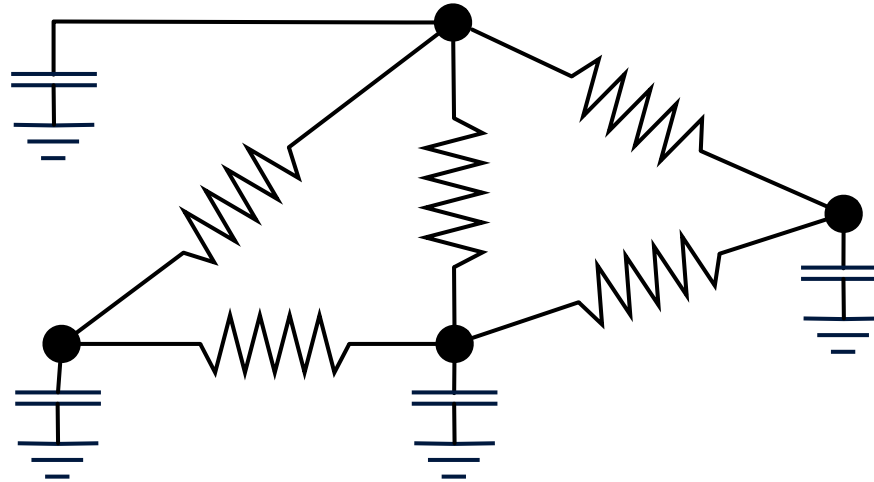
Finally, we consider an electrical network with only pure resistors and with pure capacitors connecting each node to ground. From the previous chapter, we know the vector of injected currents c_{injected} and the vector of voltages at the nodes v satisfy

$$c_{\text{injected}} = L\,v,$$

where L is the Laplacian for the graph with coefficients $a_{ij} = 1/r_{ij}$. Additionally, assuming C_i is the capacitance at node i, and keeping proper track of the current into each capacitor, we have

$$C_i \frac{d}{dt} v_i = -c_{\text{injected at } i}$$

so that, defining $C = \text{diag}(C_1, \ldots, C_n)$, we obtain

$$\frac{d}{dt} v = -C^{-1} L\, v. \tag{7.3}$$

Note: $C^{-1}L$ is an asymmetric Laplacian matrix (for a directed weighted graph).

Note: it is physically intuitive that after some transient all nodes will have the same potential. This intuition will be proved later in the chapter.

7.1.3   Example #3: Discretization of partial differential equations

The name Laplacian matrix is inherited from the *Laplacian operator* in the diffusion partial differential equation (PDEs) named after the French mathematician Pierre-Simon Laplace.
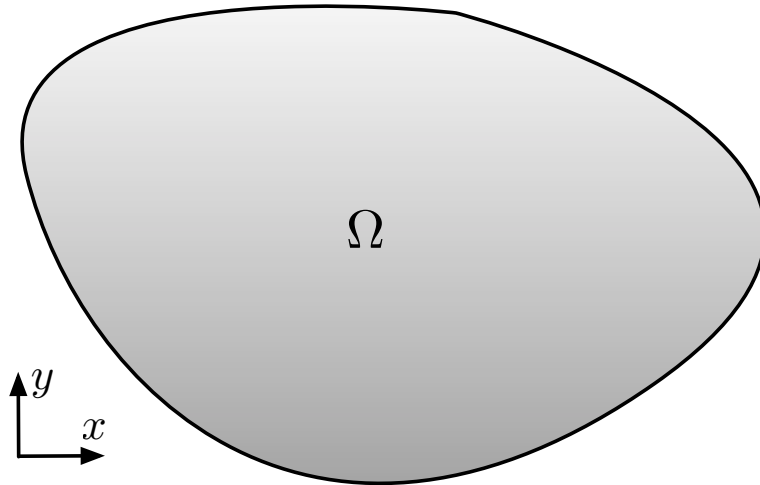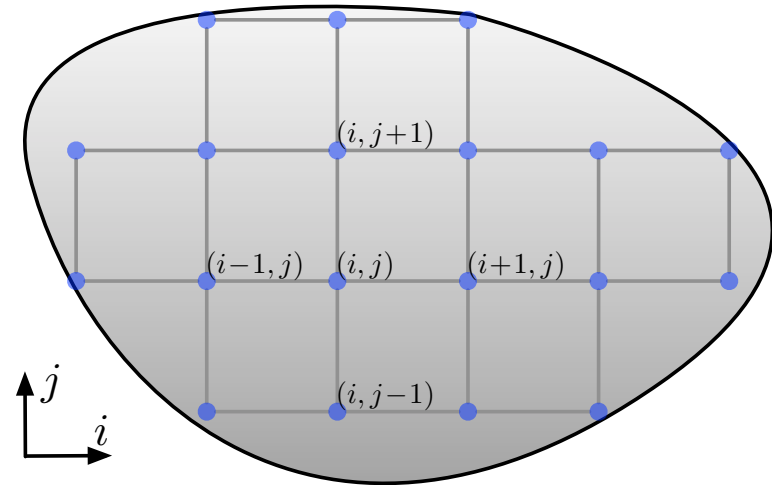
(a) Spatial domain   (b) Discretization through a mesh graph

Consider a closed bounded spatial domain $\Omega \subset \mathbb{R}^2$ and a spatio-temporal function $u(t, x, y)$ denoting the temperature at a point $(x, y) \in \Omega$ at time $t \in \mathbb{R}_{\geq 0}$. The evolution of the temperature $u(t, x, y)$ in time and space is governed by the *heat equation*

$$\frac{\partial u}{\partial t} = c\,\Delta u, \tag{7.4}$$

where $c > 0$ is the *thermal diffusivity* (which we assume constant) and the *Laplacian differential operator* is

$$\Delta u(t, x, y) = \frac{\partial^2 u}{\partial x^2}(t, x, y) + \frac{\partial^2 u}{\partial y^2}(t, x, y).$$

To approximately solve this PDE, we introduce a finite-difference approximation of (7.4). First, we discretize the spatial domain Ω through a mesh graph (i.e., a subgraph of a two-dimensional grid graph) with discrete coordinates indexed by (i, j) and where neighboring mesh points are a distance $h > 0$ apart. Second, we approximate the Laplacian operator via the finite-difference approximation:

$$\Delta u(t, x_i, y_j) \approx \frac{u(t, x_{i-1}, y_j) + u(t, x_{i+1}, y_j) + u(t, x_i, y_{j-1}) + u(t, x_i, y_{j+1}) - 4u(t, x_i, y_j)}{h^2}.$$

(Note: This is the correct expansion for an interior point; similar approximations can be written for boundary points, assuming the boundary conditions are free.)

Now, the key observation is that the finite-difference approximation renders the heat equation to a Laplacian flow. Specifically, if u_{discrete} denotes the vector of values of u at the nodes, then one can see that equation (7.4) is approximately rewritten as:

$$\frac{d}{dt} u_{\text{discrete}} = -\frac{c}{h^2} L \, u_{\text{discrete}},$$

where L is the Laplacian matrix of the unweighted mesh graph.

Another standard PDE involving the Laplacian operator is the *wave equation*

$$\frac{\partial^2 u}{\partial t^2} = s^2 \, \Delta u, \tag{7.5}$$

modeling the displacement $u(t, x, y)$ of an elastic surface on Ω with *wave propagation speed* $s > 0$. In this case, a finite-difference approximation gives rise a *second-order Laplacian flow*

$$\frac{d^2}{dt^2} u_{\text{discrete}} = -\frac{s^2}{h^2} L \, u_{\text{discrete}}. \tag{7.6}$$

We study Laplacian flows in this chapter and (general) second-order Laplacian flows in the next chapter.

7.2 Continuous-time linear systems and their convergence properties

In Section 2.1 we presented discrete-time linear systems and their convergence properties; here we present their continuous-time analogous.

A *continuous-time linear system* is

$$\dot{x}(t) = Ax(t). \tag{7.7}$$

Its *solution* $t \mapsto x(t)$, $t \in \mathbb{R}_{\geq 0}$ from an initial condition $x(0)$ satisfies $x(t) = \exp(At)x(0)$, where the *matrix exponential* of a square matrix A is defined by

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

The matrix exponential generalizes the usual exponential function $x \mapsto \mathrm{e}^x = \sum_{k=0}^{\infty} x^k/k!$. This remarkable operation enjoys numerous properties, some of which are reviewed in Exercise E7.1. A matrix $A \in \mathbb{R}^{n \times n}$ is

(i) *continuous-time semi-convergent* if $\lim_{t \to +\infty} \exp(At)$ exists, and

(ii) *Hurwitz*, or for consistency *continuous-time convergent*, if $\lim_{t \to +\infty} \exp(At) = \mathbb{0}_{n \times n}$.

The *spectral abscissa* of a square matrix A is the maximum of the real parts of the eigenvalues of A, that is,

$$\alpha(A) = \max\{\Re(\lambda) \mid \lambda \in \mathrm{spec}(A)\}.$$

Theorem 7.1 (Convergence and spectral abscissa). *For a square matrix A, the following statements hold:*

 (i) *A is continuous-time convergent (Hurwitz) if and only if $\alpha(A) < 0$,*

(ii) *A is continuous-time semi-convergent and not convergent if and only if*

    (a) 0 is an eigenvalue,

    (b) 0 is a semisimple eigenvalue, and

    (c) all other eigenvalues have negative real part.

We leave the proof of this theorem to the reader and mention that most required steps are similar to the discussion in Section 2.1 and are discussed later in this chapter.

7.3   The Laplacian flow

Let G be a weighted directed graph with n nodes and Laplacian matrix L. The *Laplacian flow* on $\mathbb{R}^n$ is the dynamical system

$$\dot{x} = -Lx, \tag{7.8}$$

or, equivalently in components,

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}(x_j - x_i) = \sum_{j \in \mathcal{N}^{\mathrm{out}}(i)} a_{ij}(x_j - x_i).$$

7.3.1 Matrix exponential of a Laplacian matrix

Before analyzing the Laplacian flow, we provide some results on the matrix exponential of (minus) a Laplacian matrix. We show how such an exponential matrix is row-stochastic and has properties analogous to those for adjacency matrices studied in Section 4.2.

Theorem 7.2 (The matrix exponential of a Laplacian matrix). *Let $L \in \mathbb{R}^{n \times n}$ be a Laplacian matrix with associated weighted digraph G and with maximum diagonal entry $\ell_{\max} = \max\{\ell_{11}, \ldots, \ell_{nn}\}$. Then*

(i) $\exp(-L) \geq e^{-\ell_{\max}} I_n \geq \mathbb{0}_{n \times n}$,
(ii) $\exp(-L)\mathbb{1}_n = \mathbb{1}_n$,
(iii) $\mathbb{1}_n^{\mathsf{T}} \exp(-L) = \mathbb{1}_n^{\mathsf{T}}$,   *if and only if*   *$G$ is weight-balanced (i.e., $\mathbb{1}_n^{\mathsf{T}} L = \mathbb{0}_n^{\mathsf{T}}$),*
(iv) $\exp(-L)\mathbb{e}_j > 0$,   *if and only if*   *the j-th node is globally reachable in G, and*
(v) $\exp(-L) > 0$,   *if and only if*   *G is strongly connected (i.e., L is irreducible).*

Note that properties (i) and (ii) together imply that $\exp(-L)$ is row-stochastic.

Proof. From the equality $L\mathbb{1}_n = \mathbb{0}_n$ and the definition of matrix exponential, we compute

$$\exp(-L)\mathbb{1}_n = \left(I_n + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!}L^k\right)\mathbb{1}_n = \mathbb{1}_n.$$

This calculation establishes statements (ii). Similarly, if $\mathbb{1}_n^{\mathsf{T}}L = \mathbb{0}_n^{\mathsf{T}}$, we compute

$$\mathbb{1}_n^{\mathsf{T}}\exp(-L) = \mathbb{1}_n^{\mathsf{T}}\left(I_n + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!}L^k\right) = \mathbb{1}_n^{\mathsf{T}}.$$

Next, we assume $\mathbb{1}_n^{\mathsf{T}}\exp(-L) = \mathbb{1}_n^{\mathsf{T}}$ and prove $\mathbb{1}_n^{\mathsf{T}}L = \mathbb{0}_n^{\mathsf{T}}$. Define $f(t) = \exp(-L^{\mathsf{T}}t)$, for $t \in [0,1]$. Note $f(0)\mathbb{1}_n = f(1)\mathbb{1}_n = \mathbb{1}_n$. From Exercise E7.1(iv) we know $\frac{d}{dt}\exp(-L^{\mathsf{T}}t) = -L\exp(-L^{\mathsf{T}}t)$ so that $\frac{d}{dt}\exp(-L^{\mathsf{T}}t)\mathbb{1}_n = -L\mathbb{1}_n$. Finally, the fundamental theorem of calculus implies

$$\mathbb{0}_n = f(1)\mathbb{1}_n - f(0)\mathbb{1}_n = \int_0^1 \frac{d}{dt}f(t)dt = -L\mathbb{1}_n.$$

This completes the proof of statement (iii).

Next, we define a non-negative matrix A_L by

$$A_L = -L + \ell_{\max} I_n \qquad \Longleftrightarrow \qquad -L = -\ell_{\max} I_n + A_L.$$

Because $A_L I_n = I_n A_L$, we know

$$\exp(-L) = \exp(-\ell_{\max} I_n) \exp(A_L) = e^{-\ell_{\max}} \exp(A_L). \tag{7.9}$$

Here we used the following properties of the matrix exponential operation: $\exp(A + B) = \exp(A) \exp(B)$ if $AB = BA$ and $\exp(aI_n) = e^a I_n$. Next, because $A_L \geq 0$, we know that $\exp(A_L) = \sum_{k=0}^{\infty} A_L^k/k!$ is lower bounded by the first $n - 1$ terms of the series so that

$$\exp(-L) = e^{-\ell_{\max}} \exp(A_L) \geq e^{-\ell_{\max}} \sum_{k=0}^{n-1} \frac{1}{k!} A_L^k. \tag{7.10}$$

Next, we derive two useful lower bounds on $\exp(-L)$ based on the inequality (7.10). First, by keeping just the first term, we establish statement (i):

$$\exp(-L) \geq e^{-\ell_{\max}} I_n \geq 0.$$

Second, we lower bound the coefficients $1/k!$ and write:

$$\exp(-L) \geq e^{-\ell_{\max}} \sum_{k=0}^{n-1} \frac{1}{k!} A_L^k \geq \frac{e^{-\ell_{\max}}}{(n-1)!} \sum_{k=0}^{n-1} A_L^k. \tag{7.11}$$

Notice now that the digraph G associated to L is the same as that associated to A_L (we do not need to worry about self-loops here). Recall now Lemma 4.4: node j is globally reachable in G if and only if the jth column of $\sum_{k=0}^{n-1} A_L^k$ is positive. But ineequality 7.11 implies that, if j is globally reachable, then the jth column of $\exp(-L)$ is positive. This establishes the "if" part of statement (iv).

Next, by equality 7.9, if the j-th column of $\exp(-L)$ is positive, so is the jth column of $\exp(A_L)$. Since A_L is non-negative, an application of the Caley-Hamilton Theorem shows that the j-th column of $\sum_{k=0}^{n-1} A_L^k$ must be positive. Therefore, by Lemma 4.4 node j is globally reachable. This concludes the proof of statement (iv). Finally, statement (v) is an immediate consequence of statement (iv). ∎

7.3.2   Equilibria and convergence of the Laplacian flow

We can now focus on the Laplacian flow dynamics.

Lemma 7.3 (Equilibrium points). *If G contains a globally reachable node, then the set of equilibrium points of the Laplacian flow* (7.8) *is* $\mathrm{span}\{\mathbb{1}_n\} = \{\beta\mathbb{1}_n \mid \beta \in \mathbb{R}\}$.

Proof. A point x is an equilibrium for the Laplacian flow if $Lx = \mathbb{0}_n$. Hence, any point in the kernel of the matrix L is an equilibrium. From Theorem 6.6, if G contains a globally reachable node, then $\mathrm{rank}(L) = n - 1$. Hence, the dimension of the kernel space is 1. The lemma follows by recalling that $L\mathbb{1}_n = \mathbb{0}_n$. ∎

7.3.3   Some intuition

We are now interested in characterizing the solution of the Laplacian flow (7.8). To build some intuition, we first consider an undirected graph G and write the modal decomposition of the solution as in Remark 2.3 for a discrete-time linear system. We proceed in two steps. First, because G is undirected, the matrix L is symmetric and has real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ with corresponding orthonormal (i.e., orthogonal and unit-length) eigenvectors $v_1, \ldots, v_n$. Define $y_i(t) = v_i^\mathsf{T} x(t)$ and left-multiply $\dot{x} = -Lx$ by v_i:

$$\frac{d}{dt} y_i(t) = -\lambda_i y_i(t), \qquad y_i(0) = v_i^\mathsf{T} x(0).$$

These n decoupled ordinary differential equations are immediately solved to give

$$x(t) = y_1(t) v_1 + y_2(t) v_2 + \ldots + y_n(t) v_n$$
$$= \mathrm{e}^{-\lambda_1 t} (v_1^\mathsf{T} x(0)) v_1 + \mathrm{e}^{-\lambda_2 t} (v_2^\mathsf{T} x(0)) v_2 + \ldots + \mathrm{e}^{-\lambda_n t} (v_n^\mathsf{T} x(0)) v_n.$$

Second, recall that $\lambda_1 = 0$ and $v_1 = \mathbb{1}_n / \sqrt{n}$ because L is a symmetric Laplacian matrix ($L\mathbb{1}_n = \mathbb{0}_n$). Therefore, we compute $(v_1^\mathsf{T} x(0)) v_1 = \mathrm{average}(x(0)) \mathbb{1}_n$ and substitute

$$x(t) = \mathrm{average}(x(0)) \mathbb{1}_n + \mathrm{e}^{-\lambda_2 t} (v_2^\mathsf{T} x(0)) v_2 + \ldots + \mathrm{e}^{-\lambda_n t} (v_n^\mathsf{T} x(0)) v_n.$$

Now, let us assume that G is connected so that its second smallest eigenvalue λ_2 is strictly positive. In this case, we can infer that

$$\lim_{t \to \infty} x(t) = \mathrm{average}(x(0)) \mathbb{1}_n,$$

or, defining a *disagreement vector* $\delta(t) = x(t) - \mathrm{average}(x(0)) \mathbb{1}_n$ as in Section 5.3, we infer

$$\delta(t) = \mathrm{e}^{-\lambda_2 t} (v_2^\mathsf{T} x(0)) v_2 + \ldots + \mathrm{e}^{-\lambda_n t} (v_n^\mathsf{T} x(0)) v_n.$$

7.3.4   Consensus with Laplacian flow with globally reachable node

In summary, we discovered that, for a connected undirected graph, the disagreement vector converges to zero with an exponential rate λ_2. In what follows, we state a more general convergence to consensus result for the continuous-time Laplacian flow. This result is parallel to Theorem 5.1.

Theorem 7.4 (Consensus for Laplacian matrices with a globally reachable node). *Let L be a Laplacian matrix and let G be its associated digraph. The following statements are equivalent:*

(AL1) *the eigenvalue 0 of $-L$ is simple and all other eigenvalues of $-L$ have negative real part;*

(AL2) *$-L$ is continuous-time semi-convergent and $\lim_{t\to\infty} \exp(-Lt) = \mathbb{1}_n w^\mathsf{T}$, where $w \in \mathbb{R}^n$ satisfies $w \geq 0$, $\mathbb{1}_n^\mathsf{T} w = 1$, and $w^\mathsf{T} L = \mathbb{0}^\mathsf{T}$; and*

(AL3) *G contains a globally reachable node.*

If any, and therefore all, of the previous conditions are satisfied, then

(i) *$w \geq 0$ is the left dominant eigenvector of $-L$ and $w_i > 0$ if and only if node i is globally reachable;*

(ii) *the solution to $\frac{d}{dt} x(t) = -Lx(t)$ satisfies*

$$\lim_{t\to\infty} x(t) = \left(w^\mathsf{T} x(0)\right)\mathbb{1}_n,$$

(iii) *if additionally G is weight-balanced, then G is strongly connected, $\mathbb{1}_n^\mathsf{T} L = \mathbb{0}_n^\mathsf{T}$, $w = \frac{1}{n}\mathbb{1}_n$, and*

$$\lim_{t\to\infty} x(t) = \frac{\mathbb{1}_n^\mathsf{T} x(0)}{n}\mathbb{1}_n = \operatorname{average}\left(x(0)\right)\mathbb{1}_n.$$

Note: Theorem 7.4 is the continuous-time version of Theorem 5.1 about discrete-time averaging systems. The only notable difference is that, in continuous time, it is not necessary to require the subgraph of globally reachable nodes to be aperiodic. Also note that it is possible to write a continuous-time version of Theorem 5.2; we leave this task to the reader.

7.3.5   Proof of Theorem 7.4

Proof. We start by noting that there are two ways to prove the theorem. Either one mimicks the proof of Theorem 5.1 or one transcribes Theorem 5.1. We take the second approach and leave the first to the interested reader. As in Exercise E6.1, pick $\varepsilon < \frac{1}{d_{\max}}$, where $d_{\max}$ is the maximum out-degree, and define

$$\mathcal{A}_{L,\varepsilon} = I_n - \varepsilon L. \tag{7.12}$$

First, note that property (A1) in Theorem 5.1 holds for $\mathcal{A}_{L,\varepsilon}$ if and only if property (AL1) holds for L. Indeed, $\mathcal{A}_{L,\varepsilon}$ has a simple strictly-dominating eigenvalue 1 if and only if L has a simple strictly-dominating eigenvalue 0 by property E6.1(v).

Second, we ask the reader to prove in Exercise E7.4 that property (A2) in Theorem 5.1 holds for $\mathcal{A}_{L,\varepsilon}$ if and only if property (AL2) holds for L, that is, $\lim_{k\to\infty} \mathcal{A}_{L,\varepsilon}^k = \mathbb{1}_n w^{\mathsf{T}}$ if and only if $\lim_{t\to\infty} \exp(-Lt) = \mathbb{1}_n w^{\mathsf{T}}$.

Third, note that property (A3) in Theorem 5.1 holds for $\mathcal{A}_{L,\varepsilon}$ if and only if property (AL3) holds for L. Indeed, $\mathcal{A}_{L,\varepsilon}$ has a strictly positive diagonal and the same pattern of zero/positive off-diagonal entries as L.

We have now established that properties (AL1), (AL2) and (AL3) are equivalent and that the consequences (i), (ii) and (iii) in Theorem 5.1 hold for the matrix $\mathcal{A}_{L,\varepsilon}$. It is easy to see that these consequences, in turn, imply properties (i), (ii) and (iii) for L. ∎

7.4   Appendix: Design of weight-balanced digraphs

Recall from Section 5.4.2 that, given a connected undirected graph, the Metropolis–Hastings algorithm computes edge weights and self-loop weights that render the resulting weighted adjacency matrix symmetric and doubly stochastic. Note that the corresponding Laplacian matrix is weight-balanced.

Problem: Given a strongly-connected weighted digraph G with adjacency matrix A, how do we rescale the weights on each edge of G such that the resulting adjacency matrix $\overline{A}$ is doubly stochastic and the resulting Laplacian matrix $\overline{L}$ is weight balanced?

Note: A solution to this problem ensures that both the discrete time and continuous time averaging systems associated to $\overline{A}$ and $\overline{L}$ converge to average consensus. Because we assume G is directed, the Metropolis–Hastings algorithm is not applicable.

Answer: Here's an algorithmic solution. Since G is strongly connected, Theorem 7.4 establishes that its left dominant eigenvector w is positive. For $\ell_{\max} = \max\{\ell_{11}, \ldots, \ell_{nn}\}$, define

$$\overline{L} = \frac{1}{\ell_{\max}} \operatorname{diag}(w)L,$$
$$\overline{A} = I_n - \overline{L}.$$

This eigenvector rescaling algorithm has the following properties:

(i) $\overline{L}$ has the same zero/negative/positive pattern as L and since L has zero row sum, it is immediate to see that also $\overline{L}$ has zero row sum. Therefore, $\overline{L}$ is a Laplacian matrix and its associated weighted digraph $\overline{G}$ has the same topology as G and has weights $\overline{a}_{ij} = w_i a_{ij} / \ell_{\max}$. In other words, the weight of each out-edge of node i is rescaled by $w_i / \ell_{\max}$.

(ii) The column sums of $\overline{L}$ are zero, since:

$$\mathbb{1}_n^\mathsf{T} \overline{L} = \mathbb{1}_n^\mathsf{T} \operatorname{diag}(w) L / \ell_{\max} = w^\mathsf{T} L / \ell_{\max} = \mathbb{0}_n^\mathsf{T}.$$

(iii) Finally, one can show that each diagonal entry of $\overline{L}$ is less than 1 and, therefore, the matrix $\overline{A}$ is non-negative. The facts that $\overline{A}$ has unit row and column sums are trivial.

7.5   Historical notes and further reading

Section 7.1.1 "Example #1: Continuous-time opinion dynamics" presents the continuous-time averaging model by (Abelson, 1964) and its relationship with the discrete-time averaging model by (French Jr., 1956; Harary, 1959; DeGroot, 1974). Abelson's work is one of the earliest on what we now call the Laplacian flow.

Regarding Example #2: "Flocking behavior for a group of animals" in Section 1.3, a classic early reference on this topic is (Reynolds, 1987). In that model, flocking behavior is controlled by three simple rules: Separation - avoid crowding neighbors (short range repulsion) Alignment - steer towards average heading of neighbors, and Cohesion - steer towards average position of neighbors (long range attraction).

The RC circuit example in Section 7.1.2 is taken from (Mesbahi and Egerstedt, 2010; Ren et al., 2007).

An early reference to Theorem 7.4 is the work by Abelson (1964) in mathematical sociology; more recent references with rigorous proofs in the control literature include (Lin et al., 2005; Ren and Beard, 2005).

A reference for the construction in Section 7.4 is (Ren et al., 2007).

7.6 Exercises

E7.1   **Properties of the matrix exponential.** Recall the definition $\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k$ for any square matrix A. Complete the following tasks:

  (i) show that $\sum_{k=0}^{\infty} \frac{1}{k!} A^k$ converges absolutely for all square matrices A,
       Hint: Recall: a matrix series $\sum_{k=1}^{\infty} B_k$ *converges absolutely* if $\sum_{k=1}^{\infty} \|B_k\|$ converges, where $\| \cdot \|$ is a matrix norm.
 (ii) show that $A = \mathrm{diag}(a_1, \ldots, a_n)$ implies $\exp(A) = \mathrm{diag}(e^{a_1}, \ldots e^{a_n})$,
(iii) show that $AB = BA$ implies $\exp(A + B) = \exp(A)\exp(B)$,
 (iv) show that $\frac{d}{dt}\exp(At) = A\exp(At) = \exp(At)A$,
  (v) show that $\exp(TAT^{-1}) = T\exp(A)T^{-1}$ for any invertible T, and
 (vi) $\lambda \in \mathrm{spec}(A)$ implies $e^{\lambda} \in \mathrm{spec}(\exp(A))$, and $\rho(\exp(A)) = e^{\alpha(A)}$, as illustrated in Figure E7.1.
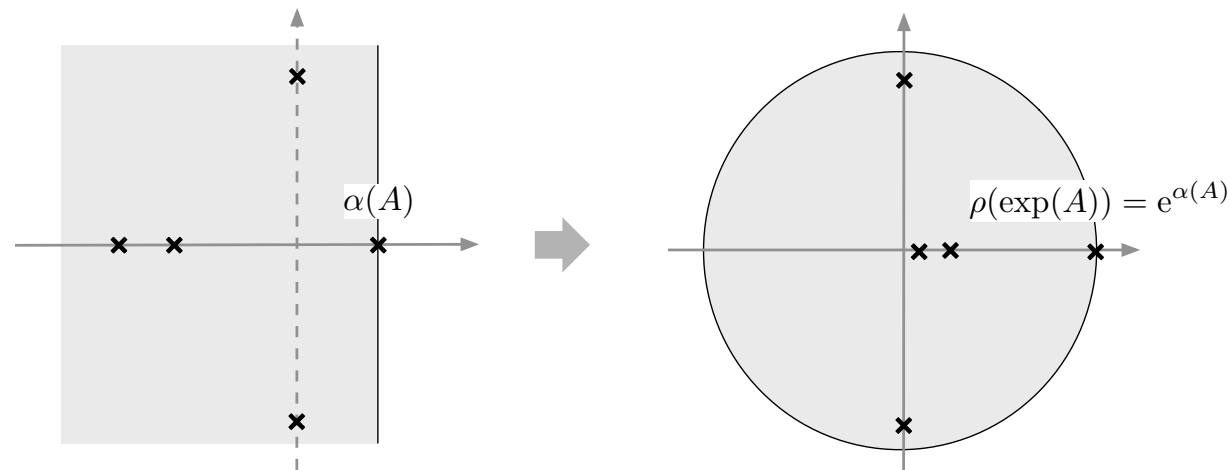
Figure E7.1: Spetrum of A and of its matrix exponential

▶ E7.2   **Continuous-time affine systems.** Given $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$, consider the continuous-time affine systems

$$\dot{x}(t) = Ax(t) + b.$$

Assume A is Hurwitz and, similarly to Exercise E2.9, show that

(i) the matrix A is invertible,
(ii) the only equilibrium point of the system is $-A^{-1}b$, and
(iii) $\lim_{t \to \infty} x(t) = -A^{-1}b$ for all initial conditions $x(0) \in \mathbb{R}^n$.

E7.3   **Semi-convergence of rank-one matrices.** Given vectors $v, w \in \mathbb{R}^n$, consider the rank-one matrix $vw^{\mathsf{T}} \in \mathbb{R}^{n \times n}$. Show

    (i) $\text{trace}(vw^{\mathsf{T}}) = v^{\mathsf{T}}w$ and $\text{spec}(vw^{\mathsf{T}}) = \{0, \dots, 0, v^{\mathsf{T}}w\}$,

    (ii) if $v^{\mathsf{T}}w \neq 0$, then $P_v = \frac{1}{v^{\mathsf{T}}w} vw^{\mathsf{T}}$ and $P_{w^{\perp}} = I_n - P_v$ are (oblique) projection matrices,

    (iii) if $v^{\mathsf{T}}w < 0$, then

       (a) vw^{T} is continuous-time semi-convergent,

       (b) $\exp(vw^{\mathsf{T}}t) = P_{w^{\perp}} + \mathrm{e}^{(v^{\mathsf{T}}w)t} P_v$ and $\lim\limits_{t \to +\infty} \exp(vw^{\mathsf{T}}t) = P_{w^{\perp}}$, and

       (c) each solution to $\dot{x} = (vw^{\mathsf{T}})x$ satisfies $x(t) = P_{w^{\perp}}x(0) + \mathrm{e}^{(v^{\mathsf{T}}w)t} P_v x(0)$ and $x(t) \to P_{w^{\perp}}x(0)$ as $t \to +\infty$.

E7.4   **Equivalent convergence properties.** In this exercise we complete the proof of Theorem 7.4 about consensus for Laplacian matrices with a globally reachable node. Let L be a Laplacian matrix and let G be its associated digraph. As in Exercise E6.1, pick $\varepsilon < \frac{1}{d_{\max}}$, where $d_{\max}$ is the maximum out-degree, and define $\mathcal{A}_{L,\varepsilon} = I_n - \varepsilon L$. Let $w \in \mathbb{R}^n$, $w \geq 0$, and $\mathbb{1}_n^\mathsf{T} w = 1$ satisfy $w^\mathsf{T} L = \mathbb{0}_n^\mathsf{T}$ and, as in Exercise E5.9, define $P_L = \mathbb{1}_n w^\mathsf{T}$. Prove that the following statements are equivalent:

(i) $\lim_{k \to \infty} \mathcal{A}_{L,\varepsilon}^k = P_L$,

(ii) $\lim_{k \to \infty} (\mathcal{A}_{L,\varepsilon} - P_L)^k = \mathbb{0}_{n \times n}$,

(iii) $\rho(\mathcal{A}_{L,\varepsilon} - P_L) < 1$,

(iv) $\alpha(-L - P_L) < 0$,

(v) $\lim_{t \to \infty} \exp((-L - P_L)t) = \mathbb{0}_{n \times n}$,

(vi) $\lim_{t \to \infty} \exp(-Lt) = P_L$.

Hint: Recall the properties of P_L established in Exercise E5.9.

▶ E7.5   **Laplacian average consensus in directed networks.** Consider the directed network in figure below with arbitrary positive weights and its associated Laplacian
flow $\dot{x}(t) = -L(x(t))$.
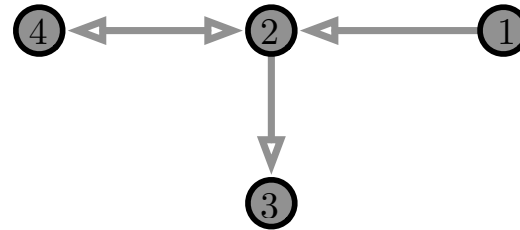
(i)   Can the network reach consensus, that is, as $t \to \infty$ does $x(t)$ converge to a limiting point in $\mathrm{span}\{\mathbb{1}_n\}$?
(ii) Does $x(t)$ achieve average consensus, that is, $\lim_{t\to\infty} x(t) = \mathrm{average}(x_0)\mathbb{1}_n$?
(iii) Will your answers change if you smartly add one directed edge and adapt the weights?

E7.6 **Convergence of discrete-time and continuous-time averaging.** Consider the following two weighted digraphs and their associated non-negative adjacency matrices A and Laplacian matrices L of appropriate dimensions. Consider the associated discrete-time iterations $x(k+1) = Ax(k)$ and continuous-time Laplacian flows $\dot{x}(t) = -Lx(t)$. For each of these two digraphs, argue about whether the discrete and/or continuous-time systems converge as time goes to infinity. If they converge, what value do they converge to?
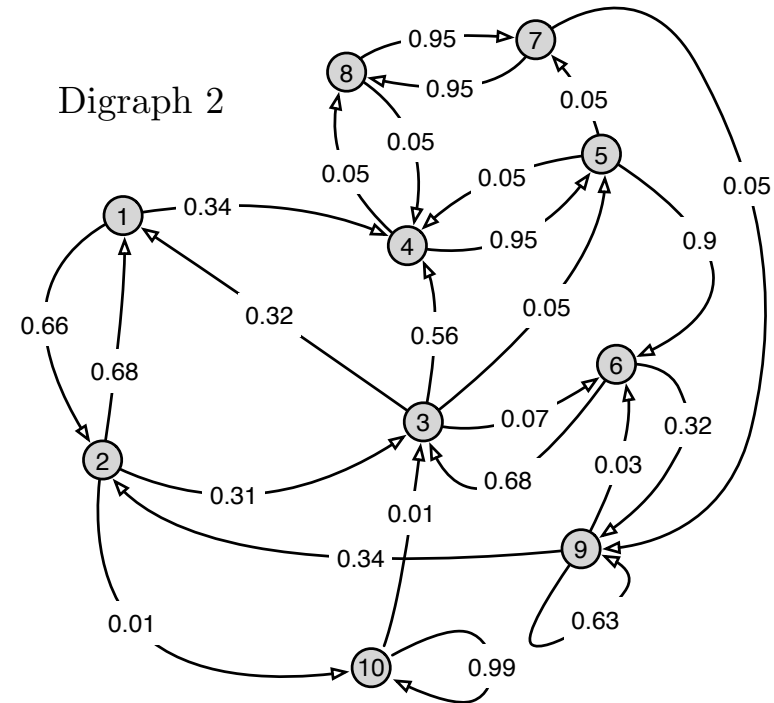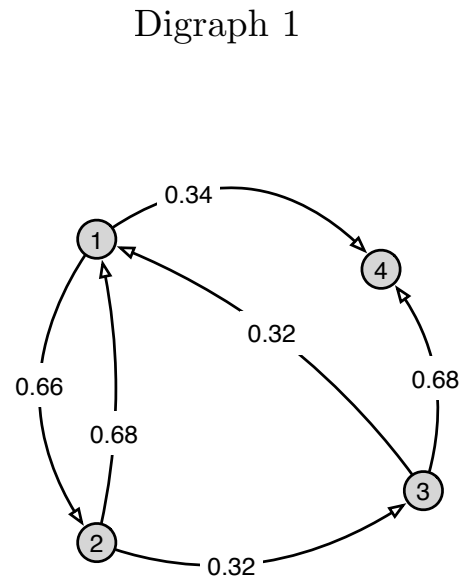
Figure E7.2: Two example weighted digraphs

E7.7   **Doubly-stochastic matrices on strongly-connected digraphs.** Given a strongly-connected unweighted digraph G, design weights along the edges of G (and possibly add self-loops) so that the weighted adjacency matrix is doubly-stochastic.

E7.8   **The Lyapunov inequality for the Laplacian matrix of a strongly-connected digraph.** Let L be the Laplacian matrix of a strongly-connected weighted digraph. Find $P = P^\mathsf{T} \succ 0$ such that

  (i) $PL + L^\mathsf{T}P \succeq 0$, and
  (ii) $(PL + L^\mathsf{T}P)\mathbb{1}_n = \mathbb{0}_n$.

Hint: Recall Section 7.4 and Exercise E6.4.

E7.9   **Delayed Laplacian flow.** Define the *delayed Laplacian flow dynamics* over a connected, weighted, and graph G by:

$$\dot{x}_i(t) = \sum_{j \in \mathcal{N}} a_{ij}(x_j(t - \tau) - x_i(t - \tau)), \qquad i \in \{1, \ldots, n\},$$

where $a_{ij} > 0$ is the weight on the edge $\{i, j\} \in E$, and $\tau > 0$ is a positive scalar delay term. The Laplace domain representation of the system is $X(s) = G(s)x(0)$ where $G(s)$ is associated transfer function

$$G(s) = (sI_n + e^{-s\tau}L)^{-1},$$

and $L = L^\mathsf{T} \in \mathbb{R}^{n \times n}$ is the network Laplacian matrix. Show that the transfer function $G(s)$ admits poles on the imaginary axis if the following *resonance* condition is true for an eigenvalue λ_i, $i \in \{1, \ldots, n\}$, of the Laplacian matrix:

$$\tau = \frac{\pi}{2\lambda_i}.$$

E7.10   **Robotic coordination: deployment, centering, and geometric optimization on the real line.** Consider $n \geq 3$ robots with dynamics $\dot{p}_i = u_i$, where $i \in \{1, \ldots, n\}$ is an index labeling each robot, $p_i \in \mathbb{R}$ is the position of robot i, and $u_i \in \mathbb{R}$ is a steering control input. For simplicity, assume that the robots are indexed according to their initial position: $p_1(0) \leq p_2(0) \leq \cdots \leq p_n(0)$. We consider the following distributed control laws to achieve some geometric configuration:

(i) *Move towards the centroid of your neighbors:* Each robot $i \in \{2, \ldots, n-1\}$ (having two neighbors) moves to the centroid of the local subset $\{p_{i-1}, p_i, p_{i+1}\}$:

$$\dot{p}_i = \frac{1}{3}(p_{i-1} + p_i + p_{i+1}) - p_i, \quad i \in \{2, \ldots, n-1\}. \tag{E7.1}$$

The robots $\{1, n\}$ (each having one neighbor) move to the centroid of the local subsets $\{p_1, p_2\}$ and $\{p_{n-1}, p_n\}$, respectively:

$$\dot{p}_1 = \frac{1}{2}(p_1 + p_2) - p_1 \quad \text{and} \quad \dot{p}_n = \frac{1}{2}(p_{n-1} + p_n) - p_n. \tag{E7.2}$$

Show that, by using the coordination laws (E7.1) and (E7.2), the robots asymptotically rendezvous.

(ii) *Move towards the centroid of your neighbors or walls:* Consider two walls at the positions $p_0 \leq p_1$ and $p_{n+1} \geq p_n$ so that all robots are contained between the walls. The walls are stationary, that is, $\dot{p}_0 = 0$ and $\dot{p}_{n+1} = 0$. Again, the robots $i \in \{2, \ldots, n-1\}$ (each having two neighbors) move to the centroid of the local subset $\{p_{i-1}, p_i, p_{i+1}\}$. The robots $\{1, n\}$ (each having one robotic neighbor and one neighboring wall) move to the centroid of the local subsets $\{p_0, p_1, p_2\}$ and $\{p_{n-1}, p_n, p_{n+1}\}$, respectively. Hence, the closed-loop robot dynamics are

$$\dot{p}_i = \frac{1}{3}(p_{i-1} + p_i + p_{i+1}) - p_i, \quad i \in \{1, \ldots, n\}. \tag{E7.3}$$

Show that, by using coordination law (E7.3), the robots become uniformly spaced on the interval $[p_0, p_{n+1}]$.

(iii) *Move away from the centroid of your neighbors or walls:* Again consider two stationary walls at $p_0 \leq p_1$ and $p_{n+1} \geq p_n$ containing the positions of all robots. We partition the interval $[p_0, p_{n+1}]$ into regions of interest, whereby each robot is assigned the territory containing all points closer to itself than to other robots. In other words, robot $i \in \{2, \ldots, n-1\}$ (having two neighbors) is assigned the region $\mathcal{V}_i = [(p_i + p_{i-1})/2, (p_{i+1} + p_i)/2]$, robot 1 is assigned the region $\mathcal{V}_1 = [p_0, (p_1 + p_2)/2]$, and robot n is assigned the region $\mathcal{V}_n = [(p_{n-1} + p_n)/2, p_{n+1}]$. We aim to design a distributed algorithm such that the robots are assigned asymptotically equal-sized regions. (This territory partition is called a *Voronoi partition*; see (Martínez et al., 2007) for further detail.) We consider the following simple coordination law, where each robot i heads for the midpoint $c_i(\mathcal{V}_i(p))$ of its partition $\mathcal{V}_i$:

$$\dot{p}_i = c_i(\mathcal{V}_i(p)) - p_i. \tag{E7.4}$$

Show that, by using the coordination law (E7.4), the robots' assigned regions asymptotically become equally large.

E7.11   **Robotic coordination and affine Laplacian flow.** Consider a group of $n = 4$ vehicles moving in the plane. Each vehicle $i \in \{1, \ldots, 4\}$ is described by its kinematics $\dot{x}_i = u_i$, where $x_i \in \mathbb{C}$ is the vehicle's position in the complex plane and $u_i \in \mathbb{C}$ is a steering command. The vehicle initial position in the complex plane is a square formation: $x(0) = \begin{bmatrix} 1 & \mathrm{i} & -1 & -\mathrm{i} \end{bmatrix}^{\mathsf{T}}$, where i is the imaginary unit. We aim to move the vehicles to a given final configuration. Specifically, we aim to achieve

$$\lim_{t \to \infty} x(t) = x_{\text{final}} = \begin{bmatrix} 0.5 + 0.5\mathrm{i} & -0.5 + 0.5\mathrm{i} & -0.5 - 0.5\mathrm{i} & 0.5 - 0.5\mathrm{i} \end{bmatrix}^{\mathsf{T}}. \tag{E7.5}$$

To achieve this goal, we propose the *complex affine averaging control law*

$$\dot{x}(t) = u(t) = -L(\alpha x(t) + \beta), \tag{E7.6}$$

where $\alpha > 0$ is a constant scalar gain, $\beta \in \mathbb{C}^n$ is a constant vector offset, and L is a Laplacian matrix of a strongly connected and weight-balanced digraph. Your tasks are the following:

  (i) Show that the affine Laplacian flow (E7.6) converges for any choice of $\alpha > 0$ and $\beta \in \mathbb{C}^n$.
 (ii) Characterize all the values of $\alpha > 0$ and $\beta \in \mathbb{C}^n$ such that the desired final configuration x_{final} is achieved by the affine Laplacian flow (E7.6).

Chapter 8

Diffusively-Coupled Linear Systems

In this chapter we study diffusive interconnection among identical linear systems and linear control systems. As example system, we study the so-called second-order Laplacian flow. The results in this chapter provide a first generalization of the consensus problem to more general classes of systems.

8.1   Diffusively-coupled linear systems

In this chapter, we consider an agent to be a continuous-time linear single-input single-output (SISO) dynamical systems with d-dimensional state, described by the matrices $\mathscr{A} \in \mathbb{R}^{d \times d}$, $\mathscr{B} \in \mathbb{R}^{d \times 1}$, and $\mathscr{C} \in \mathbb{R}^{1 \times d}$. The dynamics of the ith agent, for $i \in \{1, \ldots, n\}$, are

$$
\begin{aligned}
\dot{x}_i(t) &= \mathscr{A} x_i(t) + \mathscr{B} u_i(t), \\
y_i(t) &= \mathscr{C} x_i(t).
\end{aligned}
\tag{8.1}
$$

Here, $x_i \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^d$, $u_i \colon \mathbb{R}_{\geq 0} \to \mathbb{R}$, and $y_i \colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ are the state, input and output trajectories respectively.

The agents are interconnected through a weighted undirected graph G with edge weights $\{a_{ij}\}_{ij}$ and Laplacian matrix L; we will often assume L is symmetric. We assume that the input to each system is based on information received from only its immediate neighbors in G. Specifically, we consider the *output-dependent diffusive coupling law*

$$u_i(t) = \sum_{j=1}^{n} a_{ij}\big(y_j(t) - y_i(t)\big). \tag{8.2}$$

Note: in control theory terms, this interconnection law amounts to a static output feedback controller or, in this particular case, a proportional controller. We illustrate this interconnection in Figure 8.1. The closed-loop equations read

$$\dot{x}_i(t) = \mathscr{A}\, x_i(t) + \mathscr{B}\mathscr{C} \sum_{j=1}^{n} a_{ij}\big(x_j(t) - x_i(t)\big).$$
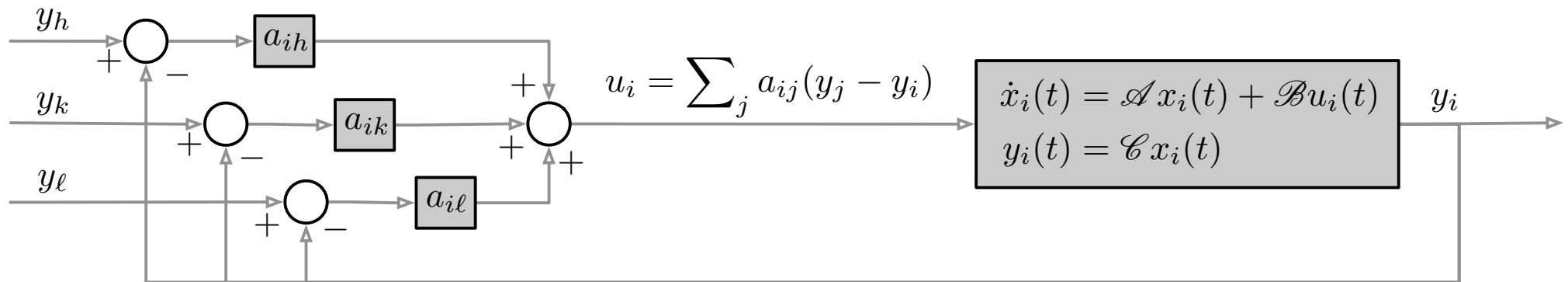
Figure 8.1: The output-dependent diffusive coupling law (8.2)

Definition 8.1. *A network of diffusively-coupled identical linear systems is composed by n identical continuous-time linear SISO systems $(\mathscr{A}, \mathscr{B}, \mathscr{C})$ and a Laplacian matrix L.*

For such interconnected systems we introduce a notion of asymptotic behavior that generalizes the asymptotic consensus achieved by the Laplacian flow, as studied in the previous chapter.

Definition 8.2. *A network of diffusively-coupled identical linear systems described by the triplet $(\mathscr{A}, \mathscr{B}, \mathscr{C})$ and the Laplacian L achieves asymptotic synchronization if, for all agents $i, j \in \{1, \ldots, n\}$ and all initial conditions, the solutions of (8.1) under feedback (8.2) satisfy*

$$\lim_{t \to \infty} \|x_i(t) - x_j(t)\|_2 = 0. \tag{8.3}$$

We note that, under mild connectivity assumptions, asymptotic synchronization is equivalent to the following property: there exists a trajectory $x_0 \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ such that, for all i,

$$\lim_{t \to \infty} \|x_i(t) - x_0(t)\|_2 = 0. \tag{8.4}$$

When this trajectory is known, we say that the system achieves *asymptotic synchronization on x_0*.

8.1.1 Second-order Laplacian flows

In this section we introduce an example of diffusively-coupled linear systems. We assume each node of the network is a so-called *double-integrator* (also referred to as *second-order dynamic*):

$$\ddot{q}_i = \bar{u}_i, \qquad \text{or, in first-order equivalent form,} \qquad \begin{cases} \dot{q}_i = v_i, \\ \dot{v}_i = \bar{u}_i, \end{cases} \tag{8.5}$$

where $\bar{u}_i$ is an appropriate control input signal to be designed.

We assume a weighted undirected graph describes the sensing and/or communication interactions among the agents with adjacency matrix A and Laplacian L. We also introduce constants $k_p, k_d \geq 0$ describing so-called *spring* and *damping* coefficients respectively, as well as constants $\gamma_p, \gamma_d \geq 0$ describing *position-averaging* and *velocity-averaging coefficients*. In summary, we consider the *proportional, derivative, position-averaging, and velocity-averaging control law*

$$\bar{u}_i = -k_p q_i - k_d \dot{q}_i + \sum_{j=1}^{n} a_{ij}\big(\gamma_p(q_j - q_i) + \gamma_d(\dot{q}_j - \dot{q}_i)\big). \tag{8.6}$$

A physical realization of this system as a spring/damper network is illustrated in Figure 8.2.
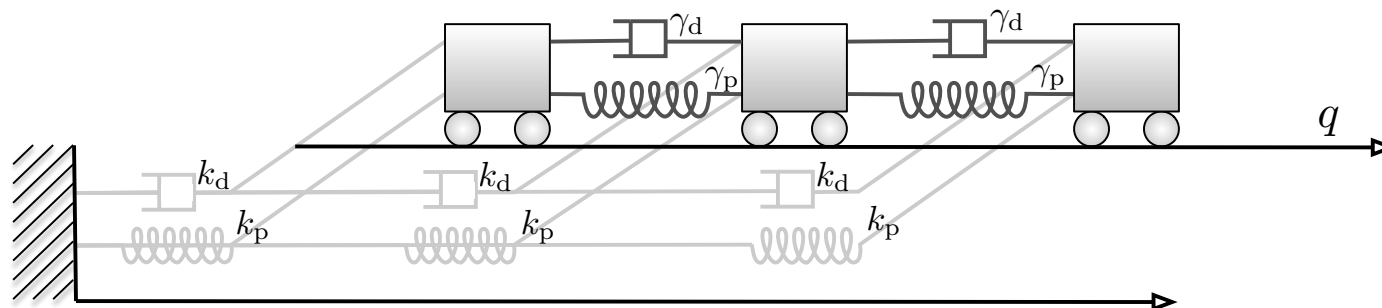
Figure 8.2: A network of unit-mass carts subject to spring and dampers gives rise to the second-order Laplacian flow (8.7). For illustration purposes, the springs and dampers connecting each cart to the left wall are drawn schematically, in light gray, as if they were overlapping and on a separate plane.

It is useful to rewrite the systems (8.5) interconnected via the law (8.6) in two useful manners. First, simply stacking each component into a vector, the corresponding closed-loop systems, called the *second-order Laplacian flow*, is

$$\ddot{q}(t) + (k_{\mathrm{d}}I_n + \gamma_{\mathrm{d}}L)\dot{q}(t) + (k_{\mathrm{p}}I_n + \gamma_{\mathrm{p}}L)q(t) = \mathbb{0}_n. \tag{8.7}$$

Second, we can rewrite $\bar{u}_i = -k_{\mathrm{p}}q_i - k_{\mathrm{d}}\dot{x}_i + u_i$, where $u_i = \sum_{j=1}^{n} a_{ij}\big(\gamma_{\mathrm{p}}(q_j - q_i) + \gamma_{\mathrm{d}}(\dot{q}_j - \dot{q}_i)\big)$ and define the matrices

$$\mathscr{A}_{\mathrm{msd}} = \begin{bmatrix} 0 & 1 \\ -k_{\mathrm{p}} & -k_{\mathrm{d}} \end{bmatrix}, \mathscr{B}_{\mathrm{msd}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \text{ and } \mathscr{C}_{\mathrm{msd}} = \begin{bmatrix} \gamma_{\mathrm{p}} & \gamma_{\mathrm{d}} \end{bmatrix}. \tag{8.8}$$

With these definitions one can see that the systems (8.5) interconnected via the law (8.6) is equivalent to:

$$\frac{d}{dt}\begin{bmatrix} q_i \\ v_i \end{bmatrix} = \mathscr{A}_{\mathrm{msd}}\begin{bmatrix} q_i \\ v_i \end{bmatrix} + \mathscr{B}_{\mathrm{msd}}u_i,$$

$$y_i = \mathscr{C}_{\mathrm{msd}}\begin{bmatrix} q_i \\ v_i \end{bmatrix}, \qquad u_i(t) = \sum_{j=1}^{n} a_{ij}\big(y_j(t) - y_i(t)\big).$$

In other words, the matrices in equation (8.8) describe the second-order Laplacian system as a network of diffusively-coupled identical linear systems.

In Table 8.1 we catalog some interesting special cases and we illustrate in Figure 8.3 the behavior of the systems corresponding to the first three rows of Table 8.1. The next sections in this chapter focus on establishing rigorously the collective emerging behavior observed in these simulations.

Name	Dynamics	Asymptotic behavior
Second-order averaging protocol	$k_{\mathrm{p}} = k_{\mathrm{d}} = 0, \gamma_{\mathrm{d}} = 1, \gamma_{\mathrm{p}} > 0$ $\implies$ $\ddot{q}(t) + L\dot{q}(t) + \gamma_{\mathrm{p}}Lq(t) = \mathbb{0}_n$	synchronization on a ramp Example: car platooning Ref: Theorem 8.7(i)
Harmonic oscillators with velocity averaging	$k_{\mathrm{d}} = \gamma_{\mathrm{p}} = 0, \gamma_{\mathrm{d}} = 1, k_{\mathrm{p}} > 0$ $\implies$ $\ddot{q}(t) + L\dot{q}(t) + k_{\mathrm{p}}q(t) = \mathbb{0}_n$	synchronization on harmonic oscillations Example E8.5: resonant inductor/capacitor circuits Ref: Theorem 8.7(ii)
Position-averaging with absolute velocity damping	$k_{\mathrm{p}} = \gamma_{\mathrm{d}} = 0, \gamma_{\mathrm{p}} = 1, k_{\mathrm{d}} > 0$ $\implies$ $\ddot{q}(t) + k_{\mathrm{d}}\dot{q}(t) + Lq(t) = \mathbb{0}_n$	synchronization on constant positions Example: rendezvous in multi-robot systems Example: swing dynamics in power networks Ref: Theorem 8.7(iii)
Laplacian oscillators	$k_{\mathrm{p}} = k_{\mathrm{d}} = \gamma_{\mathrm{d}} = 0, \gamma_{\mathrm{p}} = 1$ $\implies$ $\ddot{q}(t) + Lq(t) = \mathbb{0}_n$	superposition of ramp and harmonics Example 7.1.3: discretized wave equation Ref: Exercise E8.6

Table 8.1: Classification of second-order Laplacian flows arising from the general model in equation (8.7). Note: each of the first three examples satisfies $k_{\mathrm{p}} + \gamma_{\mathrm{p}} > 0$ and $k_{\mathrm{d}} + \gamma_{\mathrm{d}} > 0$.

Second-order averaging protocol

$$\ddot{q}(t) + L\dot{q}(t) + \gamma_{\mathrm{p}}Lq(t) = 0$$

(a) asymptotic synchronization on a ramp signal

Harmonic oscillators with velocity averaging

$$\ddot{q}(t) + L\dot{q}(t) + k_{\mathrm{p}}q(t) = 0$$

(b) asymptotic synchronization on an harmonic signal

Position-averaging with absolute velocity damping

$$\ddot{q}(t) + k_{\mathrm{d}}\dot{q}(t) + Lq(t) = 0$$

(c) asymptotic synchronization on constant weighted average
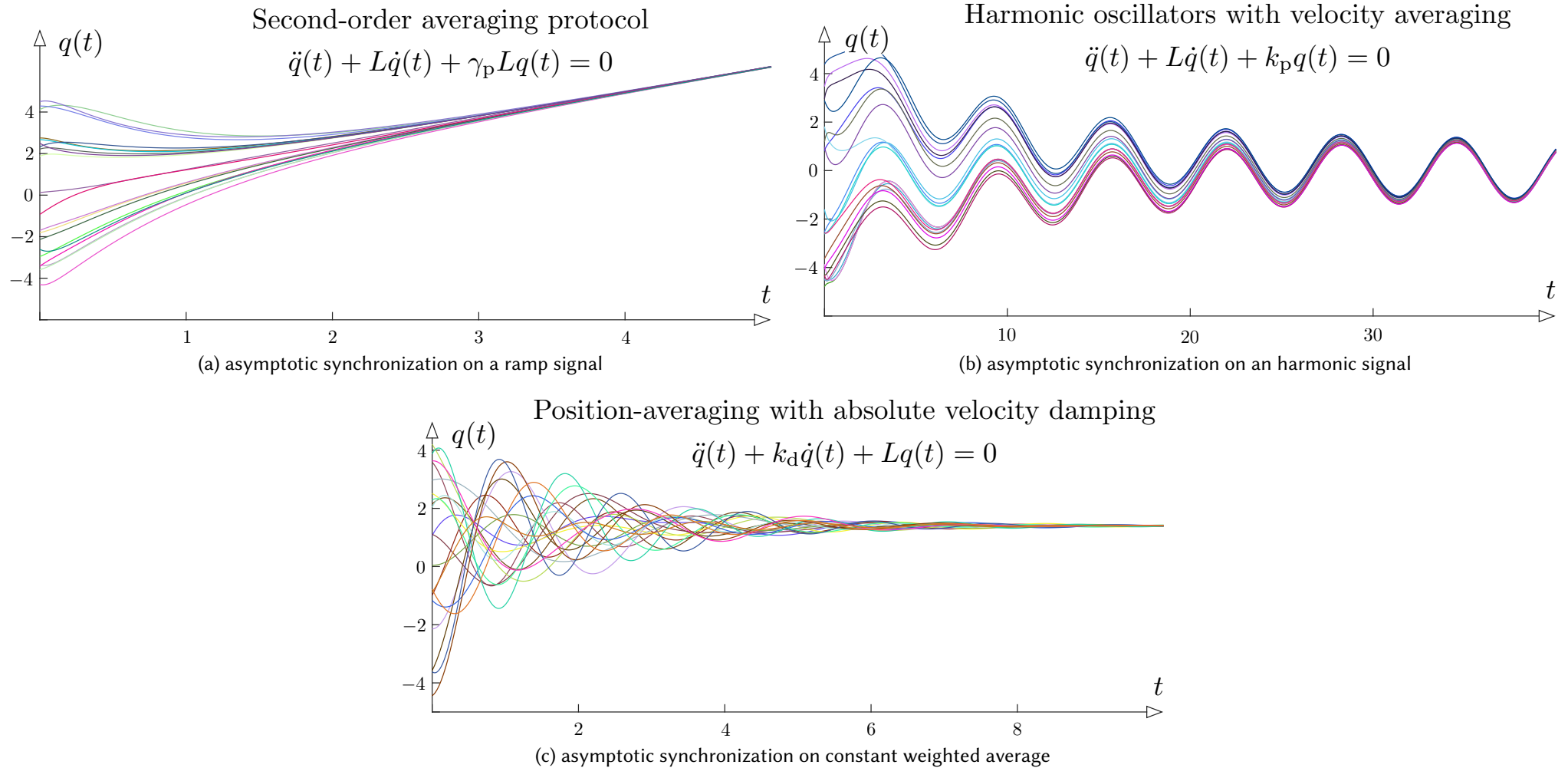
Figure 8.3: Representative trajectories of the second-order Laplacian flow (8.7) for a randomly-generated undirected graph with $n = 20$ nodes, random initial conditions, and the three choices of gains as cataloged in Table 8.1.

8.2 Modeling via Kronecker products

In this section we obtain a compact expression for the state matrix of a diffusively-coupled network of linear systems.

8.2.1   The Kronecker product

We start by introducing a useful tool. The *Kronecker product* of $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{q \times r}$ is the $nq \times mr$ matrix $A \otimes B$ given by

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \ddots & a_{nm}B \end{bmatrix}. \tag{8.9}$$

As simple example, we write

$$I_n \otimes B = \begin{bmatrix} B & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & B \end{bmatrix} \in \mathbb{R}^{nq \times nr} \quad \text{and} \quad A \otimes I_q = \begin{bmatrix} a_{11}I_q & \cdots & a_{1m}I_q \\ \vdots & \ddots & \vdots \\ a_{n1}I_q & \ddots & a_{nm}I_q \end{bmatrix} \in \mathbb{R}^{nq \times mq}. \tag{8.10}$$

Additionally, for $v, w \in \mathbb{R}^n$, we have $v \otimes w = \begin{bmatrix} v_1 w \\ \vdots \\ v_n w \end{bmatrix} \in \mathbb{R}^{n^2}$.

The Kronecker product enjoys numerous properties, including for example

the bilinearity property: $\quad (\alpha A + \beta B) \otimes (\gamma C + \delta D) = \alpha\gamma A \otimes C + \alpha\delta A \otimes D$

$$+ \beta\gamma B \otimes C + \beta\delta B \otimes D, \qquad (8.11a)$$

the associativity property: $\qquad (A \otimes B) \otimes C = A \otimes (B \otimes C), \qquad (8.11b)$

the transpose property: $\qquad (A \otimes B)^{\mathsf{T}} = A^{\mathsf{T}} \otimes B^{\mathsf{T}}, \qquad (8.11c)$

the mixed product property: $\qquad (A \otimes B)(C \otimes D) = (AC) \otimes (BD), \qquad (8.11d)$

where the A, B, C, D matrices have appropriate compatible dimensions.

The remarkable mixed product property (8.11d) leads to many useful consequences. As first example, if $Av = \lambda v$ and $Bw = \mu w$, then property (8.11d) implies

$$(A \otimes B)(v \otimes w) = (Av) \otimes (Bw) = (\lambda v) \otimes (\mu w) = \lambda\mu(v \otimes w).$$

Therefore, we know

the eigenpair property: $\quad Av = \lambda v, Bw = \mu w \implies (A \otimes B)(v \otimes w) = \lambda\mu(v \otimes w), \qquad (8.12)$

the spectrum property: $\qquad \mathrm{spec}(A \otimes B) = \{\lambda\mu \mid \lambda \in \mathrm{spec}(A), \mu \in \mathrm{spec}(B)\}. \qquad (8.13)$

A second consequence of property (8.11d) is that, for square matrices A and B, $A \otimes B$ is invertible if and only if both A and B are invertible, in which case

the inverse property: $\qquad (A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \qquad . \qquad (8.14)$

We ask the reader to prove these properties and establish other ones in Exercises E8.1 and E8.2.

8.2.2 The state matrix for a diffusively coupled system

We are now ready to provide a concise closed-form expression for the state matrix of the network.

Theorem 8.3 (Transcription of diffusively-coupled linear systems). *Consider a network of diffusively-coupled identical linear systems described by the system $(\mathscr{A}, \mathscr{B}, \mathscr{C})$ and the symmetric Laplacian matrix L. Then the following statements hold:*

(i) the open-loop system, output equation, and diffusive coupling law are, respectively,

$$
\begin{aligned}
\dot{\mathbf{x}}(t) &= (I_n \otimes \mathscr{A})\mathbf{x}(t) + (I_n \otimes \mathscr{B})\mathbf{u}(t), \\
\mathbf{y}(t) &= (I_n \otimes \mathscr{C})\mathbf{x}(t), \\
\mathbf{u}(t) &= -L\mathbf{y}(t),
\end{aligned}
$$

(ii) the closed-loop system is

$$
\dot{\mathbf{x}}(t) = (I_n \otimes \mathscr{A} - L \otimes \mathscr{BC})\mathbf{x}(t), \tag{8.15}
$$

where we adopt the notation $\mathbf{x} = \begin{bmatrix} x_1^{\mathsf{T}}, \ldots, x_n^{\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{nd}$, $\mathbf{u} = \begin{bmatrix} u_1, \ldots, u_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^n$, *and* $\mathbf{y} = \begin{bmatrix} y_1, \ldots, y_n \end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^n$.

Proof. We write the n coupled systems in a single vector-valued equation on the state space $\mathbb{R}^{dn}$ using the Kronecker product. As in equation (8.10), we stack the n dynamical systems to write

$$\dot{\mathbf{x}} = \begin{bmatrix} \mathscr{A} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & \mathscr{A} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathscr{B} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ddots & \mathscr{B} \end{bmatrix} \mathbf{u} = (I_n \otimes \mathscr{A})\mathbf{x} + (I_n \otimes \mathscr{B})\mathbf{u}.$$

Similarly we obtain $\mathbf{y} = (I_n \otimes \mathscr{C})\mathbf{x}$. Next, recalling the definition of Laplacian, we write the output-dependent diffusive coupling law (8.2) as $\mathbf{u} = -L\mathbf{y}$. Moreover, plugging in the output equation leads to

$$\dot{\mathbf{x}} = (I_n \otimes \mathscr{A})\mathbf{x} - (I_n \otimes \mathscr{B})L(I_n \otimes \mathscr{C})\mathbf{x}.$$

From the mixed product property in equation (8.11d), we obtain

$$(I_n \otimes \mathscr{B})L(I_n \otimes \mathscr{C}) = (I_n \otimes \mathscr{B})(L \otimes 1)(I_n \otimes \mathscr{C}) = (L \otimes \mathscr{B}\mathscr{C})$$

and, in turn, the closed loop (8.15). Note that $L \in \mathbb{R}^{n \times n}$, $\mathscr{B} \in \mathbb{R}^{d \times 1}$, and $\mathscr{C} \in \mathbb{R}^{1 \times d}$ together imply that $\mathscr{B}\mathscr{C}$ has dimensions $d \times d$ and that $L \otimes \mathscr{B}\mathscr{C}$ has dimensions $nd \times nd$, the same as $I_n \otimes \mathscr{A}$. Hence, equation (8.15) is dimensionally correct. This concludes the proof. ∎

8.3 The synchronization theorem

In this section we present the main result of this chapter. Define the *state average* $x_{\text{ave}} \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ by

$$\dot{x}_{\text{ave}}(t) = \mathscr{A} x_{\text{ave}}(t), \quad x_{\text{ave}}(0) = \frac{1}{n} \sum_{j=1}^{n} x_j(0),$$

and note that $x_{\text{ave}}(t) = \exp(\mathscr{A} t) x_{\text{ave}}(0)$. The following theorem characterizes when diffusively-coupled linear systems achieve asymptotic synchronization as in equation (8.3), that is,

$$\lim_{t \to \infty} \|x_i(t) - x_j(t)\|_2 = 0, \quad \text{for all } i, j \in \{1, \dots, n\}$$

and, more specifically, asymptotic synchronization on x_{ave} as in equation (8.4), that is,

$$\lim_{t \to \infty} \|x_i(t) - x_{\text{ave}}(t)\|_2 = 0, \quad \text{for all } i \in \{1, \dots, n\}.$$

Theorem 8.4 (Synchronization of output-dependent diffusively-coupled linear systems). *Consider a network of diffusively-coupled identical linear systems described by the system $(\mathscr{A}, \mathscr{B}, \mathscr{C})$, the symmetric Laplacian L, and the closed-loop dynamics (8.15). Let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of L. The following statements hold:*

(i) the system achieves asymptotic synchronization on x_{ave} if and only if each matrix $\mathscr{A} - \lambda_i \mathscr{B}\mathscr{C}$, $i \in \{2, \ldots, n\}$, is Hurwitz;

(ii) the system is exponentially stable if and only if each matrix $\mathscr{A} - \lambda_i \mathscr{B}\mathscr{C}$, $i \in \{1, \ldots, n\}$, is Hurwitz.

8.3.1   Remarks and special case

It is useful now to offer some comments and consider a special case.

Remark 8.5.   *(i) If G is connected (i.e., $\lambda_2 > 0$), then it is possible for each matrix $\mathscr{A} - \lambda_i \mathscr{B}\mathscr{C}$, $i \in \{2, \ldots, n\}$, to be Hurwitz while $\mathscr{A}$ is not. Therefore, it is possible for a system to synchronize and not be exponentially stable.*
(ii) If G is disconnected, then synchronization and exponential stability are equivalent and imply that $\mathscr{A}$ is Hurwitz.
(iii) In other words, if $\mathscr{A}$ is not Hurwitz, then synchronization is possible if and only if the graph is connected.   ●

Corollary 8.6 (State-dependent diffusive coupling).   *Consider the state-dependent diffusive-coupling case:*

$$\dot{x}_i = \mathscr{A}x_i + \sum_{j=1}^{n} a_{ij}(x_j - x_i),$$

that is, assume $\mathscr{B}\mathscr{C} = I_d$ in the setup of Theorem 8.4. Let $\alpha(\mathscr{A})$ denote the spectral abscissa of $\mathscr{A}$. Then the following statements are equivalent
(i) the system achieves asymptotic synchronization on the state average,
(ii) each $\mathscr{A} - \lambda_i I_d$, $i \in \{2, \ldots, n\}$, is Hurwitz, that is, $\mathscr{A} - \lambda_2 I_d$ is Hurwitz, and
(iii) the algebraic connectivity of L dominates the spectral abscissa of $\mathscr{A}$, that is, $\alpha(\mathscr{A}) < \lambda_2$.

8.3.2 Proof

Proof of Theorem 8.4. Let $L = U\Lambda U^\mathsf{T}$ be the eigen-decomposition of L, where U is an orthonormal matrix and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Consider the change of variable $\mathbf{z} = (U^\mathsf{T} \otimes I_d)\mathbf{x}$ and note $(U^\mathsf{T} \otimes I_d)(U \otimes I_d) = I_{nd}$. Compute:

$$\begin{aligned}
\dot{\mathbf{z}} &= (U^\mathsf{T} \otimes I_d)(I_n \otimes \mathscr{A} - L \otimes \mathscr{BC})(U \otimes I_d)\mathbf{z} \\
&= \Big((U^\mathsf{T} I_n U) \otimes (I_d \mathscr{A} I_d) - (U^\mathsf{T} L U) \otimes (I_d \mathscr{BC} I_d)\Big)\mathbf{z} \\
&= (I_n \otimes \mathscr{A} - \Lambda \otimes \mathscr{BC})\mathbf{z}.
\end{aligned}$$

Now, note that the matrix $(I_n \otimes \mathscr{A} - \Lambda \otimes \mathscr{BC})$ is block diagonal because

$$I_n \otimes \mathscr{A} = \begin{bmatrix} \mathscr{A} & & \\ & \ddots & \\ & & \mathscr{A} \end{bmatrix}, \quad \text{and} \quad \Lambda \otimes \mathscr{BC} = \begin{bmatrix} \lambda_1 \mathscr{BC} & & \\ & \ddots & \\ & & \lambda_n \mathscr{BC} \end{bmatrix}.$$

This block diagonal form immediately implies statement (ii).

Next, recalling that $\lambda_1 = 0$ and that $\mathbb{e}_1$ is the first element of the canonical basis of $\mathbb{R}^n$, we write the matrix exponential of the block-diagonal matrix $(I_n \otimes \mathscr{A} - \Lambda \otimes \mathscr{B}\mathscr{C})$ to obtain:

$$
\mathbf{z}(t) = \begin{bmatrix} \exp(\mathscr{A}t) & & & \\ & \exp\big((\mathscr{A} - \lambda_2 \mathscr{B}\mathscr{C})t\big) & & \\ & & \ddots & \\ & & & \exp\big((\mathscr{A} - \lambda_n \mathscr{B}\mathscr{C})t\big) \end{bmatrix} \mathbf{z}(0)
$$

$$
= \begin{bmatrix} \exp(\mathscr{A}t)z_1(0) \\ \mathbb{0}_d \\ \vdots \\ \mathbb{0}_d \end{bmatrix} + \begin{bmatrix} \mathbb{0}_{d\times d} & & & \\ & \exp\big((\mathscr{A} - \lambda_2 \mathscr{B}\mathscr{C})t\big) & & \\ & & \ddots & \\ & & & \exp\big((\mathscr{A} - \lambda_n \mathscr{B}\mathscr{C})t\big) \end{bmatrix} \mathbf{z}(0)
$$

$$
= \Big((\mathbb{e}_1 \mathbb{e}_1^{\mathsf{T}}) \otimes \exp(\mathscr{A}t)\Big)\mathbf{z}(0) + \mathbf{z}_{\text{transient}}(t).
$$

Here the vector-valued function $\mathbf{z}_{\text{transient}} \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^{nd}$ contains all terms of the form $\exp\big((\mathscr{A} - \lambda_i \mathscr{B}\mathscr{C})t\big)z_i(0)$, for $i \in \{2, \dots, n\}$. We note that $\mathbf{z}_{\text{transient}}$ is exponentially vanishing as $t \to \infty$ for all initial conditions $\mathbf{z}(0)$ if and only if each matrix $\mathscr{A} - \lambda_i \mathscr{B}\mathscr{C}$, $i \in \{2, \dots, n\}$ is Hurwitz.

Next, we compute

$$(U \otimes I_d)\Big((\mathbb{e}_1\mathbb{e}_1^\mathsf{T}) \otimes \exp(\mathscr{A}t)\Big)\mathbf{z}(0) = (U \otimes I_d)\Big((\mathbb{e}_1\mathbb{e}_1^\mathsf{T}) \otimes \exp(\mathscr{A}t)\Big)(U^\mathsf{T} \otimes I_d)\mathbf{x}(0)$$

$$= \Big((U\mathbb{e}_1\mathbb{e}_1^\mathsf{T}U^\mathsf{T}) \otimes \exp(\mathscr{A}t)\Big)\mathbf{x}(0).$$

From the eigen-decomposition we know that the first column of U is the first eigenvector of L normalized to have unit 2-norm. Since the first (i.e., smallest) eigenvalue 0 of L has eigenvector $\mathbb{1}_n$ (or any multiple thereof), we know that $U\mathbb{e}_1 = \mathbb{1}_n/\sqrt{n}$ and also that $\mathbb{e}_1^\mathsf{T}U^\mathsf{T} = \mathbb{1}_n^\mathsf{T}/\sqrt{n}$. In summary, with $\mathbf{x}_{\text{transient}}(t) = (U \otimes I_d)\mathbf{z}_{\text{transient}}(t)$, we obtain

$$\mathbf{x}(t) = \frac{1}{n}\Big((\mathbb{1}_n\mathbb{1}_n^\mathsf{T}) \otimes \exp(\mathscr{A}t)\Big)\mathbf{x}(0) + \mathbf{x}_{\text{transient}}(t)$$

$$= \frac{1}{n}\Big(I_n \otimes \exp(\mathscr{A}t)\Big)\Big((\mathbb{1}_n\mathbb{1}_n^\mathsf{T}) \otimes I_d\Big)\mathbf{x}(0) + \mathbf{x}_{\text{transient}}(t)$$

$$= \frac{1}{n}\begin{bmatrix} \exp(\mathscr{A}t) & & \\ & \ddots & \\ & & \exp(\mathscr{A}t) \end{bmatrix}\begin{bmatrix} I_d & \cdots & I_d \\ \vdots & \ddots & \vdots \\ I_d & \cdots & I_d \end{bmatrix}\begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix} + \mathbf{x}_{\text{transient}}(t),$$

so that the solution to each system i satisfies

$$x_i(t) = \exp(\mathscr{A}t)\Big(\frac{1}{n}\sum_{j=1}^n x_j(0)\Big) + h_i(t).$$

Finally, we recall that $\mathbf{z}_{\text{transient}}$, and therefore $\mathbf{x}_{\text{transient}}$, is exponentially vanishing for all initial conditions $\mathbf{z}(0)$, and therefore $\mathbf{x}(0)$, if and only if each matrix $\mathscr{A} - \lambda_i\mathscr{B}\mathscr{C}$, $i \in \{2, \ldots, n\}$ is Hurwitz. This concludes the proof of statement (i). ∎

8.3.3   Synchronization in second-order Laplacian systems

We now apply to second-order Laplacian systems the theoretical results obtained in the synchronization Theorem 8.4. Unlike for the general case, it is possible to obtain quite explicit results.

First, recall that second-order Laplacian systems are diffusively-coupled linear systems with matrices $(\mathscr{A}_{\mathrm{msd}}, \mathscr{B}_{\mathrm{msd}}, \mathscr{C}_{\mathrm{msd}})$ and with Laplacian interconnection matrix L. As before, let $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of L. We compute:

$$\mathscr{A}_{\mathrm{msd}} - \lambda_i \mathscr{B}_{\mathrm{msd}} \mathscr{C}_{\mathrm{msd}} = \begin{bmatrix} 0 & 1 \\ -k_{\mathrm{p}} & -k_{\mathrm{d}} \end{bmatrix} - \lambda_i \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} \gamma_{\mathrm{p}} & \gamma_{\mathrm{d}} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ -(k_{\mathrm{p}} + \lambda_i \gamma_{\mathrm{p}}) & -(k_{\mathrm{d}} + \lambda_i \gamma_{\mathrm{d}}) \end{bmatrix}.$$

In other words, the i-th subsystem $\mathscr{A}_{\mathrm{msd}} - \lambda_i \mathscr{B}_{\mathrm{msd}} \mathscr{C}_{\mathrm{msd}}$ is a spring/damper system with effective spring coefficient $k_{\mathrm{p}} + \lambda_i \gamma_{\mathrm{p}}$ and effective damper coefficient $k_{\mathrm{d}} + \lambda_i \gamma_{\mathrm{d}}$. Based on a well known result, it is easy to see that

$$\mathscr{A}_{\mathrm{msd}} - \lambda_i \mathscr{B}_{\mathrm{msd}} \mathscr{C}_{\mathrm{msd}} \quad \text{is Hurwitz} \quad \iff \quad k_{\mathrm{p}} + \lambda_i \gamma_{\mathrm{p}} > 0 \text{ and } k_{\mathrm{d}} + \lambda_i \gamma_{\mathrm{d}} > 0.$$

Next, as state average system, we define the average mass/spring/damper system by

$$\frac{d}{dt} \begin{bmatrix} q_{\mathrm{ave}}(t) \\ \dot{q}_{\mathrm{ave}}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k_{\mathrm{p}} & -k_{\mathrm{d}} \end{bmatrix} \begin{bmatrix} q_{\mathrm{ave}}(t) \\ \dot{q}_{\mathrm{ave}}(t) \end{bmatrix}, \tag{8.16}$$

where we set $q_{\mathrm{ave}}(0) = \sum_{j=1}^{n} q_j(0)$ and $\dot{q}_{\mathrm{ave}}(0) = \sum_{j=1}^{n} \dot{q}_j(0)$.

These observations lead to the main synchronization result for second-order Laplacian flows.

Theorem 8.7 (Synchronization of second-order Laplacian flows). *Consider the second-order Laplacian flow (8.7). Assume that the undirected graph associated to L is connected. If $k_p + \gamma_p > 0$ and $k_d + \gamma_d > 0$, then*

(i) the system (8.7) achieves asymptotic synchronization in the sense that

$$\lim_{t \to \infty} \|q_i(t) - q_j(t)\|_2 = \lim_{t \to \infty} \|\dot{q}_i(t) - \dot{q}_j(t)\|_2 = 0, \quad \text{for all } i, j \in \{1, \dots, n\};$$

(ii) each trajectory asymptotically converges to the state average trajectory in the sense that

$$\lim_{t \to \infty} \|q_i(t) - q_{ave}(t)\|_2 = 0, \quad \text{for all } i \in \{1, \dots, n\}.$$

Specifically:

(i) the second-order averaging protocol ($k_p = k_d = 0$, $\gamma_d = 1$, $\gamma_p > 0$, first row Table 8.1), achieves asymptotic consensus on a ramp signal, that is, as $t \to \infty$,

$$q(t) \to \left(q_{ave}(0) + \dot{q}_{ave}(0)t \right) \mathbb{1}_n;$$

(ii) the harmonic oscillators with velocity averaging ($k_d = \gamma_p = 0$, $\gamma_d = 1$, $k_p > 0$, second row Table 8.1), achieve asymptotic consensus on an harmonic signal, that is, as $t \to \infty$,

$$q(t) \to \left(q_{ave}(0) \cos(\sqrt{k_p}t) + \frac{1}{\sqrt{k_p}} \dot{q}_{ave}(0) \sin(\sqrt{k_p}t) \right) \mathbb{1}_n;$$

(iii) the position-averaging flow with absolute velocity damping ($k_p = \gamma_d = 0$, $\gamma_p = 1$, $k_d > 0$, third row Table 8.1), achieves asymptotic consensus on a weighted average value, that is, as $t \to \infty$

$$q(t) \to \left(q_{ave}(0) + \dot{q}_{ave}(0)/k_d \right) \mathbb{1}_n.$$

The asymptotic behavior of the dynamical systems as classified in the three scenarios of this theorem and defined in the first three rows of Table 8.1 is consistent with the empirical observations in Figure 8.3.

8.4 Control design for synchronization

We now generalize the study of diffusively-coupled systems in three ways: (1) we assume the interconnection graph is directed, (2) we consider a multi-input multi-output (MIMO) interconnection, and, most importantly, (3) we consider a control design problem, instead of a stability analysis problem.

For simplicity, we consider the setting of state feedback. While the transcription and stability analysis method is very similar to that in the previous sections, the method of proof for digraph interconnections relies upon a transcription into Jordan normal form instead of a diagonalization procedure. We also review various stabilizability notions from linear control theory.

8.4.1   Problem statement

In this section, we consider an agent to be a continuous-time linear control systems with d-dimensional state and p-dimensional input, described by the matrices $\mathscr{A} \in \mathbb{R}^{d \times d}$ and $\mathscr{B} \in \mathbb{R}^{d \times p}$. The dynamics of the ith agent, for $i \in \{1, \ldots, n\}$, are

$$\dot{x}_i = \mathscr{A} x_i + \mathscr{B} u_i, \tag{8.17}$$

where $x_i \in \mathbb{R}^d$ is the state and $u_i \in \mathbb{R}^p$ is the control input.

The agents communicate along the edges of a weighted directed graph G with edge weights $\{a_{ij}\}_{ij}$ and Laplacian matrix L. We assume each agent regulates its own control signal based on information received from only its immediate in-neighbors in G.

The problem statement is as follows: design a control law that, based only on the information obtained through communication, achieves asymptotic synchronization in the sense of equation (8.3), that is, for all agents i and j and all initial conditions,

$$\lim_{t \to \infty} \|x_i(t) - x_j(t)\|_2 = 0.$$

Consider the *(state-dependent) diffusive coupling law*:

$$u_i(t) = cK \sum_{j=1}^{n} a_{ij} \big(x_j(t) - x_i(t)\big), \quad \text{for } i \in \{1, \ldots, n\}, \tag{8.18}$$

where the scalar $c > 0$ is a coupling gain and $K \in \mathbb{R}^{p \times d}$ is a control gain matrix. Note that this interconnection law amounts to a static feedback controller.

Before solving this problem we generalize the transcription and synchronization Theorems 8.3 and 8.4 to this setting. The instructive proof of the following result is postponed to Section 8.4.5.

Theorem 8.8 (Transcription and synchronization of MIMO systems over digraphs). *Consider n identical continuous-time linear control systems described by the couple $(\mathscr{A}, \mathscr{B})$ and a digraph with Laplacian L and with eigenvalues $0 = \lambda_1, \lambda_2, \ldots, \lambda_n$. Let w denote the dominant left eigenvector of L satisfying $\mathbb{1}_n^\mathsf{T} w = 1$ and define the* weighted state average $x_{\mathsf{ave},w} \colon \mathbb{R}_{\geq 0} \to \mathbb{R}^d$ *by*

$$x_{\mathsf{ave},w}(t) = \exp(\mathscr{A}t)\Big(\sum_{j=1}^{n} w_j x_j(0) \Big). \tag{8.19}$$

The following statements hold:

(i) the open-loop system and the diffusive coupling law are, respectively,

$$\dot{\mathbf{x}}(t) = (I_n \otimes \mathscr{A})\mathbf{x}(t) + (I_n \otimes \mathscr{B})\mathbf{u}(t),$$
$$\mathbf{u}(t) = -c(I_n \otimes K)(L \otimes I_d)\mathbf{x}(t),$$

and the closed-loop system is

$$\dot{\mathbf{x}} = \big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K) \big)\mathbf{x}; \tag{8.20}$$

where we adopt the notation $\mathbf{x} = \big[x_1^\mathsf{T}, \ldots, x_n^\mathsf{T} \big]^\mathsf{T} \in \mathbb{R}^{nd}$ *and* $\mathbf{u} = \big[u_1^\mathsf{T}, \ldots, u_n^\mathsf{T} \big]^\mathsf{T} \in \mathbb{R}^{np}$;

(ii) the closed-loop system (8.20) achieves asymptotic synchronization on $x_{\mathsf{ave},w}$ if and only if each (possibly complex) matrix $\mathscr{A} - c\lambda_i \mathscr{B}K$, $i \in \{2, \ldots, n\}$, is Hurwitz.

Note: Assume that G contains a globally reachable node. Then one can show the following converse result: if $\mathbf{x}$ achieves asymptotic consensus for all initial conditions, then each matrix $\mathscr{A} - c\lambda_i \mathscr{B}K$, $i \in \{2, \ldots, n\}$, is Hurwitz.

8.4.2 Stabilizability of linear control systems

We now review from linear control theory the notion of stabilizability and stabilizing feedback gain design.

Given matrices $\mathscr{A} \in \mathbb{R}^{d \times d}$ and $\mathscr{B} \in \mathbb{R}^{d \times p}$, a *continuous-time linear control systems* with d-dimensional state and p-dimensional input is

$$\dot{x} = \mathscr{A}x + \mathscr{B}u, \tag{8.21}$$

where $x_i \in \mathbb{R}^d$ is the state and $u_i \in \mathbb{R}^p$ is the control input.

Given a feedback gain matrix $K \in \mathbb{R}^{d \times n}$, the feedback control signal $u = -Kx$ gives rise to the closed-loop linear system

$$\dot{x} = \mathscr{A}x + \mathscr{B}(-Kx) = (\mathscr{A} - \mathscr{B}K)x.$$

Definition 8.9. *The linear control system* $(\mathscr{A}, \mathscr{B})$ *is* *stabilizable* *if there exists a matrix* K *such that* $\mathscr{A} - \mathscr{B}K$ *is Hurwitz.*

In other words, the closed-loop system is exponentially stable.

Theorem 8.10 (Stabilizability of linear control systems). *Given matrices $\mathscr{A} \in \mathbb{R}^{d \times d}$ and $\mathscr{B} \in \mathbb{R}^{d \times p}$, the following statements are equivalent*

 (i) the linear control system (8.21) *is stabilizable,*

(ii) there exists a $d \times d$ matrix $P \succ 0$ solving the (stabilizability) Lyapunov inequality

$$\mathscr{A} P + P \mathscr{A}^{\mathsf{T}} - 2 \mathscr{B} \mathscr{B}^{\mathsf{T}} \prec 0. \tag{8.22}$$

Moreover, for any $P \succ 0$ satisfying the inequality (8.22)*, a stabilizing feedback gain matrix is $K = \mathscr{B}^{\mathsf{T}} P^{-1}$.*

We refer for example to (Hespanha, 2009) for a complete treatment of linear systems theory, including a detailed discussion of stabilizability. We recall that the Lyapunov matrix inequality can be solved easily as a *linear matrix inequality (LMI)* (LMI) problem. We refer to (Boyd et al., 1994) for a detailed treatment of control problems solved via linear matrix inequalities.

8.4.3 High-gain LMI design

Consider now the following algorithm to design the control gain matrix K and the coupling gain c. Recall the Lyapunov matrix equation (8.22) and the fact that it can be solved via an LMI solver.

High-gain LMI design

Input: the stabilizable pair $(\mathscr{A}, \mathscr{B})$
Output: a control gain matrix K and coupling gain c

  1: set $P :=$ any solution to the linear matrix equality $\mathscr{A}P + P\mathscr{A}^\mathsf{T} - 2\mathscr{B}\mathscr{B}^\mathsf{T} \prec 0$
  2: set $K := \mathscr{B}^\mathsf{T} P^{-1}$
  3: set $c := 1/\min\{\Re(\lambda_i) \mid i \in \{2, \ldots, n\}\}$

Note: the design of K depends upon only the dynamics of each agent and the design of c depends upon only the communication graph.

Theorem 8.11 (High-gain LMI design for stabilizable linear control systems). *Consider n identical continuous-time linear control systems described by the couple $(\mathscr{A}, \mathscr{B})$ and a digraph G with Laplacian L and with complex eigenvalues $0 = \lambda_1, \lambda_2, \ldots, \lambda_n$. If the pair $(\mathscr{A}, \mathscr{B})$ is stabilizable and the digraph G contains a globally reachable node, then*

(i) the high-gain LMI design algorithm is well posed in the sense that a solution matrix P exists positive definite and the scalar c is well defined, and

(ii) the resulting pair (K, c) ensures that each (possibly complex) matrix $\mathscr{A} - c\lambda_i \mathscr{B} K$, $i \in \{2, \ldots, n\}$, is Hurwitz.

Note: the last two theorems reduce the problem of analyzing a dynamical system of dimension nd to the analysis of objects of dimensions n (the Laplacian L) and d (the linear matrix equality in P).

Proof. Fact (i) is a direct consequence of Theorem 8.10 about the stabilizability of linear control systems.

Regarding fact (ii), let P be the positive definite matrix computed in the high-gain LMI design algorithm. Given a square complex matrix $A \in \mathbb{C}^{n \times n}$, recall that (i) A^{H} denotes the conjugate transpose of A, and (ii) the Lyapunov inequality ensuring that A is Hurwitz is $AP + PA^{\mathsf{H}} \prec 0$. With these concepts, the Lyapunov equation for the ith complex subsystem, $i \in \{2, \dots, n\}$, is:

$$(\mathscr{A} - c\lambda_i \mathscr{B}K)P + P(\mathscr{A} - c\lambda_i \mathscr{B}K)^{\mathsf{H}}$$
$$= \mathscr{A}P + P\mathscr{A}^{\mathsf{T}} - c\Big(\lambda_i \mathscr{B}KP + \bar{\lambda}_i PK^{\mathsf{T}}\mathscr{B}^{\mathsf{T}}\Big)$$
$$\prec 2\mathscr{B}\mathscr{B}^{\mathsf{T}} - c\Big(\lambda_i \mathscr{B}(\mathscr{B}^{\mathsf{T}}P^{-1})P + \bar{\lambda}_i P(\mathscr{B}^{\mathsf{T}}P^{-1})^{\mathsf{T}}\mathscr{B}^{\mathsf{T}}\Big),$$

where we used two statements from the high-gain LMI design: $\mathscr{A}P + P\mathscr{A}^{\mathsf{T}} - 2\mathscr{B}\mathscr{B}^{\mathsf{T}} \prec 0$ and $K = \mathscr{B}^{\mathsf{T}}P^{-1}$. Performing all simplifications, we obtain

$$(\mathscr{A} - c\lambda_i \mathscr{B}K)P + P(\mathscr{A} - c\lambda_i \mathscr{B}K)^{\mathsf{H}} \prec 2\mathscr{B}\mathscr{B}^{\mathsf{T}} - 2c\Re(\lambda_i)\mathscr{B}\mathscr{B}^{\mathsf{T}}$$
$$\prec 2\big(1 - c\Re(\lambda_i)\big)\mathscr{B}\mathscr{B}^{\mathsf{T}}.$$

For any $c \geq 1/\min\{\Re(\lambda_i) \mid i \in \{2, \dots, n\}\}$, we know that $c\Re(\lambda_i) \geq 1$ and therefore $1 - c\Re(\lambda_i) \leq 0$. In summary we have proved that

$$(\mathscr{A} - c\lambda_i \mathscr{B}K)P + P(\mathscr{A} - c\lambda_i \mathscr{B}K)^{\mathsf{H}} \prec 0.$$

Therefore the (complex) linear system $\dot{x} = (\mathscr{A} - c\lambda_i \mathscr{B}K)x$ is exponentially stable. ∎

8.4.4   Extension to output feedback design

We now present the basic concepts about the problem of output feedback synchronization. As in equation (8.1), the agent is now an input/output control system described by

$$\dot{x}_i(t) = \mathscr{A}\, x_i(t) + \mathscr{B}u_i(t),$$
$$y_i(t) = \mathscr{C}\, x_i(t). \tag{8.23}$$

Here $x_i \in \mathbb{R}^d$ is the state, $u_i \in \mathbb{R}^p$ is the control input, and $y_i \in \mathbb{R}^q$ is the output signal. Each agent receives the signal

$$\zeta_i = c \sum_{j=1}^{n} a_{ij}(y_i - y_j), \tag{8.24}$$

and executes the following *observer-based diffusive coupling law*

$$\dot{v}_i = (\mathscr{A} - \mathscr{B}K)v_i + F\Big(c \sum_{j=1}^{n} a_{ij}\mathscr{C}(v_i - v_j) - \zeta_i \Big),$$
$$u_i = -Kv_i. \tag{8.25}$$

Here c is a coupling gain, v_i is the protocol state, and K and F are control and observer gain matrices to be designed.

  One can show the following generalization of Theorem 8.8: if each matrix $\mathscr{A} - \mathscr{B}K$ and $\mathscr{A} - c\lambda_i F\mathscr{C}$, $i \in \{2, \ldots, n\}$, is Hurwitz, then the n input/output control systems (8.23) in closed loop with the observer-based diffusive coupling law (8.24)-(8.25) achieve asymptotic synchronization in the state and protocol state variables. We refer the interested reader to (Li et al., 2010; Li and Duan, 2014) for design methods to compute appropriate gain parameters c, K and F.

8.4.5   Proof of synchronization over directed graphs

Proof of Theorem 8.8. To prove statement (i), we proceed as in the proof of Theorem 8.3(i). We stack the n dynamical systems (8.17) to obtain $\dot{\mathbf{x}} = (I_n \otimes \mathscr{A})\mathbf{x} + (I_n \otimes \mathscr{B})\mathbf{u}$. We write the diffusive coupling law (8.18) as

$$\mathbf{u}(t) = -c(I_n \otimes K)\mathbf{z}(t), \quad \text{where} \quad z_i(t) = \sum_{j=1}^{n} a_{ij}\big(x_i(t) - x_j(t)\big) = \sum_{j=1}^{n} \ell_{ij}x_j(t). \qquad (8.26)$$

where ℓ_{ij} is the (ij) entry of the Laplacian L. The last equality is equivalent to $\mathbf{z}(t) = (L \otimes I_d)\mathbf{x}(t)$, so that $\mathbf{u}(t) = -c(I_n \otimes K)(L \otimes I_d)\mathbf{x}($
Finally, the mixed product property (8.11d) implies

$$\dot{\mathbf{x}} = (I_n \otimes \mathscr{A})\mathbf{x} - (I_n \otimes \mathscr{B})c(I_n \otimes K)(L \otimes I_d)\mathbf{x} = \big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K)\big)\mathbf{x}. \qquad (8.27)$$

Note that $L \in \mathbb{R}^{n \times n}$, $\mathscr{B} \in \mathbb{R}^{d \times p}$ and $K \in \mathbb{R}^{p \times d}$ together imply that $\mathscr{B}K$ has dimensions $d \times d$ and that $L \otimes \mathscr{B}K$ has dimensions $nd \times nd$ so that equation (8.27) is dimensionally correct. This concludes the proof of statement (i).

To prove statement (ii), let $\Pi_n = I_n - \mathbb{1}_n w^\mathsf{T}$ denote a projection matrix on the subspace of zero-average vectors; note that $\Pi_n^2 = \Pi_n$ and $w^\mathsf{T}\Pi_n = \mathbb{0}_n^\mathsf{T}$. As in Exercise E5.9(iii) (where A is row-stochastic), one can easily see

$$\Pi_n L = L\Pi_n = L. \tag{8.28}$$

Define the *consensus error* $\mathbf{e} \in \mathbb{R}^{nd}$ by

$$\mathbf{e} = (\Pi_n \otimes I_d)\mathbf{x}. \tag{8.29}$$

Note that $\mathbf{e} = \mathbb{0}_{nd}$ if and only if $x_1 = \cdots = x_n$. Using the mixed product property and the fact that L and Π_n commute, we compute

$$\dot{\mathbf{e}} = (\Pi_n \otimes I_d)\big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K)\big)\mathbf{x}$$
$$= \big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K)\big)(\Pi_n \otimes I_d)\mathbf{x} = \big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K)\big)\mathbf{e}.$$

Let J be the Jordan normal form of L and let T satisfy $L = TJT^{-1}$. Recall that the first column of T is $\mathbb{1}_n$ and the first row of T^{-1} is w. We define the *transformed consensus error* $\tilde{\mathbf{e}} = (T^{-1} \otimes I_d)\mathbf{e} \in \mathbb{R}^{nd}$ and, noting $(T^{-1} \otimes I_d)^{-1} = (T \otimes I_d)$, we compute

$$\dot{\tilde{\mathbf{e}}} = (T^{-1} \otimes I_d)\Big((I_n \otimes \mathscr{A}) - c(L \otimes \mathscr{B}K)\Big)(T \otimes I_d)\tilde{\mathbf{e}}$$
$$= \Big((I_n \otimes \mathscr{A}) - c(J \otimes \mathscr{B}K)\Big)\tilde{\mathbf{e}}. \tag{8.30}$$

The first d entries of the vector $\tilde{\mathbf{e}}(t)$ are identically zero at all times t, because one can show $\tilde{\mathbf{e}}_1(t) = (w^\mathsf{T} \otimes I_d)\mathbf{e}(t) = \mathbb{0}_d$. Next, since the Jordan normal form J is block diagonal, say with blocks $J_1, \ldots, J_m$ (with $J_1 = 0$), we can write the dynamics (8.30) as decoupled equations. If J_i corresponds to a simple eigenvalue λ_i and is a one dimensional block, then we have

$$\dot{\tilde{\mathbf{e}}}_i = (\mathscr{A} - c\lambda_i \mathscr{B}K)\tilde{\mathbf{e}}_i.$$

One can show that, for arbitrary dimensional Jordan blocks corresponding to eigenvalues λ_i, $i \in \{2, \ldots, n\}$, the asymptotic stability condition is that $\mathscr{A} - c\lambda_i \mathscr{B}K$ is Hurwitz. In other words, each matrix is Hurwitz if and only if $\tilde{\mathbf{e}}(t)$ and $\mathbf{e}(t)$ vanish asymptotically so that $\mathbf{x}$ achieves asymptotic consensus. This concludes the proof of statement (ii). ∎

8.5   Historical notes and further reading

Excellent reviews of the Kronecker product are given for example by (Horn and Johnson, 1994, Chapter 4), (Van Loan, 2000), and (Laub, 2005, Chapter 13).

The Kronecker formalism is related to the early work (Wu and Chua, 1995) and the textbook (Wu, 2007). Theorems 8.8 and 8.11 and the abbreviated treatment in Section 8.4.4 are due to (Li et al., 2010), see also (Xia and Scardovi, 2016, Theorem 1), (Li and Duan, 2014, Theorem 1). An early reference on the observability problem is (Tuna, 2012). A comprehensive treatment is in the text (Li and Duan, 2014).

Second-order Laplacian flows are widely studied. Early references are the works by Chow (1982) and Chow and Kokotović (1985) on slow coherency and area aggregation of power networks, modeled as first and second-order Laplacian flows; see also (Avramovic et al., 1980; Chow et al., 1984; Saksena et al., 1984) among others.

In the consensus literature, an early reference to second-order Laplacian flows is (Ren and Atkins, 2005). Relevant references include (Ren, 2008a,b; Zhu et al., 2009; Zhang and Tian, 2009; Yu et al., 2010); see also (Ren and Atkins, 2005; Ren, 2008b). We refer to (Zhu et al., 2009) for convergence results for general digraphs and gains with arbitrary signs, and to (Zhang and Tian, 2009) for the discrete-time setting.

(Montenbruck et al., 2015; van Waarde et al., 2017) discuss when diffusive coupling is necessary for optimal synchronization problems among identical linear systems with quadratic costs.

8.6   Exercises

E8.1   **Properties of the Kronecker product.** Prove properties (8.11a)–(8.11d), (8.12), (8.13) and (8.14) of the Kronecker product.

Note: Additionally, for any $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{q \times r}$, and $p \geq 1$, one can show $\|A \otimes B\|_p = \|A\|_p \|B\|_p$, see (Lancaster and Farahat, 1972) and (Wu et al., 2022, Proposition 5).

▶ E8.2   **The vectorization operator, the Kronecker product, and the Sylvester equation.** Given a matrix $X \in \mathbb{R}^{n \times m}$, the *vectorization of X* is the vector of dimension mn obtained by stacking all columns of X, that is,

$$\operatorname{vec}(X) = [x_{11}, \ldots, x_{n1}, x_{12}, \ldots, x_{n2}, \ldots, x_{1m}, \ldots, x_{nm}]^{\mathsf{T}} \in \mathbb{R}^{mn}. \tag{E8.1}$$

Show that

(i) any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ satisfy $\operatorname{vec}(xy^{\mathsf{T}}) = y \otimes x$;

(ii) for any $X, Y \in \mathbb{R}^{n \times m}$, recall their Frobenius inner product defined by $\langle\!\langle X, Y \rangle\!\rangle = \operatorname{trace}(X^{\mathsf{T}} Y)$ and show $\langle\!\langle X, Y \rangle\!\rangle = \operatorname{vec}(X)^{\mathsf{T}} \operatorname{vec}(Y)$;

(iii) any matrices A, B and C, for which the product ABC is well defined, satisfy

$$\operatorname{vec}(ABC) = (C^{\mathsf{T}} \otimes A) \operatorname{vec}(B);$$

(iv) for any matrix function $A \colon \mathbb{R}^n \to \mathbb{R}^{n \times n}$, the Jacobian of the vector field $A(x)x$ satisfies

$$\frac{\partial A(x)x}{\partial x} = (x^{\mathsf{T}} \otimes I_n) \frac{\partial \operatorname{vec}(A(x))}{\partial x} + A(x).$$

Next, for $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$, and $C \in \mathbb{R}^{n \times m}$, consider the *Sylvester equation*

$$AX + XB = C$$

in the matrix variable $X \in \mathbb{R}^{n \times m}$. Show that the Sylvester equation

(v) can be rewritten as

$$\left\{ (I_m \otimes A) + (B^{\mathsf{T}} \otimes I_n) \right\} \operatorname{vec}(X) = \operatorname{vec}(C);$$

(vi) has a unique solution for all C if and only if A and $-B$ have no common eigenvalues.

E8.3   **Second-order Laplacian matrices.** Given a Laplacian matrix $L = L^{\mathsf{T}}$ and non-negative coefficients $k_{\mathrm{p}}, k_{\mathrm{d}}, \gamma_{\mathrm{p}}, \gamma_{\mathrm{d}} \in \mathbb{R}$, define the *second-order Laplacian matrix* $\mathcal{L} \in \mathbb{R}^{2n \times 2n}$ by

$$\mathcal{L} = \begin{bmatrix} \mathbb{0}_{n \times n} & I_n \\ -k_{\mathrm{p}} I_n - \gamma_{\mathrm{p}} L & -k_{\mathrm{d}} I_n - \gamma_{\mathrm{d}} L \end{bmatrix}, \tag{E8.2}$$

and write the second-order Laplacian system (8.7) in first-order form as $\begin{bmatrix} \dot{q}(t) \\ \dot{v}(t) \end{bmatrix} = \mathcal{L} \begin{bmatrix} q(t) \\ v(t) \end{bmatrix}$. Show that

(i) the characteristic polynomial of $\mathcal{L}$ is

$$\det(\eta I_{2n} - \mathcal{L}) = \det\left(\eta^2 I_n + \eta(k_{\mathrm{d}} I_n + \gamma_{\mathrm{d}} L) + (k_{\mathrm{p}} I_n + \gamma_{\mathrm{p}} L)\right);$$

(ii) given the real eigenvalues $\lambda_1, \ldots, \lambda_n$, of L, the $2n$ eigenvalues $\eta_{1,+}, \eta_{1,-}, \ldots, \eta_{n,+}, \eta_{n,-}$ of $\mathcal{L}$ are solutions to

$$\eta^2 + (k_{\mathrm{d}} + \gamma_{\mathrm{d}} \lambda_i)\eta + (k_{\mathrm{p}} + \gamma_{\mathrm{p}} \lambda_i) = 0, \quad i \in \{1, \ldots, n\}, \tag{E8.3}$$

that is, $\eta_{1,\pm} = \frac{-k_{\mathrm{d}} \pm \sqrt{k_{\mathrm{d}}^2 - 4k_{\mathrm{p}}}}{2}$ corresponding to $\lambda_1 = 0$ and, for $i \in \{2, \ldots, n\}$,

$$\eta_{i,\pm} = \frac{-(k_{\mathrm{d}} + \gamma_{\mathrm{d}} \lambda_i) \pm \sqrt{(k_{\mathrm{d}} + \gamma_{\mathrm{d}} \lambda_i)^2 - 4(k_{\mathrm{p}} + \gamma_{\mathrm{p}} \lambda_i)}}{2};$$

(iii) if the undirected graph associated to L is connected and if $k_{\mathrm{p}} + \gamma_{\mathrm{p}} > 0$ and $k_{\mathrm{d}} + \gamma_{\mathrm{d}} > 0$, then each eigenvalue $\eta_{i,\pm}, i \in \{2, \ldots, n\}$, has negative real part;

(iv) $\mathcal{L}$ is similar to the Kronecker product expression (8.15) in Theorem 8.3 with a permutation similarity transform (i.e., a simple reordering of rows and columns).

E8.4   **Eigenvectors of the second-order Laplacian matrix.** Consider a Laplacian matrix L, scalar coefficients k_p, k_d, γ_p, $\gamma_d \in \mathbb{R}$ and the induced second-order Laplacian matrix $\mathcal{L}$ (as in (E8.2)). Let $v_{l,i}$ and $v_{r,i}$ be the left and right eigenvectors of L corresponding to the eigenvalue λ_i, show that

   (i) the right eigenvectors of $\mathcal{L}$ corresponding to the eigenvalues $\eta_{i,\pm}$ are

$$\begin{bmatrix} v_{r,i} \\ \eta_{i,\pm} v_{r,i} \end{bmatrix},$$

   (ii) for $k_p > 0$, the left eigenvectors of $\mathcal{L}$ corresponding to the eigenvalues $\eta_{i,\pm}$ are

$$\begin{bmatrix} v_{l,i} \\ \dfrac{-\eta_{i,\pm}}{k_p + \gamma_p \lambda_i} v_{l,i} \end{bmatrix}.$$

E8.5   **Synchronization of inductors/capacitors circuits.** Consider a circuit composed of n identical resonant inductor/capacitor storage nodes (i.e., a parallel interconnection of a capacitor and an inductor) coupled through a connected and undirected graph whose edges are identical resistors; see Figure E8.1. The parameters ℓ, c, r take identical values on each inductor, capacitor and resistors, respectively.
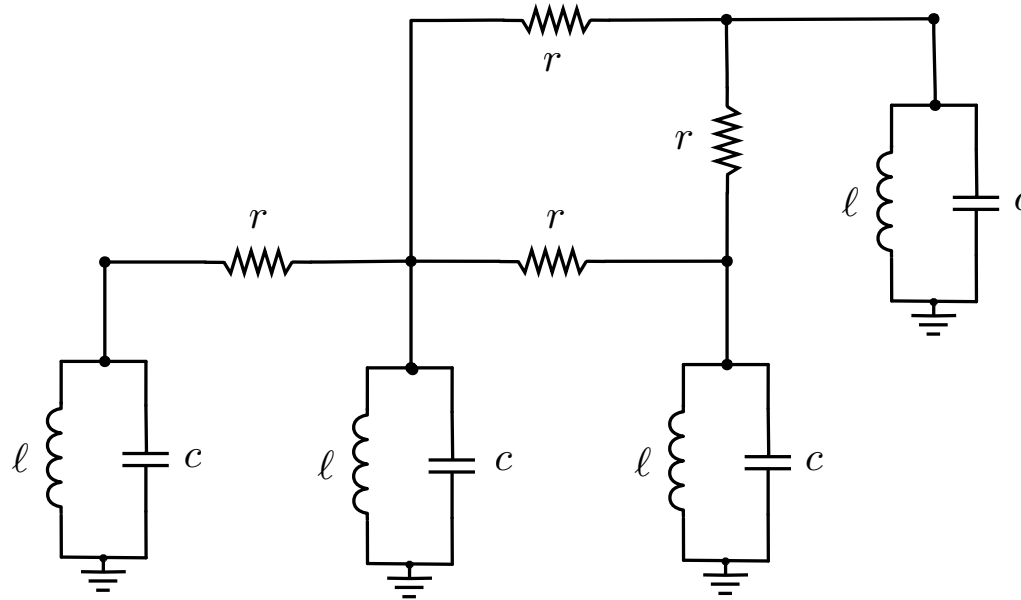
Figure E8.1: A circuit of identical inductor/capacitor storage nodes coupled through identical resistors.

 (i) Write a state-space model of the resistively-coupled inductor/capacitor storage nodes in terms of the time constant $\tau = 1/rc$, the resonant frequency $\omega_0 = 1/\sqrt{\ell c}$, and the unweighted Laplacian matrix L of the resistive network.
(ii) Characterize the asymptotic behavior of this system.

▶ E8.6   **Laplacian oscillators.** Given the Laplacian matrix $L = L^\mathsf{T} \in \mathbb{R}^{n \times n}$ of an undirected, weighted, and connected graph with edge weights a_{ij}, $i, j \in \{1, \ldots, n\}$, define the *Laplacian oscillator flow* by

$$\ddot{x}(t) + Lx(t) = \mathbb{0}_n. \tag{E8.4}$$

Recall that this equation arises for example as the discretization of the wave equation in Example 7.1.3. This flow is written as first-order differential equation as

$$\begin{bmatrix} \dot{x}(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} \mathbb{0}_{n \times n} & I_n \\ -L & \mathbb{0}_{n \times n} \end{bmatrix} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} =: \mathcal{L} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix}.$$

  (i) Write the second-order Laplacian flow in components.
 (ii) Write the characteristic polynomial of the matrix $\mathcal{L}$ using only the determinant of an $n \times n$ matrix.
(iii) Given the eigenvalues $\lambda_1 = 0, \lambda_2, \ldots, \lambda_n$ of L, show that the eigenvalues $\eta_1, \ldots, \eta_{2n}$ of $\mathcal{L}$ satisfy

$$\eta_1 = \eta_2 = 0, \qquad \eta_{2i,2i-1} = \pm\sqrt{\lambda_i}\mathrm{i}, \ \ \text{for } i \in \{2, \ldots, n\},$$

     where i is the imaginary unit.
 (iv) Show that the solution is the superposition of a ramp signal and of $n - 1$ harmonics, that is,

$$x(t) = \big(\operatorname{average}(x(0)) + \operatorname{average}(\dot{x}(0))t\big)\mathbb{1}_n + \sum_{i=2}^{n} a_i \sin(\sqrt{\lambda_i}t + \phi_i)v_i,$$

     where $\{\mathbb{1}_n/\sqrt{n}, v_2, \ldots, v_n\}$ are the orthonormal eigenvectors of L and where the amplitudes a_i and phases ϕ_i are determined by the initial conditions $\big(x(0), \dot{x}(0)\big)$.

Exercises with solution

E8.7   **The Cartesian product of graphs.** The *Cartesian product* $F\square H$ of two graph $F = (V_F, E_F)$ and $H = (V_H, E_H)$ is a graph with vertex set $V_F \times V_H$ and an edge between nodes (f_1, h_1) and (f_2, h_2) if and only if either $(f_1 = f_2$ and $\{h_1, h_2\} \in E_H)$ or $(\{f_1, f_2\} \in E_F$ and $h_1 = h_2)$. Clearly, if $|V_F| = m$ and $|V_H| = n$, then the number of edges in $F\square H$ is mn; moreover, the number of edges in $F\square H$ is $m|E_H| + n|E_F|$. In other words, the graph $F\square H$ is obtained from F by (i) replacing each of its vertices with a copy of H and (ii) each of its edges with $n = |V_H|$ edges connecting corresponding vertices of H in the two copies; as illustrated in Figure E8.1.
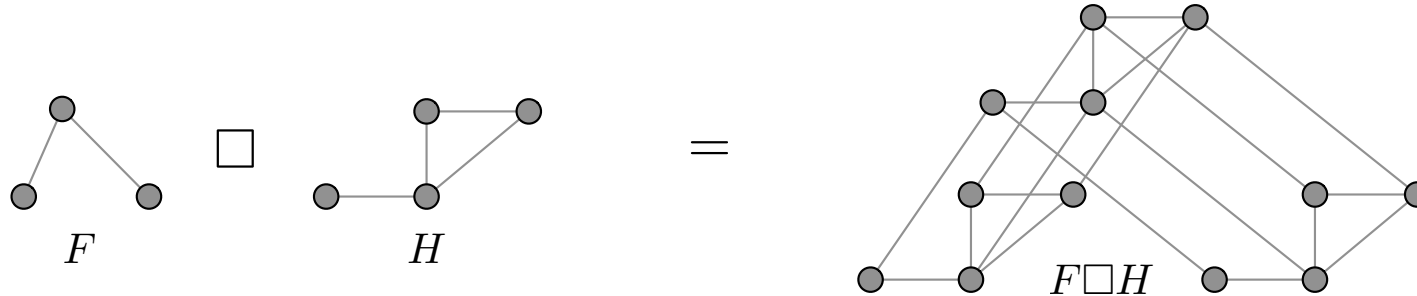
Figure E8.1: An example Cartesian product.

Let $L(F)$ and $L(G)$ denote the Laplacian matrices of F and G, respectively, with eigenvalues $\lambda_1 \leq \ldots \leq \lambda_m$ and $\mu_1 \leq \ldots \leq \mu_n$, respectively. It is known that

$$L(F\square H) = L(F) \otimes I_n + I_m \otimes L(H). \tag{E8.1}$$

Show that

 (i) if (λ_i, v_i) is an eigenpair for $L(F)$ and (μ_j, u_j) is an eigenpair for $L(H)$, for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$, then $(\lambda_i + \mu_j, v_i \otimes u_j)$ is an eigenpair for $L(F\square H)$,
(ii) the second smallest eigenvalue of $L(F\square H)$ is $\min(\lambda_2, \mu_2)$,
(iii) $F\square H$ is connected if and only if both F and H are connected, and
(iv) the $m \times n$ grid graph $G_{m,n}$ satisfies $G_{m,n} = P_m \square P_n$, where P_n is the path graph of order n, and its Laplacian spectrum is

$$4 - 2\cos(\pi i/m) - 2\cos(\pi j/n), \qquad \text{for } i \in \{0, \ldots, m-1\}, j \in \{0, \ldots, m-1\}. \tag{E8.2}$$

   Hint: See the definitions in Example 3.1 and review Exercise E6.9.

Note: Cartesian products of graphs were introduced by Sabidussi (1960). The original work on statement (i) is by Fiedler (1973).

Answer: Regarding statement (i), suppose that (λ_i, v_i) is an eigenpair for $L(F)$ and (μ_j, u_j) is an eigenpair for $L(H)$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. Then

$$L(F\square H)(v_i \otimes u_j) = (L(F) \otimes I_m + I_n \otimes L(H))(v_i \otimes u_j) = (L(F)v_i) \otimes u_j + v_i \otimes (L(H)u_j)$$
$$= \lambda_i v_i \otimes u_j + v_i \otimes \mu_j u_j = (\lambda_i + \mu_j)(v_i \otimes \mu_j).$$

Thus $v_i \otimes u_j$ is an eigenvector of $L(F \square H)$ with eigenvalue $\lambda_i + \mu_j$.

Regarding statement (ii), the matrix $L(F \square H)$ has size mn so the set of eigenvalues of $L(F \square H)$ is precisely $\{\lambda_i + \mu_j \mid 1 \le i \le n, 1 \le j \le m\}$. Since $\lambda_1 = \mu_1 = 0$, the smallest eigenvalue of $L(F \square H)$ is $\lambda_1 + \mu_1 = 0$. The second largest eigenvalue is then $\min(\lambda_1 + \mu_2, \lambda_2 + \mu_1) = \min(\lambda_2, \mu_2)$.

Regarding statement (iii), we note that $F \square H$ is connected if, and only if, its second smallest eigenvalue $\min(\lambda_2, \mu_2)$ is greater than 0. But this is only true if, and only if, $\lambda_2 > 0$ and $\mu_2 > 0$. But this happens if, and only if, F and H are connected. This completes the argument.

Statement (iv) follows from recalling that the Laplacian spectrum of P_n is $\{2(1 - \cos(\pi i/n)) \mid i \in \{0, \ldots, n-1\}\}$.

Chapter 9

The Incidence Matrix and its Applications

After studying adjacency and Laplacian matrices, in this chapter we introduce one final matrix associated with a graph: the incidence matrix. We study the properties of incidence matrices and their application to a class of estimation problems with relative measurements and to the study of cycles and cutset spaces. For simplicity we restrict our attention to undirected graphs.

9.1   The incidence matrix

Let G be an undirected unweighted graph with n nodes and m edges (and no self-loops, as by convention). Assign to each edge of G a unique identifier $e \in \{1, \ldots, m\}$ and an arbitrary direction. Given a directed edge $e = (i, j)$, we refer to node i as the *head* and node j as the *tail* of e. The *(oriented) incidence matrix* $B \in \mathbb{R}^{n \times m}$ of the graph G is defined component-wise by

$$
B_{ie} = \begin{cases} +1, & \text{if node } i \text{ is the head of edge } e, \\ -1, & \text{if node } i \text{ is the tail of edge } e, \\ 0, & \text{otherwise.} \end{cases} \tag{9.1}
$$

It is useful to consider the example graph depicted in Figure 9.1.
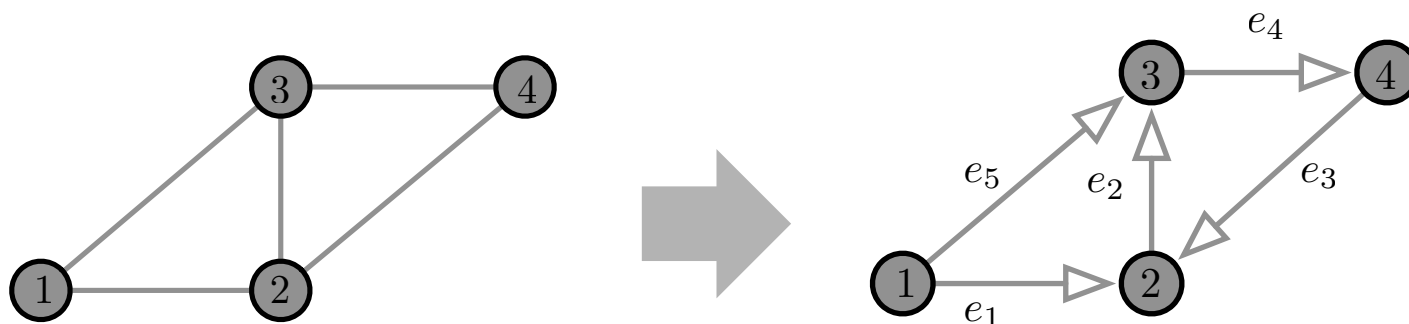
Figure 9.1: An arbitrary enumeration and orientation of the 5 edges of a graph with 4 nodes

As depicted on the right, we add an orientation to all edges, we order them and label them as follows: $e_1 = (1, 2)$, $e_2 = (2, 3)$, $e_3 = (4, 2)$, $e_4 = (3, 4)$, and $e_5 = (1, 3)$. Accordingly, the incidence matrix is

$$B = \begin{bmatrix} +1 & 0 & 0 & 0 & +1 \\ -1 & +1 & -1 & 0 & 0 \\ 0 & -1 & 0 & +1 & -1 \\ 0 & 0 & +1 & -1 & 0 \end{bmatrix} \in \mathbb{R}^{4 \times 5}. \tag{9.2}$$

Note: $\mathbb{1}_n^\mathsf{T} B = \mathbb{0}_m^\mathsf{T}$ since each column of B contains precisely one element equal to $+1$, one element equal to -1 and all other zeros.

The incidence matrix B can be regarded as a map from edge-based variables in $\mathbb{R}^m$ to node-based variables in $\mathbb{R}^n$. Specifically, given an edge-based variable $f \in \mathbb{R}^m$ and a node $i \in \{1, \ldots, n\}$,

$$(Bf)_i = \sum_{e\,:\,i\text{ is the head of }e} f_e - \sum_{e\,:\,i\text{ is the tail of }e} f_e. \tag{9.3}$$

When the edge-based variables f are flows along edges, then $(Bf)_i$ is the algebraic sum of the flows outgoing from node i. Similarly, the transpose of the incidence matrix B^T maps node-based variables in $\mathbb{R}^n$ to edge-based variables in $\mathbb{R}^m$. Specifically, given a node-based variable $x \in \mathbb{R}^n$ and an edge $e \in \{1, \ldots, m\}$ of the form $e = (i, j)$,

$$(B^\mathsf{T}x)_e = x_i - x_j. \tag{9.4}$$

When the node-based variables x are potential variables, then $(B^\mathsf{T}x)_e$ is the difference of potential across the edge e.

9.2   Properties of the incidence matrix

Given an undirected weighted graph G with edge set $\{1, \ldots, m\}$ and adjacency matrix A, recall

$$L = D - A, \qquad \text{where } D \text{ is the degree matrix.}$$

Lemma 9.1 (From the incidence to the Laplacian matrix). *Let G be an undirected graph with n nodes, m edges, and incidence matrix B. Define the weight matrix $\mathcal{A} \in \mathbb{R}^{m \times m}$ to be the diagonal matrix of edge weights $\mathcal{A} = \mathrm{diag}(\{a_e\}_{e \in \{1, \ldots, m\}})$. Then*

$$L = B\mathcal{A}B^{\mathsf{T}}.$$

Note: In the right-hand side, the matrix dimensions are $(n \times m) \times (m \times m) \times (m \times n) = n \times n$. Also note that, while the incidence matrix B depends upon the selected direction and numbering of each edge, the Laplacian matrix is independent of that.

Proof. Recall that, for matrices O, P and Q of appropriate dimensions, we have $(OPQ)_{ij} = \sum_{k,h} O_{ik} P_{kh} Q_{hj}$. Moreover, if the matrix P is diagonal, then $(OPQ)_{ij} = \sum_k O_{ik} P_{kk} Q_{kj}$.

For $i \neq j$, we compute

$$
\begin{aligned}
(B\mathcal{A}B^{\mathsf{T}})_{ij} &= \sum_{e=1}^{m} B_{ie} a_e (B^{\mathsf{T}})_{ej} \\
&= \sum_{e=1}^{m} B_{ie} B_{je} a_e \qquad (e\text{-th term} = 0 \text{ unless } e \text{ is oriented } \{i,j\}) \\
&= (+1) \cdot (-1) \cdot a_{ij} = \ell_{ij},
\end{aligned}
$$

where $L = \{\ell_{ij}\}_{i,j \in \{1,\ldots,n\}}$, and along the diagonal of B we compute

$$
(B\mathcal{A}B^{\mathsf{T}})_{ii} = \sum_{e=1}^{m} B_{ie}^2 a_e = \sum_{e=1,\ e=(i,*) \text{ or } e=(*,i)}^{m} a_e = \sum_{j=1, j \neq i}^{n} a_{ij},
$$

where, in the last equality, we counted each edge precisely once and we noted that self-loops are not allowed. ∎

Lemma 9.2 (Rank of the incidence matrix). *Let G be an undirected graph with n nodes, m edges, and incidence matrix B. Let n_{cc} be the number of connected components of G. Then*

$$\mathrm{rank}(B) = n - n_{\mathrm{cc}}.$$

Proof. We prove this result for a connected graph with $n_{\mathrm{cc}} = 1$, but the proof strategy extends to $n_{\mathrm{cc}} > 1$. Recall that the rank of the Laplacian matrix L equals $n - n_{\mathrm{cc}} = n - 1$. Since the Laplacian matrix can be factorized as $L = B\mathcal{A}B^{\mathsf{T}}$, where $\mathcal{A}$ has full rank m (and $m \geq n - 1$ due to connectivity), we have that necessarily $\mathrm{rank}(B) \geq n - 1$. On the other hand $\mathrm{rank}(B) \leq n - 1$ since $B^{\mathsf{T}}\mathbb{1}_n = \mathbb{0}_n$. It follows that B has rank $n - 1$. ∎

Here are two examples:

- if G is a tree ($m = n - 1$ and $n_{\mathrm{cc}} = 1$), we know $B \in \mathbb{R}^{n \times (n-1)}$, $\mathrm{rank}(B) = n - 1$ (that is, B is full rank and has linearly independent columns), $\mathrm{kernel}(B) = \{\mathbb{0}_{n-1}\}$ and $\mathrm{image}(B^{\mathsf{T}}) = \mathbb{R}^{n-1}$.

- if G is connected, then $\mathrm{rank}(B) = n - 1$ (that is, B is full rank), $\mathrm{image}(B^{\mathsf{T}}) = \mathbb{R}^{n-1}$, and $\mathrm{kernel}(B)$ has dimension $m - n + 1$.

9.3   Applications of the incidence matrix

The Laplacian flow as a closed-loop control system   The factorization of the Laplacian matrix as $L = B\mathcal{A}B^\mathsf{T}$ plays an important role in relative sensing networks. For example, we can decompose the Laplacian flow $\dot{x} = -Lx$ into

$$
\begin{aligned}
\text{open-loop plant:} \quad & \dot{x}_i = u_i\,, & i \in \{1,\dots,n\}\,, \quad & \text{or} \quad \dot{x} = u\,, \\
\text{measurements:} \quad & y_{ij} = x_i - x_j\,, & \{i,j\} \in E\,, \quad & \text{or} \quad y = B^\mathsf{T}x\,, \\
\text{control gains:} \quad & z_{ij} = a_{ij}y_{ij}\,, & \{i,j\} \in E\,, \quad & \text{or} \quad z = \mathcal{A}y\,, \\
\text{control inputs:} \quad & u_i = - \sum_{\{i,j\}\in E} z_{ij}\,, & i \in \{1,\dots,n\}\,, \quad & \text{or} \quad u = -Bz\,.
\end{aligned}
$$

In other words we can write

$$\dot{x} = u = -Bz = -B\mathcal{A}y = -B\mathcal{A}B^\mathsf{T}x = -Lx.$$

Indeed, this control structure, illustrated as a block-diagram in Figure 9.2, is required to implement flocking-type behavior as in Example 1.3. The control structure in Figure 9.2 has emerged as a canonical control structure in many relative sensing and flow network problems also for more complicated open-loop dynamics and possibly nonlinear control gains; e.g., see (Bai et al., 2011).
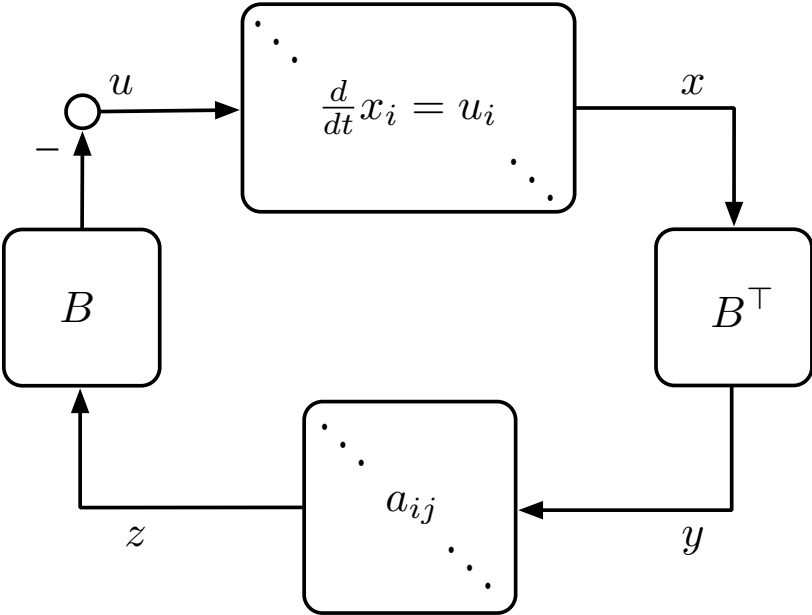
Figure 9.2: Illustration of the canonical control structure for a relative sensing network.

Kirchhoff's and Ohm's laws   We here revisit the electrical resistor network from Section 6.1.3, and re-derive its governing equations; we refer to (Dörfler et al., 2018) for a more detailed treatment.

First, we let $B \in \mathbb{R}^{n \times m}$ denote the oriented incidence matrix of electrical network (after introducing an arbitrary numbering and orientation for each edge). To each node $i \in \{1, \dots, n\}$ we associate an external current injection $\mathsf{c}_{\text{injected at } i}$. To each oriented edge $(i,j) \in E$ we associate a positive conductance (i.e., the inverse of the resistance) $a_{ij} > 0$, a current flow $\mathsf{c}_{i \to j}$, and a voltage drop u_{ij}.

The *Kirchhoff's voltage law (KVL)* states that the sum of all voltage drops around each cycle must be zero. It is well-known and easy to see that KVL implies the existence of *potential variables* v_i at each node such that $u_{ij} = \mathsf{v}_i - \mathsf{v}_j$, for each oriented edge (i,j). In other words, recalling the equality (9.4), $u_{ij} = (B^\mathsf{T}\mathsf{v})_{ij}$. In summary, given voltage drops $u \in \mathbb{R}^m$ along the edges KVL states that there exist potentials $\mathsf{v} \in \mathbb{R}^n$ at the nodes such that

$$u = B^\mathsf{T}\mathsf{v}.$$

The *Kirchhoff's current law (KCL)* states that the sum of all current injections at every node must be zero. In other words, for each node $i \in \{1, \dots, n\}$ in the network, we have that $\mathsf{c}_{\text{injected at } i} = \sum_{j=1}^n \mathsf{c}_{i \to j}$. Since all oriented edges incident to node i are described by entries in the ith row b_i of B, we can write $\mathsf{c}_{\text{injected at } i} = \sum_{j=1}^n \mathsf{c}_{i \to j} = b_i^\mathsf{T}\mathsf{c}$. In summary, KCL states that injected currents $\mathsf{c}_{\text{injected}} \in \mathbb{R}^n$ at the nodes and current flows $\mathsf{c} \in \mathbb{R}^m$ along the edges satisfy

$$\mathsf{c}_{\text{injected}} = B\mathsf{c}.$$

Finally, *Ohm's law* states that the current $\mathsf{c}_{j \to i}$ and the voltage drop u_{ij} over a resistor with resistance $1/a_{ij}$ are related as $\mathsf{c}_{j \to i} = a_{ij}u_{ij}$. By combining Kirchhoff's and Ohm's laws, we arrive at

$$\mathsf{c}_{\text{injected}} = B\mathsf{c} = B\mathcal{A}u = B\mathcal{A}B^\mathsf{T}\mathsf{v} = L\mathsf{v},$$

where we used Lemma 9.1 to recover the conductance matrix L.

9.4   Appendix: Cuts and cycles

Given an undirected unweighted graph with n nodes and m edges, its oriented incidence matrix naturally defines two useful vector subspaces of $\mathbb{R}^m$. With the customary convention to refer to $\mathbb{R}^m$ as the *edge space*, the incidence matrix induces a direct sum decomposition of the edge space based on the concepts of cycles and graph cuts. We develop these concepts in what follows.

Definition 9.3 (Cutset orientation vectors and cutset space). *Let G be an undirected graph with nodes $\{1,\ldots,n\}$ and with an arbitrary enumeration and orientation of its m edges.*

(i) *A cut χ of G is a strict non-empty subset of the nodes $\{1,\ldots,n\}$. A cut and its complement χ^c define a partition $\{\chi,\chi^c\}$ of $\{1,\ldots,n\}$, in the sense that $\chi \neq \emptyset$, $\chi^c \neq \emptyset$, $\chi \cap \chi^c = \emptyset$, and $\{1,\ldots,n\} = \chi \cup \chi^c$.*

(ii) *Given a cut $\chi \subset \{1,\ldots,n\}$ of G, the set of edges that have one endpoint in each subset of the partition is called the cutset of χ. The cutset orientation vector $v_\chi \in \{-1,0,+1\}^m$ of χ is defined component-wise, for each edge $e \in \{1,\ldots,m\}$,*

$$
(v_\chi)_e = \begin{cases} +1, & \text{if } e \text{ has its head in } \chi \text{ and tail in } \chi^c, \\ -1, & \text{if } e \text{ has its head in } \chi^c \text{ and tail in } \chi, \\ 0, & \text{otherwise.} \end{cases}
$$

Here the head (resp. tail) of a directed edge (i,j) is the node i (resp j).

(iii) *The cutset space of G is the subspace of $\mathbb{R}^m$ spanned by the cutset orientation vectors corresponding to all cuts of G, that is, $\operatorname{span}\{v_\chi \in \{-1,0,+1\}^m \mid \chi \text{ is a cut of } G\}$.*

We illustrate these concepts in Figure 9.3.
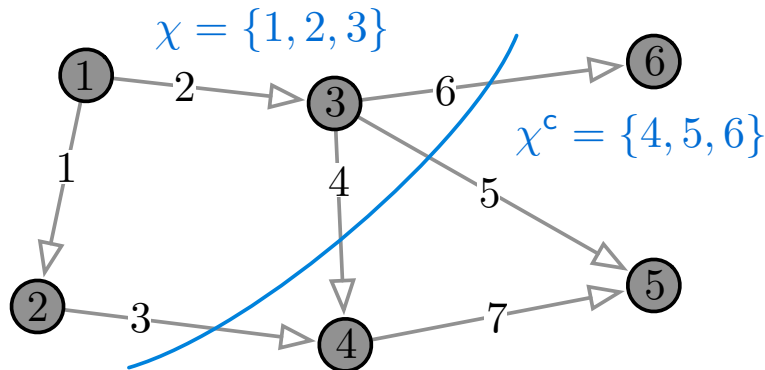
A digraph with $n = 6$ nodes and $m = 7$ edges.
A cut $\chi = \{1,2,3\}$ and its complement $\chi^c = \{4,5,6\}$.
The cutset of χ is $\{3,4,5,6\}$.
The cutset orientation vector of χ is:

$$
v_\chi = \begin{bmatrix} 0 & 0 & +1 & +1 & +1 & +1 & 0 \end{bmatrix}^{\mathsf{T}}.
$$

Figure 9.3: An undirected graph with arbitrary edge orientation. A cut with its cutset and cutset orientation vector.

Recall that, in an undirected graph, a walk is *simple* if no node appears more than once in it, except possibly for the first and last. A *cycle* is a simple walk that starts and ends at the same node and has at least three distinct nodes.

Definition 9.4 (Signed walk vectors and cycle space). *Let G be an undirected graph with n nodes, m edges, and with an arbitrary enumeration and orientation of its edges. Let γ be a simple undirected walk in G.*

(i) The signed walk vector $w_\gamma \in \{-1, 0, +1\}^m$ of γ is defined component-wise, for each edge $e \in \{1, \dots, m\}$,

$$(w_\gamma)_e = \begin{cases} +1, & \text{if } e \text{ is traversed positively by } \gamma, \\ -1, & \text{if } e \text{ is traversed negatively by } \gamma, \\ 0, & \text{otherwise.} \end{cases}$$

(ii) The cycle space of G is the subspace of $\mathbb{R}^m$ spanned by the signed walk vectors corresponding to all cycles in G, that is,

$$\text{span}\{w_\gamma \in \{-1, 0, +1\}^m \mid \gamma \text{ is a cycle in } G\}.$$

We illustrate these concepts in Figure 9.4.
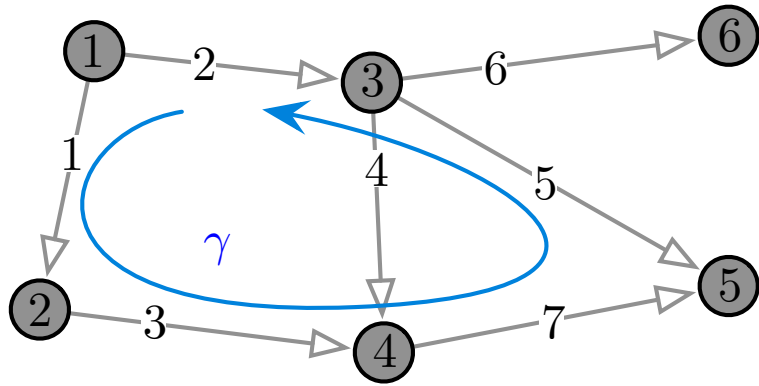
The signed walk vector is:

$$w_\gamma = \begin{bmatrix} +1 & -1 & +1 & 0 & -1 & 0 & +1 \end{bmatrix}^{\mathsf{T}}.$$

Figure 9.4: An undirected graph with arbitrary edge orientation. A directed cycle γ and its signed walk vector.

With these conventions we are now in a position to state the main result of this section.

Theorem 9.5 (Cycle and cutset spaces). *Let G be an connected undirected graph with n nodes, m edges, and incidence matrix B. The following statements hold:*

  (i) the cycle space is $\mathrm{kernel}(B)$ and has dimension $m - n + 1$,

 (ii) the cutset space is $\mathrm{image}(B^\mathsf{T})$ and has dimension $n - 1$, and

(iii) $\mathrm{kernel}(B) \perp \mathrm{image}(B^\mathsf{T})$ and $\mathrm{kernel}(B) \oplus \mathrm{image}(B^\mathsf{T}) = \mathbb{R}^m$.

Note that statement (iii) is known as a result in the fundamental theorem of linear algebra (Strang, 1993). Additionally, statement (iii) is known in circuit theory as Tellegen's Theorem (Oster and Desoer, 1971).

Proof of Theorem 9.5. The proof of statement (i) is given in Exercise E9.9.

Statement (ii) is proved as follows. For a cut χ, let $\mathbb{e}_\chi \in \{0,1\}^n$ be the *cut indicator vector* defined by $(\mathbb{e}_\chi)_i = 1$ if $i \in \chi$ and zero otherwise. Then, using the definitions, the cutset orientation vector for the cut χ is

$$v_\chi = B^\mathsf{T} \mathbb{e}_\chi.$$

This equality implies that $v_\chi \in \mathrm{image}(B^\mathsf{T})$ for all χ. Next, because G is connected, there are $n-1$ independent cutset orientation vectors corresponding to the cuts $\{\{i\} \mid i \in \{1, \ldots, n-1\}\}$. Hence these $n-1$ vectors are a basis of $\mathrm{image}(B^\mathsf{T})$ and the statement is established.

Finally, statement (iii) is proved in two steps. First, for any subspace $V \subset \mathbb{R}^m$, we have the direct sum decomposition of orthogonal subspaces $V \oplus V^\perp = \mathbb{R}^m$. Second, for any matrix B,

$$w \in \mathrm{kernel}(B) \iff \forall v \in \mathbb{R}^m \quad (Bw)^\mathsf{T} v = 0$$
$$\iff \forall v \in \mathbb{R}^m \quad w^\mathsf{T}(B^\mathsf{T} v) = 0 \iff w \in (\mathrm{image}(B^\mathsf{T})^\perp.$$

Hence, we know $\mathrm{kernel}(B) = (\mathrm{image}(B^\mathsf{T})^\perp$ and the statement follows. ∎

From the proof of the previous theorem and a bit more work, one can state the following result.

Lemma 9.6 (Bases for the cutset space and the cycle space). *Let $G = (V, E)$ be a connected unweighted undirected graph with nodes $\{1, \ldots, n\}$ and m edges.*

(i) For each node $i \in \{1, \ldots, n - 1\}$, let $v_{\{i\}} \in \{-1, 0, +1\}^m$ denote the cutset orientation vector for the cut $\{i\}$, that is, let $v_{\{i\}}$ be the transpose of the i-th row of B. Then $\{v_{\{1\}}, \ldots, v_{\{n-1\}}\}$ is a basis of the cutset space $\mathrm{image}(B^\mathsf{T})$.

(ii) Given a spanning tree $T = (V_T, E_T)$ of $G = (V, E)$, for each edge $e \in E \setminus E_T$, define the fundamental cycle associated to T and e, denoted by $\gamma_{T,e}$, to be the cycle consisting of e and the walk on T connecting the endpoints of e. Let $w_{T,e}$ be the associated signed walk vector. Then

(a) the fundamental cycle of each edge $e \in E \setminus E_T$ exists unique and is simple, and
(b) the set of signed walk vectors $\{w_{T,e} \mid e \in E \setminus E_T\}$ is a basis of the cycle space $\mathrm{kernel}(B)$.

We illustrate this lemma with the digraph in Figure 9.3, which we reproduce here with its incidence matrix for convenience.
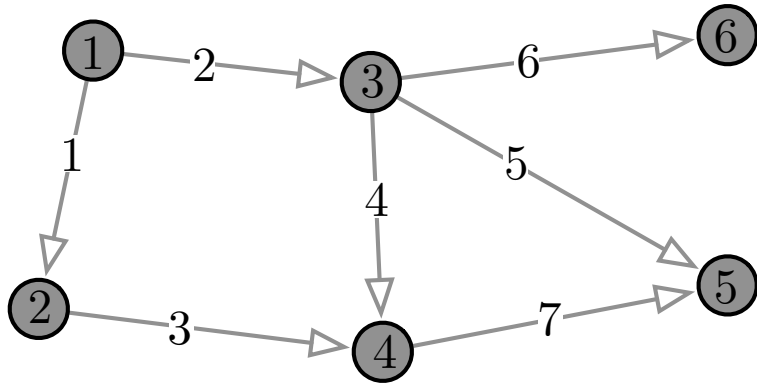
$$B = \begin{bmatrix} +1 & +1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & +1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & +1 & +1 & +1 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}.$$

Figure 9.5: The undirected graph with edge orientation from Figure 9.3 and its incidence matrix $B \in \mathbb{R}^{6 \times 7}$.

Regarding a basis for the cutset space $\mathrm{image}(B^\mathsf{T})$, it is immediate to state that (the transpose of) any 5 of the 6 rows of B form a basis of $\mathrm{image}(B^\mathsf{T})$. Indeed, since $\mathrm{rank}(B) = n - 1$, any $n - 1$ columns of the matrix B^T form a basis for the cutset space. Figure 9.6 illustrates the 5 cuts and a corresponding basis for the cutset space.
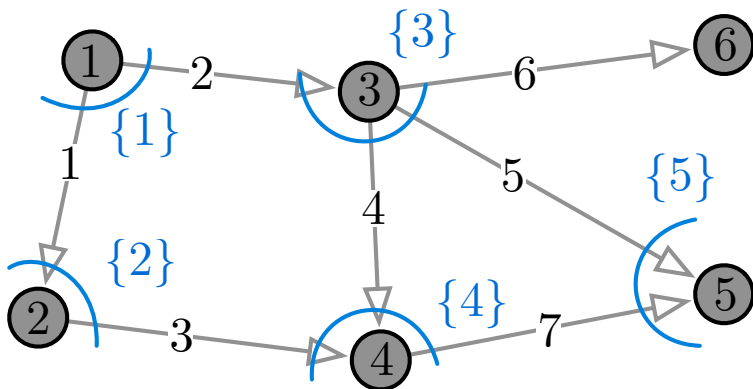
$$\begin{bmatrix} v_{\{1\}} & v_{\{2\}} & v_{\{3\}} & v_{\{4\}} & v_{\{5\}} \end{bmatrix} = \begin{bmatrix} +1 & -1 & 0 & 0 & 0 \\ +1 & 0 & -1 & 0 & 0 \\ 0 & +1 & 0 & -1 & 0 \\ 0 & 0 & +1 & -1 & 0 \\ 0 & 0 & +1 & 0 & -1 \\ 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & +1 & -1 \end{bmatrix}.$$

Figure 9.6: Five cuts, corresponding to nodes $1, \ldots, 5$, and their cutset orientation vectors generating $\mathrm{image}(B^\mathsf{T})$.

In the proof of Theorem 9.5, we also stated that, for a cut χ, $\mathbb{e}_\chi \in \{0, 1\}^n$ is the *cut indicator vector* defined by $(\mathbb{e}_\chi)_i = 1$ if $i \in \chi$

and zero otherwise, and that the cutset orientation vector for χ is given by

$$v_\chi = B^\mathsf{T} \mathbb{e}_\chi.$$

(9.5)

Indeed, one can show the following statement for the example in Figure 9.6: the cut separating nodes $\{1, 2, 3\}$ from $\{4, 5, 6\}$ has cut indicator vector $\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}^\mathsf{T}$ and cutset vector $v_{\{1\}} + v_{\{2\}} + v_{\{3\}}$ is equal to the sum of the first three columns of B^T.

Next, regarding a basis for the cycle space $\mathrm{kernel}(B)$, the spanning tree T composed of the edges $\{1, 2, 4, 5, 6\}$ and the two fundamental cycles associated to edges 3 and 7 are illustrated in Figure 9.7.
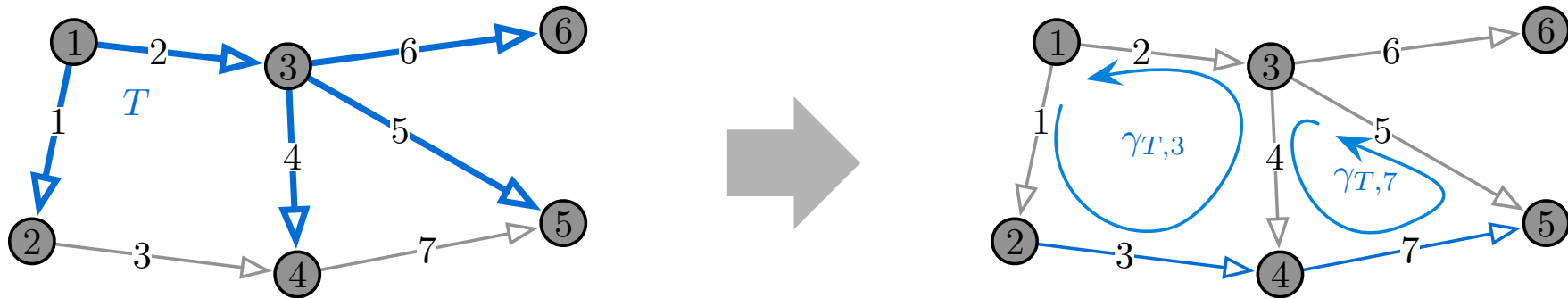
Figure 9.7: Given a graph with 6 nodes, 7 edges, and hence 2 independent cycles, the left panel depicts a spanning tree T (composed of 5 dark edges) and the right panel depicts the two resulting fundamental cycles.

The corresponding signed walk vectors are

$$w_{T,3} = \begin{bmatrix} +1 \\ -1 \\ +1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, w_{T,7} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \\ -1 \\ 0 \\ +1 \end{bmatrix}, \quad \text{and} \quad \mathrm{kernel}(B) = \mathrm{span}\{w_{T,3}, w_{T,7}\}.$$

Note that the cycle traversing the edges $(1, 3, 7, 5, 2)$ in counter-clockwise orientation has a signed walk vector given by the linear combination $w_{T,3} + w_{T,7}$.

9.5   Appendix: Distributed estimation from relative measurements

In Chapter 1 we considered estimation problems for wireless sensor networks in which each node measures a scalar "absolute" quantity (expressing some environmental variable such as temperature, vibrations, etc). In this section, we consider a second class of examples in which measurements are "relative," i.e., pairs of nodes measure the difference between their corresponding variables. Estimation problems involving relative measurements are numerous. For example, imagine a group of robots (or sensors) where no robot can sense its position in an absolute reference frame, but a robot can measure other robot's relative positions by means of on-board sensors. Similar problems arise in study of clock synchronization in networks of processors.

9.5.1   Problem statement

The optimal estimation based on relative measurement problem is stated as follows. As illustrated in Figure 9.8, we are given an undirected graph $G = (\{1, \ldots, n\}, E)$ with the following properties. First, each node $i \in \{1, \ldots, n\}$ of the network is associated with an unknown scalar quantity x_i (the x-coordinate of node i in figure). Second, the m undirected edges are given an orientation
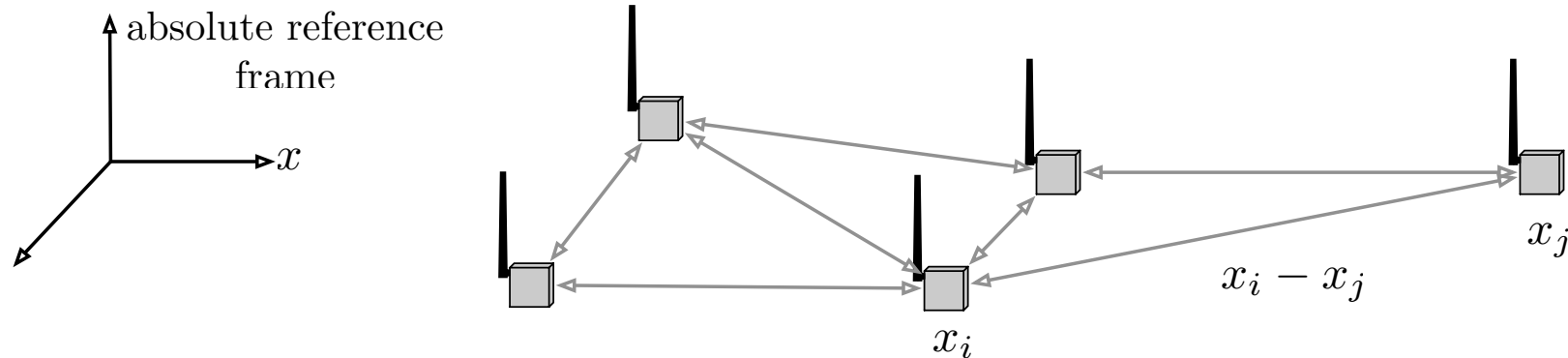
Figure 9.8: A wireless sensor network in which sensors can measure each other's relative distance and bearing. We assume that, for each link between node i and node j, the relative distance along the x-axis $x_i - x_j$ is available, where x_i is the x-coordinate of node i.

and, for each edge $e = (i, j)$, $e \in E$, the following scalar measurements are available:

$$y_{(i,j)} = x_i - x_j + v_{(i,j)} = (B^\mathsf{T} x)_e + v_{(i,j)},$$

where B is the graph incidence matrix and the measurement noises $v_{(i,j)}$, $(i, j) \in E$, are independent jointly-Gaussian variables with zero-mean $\mathbb{E}[v_{(i,j)}] = 0$ and variance $\mathbb{E}[v_{(i,j)}^2] = \sigma_{(i,j)}^2 > 0$. The joint matrix covariance is the diagonal matrix $\Sigma = \mathrm{diag}(\{\sigma_{(i,j)}^2\}_{(i,j)\in E}) \in \mathbb{R}^{m \times m}$. (For later use, it is convenient to define also $y_{(j,i)} = -y_{(j,i)} = x_j - x_i - v_{(i,j)}$.)

The optimal estimate $\widehat{x}^*$ of the unknown vector $x \in \mathbb{R}^n$ via the relative measurements $y \in \mathbb{R}^m$ is the solution to

$$\min_{\widehat{x}} \|B^\mathsf{T}\widehat{x} - y\|^2_{\Sigma^{-1}}. \tag{F1}$$

Since no absolute information is available about x, we add the additional constraint that the optimal estimate should have zero mean and summarize this discussion as follows.

Definition 9.7 (Optimal estimation problem based on relative measurements). *Given an incidence matrix B, a set of relative measurements y with covariance Σ, find $\widehat{x}$ satisfying*

$$\min_{\widehat{x} \perp \mathbb{1}_n} \|B^\mathsf{T}\widehat{x} - y\|^2_{\Sigma^{-1}}. \tag{9.6}$$

9.5.2　Optimal estimation via centralized computation

From the theory of least square estimation, the optimal solution to problem 9.6 is obtained as by differentiating the quadratic cost function with respect to the unknown variable $\widehat{x}$ and setting the derivative to zero. Specifically:

$$0 = \frac{\partial}{\partial \widehat{x}} \| B^{\mathsf{T}} \widehat{x} - y \|_{\Sigma^{-1}}^2 = 2 B \Sigma^{-1} B^{\mathsf{T}} \widehat{x}^* - 2 B \Sigma^{-1} y, \tag{F2}$$

or equivalently

$$B \Sigma^{-1} B^{\mathsf{T}} \widehat{x}^* = B \Sigma^{-1} y \quad \Longleftrightarrow \quad L \widehat{x}^* = B \Sigma^{-1} y, \tag{9.7}$$

where the Laplacian matrix L is defined by $L = B \Sigma^{-1} B^{\mathsf{T}}$. This matrix is the Laplacian for the weighted graph whose weights are the inverse noise covariances of the measurement edges.

　　We now note that equation (9.7) is a Laplacian system, as studied in Section 6.3.2. Recalling the notion of pseudoinverse Laplacian matrix and the constraint $\widehat{x} \perp \mathbb{1}_n$, Lemma 6.12 in Section 6.3.2 leads to the following result.

Lemma 9.8 (Unique optimal estimate).　*If the undirected graph G is connected, then there exists a unique solution to the optimization problem (9.6) given by*

$$\widehat{x}^* = L^{\dagger} B \Sigma^{-1} y. \tag{F3}$$

9.5.3   Optimal estimation via decentralized computation

To compute $\widehat{x}^*$ in a distributed way, we propose the following distributed algorithm. Pick a small $\alpha > 0$ and let each node implement the affine averaging algorithm:

$$\widehat{x}_i(k+1) = \widehat{x}_i(k) - \alpha \sum_{j \in \mathcal{N}(i)} \frac{1}{\sigma_{(i,j)}^2} \Big(\widehat{x}_i(k) - \widehat{x}_j(k) - y_{(i,j)} \Big),$$

$$\widehat{x}_i(0) = 0. \tag{9.8}$$

This algorithm is interpreted as follows: the estimate at node i is adjusted at each iteration as a function of edge errors, each edge error (difference between estimated and measured edge difference) contributes to a weighted small correction in the node value.

Lemma 9.9 (Convergence of the affine averaging algorithm). *Given a graph G describing a relative measurement problem for the unknown variables $x \in \mathbb{R}^n$, with measurements $y \in \mathbb{R}^m$, and measurement covariance matrix $\Sigma = \mathrm{diag}(\{\sigma^2_{(i,j)}\}_{(i,j) \in E}) \in \mathbb{R}^{m \times m}$. The following statements hold:*

(i) the affine averaging algorithm can be written as

$$\widehat{x}(k+1) = (I_n - \alpha L)\widehat{x}(k) + \alpha B \Sigma^{-1} y,$$
$$\widehat{x}(0) = \mathbb{0}_n;$$

(9.9)

(ii) if G is connected and if $\alpha < 1/d_{\max}$ where $d_{\max}$ is the maximum weighted out-degree of G, then the solution $k \mapsto \widehat{x}(k)$ of the affine averaging algorithm (9.8) converges to the unique solution $\widehat{x}^*$ of the optimization problem (9.6).

Proof. To show fact (i), note that the algorithm can be written in vector form as

$$\widehat{x}(k+1) = \widehat{x}(k) - \alpha B \Sigma^{-1}(B^\mathsf{T}\widehat{x}(k) - y),$$

and, using $L = B\Sigma^{-1}B^\mathsf{T}$, as equation (9.9).

To show fact (ii), define the error signal $\eta(k) = \widehat{x}^* - \widehat{x}(k)$. Note that $\eta(0) = \widehat{x}^*$ and that $\mathrm{average}(\eta(0)) = 0$ because $\mathbb{1}_n^\mathsf{T}\widehat{x}^* = 0$. Compute

$$\eta(k+1) = (I_n - \alpha L + \alpha L)\widehat{x}^* - (I_n - \alpha L)\widehat{x}(k) - \alpha B \Sigma^{-1} y$$
$$= (I_n - \alpha L)\eta(k) + \alpha(L\widehat{x}^* - B\Sigma^{-1}y)$$
$$= (I_n - \alpha L)\eta(k).$$

Now, according to Exercise E6.1, α is sufficiently small so that $I_n - \alpha L$ is non-negative. Moreover, $(I_n - \alpha L)$ is doubly-stochastic and symmetric, and its corresponding undirected graph is connected and aperiodic. Therefore, Theorem 5.1 implies that, as k diverges, $\eta(k) \to \mathrm{average}(\eta(0))\mathbb{1}_n = \mathbb{0}_n$. ∎

9.6   Historical notes and further reading

Standard references on incidence matrices include texts on algebraic graph theory such as (Biggs, 1994; Foulds, 1995; Godsil and Royle, 2001). An extensive discussion about algebraic potential theory on graphs is given by Biggs (1997).

The algorithm in Section 9.5.3 is taken from (Bolognani et al., 2010). For the notion of edge Laplacian and its properties, we refer to (Zelazo, 2009; Zelazo and Mesbahi, 2011; Zelazo et al., 2013). Additional references on distributed estimation for relative sensing networks include (Barooah and Hespanha, 2007, 2008; Bolognani et al., 2010; Piovan et al., 2013).

A comprehensive survey on cycle bases, their rich structure, and related algorithms is given by Kavitha et al. (2009).

9.7   Exercises

▶ E9.1   **Relations between incidence and Laplacian matrix.** Let G be a weighted undirected graph with n nodes and edge weights $\{a_e\}_{e\in\{1,\ldots,m\}}$. Select an enumeration and orientation for the edges. Let L, B and $\mathcal{A}$ denote, respectively, the Laplacian matrix, incidence matrix, and the weight matrix of G. Show that

(i) for any x and $y \in \mathbb{R}^n$,

$$y^\mathsf{T} L x = (B^\mathsf{T} y)^\mathsf{T} \mathcal{A} (B^\mathsf{T} x) = \sum_{\{i,j\}\in E} a_{ij}(x_i - x_j)(y_i - y_j); \tag{E9.1}$$

(ii) for any $x \in \mathbb{R}^n$, (as in equation (6.4))

$$x^\mathsf{T} L x = \|B^\mathsf{T} x\|_\mathcal{A}^2 = \sum_{\{i,j\}\in E} a_{ij}(x_i - x_j)^2; \tag{E9.2}$$

(iii) any $x \in \mathbb{R}^n$ satisfies

$$\mathrm{sign}(x)^\mathsf{T} L x \geq 0, \tag{E9.3}$$

and the inequality is strict if and only if there exists an edge $\{i, j\}$ such that either x_i and x_j have opposite sign or one is zero and the other is non-zero. (Here $\mathrm{sign}\colon \mathbb{R}^n \to \{-1, 0, +1\}^n$ is the entry-wise sign function.)

E9.2   **The edge Laplacian matrix (Zelazo and Mesbahi, 2011).** For an unweighted undirected graph with n nodes and m edges, introduce an arbitrary orientation for the edges. Recall the notions of incidence matrix $B \in \mathbb{R}^{n \times m}$ and Laplacian matrix $L = BB^\mathsf{T} \in \mathbb{R}^{n \times n}$ and define the *edge Laplacian matrix* by

$$L_{\text{edge}} = B^\mathsf{T} B \in \mathbb{R}^{m \times m}.$$

(Note that, in general, the edge Laplacian matrix is not a Laplacian matrix.) Select an edge orientation and compute B, L and L_{edge} for

  (i) a line graph with three nodes, and
 (ii) for the graph with four nodes in Figure 9.1.

Show that, for an arbitrary undirected graph,

(iii) $\text{kernel}(L_{\text{edge}}) = \text{kernel}(B)$;
 (iv) $\text{rank}(L) = \text{rank}(L_{\text{edge}})$;
  (v) for an acyclic graph L_{edge} is nonsingular (do not assume G is connected); and
 (vi) the non-zero eigenvalues of L_{edge} are equal to the non-zero eigenvalues of L.

E9.3   **Evolution of the relative disagreement error (Zelazo and Mesbahi, 2011).** Consider the Laplacian flow $\dot{x} = -Lx$, defined over an undirected, unweighted, and connected graph with n nodes and m edges. Beside the usual disagreement vector $\delta(t) = x(t) - \text{average}(x(t))\mathbb{1}_n \in \mathbb{R}^n$ (defined in Section 5.3), we can also analyze the *relative disagreement vector* $e_{ij}(t) = x_i(t) - x_j(t)$, for every edge $\{i, j\}$.

   (i) Write a differential equation for the relative disagreement errors $t \mapsto e(t) \in \mathbb{R}^m$.
   (ii) Based on Exercise E9.2, show that the relative disagreement errors converge to zero with exponential convergence rate given by the algebraic connectivity $\lambda_2(L)$.

E9.4   **Out-incidence and in-incidence matrices for digraphs.** Consider a weighted directed graph G with nodes $\{1, \ldots, n\}$ and edges $\{1, \ldots, m\}$. As usual, let A, D_{out}, D_{in}, and L denote the $n \times n$ adjacency, weighted out-degree, weighted in-degree, and Laplacian matrix, respectively, and let $\mathcal{A}$ denote the $m \times m$ weight matrix. Define the *out-incidence matrix* $B_{\text{out}} \in \{0, 1\}^{n \times m}$ and *in-incidence matrix* $B_{\text{in}} \in \{0, 1\}^{n \times m}$ by

$$(B_{\text{out}})_{ie} = \begin{cases} 1, & \text{if node } i \text{ is the head of edge } e, \\ 0, & \text{otherwise,} \end{cases} \tag{E9.4}$$

$$(B_{\text{in}})_{ie} = \begin{cases} 1, & \text{if node } i \text{ is the tail of edge } e, \\ 0, & \text{otherwise,} \end{cases} \tag{E9.5}$$

for each node $i \in \{1, \ldots, n\}$ and edge $e \in \{1, \ldots, m\}$.
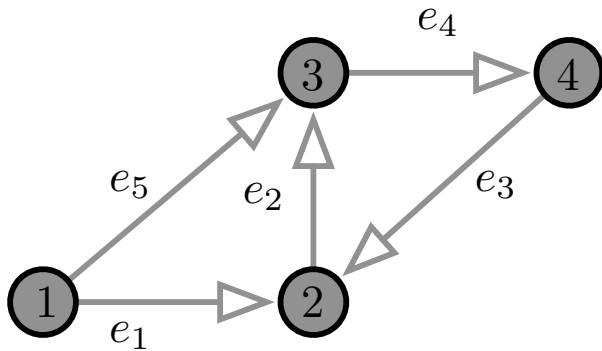
Figure E9.1: For this example digraph, we compute the 4×5 matrices

$$B_{\text{out}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B_{\text{in}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$ For comparison's sake, if we regard the edges of G as the oriented edges of an undirected graph G', then the oriented incidence matrix B of G' satisfies $B = B_{\text{out}} - B_{\text{in}}$; see equation (9.2).

Show that

(i) for each $x \in \mathbb{R}^n$ and each edge e of the form $e = (i, j)$,

$$(B_{\text{out}}^{\mathsf{T}} x)_e = x_i, \quad \text{and} \quad (B_{\text{in}}^{\mathsf{T}} x)_e = x_j, \tag{E9.6}$$

(ii) the following relationships hold:

$$D_{\text{out}} = B_{\text{out}} \mathcal{A} B_{\text{out}}^{\mathsf{T}}, \qquad\qquad\qquad D_{\text{in}} = B_{\text{in}} \mathcal{A} B_{\text{in}}^{\mathsf{T}}, \tag{E9.7}$$

$$A = B_{\text{out}} \mathcal{A} B_{\text{in}}^{\mathsf{T}}, \qquad\qquad\qquad L = B_{\text{out}} \mathcal{A} (B_{\text{out}} - B_{\text{in}})^{\mathsf{T}}, \tag{E9.8}$$

(iii) $\|B_{\text{out}}^{\mathsf{T}}\|_\infty = \|B_{\text{in}}^{\mathsf{T}}\|_\infty = 1$ and $\|B_{\text{out}}\|_\infty$ and $\|B_{\text{in}}\|_\infty$ are the maximum topological out-degree and in-degree of G, respectively.

Note: Balbuena et al. (2003) use these incidence matrices to study the line digraph.

E9.5   **Unoriented incidence matrix and signless Laplacian.** Consider a weighted undirected graph G with nodes $\{1, \ldots, n\}$, edges $\{1, \ldots, m\}$, adjacency matrix A and degree matrix $D = \mathrm{diag}(A\mathbb{1}_n)$. Define the *signless Laplacian matrix* $Q \in \mathbb{R}_{\geq 0}^{n \times n}$ by

$$Q = A + D \tag{E9.9}$$

and the *unoriented incidence matrix* $K \in \{0,1\}^{n \times m}$ by

$$K_{ie} = \begin{cases} 1, & \text{if node } i \text{ is incident to the undirected edge } e, \\ 0, & \text{otherwise.} \end{cases} \tag{E9.10}$$

Show that

(i) the degree of node i is equal to $(A\mathbb{1}_n)_i = (K\mathcal{A}\mathbb{1}_m)_i$,

(ii) $Q = K\mathcal{A}K^{\mathsf{T}} \succeq 0$, where $\mathcal{A} \in \mathbb{R}^{m \times m}$ denotes the weight matrix of G,

(iii) $x^{\mathsf{T}}Qx = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x_i + x_j)^2$ for any $x \in \mathbb{R}^n$, and

(iv) Q has an eigenvalue equal to 0 if and only if G consists of isolated vertices and bipartite connected components.

Note: We refer to (Desai and Rao, 1994; Cvetković and Simić, 2009) for early references.

▶ E9.6   **Averaging with distributed integral control.** Consider a Laplacian flow implemented as a relative sensing network over a connected and undirected graph with n nodes, m edges, incidence matrix $B \in \mathbb{R}^{n \times m}$, and weights $a_{ij} > 0$ for $(i, j) \in E$, and subject to a constant disturbance term $\eta \in \mathbb{R}^m$, as shown in Figure E9.2.
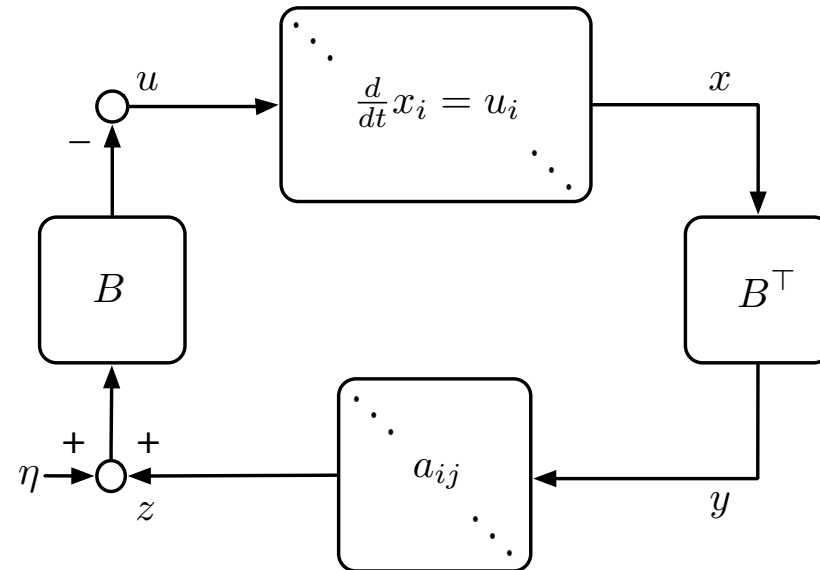
Figure E9.2: A relative sensing network with a constant disturbance input $\eta \in \mathbb{R}^m$.

(i) Derive the dynamic closed-loop equations describing the model in Figure E9.2.
(ii) Show that the state $x(t)$ converges asymptotically to a vector $x^* \in \mathbb{R}^n$ that depends on the value of the disturbance η and that is not necessarily a consensus state.

Consider the system in Figure E9.2 with a *distributed integral controller* forcing convergence to consensus, as shown in Figure E9.3.
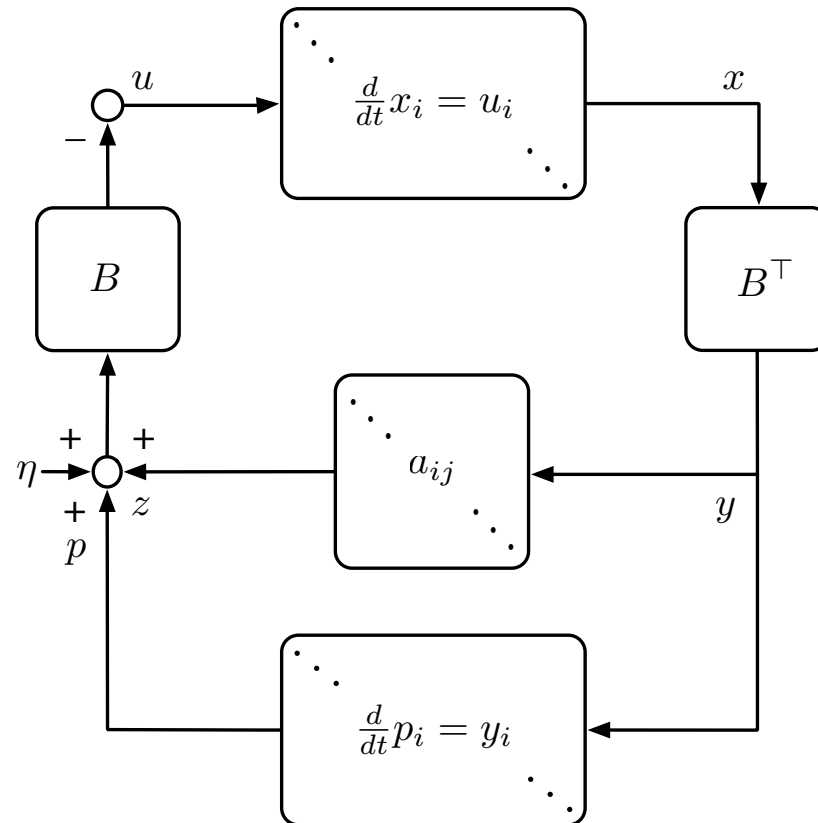
Figure E9.3: Relative sensing network with a disturbance $\eta \in \mathbb{R}^m$ and distributed integral action.

(iii) Derive the dynamic closed-loop equations describing the model in Figure E9.3.

(iv) Show that the distributed integral controller in Figure E9.3 asymptotically stabilizes the set of steady states (x^*, p^*), with $x^* \in \mathrm{span}\{\mathbb{1}_n\}$ corresponding to consensus.

Hint: Study the properties of saddle point matrices in Exercise E9.13.

E9.7   **Incidence matrix, cutset and cycle spaces for a triangle.** Consider an undirected triangle with nodes $\{1, 2, 3\}$. Let $\gamma = (1, 2, 3)$ be the only simple cycle (i.e.,
closed walk with three or more nodes). Select the number and orientation of the three edges as well as three possible cuts as in Figure E9.4. Perform the following
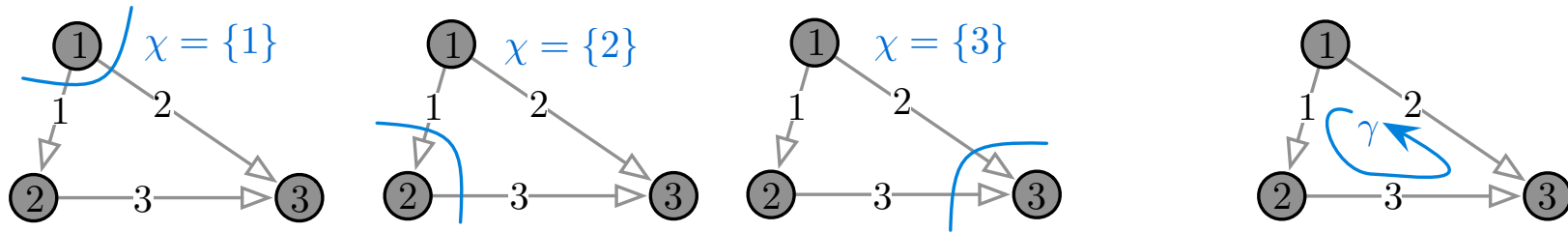
Figure E9.4: Three cuts and a cycle in a triangle.

steps:

(i)   compute the incidence matrix B, the cutset orientation vector $v_{\{i\}}$ for each cut $\{i\}$, and the signed walk vector for γ,

(ii) show $v_{\{1\}} + v_{\{2\}} + v_{\{3\}} = \mathbb{0}_3$ and $\mathrm{span}\{v_{\{1\}}, v_{\{2\}}, v_{\{3\}}\} \perp \mathrm{span}\{w_\gamma\}$, and

(iii) verify the equality (9.5) between incidence matrix and cutset orientation vectors.

E9.8   **Incidence matrix, cutset and cycle spaces for basic graphs.** Recall Examples 3.1 and 4.1, and consider the following unweighted undirected graphs with node set
$\{1, \ldots, 4\}$:

  (i) the path graph P_4;
 (ii) the cycle graph C_4;
(iii) the star graph S_4; and
(iv)   the complete graph K_4.

For each graph, select an arbitrary orientation of the edges, compute the incidence matrix, compute a basis for the cutset space, and compute a basis for the cycle space.

E9.9   **Incidence matrix and signed walk vectors.** Given an undirected graph G, consider an arbitrary orientation of its edges, its incidence matrix $B \in \mathbb{R}^{n \times m}$, and a simple walk γ with distinct initial and final nodes described by a signed walk vector $w^\gamma \in \mathbb{R}^m$.

(i) Show that the vector $y = Bw^\gamma \in \mathbb{R}^n$ has components

$$
y_i = \begin{cases} +1, & \text{if node } i \text{ is the initial node of } \gamma, \\ -1, & \text{if node } i \text{ is the final node of } \gamma, \\ 0, & \text{otherwise.} \end{cases}
$$

(ii) Prove statement (i) in Theorem 9.5.

E9.10   **Properties of signed walk vectors.** Let G be an undirected unweighted graph and let χ and ψ be two cuts on G. Show that:

(i) $v_{\chi^c} = -v_{\chi}$,
(ii) if $\chi \cap \psi = \emptyset$, then $v_{\chi} + v_{\psi} = v_{\chi \cup \psi}$, and
(iii) if G has d connected components, then there exist $n - d$ independent cutset orientation vectors.

E9.11 **The orthogonal projection onto the cutset space (Jafarpour and Bullo, 2019).** Recall the following well-known facts from linear algebra: a square matrix $P \in \mathbb{R}^{m \times m}$ is an *orthogonal projection* if $P = P^\mathsf{T}$ and $P^2 = P$; given a full-rank matrix $X \in \mathbb{R}^{m \times n}$, $n < m$, the matrix $P = X(X^\mathsf{T}X)^{-1}X^\mathsf{T}$ is the orthogonal projection onto the $\mathrm{image}(X)$. Prove that

(i) if X is not full rank (i.e., it has a trivial kernel), the matrix $P = X(X^\mathsf{T}X)^\dagger X^\mathsf{T}$ is the orthogonal projection onto $\mathrm{image}(X)$, where $(X^\mathsf{T}X)^\dagger$ is the pseudoinverse of $X^\mathsf{T}X$, as defined in Exercise E2.20.

Given an unweighted undirected graph with an oriented incidence matrix B, Laplacian matrix $L = BB^\mathsf{T}$, and pseudoinverse Laplacian matrix $L^\dagger$, recall that $\mathbb{R}^m = \mathrm{image}(B^\mathsf{T}) \oplus \mathrm{kernel}(B)$ is the orthogonal decomposition into cutset space and cycle space. Show that

(ii) $P = B^\mathsf{T}L^\dagger B$ is an orthogonal projection matrix, and

(iii) $P = B^\mathsf{T}L^\dagger B$ is the orthogonal projection onto the cutset space $\mathrm{image}(B^\mathsf{T})$.

E9.12 **Sensitivity of Laplacian eigenvalues.** Consider an unweighted undirected graph $G = (V, E)$ with incidence matrix $B \in \mathbb{R}^{n \times m}$, and Laplacian matrix $L = BB^{\mathsf{T}} \in \mathbb{R}^{n \times n}$. Define a undirected graph G' by adding one unweighted edge $e \notin E$ to G, that is, $G' = (V, E \cup e)$. Show that

$$\lambda_{\max}(L_G) \leq \lambda_{\max}(L_{G'}) \leq \lambda_{\max}(L_G) + 2.$$

Hint: Use the edge Laplacian matrix $L_{\text{edge}} = B^{\mathsf{T}} B \in \mathbb{R}^{m \times m}$ in Exercise E9.2 and Cauchy's Interlacing Theorem (e.g., see (Horn and Johnson, 1985, Theorem 4.3.17)): Let A denote a symmetric matrix with ordered eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and B denote a principal submatrix of A with ordered eigenvalues $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_{n-1}$; then the eigenvalues of A and B interlace, that is, $\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \ldots \leq \mu_{n-1} \leq \lambda_n$.

E9.13 **Spectrum of saddle point matrices.** Given a positive semidefinite matrix $\mathcal{S} = \mathcal{S}^{\mathsf{T}} \succeq 0$ in $\mathbb{R}^{n \times n}$ and a matrix $\mathcal{C}$ in $\mathbb{R}^{m \times n}$, define the *saddle point matrix* $\mathcal{A}_{\mathrm{sp}} \in \mathbb{R}^{(n+m) \times (n+m)}$ by

$$\mathcal{A}_{\mathrm{sp}} = \begin{bmatrix} \mathcal{S} & \mathcal{C}^{\mathsf{T}} \\ -\mathcal{C} & \mathbb{0}_{m \times m} \end{bmatrix}.$$

Then each eigenvalue λ of $\mathcal{A}_{\mathrm{sp}}$ satisfies

(i) $\Re(\lambda) \geq 0$,
(ii) if $\mathrm{kernel}(\mathcal{S}) \cap \mathrm{image}(\mathcal{C}^{\mathsf{T}}) = \{\mathbb{0}_n\}$, then either $\Re(\lambda) > 0$ or $\lambda = 0$; moreover, if $\lambda = 0$, then λ is semisimple, and
(iii) if $\mathcal{S}$ is positive definite and $\mathrm{kernel}(\mathcal{C}^{\mathsf{T}}) = \{\mathbb{0}_m\}$, then $\Re(\lambda) > 0$.

Note: Statements (i) and (iii) are (Benzi et al., 2005, Theorem 3.6). Statement (ii) is (Cherukuri et al., 2017, Lemma 5.3). Additional results on saddle point matrices are given in (Dörfler et al., 2018, Proposition 5.13).

E9.14   **Monotonicity of the Laplacian as a positive semidefinite matrix.** As in Lemma 6.9, consider a weighted undirected graph with symmetric adjacency matrix A and symmetric Laplacian matrix L. Show that

$$A \leq A' \quad \implies \quad L \preceq L',$$

where A' is a symmetric adjacency matrix with corresponding Laplacian L' and where $L \preceq L'$ means that $L - L'$ is negative semidefinite.

E9.15   **The pseudoinverse of the incidence matrix of a spanning tree.** Let T be a spanning tree of the complete graph K_n and H be a subgraph of K_n with h edges. Given an edge enumeration and orientation, let $B_T \in \mathbb{R}^{n \times (n-1)}$ and $B_H \in \mathbb{R}^{n \times h}$ denote the incidence matrices of T and H, respectively. Let $B_T^\dagger \in \mathbb{R}^{(n-1) \times n}$ denote the pseudoinverse of B_T. Show

(i) $B_T^\dagger B_T = I_{n-1}$ and $B_T B_T^\dagger = I_n - \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T}$,

(ii) for any $x \in \mathbb{R}^n$ with $x_{\text{ave}} = \mathbb{1}_n^\mathsf{T} x / n$ and with the notation $f_T = B_T^\mathsf{T} x \in \mathbb{R}^{n-1}$ and $f_H = B_H^\mathsf{T} x \in \mathbb{R}^h$,

$$(B_T^\dagger)^\mathsf{T} f_T = x - x_{\text{ave}} \mathbb{1}_n, \quad \text{and} \quad f_H = B_H^\mathsf{T} (B_T^\dagger)^\mathsf{T} f_T.$$

Note: Bapat (1997) gives further properties of the pseudoinverse of the incidence matrix of a tree.

Chapter 10

Metzler Matrices and Dynamical Flow Systems

In this chapter we study positive systems, that is, dynamical systems with state variables that are typically non-negative, and dynamical flow systems (also called compartmental systems), that is, dynamical processes characterized by conservation laws (e.g., mass, fluid, energy) and by the flow of commodities between units known as compartments. For simplicity we focus on continuous-time models, though a comparable theory exists for discrete-time systems. Example dynamical flow systems are transportation networks, queueing networks, communication networks, epidemic propagation models in social contact networks, as well as ecological and biological networks. Linear dynamical flow systems and positive systems in continuous time are described by so-called Metzler matrices; we define and study such matrices in this chapter.

10.1   Example systems

In this section we review some examples of dynamical flow systems.

Ecological and environmental systems   The flow of energy and nutrients (water, nitrates, phosphates, etc) in *ecosystems* is typically studied using compartmental modeling. For example, Figure 10.1 illustrates a widely-cited water flow model for a desert ecosystem (Noy-Meir, 1973). Other classic ecological network systems include models for dissolved oxygen in stream, nutrient flow in forest growth and biomass flow in fisheries (Walter and Contreras, 1999).
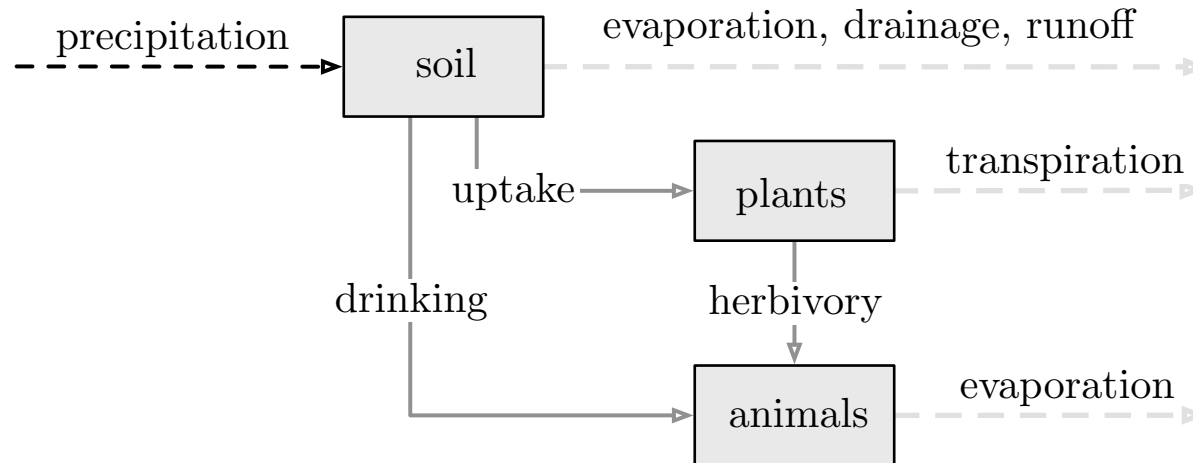
Figure 10.1: Water flow model for a desert ecosystem. The black dashed line denotes an inflow from the outside environment. The light-gray dashed lines denote outflows into the outside environment.

Epidemiology of infectious diseases   To study the *propagation of infectious diseases*, the population at risk is typically divided into compartments consisting of individuals who are susceptible (S), infected (I), and, possibly, recovered and no longer susceptible (R). As illustrated in Figure 10.2, the three basic epidemiological models are (Hethcote, 2000) called SI, SIS, SIR, depending upon how the disease spreads. For a review article in the spirit of these lecture notes, we refer the interested reader to (Mei et al., 2017).
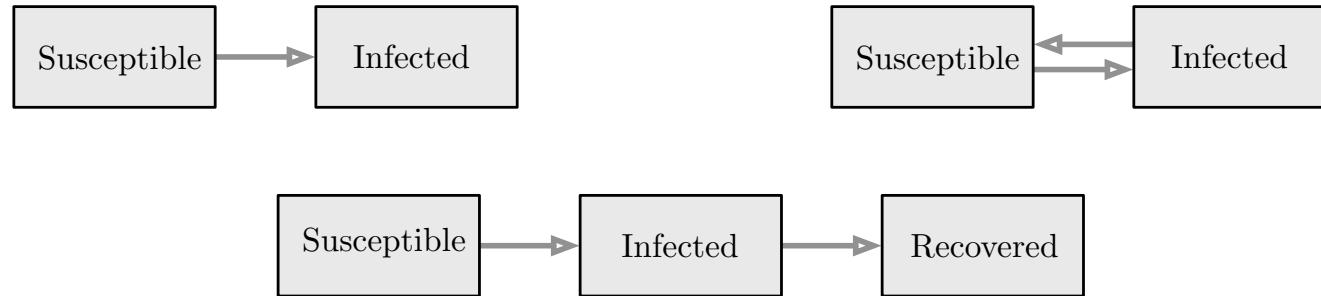
Figure 10.2: The three basic models SI, SIS and SIR for the propagation of an infectious disease

Drug and chemical kinetics in biomedical systems   Compartmental model are also widely adopted to characterize the kinetics of drugs and chemicals in biomedical systems. Here is a classic example (Charkes et al., 1978) from nuclear medicine: bone scintigraphy (also called bone scan) is a medical test in which the patient is injected with a small amount of radioactive material and then scanned with an appropriate radiation camera.
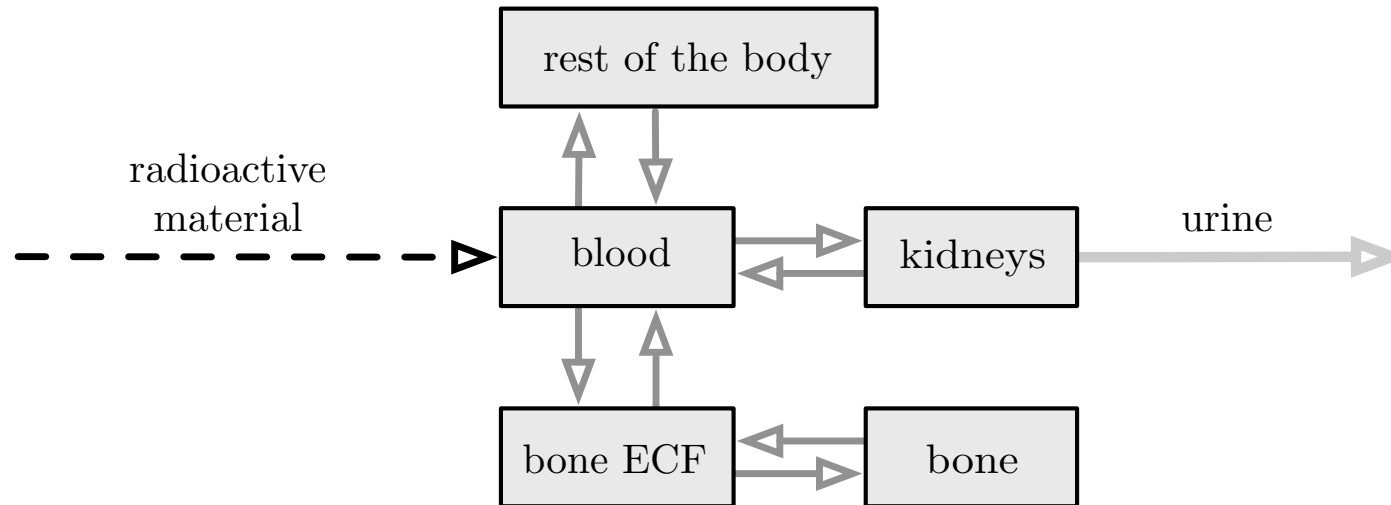
Figure 10.3: The kinetics of a radioactive isotope through the human body (ECF = extra-cellular fluid).

10.2　Metzler matrices and positive systems

We start by introducing a new class of matrices.

Definition 10.1 (Metzler matrix). *For a matrix $M \in \mathbb{R}^{n \times n}$, $n \geq 2$,*

 (i) *M is Metzler if all its off-diagonal elements are non-negative;*

(ii) *if M is Metzler, its associated digraph is a weighted digraph defined as follows: $\{1, \dots, n\}$ are the nodes, there are no self-loops, (i, j), $i \neq j$ is an edge with weight m_{ij} if and only if $M_{ij} > 0$; and*

(iii) *if A is Metzler, A is irreducible if its associated digraph is strongly connected.*

$$\begin{bmatrix} * & + & + & + \\ + & * & + & + \\ + & + & * & + \\ + & + & + & * \end{bmatrix}$$

Figure 10.4: The sign pattern of a Metzler matrix: $+$ stands for non-negative and $*$ stands for arbitrary.

The sign pattern of a Metzler matrix is illustrated in Figure 10.4. Metzler matrices are sometimes also referred to as *quasi-positive* or *essentially non-negative*.

Note: M is Metzler if and only if there exists a scalar $\gamma > 0$ such that $M + \gamma I_n$ is non-negative.

Note: if L is a Laplacian matrix, then $-L$ is a Metzler matrix with zero row-sums.

Metzler matrices have numerous properties. We start by writing a version of Perron–Frobenius Theorem 2.12 and illustrating it in Figure 10.5.

Theorem 10.2 (Perron–Frobenius Theorem for Metzler matrices). *If $M \in \mathbb{R}^{n \times n}$, $n \geq 2$, is Metzler, then*

 (i) *there exists a real eigenvalue λ such that $\lambda \geq \Re(\mu)$ for all other eigenvalues μ, and*

 (ii) *the right and left eigenvectors of λ can be selected non-negative.*

 If additionally M is irreducible, then

(iii) *there exists a real simple eigenvalue λ such that $\lambda > \Re(\mu)$ for all other eigenvalues μ, and*

(iv) *the right and left eigenvectors of λ are unique and positive (up to rescaling).*
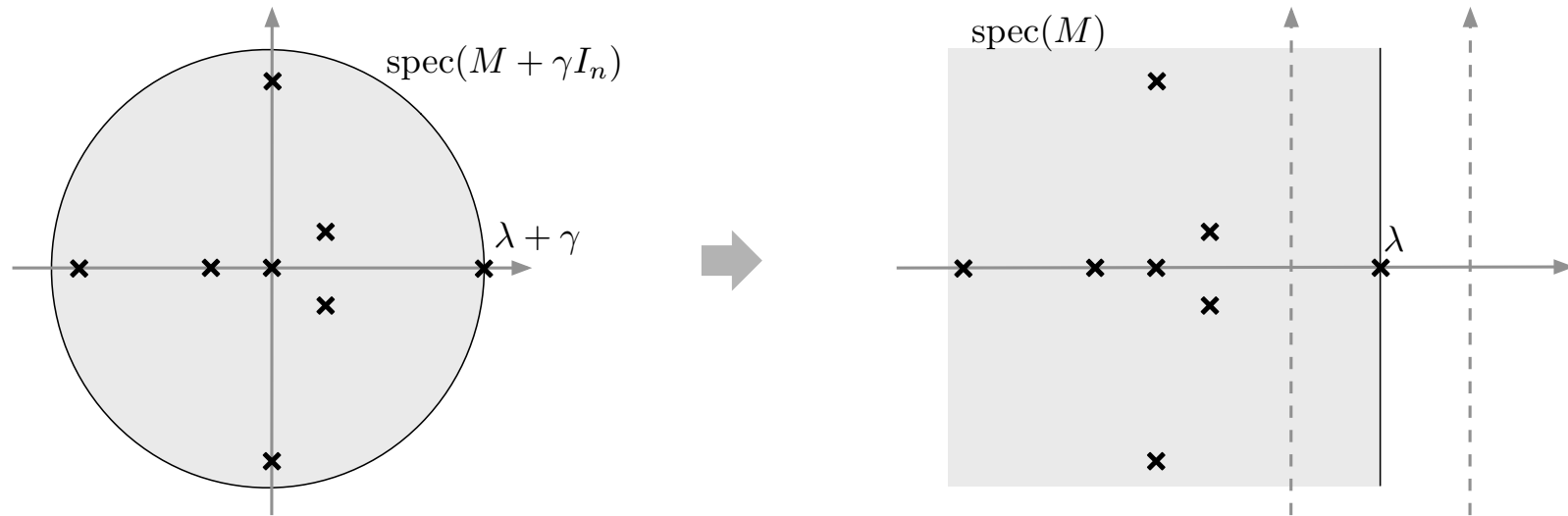
Figure 10.5: Illustrating the Perron–Frobenius Theorem 10.2 for a Metzler matrix M. Left image: for sufficiently large γ, $M + \gamma I_n$ is non-negative and $\lambda + \gamma$ is its dominant Perron eigenvalue. Right image: the spectrum of M is equal to that of $M + \gamma I_n$ translated by $-\gamma$; λ is dominant in the sense that $\lambda \geq \Re(\mu)$ for all other eigenvalues μ; it is not determined whether $\lambda < 0$ (the imaginary axis is to the right of λ) or $\lambda > 0$ (the imaginary axis is to the left of λ).

Note: As in the case of non-negative matrices, we refer to λ as to the dominant eigenvalue. For a Metzler matrix M, the dominant eigenvalue is equal to the spectral abscissa $\alpha(M)$ (whereas the dominant eigenvalue of a non-negative matrix A is its spectral radius $\rho(A)$). We invite the reader to work out the details of the proof in Exercise E10.2.

Note: this theorem is consistent with and generalizes the treatment of Laplacian matrices. Specifically we know that, if L is a

Laplacian, the Metzler matrix $-L$ has dominant eigenvalue $\lambda = 0$.

Next, we give necessary and sufficient conditions for the dominant eigenvalue of a Metzler matrix to be strictly negative.

Theorem 10.3 (Metzler Hurwitz Theorem). *For a Metzler matrix M, the following statements are equivalent:*

 (i) M *is Hurwitz,*

 (ii) M *is invertible and* $-M^{-1} \geq 0$, *and*

(iii) for all $b \geq \mathbb{0}_n$, *there exists a unique* $x^* \geq \mathbb{0}_n$ *solving* $Mx^* + b = \mathbb{0}_n$.

Moreover, if M is Metzler, Hurwitz and irreducible, then $-M^{-1} > 0$.

Proof. We start by showing that (i) implies (ii). Clearly, if M is Hurwitz, then it is also invertible. So it suffices to show that $-M^{-1}$ is non-negative. As in Exercise E6.1 and in the proof of Theorem 7.4, pick $\varepsilon > 0$ and define $\mathcal{A}_{M,\varepsilon} = I_n + \varepsilon M$, that is, $(-\varepsilon M) = (I_n - \mathcal{A}_{M,\varepsilon})$. Because M is Metzler, ε can be selected small enough so that $\mathcal{A}_{M,\varepsilon} \geq 0$. Moreover, because the spectrum of M is strictly in the left half plane, one can verify that, for ε small enough, $\operatorname{spec}(\varepsilon M)$ is inside the disk of unit radius centered at the point -1; as illustrated in Figure 10.6. In turn, this last property implies that $\operatorname{spec}(I_n + \varepsilon M)$ is strictly inside the disk of unit radius
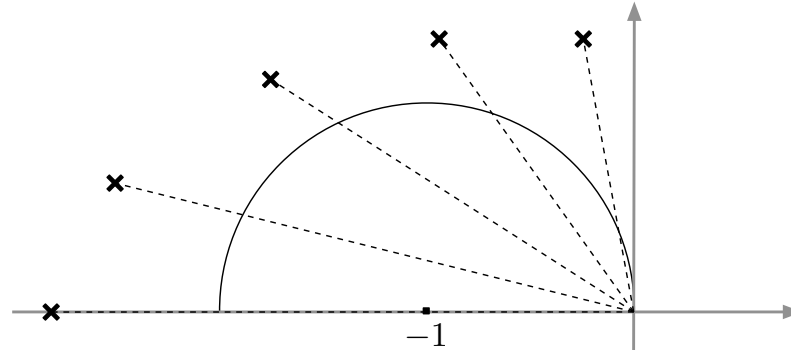
Figure 10.6: For any $\lambda \in \mathbb{C}$ with strictly negative real part, there exists ε such that the segment from the origin to $\varepsilon \lambda$ is inside the disk of unit radius centered at the point -1.

centered at the origin, that is, $\rho(\mathcal{A}_{M,\varepsilon}) < 1$. We now adopt the Neumann series as defined in Exercise E2.11: because $\rho(\mathcal{A}_{M,\varepsilon}) < 1$, we know that $(I_n - \mathcal{A}_{M,\varepsilon}) = (-\varepsilon M)$ is invertible and that

$$(-\varepsilon M)^{-1} = (I_n - \mathcal{A}_{M,\varepsilon})^{-1} = \sum_{k=0}^{\infty} \mathcal{A}_{M,\varepsilon}^k. \tag{10.1}$$

Note now that the right-hand side is non-negative because it is the sum of non-negative matrices. In summary, we have shown that M is invertible and that $-M^{-1} \geq 0$. This statement proves that (i) implies (ii).

Next we show that (ii) implies (i). We know M is Metzler, invertible and satisfies $-M^{-1} \geq 0$. By the Perron–Frobenius Theorem 10.2 for Metzler matrices, we know there exists $v \geq \mathbb{0}_n$, $v \neq \mathbb{0}_n$, satisfying $Mv = \lambda_{\text{Metzler}}(M)v$, where $\lambda_{\text{Metzler}}(M) = \alpha(M) = \max\{\Re(\lambda) \mid \lambda \in \operatorname{spec}(M)\}$. Clearly, M invertible implies $\lambda_{\text{Metzler}}(M) \neq 0$ and, moreover, $v = \lambda_{\text{Metzler}}(M)M^{-1}v$. Now, we

know v is non-negative and $M^{-1}v$ is non-positive. Hence, $\lambda_{\text{Metzler}}(M)$ must be negative and, in turn, M is Hurwitz. This statement establishes the equivalence between (ii) implies (i)

Finally, regarding the equivalence between statement (ii) and statement (iii), note that, if $-M^{-1} \geq 0$ and $b \geq \mathbb{0}_n$, then clearly $x^* = -M^{-1}b \geq \mathbb{0}_n$ is unique and solves $Mx^* + b = \mathbb{0}_n$. This proves that (ii) implies (iii). Vice versa, if statement (iii) holds, then let x_i^* be the non-negative solution of $Mx_i^* = -\mathbb{e}_i$ and let X be the non-negative matrix with columns $x_1^*, \ldots, x_n^*$. Therefore, we know $MX = -I_n$ so that M is invertible, $-X$ is its inverse, and $-M^{-1} = -(-X) = X$ is non-negative. This statement proves that (iii) implies (ii).

Finally, the statement that $-M^{-1} > 0$ for each Metzler, Hurwitz and irreducible matrix M is proved as follows. Because M is irreducible, the matrix $\mathcal{A}_{M,\varepsilon} = I_n + \varepsilon M$ is non-negative (for ε sufficiently small) and primitive. Therefore, the right-hand side of equation (10.1) is strictly positive. ∎

Remark 10.4 (Hurwitz Metzler and M-matrices). *The following notion is often adopted in the literature, e.g., see (Horn and Johnson, 1994, Section 2.5). A matrix $U \in \mathbb{R}^{n \times n}$ is an M-matrix if $u_{ii} > 0$, $u_{ij} \leq 0$ for all $i \neq j$, U is invertible, and $U^{-1} \geq 0$. It is easy to see that U is an M-matrix if and only if $-U$ is Hurwitz Metzler.*

10.2.1 Continuous-time positive systems

Motivated by the examples of dynamical flow systems in the previous section and by the treatment of Metzler matrices, we are now ready to introduce the class of positive systems.

Definition 10.5 (Positive systems in continuous time). *A dynamical system $\dot{x}(t) = f(x(t), t)$, $x \in \mathbb{R}^n$, is positive if $x(0) \geq \mathbb{0}_n$ implies $x(t) \geq \mathbb{0}_n$ for all $t \in \mathbb{R}_{\geq 0}$.*

We are especially interested in linear and affine systems, described by

$$\dot{x}(t) = Mx(t), \qquad \text{and} \quad \dot{x}(t) = Mx(t) + b,$$

where $M \in \mathbb{R}^{n \times n}$. The following theorem characterizes the importance of Metzler matrices for continuous-time positive systems and extends the results in Exercise E7.2 about convergence of affine systems.

Theorem 10.6 (Positive affine systems and Metzler matrices). *For the affine system $\dot{x}(t) = Mx(t) + b$, the following statements are equivalent:*

(i) the system is positive, that is, $x(t) \geq \mathbb{0}_n$ for all $t \in \mathbb{R}_{\geq 0}$ and all $x(0) \geq \mathbb{0}_n$,

(ii) M is Metzler and $b \geq \mathbb{0}_n$.

Moreover, if the matrix M is Metzler and Hurwitz, then

(iii) the system has a unique non-negative equilibrium point $x^* \in \mathbb{R}^n_{\geq 0}$, to which all trajectories converge asymptotically.

Proof. We start by showing that statement (i) implies statement (ii). If $x(0) = \mathbb{0}_n$, then $\dot{x}$ cannot have any negative components, hence $b \geq \mathbb{0}_n$. If any off-diagonal entry (i, j), $i \neq j$, of M is strictly negative, then consider an initial condition $x(0)$ with all zero entries except for $x(j) > b_i/|M_{ij}|$. It is easy to see that $\dot{x}_i(0) < 0$ which is a contradiction.

Next, we show that statement (ii) implies statement (i). It suffices to note that, anytime there exists i such that $x_i(t) = 0$, the conditions $x(t) \geq \mathbb{0}_n$, M Metzler and $b \geq \mathbb{0}_n$ together imply $\dot{x}_i(t) = \sum_{i \neq j} M_{ij} x_j(t) + b_i \geq 0$.

Statement (iii) follows from the Metzler Hurwitz Theorem 10.3 and Exercise E7.2. ■

10.2.2 Table of correspondences between non-negative and Metzler matrices

We conclude this section by highlighting the correspondences between non-negative and Metzler matrices in Table 10.1.

Non-negative matrices	Metzler matrices
	Exercise E10.1: M Metzler $\iff \exp(tM) \geq 0$ for all $t \geq 0$
Perron-Frobenius Theorem 2.12: (reducible and irreducible) non-negative matrix A has dominant eigenpair	Perron-Frobenius Theorem 10.2: (reducible and irreducible) Metzler matrix M has dominant eigenpair
Exercise E10.3: A is non-negative and convergent ($\rho(A) < 1$) $\iff (I_n - A)$ invertible and $(I_n - A)^{-1} \geq 0$	Theorem 10.3: M is Metzler and Hurwitz ($\alpha(M) < 0$) $\iff M$ invertible and $-M^{-1} \geq 0$
Exercise E10.4: A is non-negative $\iff x(k+1) = Ax(k)$ is positive	Theorem 10.6: M is Metzler $\iff \dot{x}(t) = Mx(t)$ is positive
Lemmas 4.9 and 4.10: Bounds and monotonicity of $\rho(A)$ for non-negative A	Exercise E10.5: Bounds and monotonicity of $\alpha(M)$ for Metzler M

Table 10.1: Table of correspondences between non-negative and Metzler matrices

10.3 Dynamical flow systems

In this section, motivated by the examples in Section 10.1 and by the treatment in Section 1.4, we study an important class of positive affine systems.

A *dynamical flow system* is a dynamical system in which material is stored at individual locations and is transferred along the edges of directed graph, called the *compartmental digraph*; see Figure 10.7b. Dynamical flow systems are also referred to as compartmental systems. The "storage" nodes are referred to as *compartments*; each compartment contains a time-varying quantity $q_i(t)$. Each
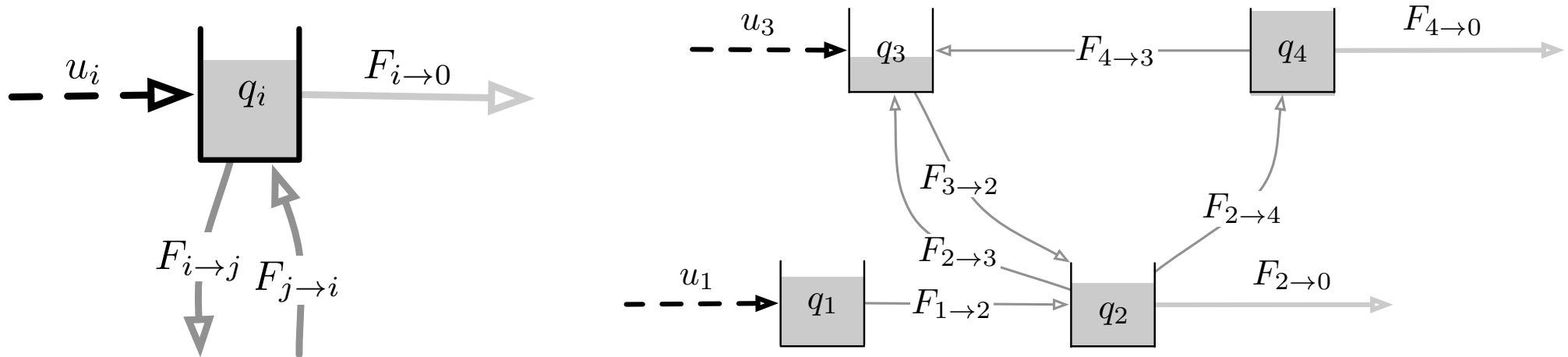
(a) A compartment with inflow u_i, outflow $F_{i \to 0}$, and inter-compartmental flows $F_{i \to j}$

(b) A compartmental system with two inflows and two outflows

Figure 10.7: Example of a single compartment and of a dynamical flow system

directed arc (i, j) represents a *mass flow* (or *flux*), denoted $F_{i \to j}$, from compartment i to compartment j. The dynamical flow system interacts with its surrounding environment via inputs and output flows, denoted in figure by black dashed and light-gray solid arcs respectively: the inflow from the environment into compartment i is denoted by u_i and the outflow from compartment i into the environment is denoted by $F_{i \to 0}$.

In summary, a (nonlinear) dynamical flow system is described by an unweighted digraph, by maps $F_{i \to j}$ for all edges (i, j) of

the digraph, and by inflow and outflow maps. (The compartmental digraph has no self-loops.) The dynamic equations of the dynamical flow system are obtained by the *instantaneous flow balance* at each compartment. In other words, asking that the rate of accumulation at each compartment equals the net inflow rate we obtain:

$$\dot{q}_i(t) = \sum_{j=1, j \neq i}^{n} (F_{j \to i} - F_{i \to j}) - F_{i \to 0} + u_i. \tag{10.2}$$

In general, the flow along (i, j) is a function of the entire system state so that $F_{i \to j} = F_{i \to j}(q)$.

Remarks 10.7 (Basic properties of dynamical flow systems). *(i) The mass in each of the compartments as well as the mass flowing along each of the edges must be non-negative at all times (recall we assume $u_i \geq 0$). Specifically, we require the mass flow functions to satisfy*

$$F_{i \to j}(q) \geq 0 \text{ for all } (q), \quad \text{and} \quad F_{i \to j}(q) = 0 \text{ for all } (q) \text{ such that } q_i = 0. \tag{10.3}$$

Under these conditions, if at some time t_0 one of the compartments has no mass, that is, $q_i(t_0) = 0$ and $q(t_0) \in \mathbb{R}^n_{\geq 0}$, it follows that $\dot{q}_i(t_0) = \sum_{j=1, j \neq i}^n F_{j \to i}(q(t_0)) + u_i \geq 0$ so that q_i does not become negative. The dynamical flow system (10.2) is therefore a positive system, as introduced in Definition 10.5.

(ii) Let $F_{\text{flows}} \colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}^{n \times n}_{\geq 0}$ denote the compartment-to-compartment flow matrix with entries $(F_{\text{flows}})_{ij}(q) = F_{i \to j}(q)$ (and zero diagonal) and $F_{\text{outflows}} \colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n_{\geq 0}$ denote the outflow vector. Then the equation (10.2) is written in vector form as

$$\dot{q} = F_{\text{flows}}(q)^\mathsf{T} \mathbb{1}_n - F_{\text{flows}}(q) \mathbb{1}_n - F_{\text{outflows}}(q) + u.$$

(iii) If $M(q) = \sum_{i=1}^n q_i = \mathbb{1}_n^\mathsf{T} q$ denotes the total mass in the system, then along the solutions of (10.2)

$$\frac{d}{dt} M(q(t)) = -\underbrace{\sum_{i=1}^n F_{i \to 0}(q(t))}_{\text{outflow into environment}} + \underbrace{\sum_{i=1}^n u_i}_{\text{inflow from environment}}. \tag{10.4}$$

This equality implies that the total mass $t \mapsto M(q(t))$ is constant in systems without inflows and outflows. ●

Linear dynamical flow systems

Loosely speaking, a dynamical flow system is linear if (i) all flows depend linearly upon the mass in the originating compartment, except (ii) the inflow from the environment is constant and non-negative. In other words, we assume that in a linear dynamical flow system,

$$F_{i \to j}(q) = f_{ij} q_i, \quad \text{for } j \in \{1, \dots, n\},$$
$$F_{i \to 0}(q) = f_{i0} q_i, \quad \text{and}$$
$$u_i(q) = u_i,$$

where the f_{ij} and f_{i0} coefficients are called *flow rates*. Indeed, this model is also referred to as *donor-controlled flow*. Note that this model satisfies the physically-meaningful constraints (10.3). We illustrate these assumptions in Figure 10.8.
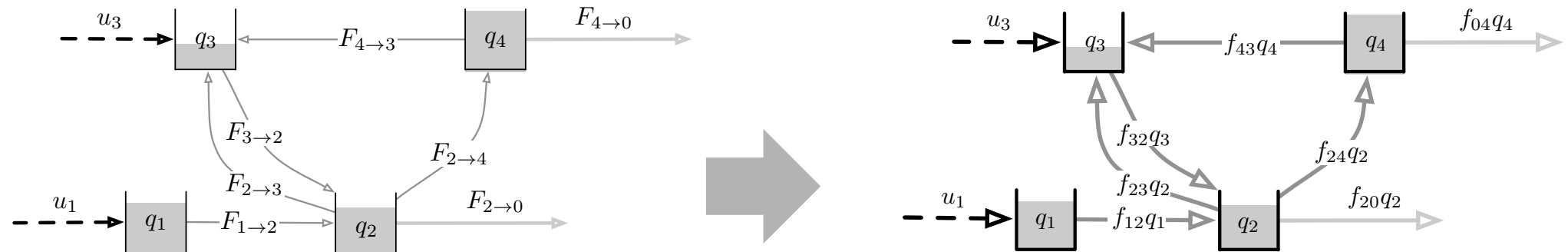
Figure 10.8: In a linear dynamical flow system the flows among the compartments and onto the environment are linear functions of the donor compartment. The compartmental digraph has weights on the edges given by the corresponding flow rates and, accordingly, its adjacency matrix is called the *flow rate matrix*.

Definition 10.8 (Linear dynamical flow systems). *A linear dynamical flow system with n compartments is a triplet (F, f_0, u) consisting of*

(i) a non-negative $n \times n$ matrix $F = (f_{ij})_{i,j \in \{1,\dots,n\}}$ with zero diagonal, called the flow rate matrix,

(ii) a vector $f_0 \geq \mathbb{0}_n$, called the outflow rates vector, and

(iii) a vector $u \geq \mathbb{0}_n$, called the inflow vector.

For a linear dynamical flow system, it is customary to regard the flow rate matrix F as the adjacency matrix of the compartmental digraph (which is now therefore a weighted digraph without self-loops) and denote the compartmental digraph by G_F.

With the notion of compartmental matrix, the dynamics of the linear dynamical flow system are

$$\dot{q}_i(t) = -\left(f_{i0} + \sum_{j=1,j\neq i}^n f_{ij}\right) q_i(t) + \sum_{j=1,j\neq i}^n f_{ji} q_j(t) + u_i, \tag{10.5}$$

or, in vector notation,

$$\dot{q}(t) = Cq(t) + u, \tag{10.6}$$

where the *compartmental matrix* $C = (c_{ij})_{i,j\in\{1,\ldots,n\}}$ of a dynamical flow system (F, f_0, u) is defined by

$$c_{ij} = \begin{cases} f_{ji}, & \text{if } i \neq j, \\ -f_{i0} - \sum_{h=1,h\neq i}^n f_{ih}, & \text{if } i = j. \end{cases}$$

Note that here the components of the vector f_0 are $(f_0)_i = f_{i0}$, for consistency with the notation above. Equivalently, if $L_F = \mathrm{diag}(F\mathbb{1}_n) - F$ is the Laplacian matrix of the compartmental digraph, then the compartmental matrix satisfies

$$C = -L_F^\mathsf{T} - \mathrm{diag}(f_0) = F^\mathsf{T} - \mathrm{diag}(F\mathbb{1}_n + f_0). \tag{10.7}$$

Note: since $L_F \mathbb{1}_n = \mathbb{0}_n$, we know $\mathbb{1}_n^\mathsf{T} C = -f_0^\mathsf{T}$ and, consistently with equation (10.4), we know $\frac{d}{dt} M(q(t)) = -f_0^\mathsf{T} q(t) + \mathbb{1}_n^\mathsf{T} u$.

In what follows it is convenient to introduce the following definition.

Definition 10.9 (Compartmental matrices). *A matrix $C \in \mathbb{R}^{n \times n}$ is compartmental if*

(i) *the off-diagonal entries are non-negative: $c_{ij} \geq 0$, for $i \neq j$, (i.e., C is Metzler)*
(ii) *the column sums are non-positive: $\sum_{i=1}^{n} c_{ij} \leq 0$, for all $j \in \{1, \ldots, n\}$.*

In equivalent words, a compartmental matrix C is Metzler, has non-positive diagonal entries, and is *weakly column diagonally dominant* in the sense that $|c_{jj}| \geq \sum_{i=1, i \neq j}^{n} c_{ij}$, for all columns $j \in \{1, \ldots, n\}$.

Remark 10.10 (Symmetric flows). *The donor-controlled model entails a flow $f_{ij} q_i$ from i to j and a flow $f_{ji} q_j$ from j to i. If the flow rates are equal $f_{ij} = f_{ji}$, then the resultant flow as measured from i to j is $f_{ij}(q_i - q_j)$, i.e., proportional to the difference in stored quantities. The flow rate matrix F is often symmetric in physical networks.* •

Algebraic and graphical properties of linear dynamical flow systems

In this section we study the algebraic and spectral graph theory of present useful properties of compartmental matrices. We start with some useful graph-theoretical notions, illustrated in Figure 10.9. In the compartmental digraph, a set of compartments S is

(i) *outflow-connected* if there exists a directed walk from every compartment in S to the environment, that is, to a compartment j with a positive flow rate constant $f_{0j} > 0$,

(ii) *inflow-connected* if there exists a directed walk from the environment to every compartment in S, that is, from a compartment i with a positive inflow $u_i > 0$,

(iii) a *trap* if there is no directed walk from any of the compartments in S to the environment or to any compartment outside S, and

(iv) a *simple trap* is a trap that has no traps inside it.

It is immediate to realize the following equivalence: the system is outflow connected (i.e., all compartments are outflow-connected) if and only if the system contains no trap.
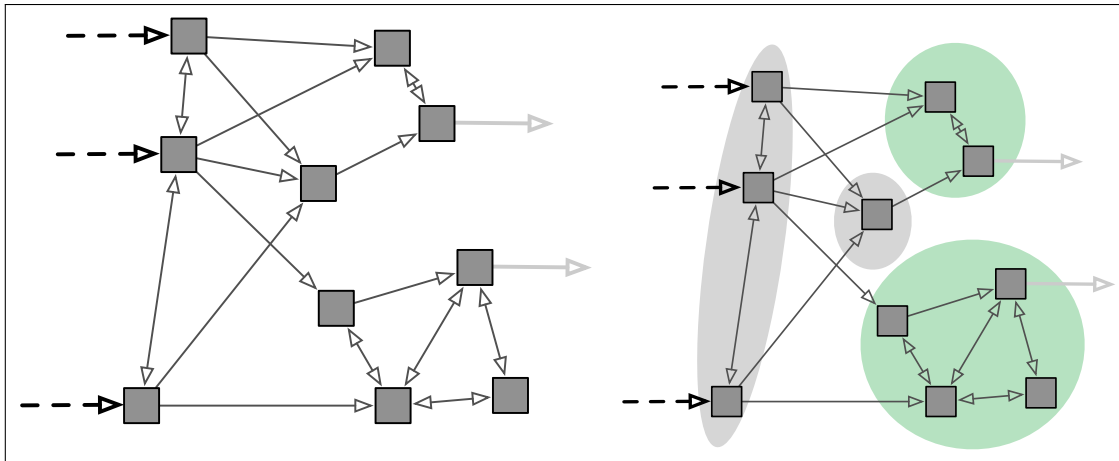
(a) An example dynamical flow system and its strongly connected components: this system is outflow-connected because its two sinks in the condensation digraph are outflow-connected.
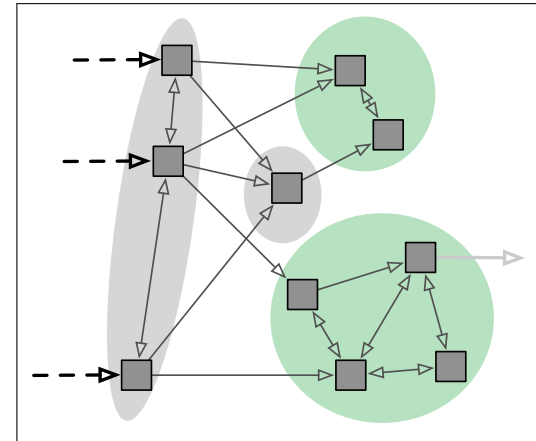
(b) This dynamical flow system is not outflow-connected because one of its sink strongly-connected components is a trap.

Figure 10.9: Outflow-connectivity and traps in dynamical flow system

Theorem 10.11 (Algebraic graph theory of dynamical flow systems). *Consider the linear dynamical flow system* (F, f_0, u) *with compartmental matrix* C *and condensation of the compartmental digraph* $C(G_F)$*. Then*

(i) *each eigenvalue of* C *either is equal to* 0 *or has negative real part;*

(ii) *the following statements are equivalent:*

   (a) C *is Hurwitz,*

   (b) *there are no traps, that is, the system is outflow-connected, and*

   (c) *each sink of* $C(G_F)$ *is outflow-connected;*

(iii) *if* C *is not Hurwitz, then* 0 *is semisimple with multiplicity equal to the number of simple traps (that is, the number of sinks of* $C(G_F)$ *that are not outflow-connected).*

Proof. The fact that each eigenvalue is either 0 or has strictly negative real part is similar to the result in Lemma 6.5 and can be proved by an application of the Geršgorin Disks Theorem 2.8. We invite the reader to fill out the details in Exercise E10.7.

The equivalence between statements (ii)b and (ii)c is immediate. To establish the equivalence between (ii)c and (ii)a, we first consider the case in which G_F is strongly connected and at least one compartment has a strictly positive outflow rate. Therefore, the compartmental matrix $C = -L_F^\mathsf{T} - \mathrm{diag}(f_0)$ is irreducible. As in Exercise E6.1 and in the proof of Theorems 7.4 and 10.3, pick $0 < \varepsilon < 1/\max_i |c_{ii}|$, and define the non-negative irreducible $\mathcal{A}_{C,\varepsilon} = I_n + \varepsilon C^\mathsf{T}$. The row-sums of $\mathcal{A}_{C,\varepsilon}$ are:

$$\mathcal{A}_{C,\varepsilon} \mathbb{1}_n = \mathbb{1}_n + \varepsilon(-L_F - \mathrm{diag}(f_0))\mathbb{1}_n = \mathbb{1}_n - \varepsilon f_0.$$

Therefore, $\mathcal{A}_{C,\varepsilon}$ is row-substochastic and, because $\mathcal{A}_{C,\varepsilon}$ is irreducible, Corollary 4.13 implies that $\rho(A) < 1$. Now, let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of $\mathcal{A}_{C,\varepsilon}$ and recall from Exercise E6.1 that the eigenvalues $\eta_1, \dots, \eta_n$ of C satisfy $\lambda_i = 1 + \varepsilon\eta_i$ so that $\max_i \Re(\lambda_i) = 1 + \varepsilon \max_i \Re(\eta_i)$. Finally, $\rho(\mathcal{A}_{C,\varepsilon}) < 1$ implies $\max_i \Re(\lambda_i) < 1$ so that

$$\max_i \Re(\eta_i) = \frac{1}{\varepsilon}\left(\max_i \Re(\lambda_i) - 1\right) < 0.$$

This concludes the proof that if G is strongly connected, then C has eigenvalues with strictly negative real part. The converse is easy to prove by contradiction: if $f_0 = \mathbb{0}_n$, then the matrix C has zero column-sums, but this is a contradiction with the assumption that C is invertible.

Next, to prove the equivalence between (ii)c and (ii)a for a graph G_F whose condensation digraph has an arbitrary number of sinks, we proceed as in the proof of Theorem 6.6 and, more precisely, Theorem 5.2: we reorder the compartments as described in Exercise E3.2 so that the Laplacian matrix L_F is block lower-triangular. We then consider the matrix $\mathcal{A}_{C,\varepsilon} = I_n + \varepsilon C^\mathsf{T}$ and proceed as above. We leave the remaining details to the reader. ∎

Remark 10.12 (Augmenting the compartmental digraph). *An alternative clever proof strategy for the equivalence between (ii)c and (ii)a is given as follows. Define the matrix*

$$C_{\text{augmented}} = \begin{bmatrix} C & \mathbb{0}_n \\ f_0^{\mathsf{T}} & 0 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} ,$$

and consider the augmented linear system $\dot{x} = C_{\text{augmented}}x$ with $x \in \mathbb{R}^{n+1}$. Note that $L_{\text{augmented}} = -C_{\text{augmented}}^{\mathsf{T}}$ is the Laplacian matrix of the augmented graph $G_{\text{augmented}}$, whose nodes $\{1, \dots, n, n+1\}$ include the n compartments and the environment as $(n+1)$st node, and whose edges are the edges of the compartmental graph G_F as well as the outflow edges to the environment node. Note that the environment node $n+1$ in the digraph $G_{\text{augmented}}$ is the only globally reachable node of $G_{\text{augmented}}$ if and only if the compartmental digraph G_F is outflow connected. Assume now that statement (ii)c is true. Then, Theorem 7.4 implies

$$\lim_{t\to\infty} e^{-L_{\text{augmented}}t} = \mathbb{1}_{n+1}\mathbb{e}_{n+1}^{\mathsf{T}},$$

which, taking a transpose operation, immediately implies $\lim_{t\to\infty} e^{C_{\text{augmented}}t} = \mathbb{e}_{n+1}\mathbb{1}_{n+1}^{\mathsf{T}}$. We now can easily compute

$$\lim_{t\to\infty} \begin{bmatrix} q(t) \\ x_{n+1}(t) \end{bmatrix} = \mathbb{e}_{n+1}\mathbb{1}_{n+1}^{\mathsf{T}} \begin{bmatrix} q(0) \\ x_{n+1}(0) \end{bmatrix}$$
$$\implies \lim_{t\to\infty} q(t) = \mathbb{0}_n \quad \lim_{t\to\infty} x_{n+1}(t) = \mathbb{1}_n^{\mathsf{T}}q(0) + x_{n+1}(0).$$

In other words, all mass in the system reaches asymptotically the environment and the mass in all compartments converge exponentially fast to zero. This occurs for all initial conditions if and only if the matrix C is Hurwitz. Hence we have established that statement (ii)c implies statement (ii)a. We leave the converse to the reader.

Dynamic properties of linear dynamical flow systems

Consider a linear dynamical flow system (F, f_0, u) with compartmental matrix C and compartmental digraph G_F. Assuming the system has at least one trap, we define the *reduced compartmental system* $(F_{\text{rd}}, f_{0,\text{rd}}, u_{\text{rd}})$ as follows: remove all simple traps from G_F and regard the edges into the removed compartments as outflow edges into the environment, e.g., see Figure 10.10.
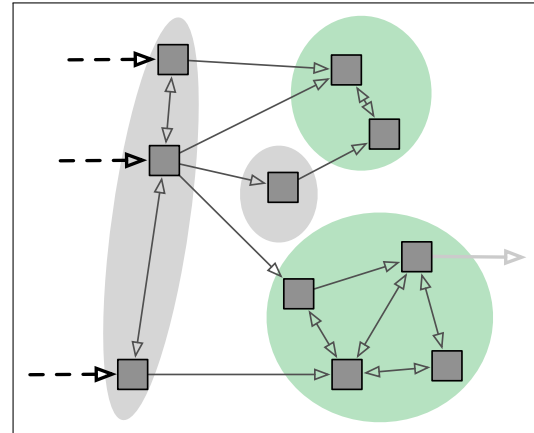
(a) A dynamical flow system that is not outflow-connected
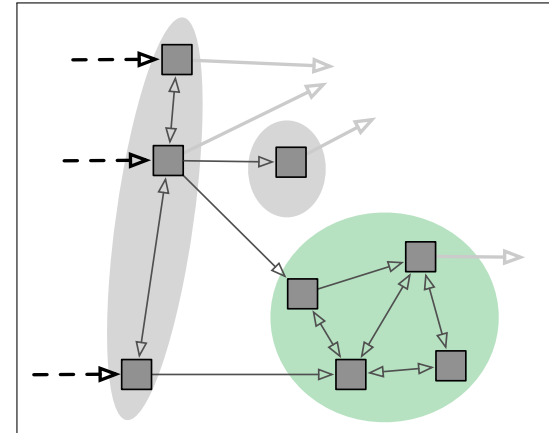
(b) The corresponding reduced dynamical flow system

Figure 10.10: An example reduced dynamical flow system

We now state our main result about the asymptotic behavior of linear dynamical flow systems.

Theorem 10.13 (Asymptotic behavior of dynamical flow systems). *The linear dynamical flow system (F, f_0, u) with compartmental matrix C and compartmental digraph G_F has the following possible asymptotic behaviors:*

(i) *if the system is outflow-connected, then the compartmental matrix C is invertible, every solution tends exponentially to the unique equilibrium $q^* = -C^{-1}u \geq \mathbb{0}_n$, and in the ith compartment $q_i^* > 0$ if and only if the ith compartment is inflow-connected to a positive inflow;*

(ii) *if the system contains one or more simple traps, then:*

  (a) *the reduced compartmental system $(F_{\mathrm{rd}}, f_{0,\mathrm{rd}}, u_{\mathrm{rd}})$ is outflow-connected and all its solutions converge exponentially fast to the unique non-negative equilibrium $-C_{\mathrm{rd}}^{-1}u_{\mathrm{rd}}$, for $C_{\mathrm{rd}} = F_{\mathrm{rd}}^{\mathsf{T}} - \mathrm{diag}(F_{\mathrm{rd}}\mathbb{1}_n + f_{0,\mathrm{rd}})$;*

  (b) *any simple trap H contains non-decreasing mass along time. If H is inflow-connected to a positive inflow, then the mass inside H grows linearly with time. Otherwise, the mass inside H converges asymptotically to a scalar multiple of the right eigenvector corresponding to the eigenvalue 0 of the compartmental submatrix for H.*

Proof. Statement (i) is an immediate consequence of Theorem 10.6. We leave the proof of statement (ii) to the reader. ∎

Closed dynamical flow systems

Finally, we elaborate on the last case considered in Theorem 10.13. Specifically, we consider a *closed* dynamical flow system, that is, a system without inflows and outflows. The governing equation of such a system is

$$\dot{q} = -L_F^{\mathsf{T}} q, \quad q(0) \geq 0. \tag{10.8}$$

For example, these are the equations of a continuous-time Markov chain.

Clearly, the total mass of commodity is conserved in a closed system since the equality $\mathbb{1}_n^{\mathsf{T}} L_F^{\mathsf{T}} = \mathbb{0}_n^{\mathsf{T}}$ implies that the function $t \mapsto \mathbb{1}_n^{\mathsf{T}} q(t)$ is constant.

If the compartmental digraph of a closed dynamical flow system contains a globally reachable node, then Theorem 7.4 implies $\lim_{t \to \infty} \exp(-L_F t) = \mathbb{1}_n w^{\mathsf{T}}$ and so

$$\lim_{t \to \infty} q(t) = \left(\mathbb{1}_n^{\mathsf{T}} q(0)\right) w,$$

where $w \geq 0$ is the left dominant eigenvector of L and $w_i > 0$ if and only if i is globally reachable. In other words, the commodity present in the system at initial time $\left(\mathbb{1}_n^{\mathsf{T}} q(0)\right)$ concentrates asymptotically in the globally reachable nodes. For the continuous-time Markov chain case, the probability vector converges to asymptotic positive values only at the so-called absorbing states.

10.4   Appendix: Metzler Hurwitz matrices

In this appendix we present various properties and additional characterization of Metzler Hurwitz matrices. These selected results have found application for example in the study of epidemic, economic, and network control problems. For convenience we start by reporting and extending the results in Metzler Hurwitz Theorem 10.3.

Theorem 10.14 (Metzler Hurwitz Theorem: Theorem 10.3 continued). *For a Metzler matrix $M \in \mathbb{R}^{n \times n}$, the following statements are equivalent:*

 (i) M is Hurwitz,
 (ii) M is invertible and $-M^{-1} \geq 0$,
 (iii) for all $b \geq \mathbb{0}_n$, there exists a unique $x^* \geq \mathbb{0}_n$ solving $Mx^* + b = \mathbb{0}_n$,
 (iv) there exists $\xi \in \mathbb{R}^n$ such that $\xi > \mathbb{0}_n$ and $M\xi < \mathbb{0}_n$,
 (v) there exists $\eta \in \mathbb{R}^n$ such that $\eta > \mathbb{0}_n$ and $\eta^{\mathsf{T}} M < \mathbb{0}_n^{\mathsf{T}}$,
 (vi) there exists a diagonal matrix $P \succ 0$ such that $M^{\mathsf{T}} P + PM \prec 0$, and
 (vii) for any $\zeta \geq \mathbb{0}_n$ and $\zeta \neq \mathbb{0}_n$, the vector $M\zeta$ has at least one negative entry.

Note: if the vectors ξ and η satisfy the conditions of statements (iv) and (v) respectively, then the matrix $P = \operatorname{diag}(\eta_1/\xi_1, \dots, \eta_n/\xi_n)$ satisfies the conditions of statement (vi).

Note: a matrix M with a diagonal matrix P as in statement (vi) is said to be *diagonally stable*.

Note: condition (iv) is equivalently rewritten as $a_{ii} < 0$ and $\xi_i |a_{ii}| > \sum_{j=1, j\neq i}^n \xi_j |a_{ij}|$. The latter condition is sometimes referred to *(strict row) quasi diagonal dominance*. Similar statements hold for condition (v).

Proof. The equivalence between statements (i), (ii), and (iii) is established in the Metzler Hurwitz Theorem 10.3. We omit the proof of statement (vii) in the interest of brevity.

Statements (iv) and (v) are equivalent because of the following argument and its converse: if statement (iv) holds with $\xi = \xi(M)$, then statement (v) holds with $\eta = \xi(M^{\mathsf{T}})$.

We first prove that (ii) implies (iv). Set $\xi = -M^{-1}\mathbb{1}_n$. Because $-M^{-1} \geq 0$ is invertible, it can have no row identically equal to zero. Hence $\xi = -M^{-1}\mathbb{1}_n > \mathbb{0}_n$. Moreover $M\xi = -\mathbb{1}_n < \mathbb{0}_n$.

Next, we prove that (iv) implies (i). Let λ be an eigenvalue of M with eigenvector v. Define $w \in \mathbb{R}^n$ by $w_i = v_i/\xi_i$, for $i \in \{1, \ldots, n\}$, where ξ is as in statement (iv). We have therefore $\lambda \xi_i w_i = \sum_{j=1}^n a_{ij} \xi_j w_j$. If ℓ is the index satisfying $|w_\ell| = \max_i |w_i| > 0$, then

$$\lambda \xi_\ell = a_{\ell\ell} \xi_\ell + \sum_{j=1, j\neq\ell}^n a_{\ell j} \xi_j \frac{w_j}{w_\ell},$$

which, in turn, implies

$$|\lambda \xi_\ell - a_{\ell\ell}\xi_\ell| \leq \sum_{j=1, j\neq\ell}^n a_{\ell j}\xi_j \left|\frac{w_j}{w_\ell}\right| \leq \sum_{j=1, j\neq\ell}^n a_{\ell j}\xi_j < -a_{\ell\ell}\xi_\ell,$$

where the last equality follows from the ℓ-th row of the inequality $M\xi < \mathbb{0}_n$. Therefore, $|\lambda - a_{\ell\ell}| < -a_{\ell\ell}$. This inequality implies that the eigenvalue λ must belong to an open disc in the complex plan with center $a_{\ell\ell} < 0$ and radius $|a_{\ell\ell}|$. Hence, λ, together with all other eigenvalues of M, must have negative real part.

We now prove that (iv) implies (vi). From statement (iv) applied to M and M^{T}, let $\xi > \mathbb{0}_n$ satisfy $M\xi < \mathbb{0}_n$ and $\eta > \mathbb{0}_n$ satisfy $M^{\mathsf{T}}\eta < \mathbb{0}_n$. Define $P = \mathrm{diag}(\eta_1/\xi_1, \ldots, \eta_n/\xi_n)$ and consider the symmetric matrix $M^{\mathsf{T}}P + PM$. This matrix is Metzler and satisfies $(M^{\mathsf{T}}P + PM)\xi = M^{\mathsf{T}}\eta + PM\xi < \mathbb{0}_n$. Hence, $M^{\mathsf{T}}P + PM$ is negative diagonally dominant and, because (iv) $\implies$ (i), Hurwitz. In summary, $M^{\mathsf{T}}P + PM$ is symmetric and Hurwitz, hence, it is negative definite.

Finally, the implication (vi) $\implies$ (i) is established in Theorem 15.9. ∎

Next, we present a result on matrix splitting that is useful for example to understand the concept of reproduction number in epidemic models, e.g., see (den Driessche and Watmough, 2002). Additional results on matrix splitting are given, e.g., in (Varga, 1962, Theorem 3.13), and (Dashkovskiy et al., 2011, Lemma 3.1).

Lemma 10.15 (Stability tests for Metzler matrices based on matrix splitting). *Let $M \in \mathbb{R}^{n \times n}$ be Metzler irreducible and have negative diagonal entries. Define a nonnegative matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$ with zero diagonal and a diagonal nonnegative matrix $\Lambda \in \mathbb{R}_{\geq 0}^{n \times n}$ in such a way that $M = -\Lambda + A$. Then*

(i) $\alpha(M) < 0$ if and only if $\rho(\Lambda^{-1}A) < 1$,
(ii) $\alpha(M) = 0$ if and only if $\rho(\Lambda^{-1}A) = 1$, and
(iii) $\alpha(M) > 0$ if and only if $\rho(\Lambda^{-1}A) > 1$.

Proof. First, we claim that is suffices to show that

(F1) $\alpha(M) < 0$ if $\rho(\Lambda^{-1}A) < 1$,

(F2) $\alpha(M) = 0$ if $\rho(\Lambda^{-1}A) = 1$, and

(F3) $\alpha(M) > 0$ if $\rho(\Lambda^{-1}A) > 1$.

Indeed, assume for example that $\alpha(M) < 0$ and $\rho(\Lambda^{-1}A) \geq 1$. But statement (F2) for $\rho(\Lambda^{-1}A) = 1$ or statement (F3) for $\rho(\Lambda^{-1}A) > 1$ immediately lead to a contradiction.

Second, the Perron–Frobenius Theorem applied to the irreducible non-negative matrix $\Lambda^{-1}A$ implies the existence of $v \in \mathbb{R}^n_{>0}$ such that $v^{\mathsf{T}}(\Lambda^{-1}A) = \lambda v^{\mathsf{T}}$, where $\lambda = \rho(\Lambda^{-1}A)$. We then compute

$$
\begin{aligned}
v^{\mathsf{T}}(\Lambda^{-1}A) = \lambda v^{\mathsf{T}} \iff & \quad v^{\mathsf{T}}(\Lambda^{-1}A) = (\lambda - 1)v^{\mathsf{T}} + v^{\mathsf{T}} \\
\iff & \quad v^{\mathsf{T}}(-I_n + \Lambda^{-1}A) = (\lambda - 1)v^{\mathsf{T}} \\
\iff & \quad v^{\mathsf{T}}\Lambda^{-1}(-\Lambda + A) = (\lambda - 1)v^{\mathsf{T}}.
\end{aligned}
$$

In summary, there exists a positive vector v such that

$$(\Lambda^{-1}v)^{\mathsf{T}}M = \big(\rho(\Lambda^{-1}A) - 1\big)v^{\mathsf{T}}. \tag{10.9}$$

Note that $\Lambda^{-1}v$ is also a positive vector.

Finally, we conclude the proof as follows:

(F1) if $\rho(\Lambda^{-1}A) < 1$, then equation (10.9) implies $(\Lambda^{-1}v)^{\mathsf{T}}M < \mathbb{0}_n^{\mathsf{T}}$ or, equivalently, the inequality $M^{\mathsf{T}}(\Lambda^{-1}v) < \mathbb{0}_n$. We then apply Exercise E10.5(iii) to the irreducible Metzler matrix M^{T} with $a_2 = 0$ and $x = (\Lambda^{-1}v)$ to obtain $\alpha(M^{\mathsf{T}}) = \alpha(M) < 0$;

(F2) if $\rho(\Lambda^{-1}A) = 1$, then equation (10.9) implies $(\Lambda^{-1}v)^{\mathsf{T}}M = \mathbb{0}_n^{\mathsf{T}}$. But then Exercise E2.4 on the "Uniqueness of the non-negative eigenvector in irreducible non-negative matrices" implies $\alpha(M) = 0$; and

(F3) if $\rho(\Lambda^{-1}A) > 1$, then equation (10.9) implies $(\Lambda^{-1}v)^{\mathsf{T}}M > \mathbb{0}_n^{\mathsf{T}}$ or, equivalently, the inequality $M^{\mathsf{T}}(\Lambda^{-1}v) > \mathbb{0}_n$. We then apply Exercise E10.5(iii) to the irreducible Metzler matrix M^{T} with $a_1 = 0$ and $x = (\Lambda^{-1}v)$ to obtain $\alpha(M^{\mathsf{T}}) = \alpha(M) > 0$. ∎

The following properties are also relevant in epidemic problems and are extensions of (Berman and Plemmons, 1994, Exercise (5.2) at page 159), (Horn and Johnson, 1994, Exercise 6b at page 127), (den Driessche and Watmough, 2002, Lemma 5 and Theorem 2) and (Smith and Bullo, 2021).

Lemma 10.16 (More properties of Hurwitz Metzler matrices). *Let M be Metzler and Hurwitz and H be Metzler. Then*

(i) *if $-HM^{-1}$ is Metzler, then H is Hurwitz if and only if $-HM^{-1}$ is Hurwitz, and*

(ii) *if $-M^{-1}H$ is Metzler, then H is Hurwitz if and only if $-M^{-1}H$ is Hurwitz.*

Next, let $E \in \mathbb{R}_{\geq 0}^{n \times n}$ be a non-negative perturbation matrix. The following statements are equivalent:

(iii) *$M + E$ is Hurwitz,*

(iv) *$\rho(-EM^{-1}) < 1$, and*

(v) *$\rho(-M^{-1}E) < 1$.*

Proof. Since M is Metzler Hurwitz, Theorem 10.3(ii) implies $-M^{-1} \geq 0$ and, moreover, no column and no row of $-M^{-1}$ can be zero (otherwise M^{-1} would be singular). Regarding statement (i), we reason as follows

$$H \text{ is Hurwitz } \iff \exists \eta \in \mathbb{R}^n_{>0} \text{ such that } \eta^\mathsf{T} H < 0 \tag{10.10}$$
$$\implies \exists \eta \in \mathbb{R}^n_{>0} \text{ such that } \eta^\mathsf{T} H(-M^{-1}) < 0 \tag{10.11}$$
$$\iff H(-M^{-1}) \text{ is Hurwitz.} \tag{10.12}$$

Additionally

$$H(-M^{-1}) \text{ is Hurwitz } \iff \exists \xi \in \mathbb{R}^n_{>0} \text{ such that } H(-M^{-1})\xi < 0 \tag{10.13}$$
$$\implies \exists \xi' = (-M^{-1})\xi \in \mathbb{R}^n_{>0} \text{ such that } H\xi' < 0 \tag{10.14}$$
$$\iff H \text{ is Hurwitz.} \tag{10.15}$$

This proves the statement and (i). We leave statement (ii) to the reader.

Next, we show that statement (iii) is equivalent to (iv) and (v), define $H' := M + E$ and

$$A_1 = -H'M^{-1} = -(M+E)M^{-1} = -I_n - EM^{-1}, \tag{10.16}$$
$$A_2 = -M^{-1}H' = -M^{-1}(M+E) = -I_n - M^{-1}E. \tag{10.17}$$

Since M is Metzler Hurwitz, Theorem 10.3(ii) implies $-M^{-1} \geq 0$ and so $A_1 = -I_n + \text{(nonnegative matrix)}$ and $A_2 = -I_n + \text{(nonnegative matrix)}$ are Metzler. Then by the statements (i) and (ii), we know $H' = M + E$ is Hurwitz if and only if A_1 is Hurwitz if and only if A_2 is Hurwitz.

We proceed as follows. First, if $\rho(-EM^{-1}) < 1$, then A_1 is clearly Hurwitz. One the other hand, if $\rho(-EM^{-1}) \geq 1$, then A_1 cannot be Hurwitz. Indeed, since $-EM^{-1} \geq 0$, the Perron-Frobenius Theorem guarantees that its dominant eigenvalue is real and non-negative, so $A_1 = -I_n - EM^{-1}$ has an eigenvalue with non-negative real part. This completes the proof of the equivalence (iii) $\iff$ (iv). We leave the equivalence with (v) to the reader. ∎

   Next, we consider stability tests based upon Schur complements. The following lemma improves upon (Shorten and Narendra, 2009, Theorem 4.1) and is a special case of (Ebihara et al., 2017, Lemma 2). We present the more general treatment in Exercise E10.15.

Lemma 10.17 (Stability tests for Metzler matrices based on Schur complements). *Let $M \in \mathbb{R}^{n \times n}$ be Metzler and decompose it in the block form*

$$M = \begin{bmatrix} A & b \\ c^{\mathsf{T}} & d \end{bmatrix}, \quad \text{where } A \in \mathbb{R}^{(n-1) \times (n-1)} \text{ is Metzler}, b, c \in \mathbb{R}_{\geq 0}^{(n-1) \times 1}, \text{ and } d \in \mathbb{R}. \tag{10.18}$$

When $d < 0$, the Schur complement of the block A is the matrix $A - dbc^{\mathsf{T}} \in \mathbb{R}^{(n-1) \times (n-1)}$. It is easy to see that $A - dbc^{\mathsf{T}}$ is Metzler and satisfies:

$$\begin{bmatrix} A & b \\ c^{\mathsf{T}} & d \end{bmatrix} = \begin{bmatrix} I_{n-1} & d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} A - dbc^{\mathsf{T}} & \mathbb{0}_{n-1} \\ \mathbb{0}_{n-1}^{\mathsf{T}} & d \end{bmatrix} \begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix}. \tag{10.19}$$

The following statements are equivalent:

(i) M is Hurwitz,

(ii) $d < 0$ and $A - dbc^{\mathsf{T}}$ is Hurwitz.

Proof. We start by rewriting (10.19) in two equivalent format:

$$\begin{bmatrix} I_{n-1} & -d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} A & b \\ c^{\mathsf{T}} & d \end{bmatrix} = \begin{bmatrix} A - dbc^{\mathsf{T}} & \mathbb{0}_{n-1} \\ \mathbb{0}_{n-1}^{\mathsf{T}} & d \end{bmatrix} \begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix}, \tag{10.20}$$

$$\begin{bmatrix} A & b \\ c^{\mathsf{T}} & d \end{bmatrix} \begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ -d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix} = \begin{bmatrix} I_{n-1} & d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} A - dbc^{\mathsf{T}} & \mathbb{0}_{n-1} \\ \mathbb{0}_{n-1}^{\mathsf{T}} & d \end{bmatrix}, \tag{10.21}$$

and noting $\begin{bmatrix} I_{n-1} & -d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \geq 0$ and $\begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ -d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix} \geq 0.$

　Regarding (ii) $\implies$ (i), since the Metzler matrix $A - dbc^{\mathsf{T}}$ is Hurwitz, there exists $\xi \in \mathbb{R}_{>0}^{n-1}$ such that $(A - dbc^{\mathsf{T}})\xi \leq -\mathbb{1}_{n-1}$. Pick $0 < \varepsilon$ such that $\varepsilon b < \mathbb{1}_{n-1}$ and define

$$\bar{\xi} = \begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ -d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} \xi \\ \varepsilon \end{bmatrix} \geq \begin{bmatrix} \xi \\ \varepsilon \end{bmatrix} > \mathbb{0}_n. \tag{10.22}$$

Using (10.21) we now compute:

$$\begin{aligned} M\bar{\xi} &= \begin{bmatrix} A & b \\ c^{\mathsf{T}} & d \end{bmatrix} \begin{bmatrix} I_{n-1} & \mathbb{0}_{n-1} \\ -d^{-1}c^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} \xi \\ \varepsilon \end{bmatrix} \\ &= \begin{bmatrix} I_{n-1} & d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} A - dbc^{\mathsf{T}} & \mathbb{0}_{n-1} \\ \mathbb{0}_{n-1}^{\mathsf{T}} & d \end{bmatrix} \begin{bmatrix} \xi \\ \varepsilon \end{bmatrix} \leq \begin{bmatrix} I_{n-1} & d^{-1}b \\ \mathbb{0}_{n-1}^{\mathsf{T}} & 1 \end{bmatrix} \begin{bmatrix} -\mathbb{1}_{n-1} \\ \varepsilon d \end{bmatrix} \\ &= \begin{bmatrix} -\mathbb{1}_{n-1} + \varepsilon b \\ \varepsilon d \end{bmatrix} < \mathbb{0}_n. \end{aligned}$$

The Metzler Hurwitz Theorem 10.14(iv) now implies that M is Hurwitz. This completes the proof of (ii) $\implies$ (i). The proof of the converse implication is analogous and left to the reader. ∎

We conclude with the so-called *Hicksian stability condition* from the classic economics literature (Habibagahi and Quirk, 1973); to understand the relationship with the nomenclature in the literature we recall that, in the language of Remark 10.4, the matrix $-M$ is an M-matrix. Also recall that the *leading principal submatrices* of a matrix $A \in \mathbb{R}^{n \times n}$ are the square submatrices $A_i \in \mathbb{R}^{i \times i}$, $i \in \{1, \dots, n\}$, containing the first i rows and i columns of A. Also recall that the *leading principal minors* of A are the determinants of the leading principal submatrices.

Lemma 10.18 (Stability tests for Metzler matrices based on leading principal minors). *Given a Metzler matrix M, the following statements are equivalent:*

 (i) *M is Hurwitz, and*

(ii) *each leading principal minor of $-M$ is positive.*

We omit the proof of Lemma 10.18 and refer the reader to (Berman and Plemmons, 1994, Chapter 6) and (Farina and Rinaldi, 2000, Theorem 13).

10.5   Appendix: Examples of nonlinear flow systems

10.5.1 Symmetric physical flow systems

Many physical dynamical flow systems are described by symmetric flows that depend upon effort variables and energy stored at nodes. For an insightful treatment of physical and port-Hamiltonian network systems we refer to (van der Schaft, 2015; van der Schaft and Wei, 2012). We here present a brief introduction without outflows and inflows, for simplicity.

Following (van der Schaft, 2015), we let G be an undirected graph with n nodes and m edges and with oriented incidence matrix $B \in \mathbb{R}^{n \times m}$ and proceed as follows:

(i) for an oriented edge (i, j), let u_{ij} denote the total flow from i to j (that is, $u_{ij} = F_{i \to j} - F_{j \to i}$) so that the flow vector is $u \in \mathbb{R}^m$. Given storage q_i at each node i, mass conservation implies $\dot{q} = Bu \in \mathbb{R}^n$; (if instead the nodes have no storage, then mass conservation implies $Bu = \mathbb{0}_n$, which is consistent with Kirchhoff's current law as stated in Section 9.3.)

(ii) typically, the flow through an edge u_{ij} is proportional to an "effort on the edge" e_{ij}, that is, $u_{ij} = -c_{ij}e_{ij}$, for a "conductance constant" $c_{ij} > 0$. In vector form, $u = -Ce \in \mathbb{R}^m$;

(iii) typically, the edge effort e_{ij} is the difference between node effort variables, that is, $e = B^\mathsf{T} e_{\mathrm{nodes}} \in \mathbb{R}^m$, for nodal effort variables $e_{\mathrm{nodes}} \in \mathbb{R}^n$;

(iv) finally, node efforts are determined by the storage variables according to:

$$e_{\mathrm{nodes}} = \frac{\partial H}{\partial q}(q) \in \mathbb{R}^n, \tag{10.23}$$

where $H(q)$ is the total stored energy. Typically, $H(q) = \sum_{i=1}^n H_i(q_i)$, where $H_i(q_i)$ denotes the energy stored at node i.

In summary, the symmetric physical dynamical flow system obeys

$$\dot{q} = Bu = -BCe = -BCB^\mathsf{T} e_{\mathrm{nodes}} = -BCB^\mathsf{T}\frac{\partial H}{\partial q}(q) = -L\frac{\partial H}{\partial q}(q), \tag{10.24}$$

where L is the conductance-weighted Laplacian matrix of the compartmental graph.

For example, consider a hydraulic flow network among n fluid reservoirs. The liquid stored at the reservoirs is given by a vector $q \in \mathbb{R}^n_{\geq 0}$. Assume there exists an energy function H_i (possibly the same function at all locations) such that $\frac{\partial H_i}{\partial q_i}(q_i)$ is the pressure at reservoir i. Assume that the liquid flow along the pipe from head reservoir i to tail reservoir j is proportional to the difference between the pressure at i and the pressure at j. Then equation (10.24) describes the mass balance equation among the reservoirs.

10.5.2   A static nonlinear flow problem

In this appendix, we consider a static compartmental flow system, where a commodity (e.g., power or water) is transported through a network (e.g., a power grid or a piping system). We model this scenario with an undirected and connected graph with n nodes and m edges. With each node we associate an external supply/demand variable (positive for a source and negative for a sink) y_i and assume that the overall network is balanced: $\sum_{i=1}^{n} y_i = 0$. We also associate a potential variable x_i with every node (e.g., voltage or pressure) and, for each undirected edge $\{i, j\}$, we assume that the flow of commodity from node i to node j depends on the potential difference $(x_i - x_j)$. Specifically, we assume that the total flow from i to j satisfies

$$F_{i \to j}(q) - F_{j \to i}(q) = a_{ij} h(x_i - x_j),$$

where a_{ij} is akin to a conductance weight for the edge $\{i, j\}$ and where the function $h \colon \mathbb{R} \to \mathbb{R}$ is odd, differentiable, and satisfies $h(0) = 0$ and $h'(0) = 1$. For example, for piping systems and power grids the function h is given by the empirical Hazen-Williams law and the trigonometric power flow equation, respectively. In both cases the function h is monotone in the region of interest. By balancing the flow at each node (as we do for network with nodal storage in equation (10.2)), we obtain at node i

$$y_i = \sum_{j=1}^{n} a_{ij} h(x_i - x_j), \quad i \in \{1, \ldots, n\}.$$

In vector notation, letting f denote the vector of all flows, the combined physical flow model and flow balance equations read

$$\begin{aligned} y &= Bf, \\ f &= \mathcal{A}h(B^\mathsf{T} x), \end{aligned} \tag{10.25}$$

or, equivalently, $y = B\mathcal{A}h(B^\mathsf{T} x)$.

In what follows, we are given the graph topology B, the edge conductances $\mathcal{A}$, the nonlinearity h, and the supply/demand vector y. With this information, we are interested in computing the solution equilibrium flows f and potential variables x.

Linear setting.   We consider now the associated linear problem, where $h(\delta) = \delta$ and $y = B\mathcal{A}B^\mathsf{T} x = Lx$, where L is the network Laplacian matrix. In other words, the linearized problem is a Laplacian system, as studied in Section 6.3.2. Letting $L^\dagger$ denote the

pseudoinverse of L, the equilibrium potential variables and equilibrium flows in the linear problem are, respectively,

$$x^\star = L^\dagger y, \qquad f^\star = \mathcal{A}B^\mathsf{T}L^\dagger y.$$

(Recall from Lemma 6.12(ii) that the potential variable $x^\star = L^\dagger y$ is only one solution to $Lx = y$; for a connected graph and a balanced supply/demand vector, all solutions are of the form $x^\star + \beta\mathbb{1}_n$, for $\beta \in \mathbb{R}$.)

Nonlinear acyclic setting.   Next, we consider an acyclic network, i.e., the graph is a tree, and show that the nonlinear equilibrium solution has similarities with the solution of the linear problem. We also assume, for simplicity, that h is monotonic increasing and unbounded. We introduce the normalized flow variable $v = h(B^\mathsf{T}x)$ and rewrite the physical flow balance equation (10.25) as

$$y = B\mathcal{A}v, \tag{10.26a}$$
$$v = h\left(B^\mathsf{T}x\right). \tag{10.26b}$$

In the acyclic case, we know $\mathrm{kernel}(B) = \{\mathbb{0}_{n-1}\}$ and necessarily $v \in \mathrm{image}(B^\mathsf{T})$ since $\mathrm{image}(B^\mathsf{T}) = \mathbb{R}^{n-1}$. In turn, there must exist $w \in \mathbb{R}^n$ such that or $v = B^\mathsf{T}w$. Thus, equation (10.26a) reads $y = B\mathcal{A}v = B\mathcal{A}B^\mathsf{T}w = Lw$ and its solution is $w^* = L^\dagger y$. Equation (10.26b) then reads $h(B^\mathsf{T}x) = v = B^\mathsf{T}w = B^\mathsf{T}L^\dagger y$, and its unique solution (due to the monotonicity of h) is

$$B^\mathsf{T}x^\star = h^{-1}(B^\mathsf{T}L^\dagger y).$$

Left-multiplying by $B\mathcal{A}$, we obtain $B\mathcal{A}B^\mathsf{T}x^\star = B\mathcal{A}h^{-1}(B^\mathsf{T}L^\dagger y)$. In summary, the equilibrium potential variables and equilibrium flows in the nonlinear acyclic problem are, respectively,

$$x^\star = L^\dagger B\mathcal{A}h^{-1}(B^\mathsf{T}L^\dagger y), \qquad f^\star = \mathcal{A}B^\mathsf{T}L^\dagger y.$$

10.6   Appendix: Tables of asymptotic behaviors for averaging and positive systems

We conclude this chapter with Tables 10.2 and 10.3 summarizing numerous results presented in this and previous chapters.

Dynamics	Assumptions & Asymptotic Behavior	References
averaging system $x(k+1) = Ax(k)$ A row-stochastic	the associated digraph has a globally reachable node $\implies$ $\lim_{k \to \infty} x(k) = (w^{\mathsf{T}} x(0))\mathbb{1}_n$ where $w \geq 0$ is the left eigenvector of A with eigenvalue 1 satisfying $\mathbb{1}_n^{\mathsf{T}} w = 1$	Convergence properties: Theorem 5.1 Examples: opinion dynamics & averaging in Chapter 1
affine system $x(k+1) = Ax(k) + b$	A convergent (that is, its spectral radius is less than 1) $\implies \quad \lim_{k \to \infty} x(k) = (I_n - A)^{-1} b$	Convergence properties: Exercise E2.9 Examples: Friedkin-Johnsen system in Exercise E5.26
positive affine system $x(k+1) = Ax(k) + b$ $A \geq 0, b \geq \mathbb{0}_n$	$x(0) \geq \mathbb{0}_n \implies x(k) \geq \mathbb{0}_n$ for all k, and A convergent (that is, $|\lambda| < 1$ for all $\lambda \in \mathrm{spec}(A)$) $\implies \quad \lim_{k \to \infty} x(k) = (I_n - A)^{-1} b \geq \mathbb{0}_n$	Positivity properties: Exercise E10.4 Examples: Leslie population model in Exercise E4.19

Table 10.2: Discrete-time systems

Dynamics	Assumptions & Asymptotic Behavior	References
averaging system $\dot{x}(t) = -Lx(t)$ L Laplacian matrix	the associated digraph has a globally reachable node $\Longrightarrow$ $\lim_{t\to\infty} x(t) = (w^{\mathsf{T}}x(0))\mathbb{1}_n$ where $w \geq 0$ is the left eigenvector of L with eigenvalue 0 satisfying $\mathbb{1}_n^{\mathsf{T}}w = 1$	Convergence properties: Theorem 7.4 Examples: Flocking system in Section 1.3
affine system $\dot{x}(t) = Ax(t) + b$	A Hurwitz (that is, its spectral abscissa is negative) $\Longrightarrow \quad \lim_{t\to\infty} x(t) = -A^{-1}b$	Convergence properties: Exercise E7.2
positive affine system $\dot{x}(t) = Mx(t) + b$ M Metzler, $b \geq \mathbb{0}_n$	$x(0) \geq \mathbb{0}_n \implies x(t) \geq \mathbb{0}_n$ for all t, and M Hurwitz (that is, $\Re(\lambda) < 0$ for all $\lambda \in \mathrm{spec}(M)$) $\Longrightarrow \quad \lim_{t\to\infty} x(t) = -M^{-1}b \geq \mathbb{0}_n$	Positivity properties: Theorem 10.6 Example: dynamical flow systems in Section 10.1
dynamical flow system $\dot{q}(t) = Cq(t) + u$ C compartmental, $u \geq \mathbb{0}_n$	$q(0) \geq \mathbb{0}_n \implies q(t) \geq \mathbb{0}_n$ for all t, and system is outflow-connected $\Longrightarrow \quad \lim_{t\to\infty} q(t) = -C^{-1}u \geq \mathbb{0}_n$	Algebraic graph theory: Theorem 10.11 Asymptotic behavior: Theorem 10.13

Table 10.3: Continuous-time systems

10.7   Historical notes and further reading

Metzler matrices are widely studied. For a comprehensive treatment of M-matrices (i.e., minus Metzler Hurwitz matrices), we refer to (Berman and Plemmons, 1994, Chapter 6) and (Horn and Johnson, 1994, Section 2.5). We refer the interested readers to (Farina and Rinaldi, 2000; Kaczorek, 2001) for a detailed study of linear positive systems. As example recent extensions, (Narendra and Shorten, 2010) studies an iterative method based on the Schur complement to check the Hurwitzness of Metzler matrices and (Ebihara et al., 2017) provides a comprehensive analysis of interconnected positive systems. Additionally, Blanchini et al. (2012) study switched Metzler systems and convex combinations of Metzler Hurwitz matrices; see also in (Meng et al., 2017). Finally, Duan et al. (2021) provide graph-theoretic stability conditions based on small-gain concepts.

The treatment of compartmental systems is inspired by the excellent text by Walter and Contreras (1999) and the tutorial treatment by Jacquez and Simon (1993); see also the texts (Luenberger, 1979; Farina and Rinaldi, 2000; Haddad et al., 2010).

For nonlinear extensions of the material in this chapter, including recent studies of traffic networks, we refer to (Como et al., 2013; Coogan and Arcak, 2015). The survey by Sontag (2007) reviews theoretical results and applications of interconnected monotone systems.

10.8   Exercises

E10.1   **The matrix exponential of a Metzler matrix.** In this exercise we extend and adapt Theorem 7.2 about the matrix exponential of a Laplacian matrix to the setting of Metzler matrices. Let M be an $n \times n$ Metzler matrix with minimum diagonal entry $m_{\min} = \min\{m_{11}, \ldots, m_{nn}\}$. As usual, associate to M a digraph G without self-loops in the natural way, that is, (i, j) is an edge if and only if $m_{ij} > 0$. Prove that

   (i) $\exp(M) \geq e^{m_{\min}} I_n \geq 0$,   for any digraph G,
   (ii) $\exp(M)\mathbb{e}_j > 0$,   for a digraph G whose j-th node is globally reachable,
   (iii) $\exp(M) > 0$,   for a strongly connected digraph G (i.e., for an irreducible M).

   Moreover, prove that, for any square matrix A,

   (iv) $\exp(At) \geq 0$ for all $t \geq 0$ if and only if A is Metzler.

▶ E10.2   **Proof of the Perron–Frobenius Theorem for Metzler matrices.** Prove Theorem 10.2.

E10.3   **Non-negative convergent matrices and inverse positivity.** This exercise is the discrete-time equivalent of Metzler Hurwitz Theorem 10.3. For a non-negative matrix A, show that the following statements are equivalent:

(i) A is convergent ($\rho(A) < 1$),

(ii) $(I_n - A)$ is invertible and $(I_n - A)^{-1} \geq 0$, and

(iii) for all $b \geq \mathbb{0}_n$, there exists a unique $x^* \geq \mathbb{0}_n$ solving $x^* = Ax^* + b$.

Moreover, show that

(iv) if A is non-negative, convergent and irreducible, then $(I_n - A)^{-1} > 0$.

E10.4   **Discrete-time positive affine systems and non-negative matrices.** This exercise is the discrete-time equivalent of Theorem 10.6. For the affine system $x(k+1) = Ax(t) + b$, the following statements are equivalent:

(i) the system is positive, that is, $x(k) \geq \mathbb{0}_n$ for all $k \in \mathbb{N}$ and all $x(0) \geq \mathbb{0}_n$,

(ii) A is non-negative and $b \geq \mathbb{0}_n$.

Moreover, if the matrix A is non-negative and convergent, then

(iii) the system has a unique non-negative equilibrium point $x^* \in \mathbb{R}^n_{\geq 0}$, to which all trajectories converge asymptotically.

▶ E10.5 **Bounds and monotonicity of the spectral abscissa of Metzler matrices.** Let $M \in \mathbb{R}^n$ be Metzler. For $a_1, a_2 > 0$ and $x \in \mathbb{R}^n_{\geq 0}$, $x \neq \mathbb{0}_n$, show

(i) if $a_1 x \leq Mx$, then $a_1 \leq \alpha(M)$,
(ii) if $Mx \leq a_2 x$ and $x \in \mathbb{R}^n_{>0}$, then $\alpha(M) \leq a_2$,
(iii) if $a_1 x \leq Mx \leq a_2 x$, $a_1 x \neq Mx \neq a_2 x$, and M is irreducible, then $a_1 < \alpha(M) < a_2$ and $x \in \mathbb{R}^n_{>0}$.

Moreover, let M' be a Metzler matrix of the same dimension as M. Show

(iv) if $M \leq M'$, then $\alpha(M) \leq \alpha(M')$,
(v) if additionally $M \neq M'$ and M' is irreducible, then $\alpha(M) < \alpha(M')$.
    Hint: Recall Lemmas 4.9 and 4.10 for non-negative matrices.

E10.6   **Monotonicity properties of positive systems.** Consider the two continuous-time positive affine system

$$\dot{x} = Mx + b, \quad \text{and} \quad \dot{x}' = M'x' + b',$$

where M and M' are Metzler and b and b' are non-negative. Let $x(t)$ and $x'(t)$ denote the solutions of the respective systems from initial conditions $x_0 \in \mathbb{R}^n_{\geq 0}$ and $x'_0 \in \mathbb{R}^n_{\geq 0}$ at time 0. Assume both matrices are Hurwitz and let x_* and x'_* denote the equilibrium points of the two systems. Show that

$$x_0 \leq x'_0, \quad M \leq M', \quad \text{and} \quad b \leq b'$$

implies

$$x(t) \leq x'(t), \quad \text{and} \quad x_* \leq x'_*.$$

E10.7 **Establishing the spectral properties of compartmental matrices.** Given a compartmental matrix C, show that if $\lambda \in \mathrm{spec}(C)$, then either $\lambda = 0$ or $\Re(\lambda) < 0$.

E10.8   **Simple traps and strong connectivity.** Show that a dynamical flow system that has no outflows and that is a simple trap, is strongly connected.

E10.9 **Decompositions of a Metzler matrix and sufficient conditions to be Hurwitz.** Given a Metzler matrix $M \in \mathbb{R}^{n \times n}$, $n \geq 2$, show that

(i) there exists a unique Laplacian matrix L_r and a vector $v_r = M \mathbb{1}_n$ such that $M = -L_r + \operatorname{diag}(v_r)$,

(ii) there exists a unique Laplacian matrix L_c and a vector $v_c = M^\mathsf{T} \mathbb{1}_n$ such that $M = -L_c^\mathsf{T} + \operatorname{diag}(v_c)$,

(iii) if M is irreducible and the vector $v_r = M \mathbb{1}_n$ satisfies $v_r \leq \mathbb{0}_n$ and $v_r \neq \mathbb{0}_n$ (or $v_c = M^\mathsf{T} \mathbb{1}_n$ satisfies $v_c \leq \mathbb{0}_n$ and $v_c \neq \mathbb{0}_n$), then M is Hurwitz, and

(iv) if M is an irreducible Metzler matrix with $M \mathbb{1}_n = \mathbb{0}_n$, then, for any $i \in \{1, \dots, n\}$ and $\varepsilon > 0$, all eigenvalues of $M - \varepsilon e_i e_i^\mathsf{T}$ are negative.

E10.10   **On Metzler matrices and dynamical flow systems with growth and decay.** Let M be an $n \times n$ symmetric Metzler matrix. As in Exercise E10.9, decompose M
into $M = -L + \mathrm{diag}(v)$, where $v = M\mathbb{1}_n \in \mathbb{R}^n$ and L is a symmetric Laplacian matrix. Show that:

(i) if M is Hurwitz, then $\mathbb{1}_n^\mathsf{T} v < 0$.

Next, assume $n = 2$ and assume v has both non-negative and non-positive entries. (If v is non-negative, lack of stability can be established from statement (i); if v is
non-positive, stability can be established via Theorem 10.11.) Show that

(ii) there exist non-negative numbers f, d and g such that, modulo a permutation, M can be written in the form:

$$M = -f \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} g & 0 \\ 0 & -d \end{bmatrix} = \begin{bmatrix} (g-f) & f \\ f & (-d-f) \end{bmatrix},$$

(iii) M is Hurwitz if and only if

$$d > g \quad \text{and} \quad f > \frac{gd}{d-g}.$$

*Note: The inequality $d > g$ (for $n = 2$) is equivalent to the inequality $\mathbb{1}_n^\mathsf{T} v < 0$ in statement (i). In the interpretation of dynamical flow systems with growth and decay
rates, f is a flow rate, d is a decay rate and g is a growth rate. Statement (iii) is then interpreted as follows: M is Hurwitz if and only if the decay rate is larger than the
growth rate and the flow rate is sufficiently large.*

E10.11   **Grounded Laplacian matrices.** This exercises is a followup and generalization of Exercise E6.15 about grounded spring networks. Let G be a weighted undirected graph with Laplacian $L \in \mathbb{R}^{n \times n}$. Select a set S of $s \geq 1$ nodes and call them *grounded nodes*. Given S, the *grounded Laplacian matrix* $L_{\text{grounded}} \in \mathbb{R}^{(n-s) \times (n-s)}$ is the principal submatrix of L obtained by removing the s rows and columns corresponding to the grounded nodes. In other words, if the grounded nodes are nodes $\{n - s + 1, \ldots, n\}$ and L is partitioned in block matrix form

$$L = \begin{bmatrix} L_{11} & L_{12} \\ L_{12}^{\mathsf{T}} & L_{22} \end{bmatrix}, \qquad \text{with } L_{11} \in \mathbb{R}^{(n-s) \times (n-s)} \text{ and } L_{22} \in \mathbb{R}^{s \times s},$$

then $L_{\text{grounded}} = L_{11}$. Show the following statements:

(i) If G is connected, then

   (a) L_{grounded} is positive definite,
   (b) L_{grounded}^{-1} is non-negative, and
   (c) the eigenvector associated with the smallest eigenvalue of L_{grounded} can be selected non-negative.

(ii) If additionally the graph obtained by removing from G the nodes in S and all the corresponding edges is connected, then

   (d) L_{grounded}^{-1} is positive, and
   (e) the eigenvector associated with the smallest eigenvalue of L_{grounded} is unique and positive (up to rescaling).

   Hint: Show that $-L_{\text{grounded}}$ is a compartmental matrix.

Note: For more information on grounded Laplacian matrices we refer to (Dörfler and Bullo, 2013; Pirani and Sundaram, 2016; Xia and Cao, 2017).

E10.12   **Mean residence time for a particle in a dynamical flow system.** Consider an outflow-connected dynamical flow system with irreducible matrix C and spectral abscissa $\alpha(C) < 0$. Let v is the dominant eigenvector of C, that is, $Cv = \alpha(C)v$, $\mathbb{1}_n^\mathsf{T} v = 1$, and $v > 0$.

Assume a tagged particle is randomly located inside the compartmental system at time 0 with probability mass function v. The *mean residence time (mrt)* of the tagged particle is the expected time that the particle remains inside the dynamical flow system. Using the definition of expectation, the mean residence time is

$$\mathrm{mrt} = \int_0^\infty t\, \mathbb{P}[\text{particle leaves at time } t]\, dt.$$

Let us also take for granted that: $\mathbb{P}[\text{particle leaves at time } t] = -\left(\frac{d}{dt} \mathbb{P}[\text{particle inside at time } t] \right)$. Show that

$$\mathrm{mrt} = -\frac{1}{\alpha(C)}.$$

E10.13 **Resistive circuits as dynamical flow systems (Dörfler et al., 2018).** Consider a resistive circuit with shunt capacitors at each node as in figure below (see also in Section 7.1.2). Assume that the circuit is connected. Attach to at least one node $j \in \{1, \ldots, n\}$ a current source generating an injected current $c_{\text{injected at } j} > 0$, and connect to at least one node $i \in \{1, \ldots, n\}$ a positive resistor to ground.

(i) Model the resulting system as a dynamical flow system, i.e., identify the conserved quantity and write the compartmental matrix, the inflow vector and the outflow rate vector, and

(ii) show that there exists a unique steady state that is positive and globally-asymptotically stable.
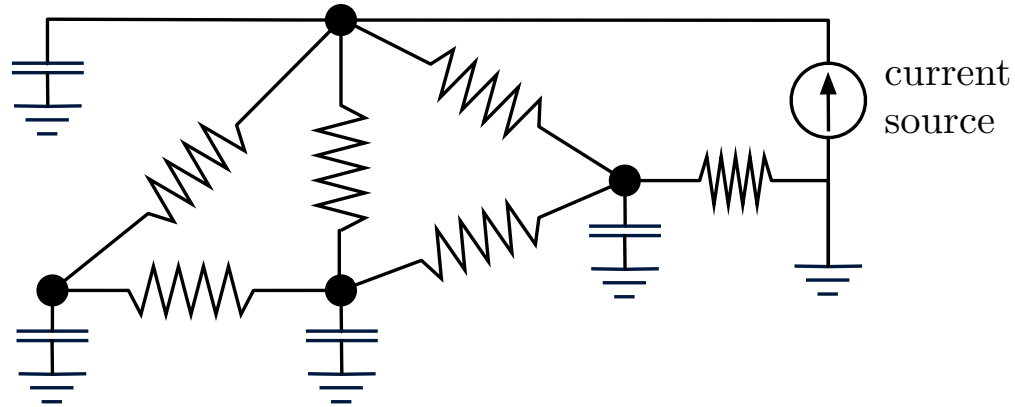
E10.14   **Discretization of the Laplace partial differential equation (Luenberger, 1979, Chapter 6).** The electric potential V within a two-dimensional domain is governed by the partial differential equation known as the *Laplace's equation*:

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0, \tag{E10.1}$$

combined with the value of V along the boundary of the enclosure; see the left image in Figure E10.1. (A similar setup with a time-varying spatial quantity and free boundary conditions was described in Section 7.1.3.)
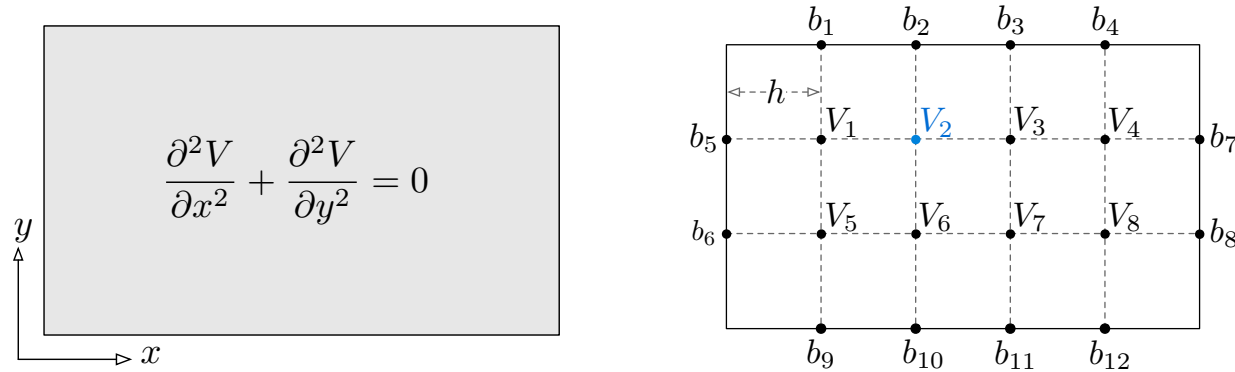
Figure E10.1: Laplace's equation over a rectangular enclosure and a mesh graph.

For arbitrary enclosures and boundary conditions, it is impossible to solve the Laplace's equation in closed form. An approximate solution is computed by (i) introducing an appropriate mesh graph (i.e., a two-dimensional grid graph without corner nodes) whose nodes are physical locations with spacing h, e.g., see the right image in Figure E10.1, and (ii) approximating the second-order derivatives by second-order finite differences. Specifically, at node 2 of the mesh, we have along the x direction

$$\frac{\partial^2 V}{\partial x^2}(V_2) \approx \frac{1}{h^2}(V_3 - V_2) - \frac{1}{h^2}(V_2 - V_1) = \frac{1}{h^2}(V_3 + V_1 - 2V_2),$$

so that equation (E10.1) is approximated as follows:

$$0 = \frac{\partial^2 V}{\partial x^2}(V_2) + \frac{\partial^2 V}{\partial y^2}(V_2) \approx \frac{1}{h^2}(V_1 + V_3 + V_6 + b_2 - 4V_2) \quad \Longrightarrow \quad 4V_2 = V_1 + V_3 + V_6 + b_2.$$

This approximation translates into the matrix equation:

$$4V = A_{\text{mesh}}V + C_{\text{mesh-boundary}}b, \tag{E10.2}$$

where $V \in \mathbb{R}^n$ is the vector of unknown potentials, $b \in \mathbb{R}^m$ is the vector of boundary conditions, $A_{\text{mesh}} \in \{0,1\}^{n \times n}$ is the binary adjacency matrix of the *(interior) mesh graph* (that is, $(A_{\text{mesh}})_{ij} = 1$ if and only if the interior nodes i and j are connected by an edge), and $C_{\text{mesh-boundary}} \in \{0,1\}^{n \times m}$ is the connection matrix between interior and boundary nodes (that is, $(C_{\text{mesh-boundary}})_{i\alpha} = 1$ if and only if mesh interior node i is connected with boundary node α). Show that

(i) A_{mesh} is irreducible but not primitive,

(ii) $\rho(A_{\text{mesh}}) < 4$,

   Hint: Recall Theorem 4.11.

(iii) there exists a unique solution V^* to equation (E10.2),

(iv) the unique solution V^* satisfies $V^* > \mathbb{0}_n$ if $b \geq \mathbb{0}_m$ and $b \neq \mathbb{0}_m$, and

(v) each solution to the following iteration converges to V^*:

$$4V(k+1) = A_{\text{mesh}}V(k) + C_{\text{mesh-boundary}}b.$$

Note that, at each step of this iteration, the value of V at each node is updated to the average of the values at its neighboring nodes.

E10.15   **Stability tests for Metzler matrices based on Schur complements (Ebihara et al., 2017, Lemma 2).** Let M be Metzler and block partitioned as $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$. Clearly M_{11} and M_{22} are Metzler and M_{12} and M_{21} are nonnegative. Show that the following statements are equivalent:

(i) M is Hurwitz,
(ii) the Metzler matrices M_{11} and $M_{22} - M_{21} M_{11}^{-1} M_{12}$ are Hurwitz,
(iii) the Metzler matrices M_{22} and $M_{11} - M_{12} M_{22}^{-1} M_{21}$ are Hurwitz.

E10.16 **Stabilizing Metzler matrices via balancing (Ma et al., 2022).** Consider an irreducible Metzler matrix $A \in \mathbb{R}^{n \times n}$, a target spectral abscissa $\eta \in \mathbb{R}$, and positive weights $w \in \mathbb{R}^n_{>0}$. Let $d^* \in \mathring{\Delta}_n$ denote the balancing vector for the matrix $\operatorname{diag}(w)A$, as given in Theorem 4.14, and define

$$\ell^* = \operatorname{diag}(d^*)^{-1} A d^* - \eta \mathbb{1}_n \ \in \mathbb{R}^n. \tag{E10.3}$$

Show that

(i) the Metzler matrix $A - \operatorname{diag}(\ell^*)$ has spectral abscissa η and right Perron eigenvector d^*,

(ii) ℓ^* is the solution to the optimization problem

$$\begin{aligned} \min_{\ell \in \mathbb{R}^n} \quad & w^\mathsf{T} \ell, \\ \text{s.t.} \quad & \alpha(A - \operatorname{diag}(\ell)) \le \eta \end{aligned} \tag{E10.4}$$

(iii) if $A - \eta I_n \ge 0$, then $\ell^* \ge 0$.

Part II

Topics in Averaging Systems

Chapter 11

Convergence Rates, Scalability and Optimization

In this chapter we discuss the convergence rate of averaging algorithms. We focus on discrete-time systems and their convergence factors. The study of continuous-time systems is analogous. We also perform a scalability analysis for an example system and discuss some interesting optimization problems.

Before proceeding, we recall a few basic facts from Chapter 2, Exercise E4.17 and Section 5.3. Given a square matrix A,

(i) the spectral radius of A is $\rho(A) = \max\{|\lambda| \mid \lambda \in \mathrm{spec}(A)\}$;

(ii) the p-induced norm of A, for $p \in \mathbb{N}\cup\{\infty\}$, is

$$\|A\|_p = \max\left\{\|Ax\|_p \mid x \in \mathbb{R}^n \text{ and } \|x\|_p = 1\right\} = \max_{x\neq \mathbb{0}_n} \frac{\|Ax\|_p}{\|x\|_p},$$

and, specifically, the induced 2-norm of A is $\|A\|_2 = \max\{\sqrt{\lambda} \mid \lambda \in \mathrm{spec}(A^\mathsf{T} A)\}$;

(iii) for any p, $\rho(A) \leq \|A\|_p$;

(iv) if $A = A^\mathsf{T}$, then $\|A\|_2 = \rho(A)$; and

(v) the *essential spectral radius* of a row-stochastic matrix A is

$$\rho_{\mathrm{ess}}(A) = \begin{cases} 0, & \text{if } \mathrm{spec}(A) = \{1, \ldots, 1\}, \\ \max\{|\lambda| \mid \lambda \in \mathrm{spec}(A) \setminus \{1\}\}, & \text{otherwise.} \end{cases}$$

11.1   Some preliminary calculations and observations

The convergence factor for symmetric row-stochastic matrices   To build some intuition about the general case, we review the calculations perfomed in Section 5.3 on the quadratic disagreement error. We consider a weighted undirected graph G with symmetric adjacency matrix A that is row-stochastic and primitive. We consider the corresponding discrete-time averaging algorithm

$$x(k+1) = Ax(k).$$

Since A is symmetric, A has real eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ and corresponding orthonormal eigenvectors $v_1, \ldots, v_n$. Because A is row-stochastic, $\lambda_1 = 1$ and $v_1 = \mathbb{1}_n/\sqrt{n}$. As discussed in Remark 2.3 and Section 5.3, the solution satisfies

$$x(k) = \text{average}(x(0))\mathbb{1}_n + \lambda_2^k(v_2^\mathsf{T}x(0))v_2 + \ldots + \lambda_n^k(v_n^\mathsf{T}x(0))v_n.$$

Moreover, A being primitive implies that $\max\{|\lambda_2|, \ldots, |\lambda_n|\} < 1$. Specifically, for a symmetric and primitive A, we have $\rho_{\text{ess}}(A) = \max\{|\lambda_2|, |\lambda_n|\} < 1$. Therefore, as predicted by Theorem 5.1

$$\lim_{k\to\infty} x(k) = \mathbb{1}_n\mathbb{1}_n^\mathsf{T}x(0)/n = \text{average}(x(0))\mathbb{1}_n.$$

To upper bound the error, since the vectors $v_1, \ldots, v_n$ are orthonormal, we compute

$$\left\|x(k) - \text{average}(x(0))\mathbb{1}_n\right\|_2 = \Big\|\sum_{j=2}^n \lambda_j^k(v_j^\mathsf{T}x(0))v_j\Big\|_2 = \sqrt{\sum_{j=2}^n |\lambda_j|^{2k}\big\|(v_j^\mathsf{T}x(0))v_j\big\|_2^2}$$

$$\leq \rho_{\text{ess}}(A)^k\sqrt{\sum_{j=2}^n \big\|(v_j^\mathsf{T}x(0))v_j\big\|_2^2} = \rho_{\text{ess}}(A)^k\big\|x(0) - \text{average}(x(0))\mathbb{1}_n\big\|_2, \quad (11.1)$$

where the second and last equalities are Pythagoras Theorem.

In summary, we have learned that, for symmetric matrices, the essential spectral radius $\rho_{\text{ess}}(A) < 1$ is the *convergence factor* to average consensus, i.e., the factor determining the exponential convergence of the error to zero. (The wording "convergence factor" is for discrete-time systems, whereas the wording "convergence rate" is for continuous-time systems.)

A note on convergence factors for asymmetric matrices   The behavior of asymmetric row-stochastic matrices is more complex than of symmetric ones. For large even n, consider the asymmetric positive matrix

$$A_{\text{large-gain}} = \frac{1}{2n}\mathbb{1}_n\mathbb{1}_n^{\mathsf{T}} + \frac{1}{2}\left(\mathbb{1}_{1:n/2}\mathbb{e}_1^{\mathsf{T}} + \mathbb{1}_{n/2:n}\mathbb{e}_n^{\mathsf{T}}\right),$$

where $\mathbb{1}_{1:n/2}$ (resp. $\mathbb{1}_{n/2:n}$) is the vector whose first (resp. second) $n/2$ entries are equal to 1 and whose second (resp. first) $n/2$ entries are equal to 0. The digraph associated to $\mathbb{1}_{1:n/2}\mathbb{e}_1^{\mathsf{T}} + \mathbb{1}_{n/2:n}\mathbb{e}_n^{\mathsf{T}}$ is depicted in Figure 11.1.

$$\mathbb{1}_{1:3}\mathbb{e}_1^{\mathsf{T}} + \mathbb{1}_{3:6}\mathbb{e}_6^{\mathsf{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$
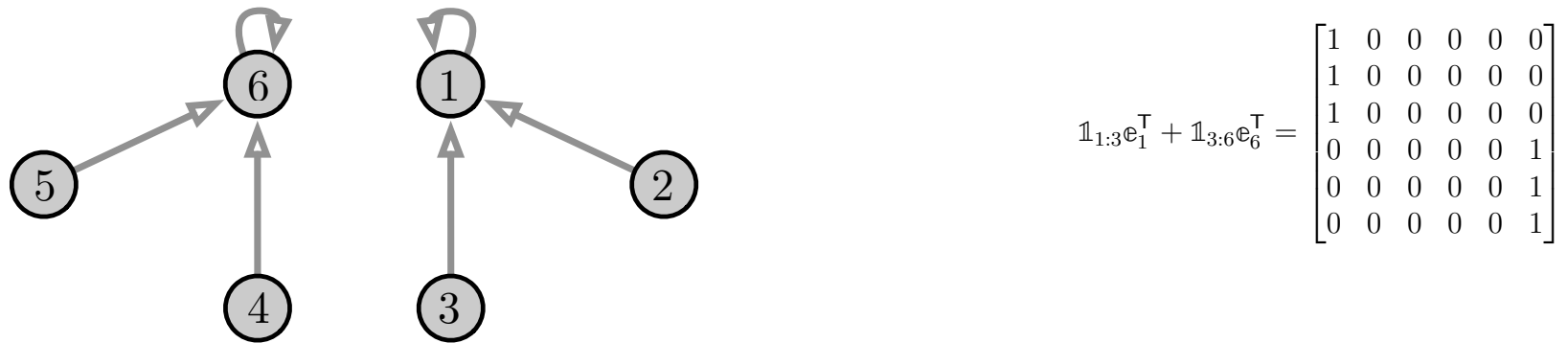
Figure 11.1: The unweighted digraph associated to the matrix $\mathbb{1}_{1:n/2}\mathbb{e}_1^{\mathsf{T}} + \mathbb{1}_{n/2:n}\mathbb{e}_n^{\mathsf{T}}$, for $n = 6$. This digraph is the union of two disjoint stars. The weighted digraph associated to $A_{\text{large-gain}}$ is the superposition of these two stars with a complete digraph.

The matrix $A_{\text{large-gain}}$ is row-stochastic because, given $\mathbb{1}_n^{\mathsf{T}}\mathbb{1}_n = n$ and $\mathbb{e}_j^{\mathsf{T}}\mathbb{1}_n = 1$ for all j, we compute

$$A_{\text{large-gain}}\mathbb{1}_n = \frac{1}{2}\mathbb{1}_n + \frac{1}{2}(\mathbb{1}_{1:n/2} + \mathbb{1}_{n/2:n})1 = \mathbb{1}_n.$$

Therefore, Theorem 5.1 implies that every solution to $x(k+1) = A_{\text{large-gain}}x(k)$ converges to consensus and Exercise E1.1 implies that $k \mapsto V_{\text{max-min}}(x(k))$ is non-increasing. Nevertheless, the quadratic disagreement (that it, the 2-norm of the disagreement vectors) can easily increase. For example, take $x(0) = \mathbb{e}_1 - \mathbb{e}_n$ and compute

$$x(1) = A_{\text{large-gain}}x(0) = \frac{1}{2}\mathbb{1}_{1:n/2} - \frac{1}{2}\mathbb{1}_{n/2:n}.$$

Because $\operatorname{average}(x(0)) = \operatorname{average}(x(1)) = 0$, we compute

$$\|x(0) - \operatorname{average}(x(0))\mathbb{1}_n\|_2 = \sqrt{2},$$

$$\|x(1) - \operatorname{average}(x(1))\mathbb{1}_n\|_2 = \frac{1}{2}\|\mathbb{1}_{1:n/2} - \mathbb{1}_{n/2:n}\|_2 = \frac{1}{2}\sqrt{n}.$$

In other words, the quadratic disagreement (2-norm of $x(k) - \operatorname{average}(x(k))\mathbb{1}_n$) along the averaging system defined by $A_{\text{large-gain}}$ grows to be at least of order $\sqrt{n}$ (starting from $\mathcal{O}(1)$);[1] compare this behavior with equation (11.1) for the symmetric case. The problem is that the eigenvalues (alone) of an asymmetric matrix do not fully describe the state amplification that may take place during a transient period of time.

---

[1] Here and in what follows, $\mathcal{O}(x)$ is a scalar function upper bounded by a constant times x.

11.2   Convergence factors for row-stochastic matrices

Consider a discrete-time averaging algorithm (distributed linear averaging)

$$x(k+1) = Ax(k),$$

where A is doubly-stochastic and not necessarily symmetric. If A is primitive (i.e., the associated digraph is aperiodic and strongly connected), we know

$$\lim_{k \to \infty} x(k) = \operatorname{average}(x(0))\mathbb{1}_n = \left(\mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n\right)x(0).$$

We now define two possible notions of convergence factors. The *per-step convergence factor* is

$$r_{\text{step}}(A) = \sup_{x(k) \neq x_{\text{final}}} \frac{\|x(k+1) - x_{\text{final}}\|_2}{\|x(k) - x_{\text{final}}\|_2},$$

where $x_{\text{final}} = \operatorname{average}(x(0))\mathbb{1}_n = \operatorname{average}(x(k))\mathbb{1}_n$ and where the supremum is taken over any possible sequence. Moreover, the *asymptotic convergence factor* is

$$r_{\text{asym}}(A) = \sup_{x(0) \neq x_{\text{final}}} \lim_{k \to \infty} \left(\frac{\|x(k) - x_{\text{final}}\|_2}{\|x(0) - x_{\text{final}}\|_2}\right)^{1/k}.$$

Given these definitions and the preliminary calculations in the previous Section 11.1, we can now state our main results.

Theorem 11.1 (Convergence factor and solution bounds). *Let A be doubly-stochastic and primitive.*

(i) The convergence factors of A satisfy

$$r_{\text{step}}(A) = \|A - \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n\|_2,$$
$$r_{\text{asym}}(A) = \rho_{\text{ess}}(A) = \rho(A - \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n) < 1. \qquad (11.2)$$

Moreover, $r_{\text{step}}(A) \geq r_{\text{asym}}(A)$, and $r_{\text{step}}(A) = r_{\text{asym}}(A)$ if A is symmetric.

(ii) For any initial condition $x(0)$ with corresponding $x_{\text{final}} = \text{average}(x(0))\mathbb{1}_n$,

$$\left\|x(k) - x_{\text{final}}\right\|_2 \leq r_{\text{step}}(A)^k\left\|x(0) - x_{\text{final}}\right\|_2, \qquad (11.3)$$
$$\left\|x(k) - x_{\text{final}}\right\|_2 \leq c_\varepsilon(r_{\text{asym}}(A) + \varepsilon)^k\left\|x(0) - x_{\text{final}}\right\|_2, \qquad (11.4)$$

where $\varepsilon > 0$ is an arbitrarily small constant and c_ε is a sufficiently large constant independent of $x(0)$.

Note: A sufficient condition for $r_{\text{step}}(A) < 1$ is given in Exercise E11.1.

Before proving Theorem 11.1, we review the notion of disagreement vector and some relevant properties from Section 5.3. For $x_{\text{final}} = \text{average}(x(0))\mathbb{1}_n$, the *disagreement vector* is the error signal

$$\delta(k) = x(k) - x_{\text{final}}. \tag{11.5}$$

The following lemma collects various useful properties; we do not include its proof since it only slightly generalizes Lemma 5.3 in Section 5.3.

Lemma 11.2 (Convergence of disagreement vector). *Given a doubly-stochastic matrix A, the disagreement vector $\delta(k)$ satisfies*
(i) $\delta(k) \perp \mathbb{1}_n$ for all k,
(ii) $\delta(k+1) = \left(A - \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n\right)\delta(k)$,
(iii) the following properties are equivalent:

  (a) $\lim_{k \to \infty} A^k = \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n$, (that is, the averaging algorithm achieves average consensus)
  (b) A is primitive, (that is, the digraph is aperiodic and strongly connected)
  (c) $\rho(A - \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n) < 1$. (that is, the error dynamics is convergent)

11.2.1 Proof of Theorem 11.1

We are now ready to prove the main theorem in this section.

Proof of Theorem 11.1. Regarding the equalities (11.2), the formula for r_{step} is an consequence of the definition of induced 2-norm:

$$r_{\text{step}}(A) = \sup_{x(k) \neq x_{\text{final}}} \frac{\|x(k+1) - x_{\text{final}}\|_2}{\|x(k) - x_{\text{final}}\|_2} = \sup_{\delta(k) \perp \mathbb{1}_n} \frac{\|\delta(k+1)\|_2}{\|\delta(k)\|_2}$$

$$= \sup_{\delta(k) \perp \mathbb{1}_n} \frac{\|(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n)\delta(k)\|_2}{\|\delta(k)\|_2} = \sup_{y \neq \mathbb{0}_n} \frac{\|(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n)y\|_2}{\|y\|_2},$$

where the last equality follows from $(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n)\mathbb{1}_n = \mathbb{0}_n$.

The equality $r_{\text{asym}}(A) = \rho(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n)$ is a consequence of the error dynamics in Lemma 11.2, statement (ii), and of Gelfand's formula $\rho(A) = \lim_{k \to \infty} \|A^k\|^{1/k}$; see Exercise E4.17.

Next, note that $\rho(A) = 1$ is a simple eigenvalue and A is semi-convergent. Hence, by Exercise E2.2 on the Jordan normal form of A, there exists a nonsingular T such that

$$A = T \begin{bmatrix} 1 & \mathbb{0}_{n-1}^\mathsf{T} \\ \mathbb{0}_{n-1} & B \end{bmatrix} T^{-1},$$

where $B \in \mathbb{R}^{(n-1) \times (n-1)}$ is convergent, that is, $\rho(B) < 1$. Moreover we know $\rho_{\text{ess}}(A) = \rho(B)$.

Usual properties of similarity transformations imply

$$A^k = T \begin{bmatrix} 1 & \mathbb{0}_{n-1}^\mathsf{T} \\ \mathbb{0}_{n-1} & B^k \end{bmatrix} T^{-1}, \qquad \implies \qquad \lim_{k \to \infty} A^k = T \begin{bmatrix} 1 & \mathbb{0}_{n-1}^\mathsf{T} \\ \mathbb{0}_{n-1} & \mathbb{0}_{(n-1) \times (n-1)} \end{bmatrix} T^{-1}.$$

Because A is doubly-stochastic and primitive, we know $\lim_{k \to \infty} A^k = \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n$ so that A can be decomposed as

$$A = \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n + T \begin{bmatrix} 0 & \mathbb{0}_{n-1}^\mathsf{T} \\ \mathbb{0}_{n-1} & B \end{bmatrix} T^{-1},$$

and conclude with $\rho_{\text{ess}}(A) = \rho(B) = \rho(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n)$. This concludes the proof of the equalities (11.2).

The bound (11.3) is an immediate consequence of the definition of induced norm.

Finally, we leave to the reader the proof of the bound (11.4) in Exercise E5.10. Note that the arbitrarily-small positive parameter ε is required because the eigenvalue corresponding to the essential spectral radius may have an algebraic multiplicity strictly larger than its geometric multiplicity. ∎

11.3   Cumulative quadratic disagreement for symmetric matrices

The previous convergence metrics (per-step convergence factor and asymptotic convergence factor) are *worst-case* convergence metrics (both are defined with a supremum operation) that are achieved only for particular initial conditions, e.g., the performance predicted by the asymptotic metric $r_{\text{asym}}(A)$ is achieved when $x(0) - x_{\text{final}}$ is aligned with the eigenvector associated to $\rho_{\text{ess}}(A) = \rho(A - \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}}/n)$.

   In what follows we study an appropriate average transient performance. We consider an averaging algorithm

$$x(k + 1) = Ax(k),$$

defined by a row-stochastic matrix A and subject to random initial conditions x_0 satisfying

$$\mathbb{E}[x_0] = \mathbb{0}_n, \quad \text{and} \quad \mathbb{E}[x_0 x_0^{\mathsf{T}}] = I_n.$$

Recall the disagreement vector $\delta(k)$ defined in (11.5) and the associated disagreement dynamics

$$\delta(k + 1) = \left(A - \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}}/n\right)\delta(k),$$

and observe that the initial conditions of the disagreement vector $\delta(0)$ satisfy

$$\mathbb{E}[\delta(0)] = \mathbb{0}_n \quad \text{and} \quad \mathbb{E}[\delta(0)\delta(0)^{\mathsf{T}}] = I_n - \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}}/n.$$

To define an average transient and asymptotic performance of this averaging algorithm, we define the *cumulative quadratic disagreement* of the matrix A by

$$\mathcal{J}_{\text{cum}}(A) = \lim_{K \to \infty} \frac{1}{n} \sum_{k=0}^{K} \mathbb{E}\big[\|\delta(k)\|_2^2\big]. \tag{11.6}$$

Theorem 11.3 (Cumulative quadratic disagreement for symmetric matrices). *The cumulative quadratic disagreement* (11.6) *of a row-stochastic, primitive, and symmetric matrix A satisfies*

$$\mathcal{J}_{\mathrm{cum}}(A) = \frac{1}{n} \sum_{\lambda \in \mathrm{spec}(A) \setminus \{1\}} \frac{1}{1 - \lambda^2} \, .$$

Proof. Pick a terminal time $K \in \mathbb{N}$ and define $\mathcal{J}_K(A) = \frac{1}{n} \sum_{k=0}^{K} \mathbb{E}\big[\|\delta(k)\|_2^2\big]$. From the definition (11.6) and the disagreement dynamics, we compute

$$\mathcal{J}_K(A) = \frac{1}{n} \sum_{k=0}^{K} \mathrm{trace}\big(\mathbb{E}\big[\delta(k)\delta(k)^\mathsf{T}\big]\big)$$

$$= \frac{1}{n} \sum_{k=0}^{K} \mathrm{trace}\Big(\big(A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big)^k \mathbb{E}\big[\delta(0)\delta(0)^\mathsf{T}\big] \big(\big(A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big)^k\big)^\mathsf{T} \Big)$$

$$= \frac{1}{n} \sum_{k=0}^{K} \mathrm{trace}\Big(\big(A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big)^k \big(\big(A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big)^k\big)^\mathsf{T} \Big) .$$

Because A is symmetric, also the matrix $A - \mathbb{1}_n \mathbb{1}_n^\top / n$ is symmetric and can be diagonalized as $A - \mathbb{1}_n \mathbb{1}_n^\top / n = Q \Lambda Q^\top$, where Q is orthonormal and Λ is a diagonal matrix whose diagonal entries are the elements of $\operatorname{spec}(A - \mathbb{1}_n \mathbb{1}_n^\top / n) = \{0\} \cup \operatorname{spec}(A) \setminus \{1\}$. It follows that

$$
\begin{aligned}
\mathcal{J}_K(A) &= \frac{1}{n} \sum_{k=0}^{K} \operatorname{trace}\left(Q \Lambda^k Q^\top \left(Q \Lambda^k Q^\top \right)^\top \right) \\
&= \frac{1}{n} \sum_{k=0}^{K} \operatorname{trace}\left(\Lambda^k \cdot \Lambda^k \right) && (\text{because } \operatorname{trace}(AB) = \operatorname{trace}(BA)) \\
&= \frac{1}{n} \sum_{k=0}^{K} \sum_{\lambda \in \operatorname{spec}(A) \setminus \{1\}} \lambda^{2k} \\
&= \frac{1}{n} \sum_{\lambda \in \operatorname{spec}(A) \setminus \{1\}} \frac{1 - \lambda^{2(K-1)}}{1 - \lambda^2}. && (\text{because of the geometric series})
\end{aligned}
$$

The formula for $\mathcal{J}_{\text{cum}}$ follows from taking the limit as $K \to \infty$ and recalling that A primitive implies $\rho_{\text{ess}}(A) < 1$. ∎

Note: All eigenvalues of A appear in the computation of the cumulative quadratic disagreement (11.6), not only the dominant eigenvalue as in the asymptotic convergence factor.

11.4 Circulant network examples and scalability analysis

In general it is difficult to compute explicitly the second largest eigenvalue magnitude for an arbitrary matrix. There are some graphs with *constant* essential spectral radius, independent of the network size n. For example, a complete graph with identical weights and doubly stochastic adjacency matrix $A = \mathbb{1}_n \mathbb{1}_n^\top / n$ has $\rho_{\text{ess}}(A) = 0$. In this case, the associated averaging algorithm converges in a single step.

Next, we present an interesting family of examples where all eigenvalues are known. Recall the cyclic balancing problem from Section 1.6, where each bug feels an attraction towards the closest counterclockwise and clockwise neighbors, Exercise E4.3 on circulant matrices, and the results in Table 4.1. Given the angular distances between bugs $d_i = \theta_{i+1} - \theta_i$, for $i \in \{1, \dots, n\}$ (with the usual convention that $d_{n+1} = d_1$ and $d_0 = d_n$), the closed-loop system is $d(k+1) = A_{n,\kappa} d(k)$, where $\kappa \in [0, 1/2[$, and

$$
A_{n,\kappa} = \begin{bmatrix}
1 - 2\kappa & \kappa & 0 & \dots & 0 & \kappa \\
\kappa & 1 - 2\kappa & \kappa & \ddots & \ddots & 0 \\
0 & \kappa & 1 - 2\kappa & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \ddots & \ddots & \kappa & 1 - 2\kappa & \kappa \\
\kappa & 0 & \dots & 0 & \kappa & 1 - 2\kappa
\end{bmatrix}.
$$
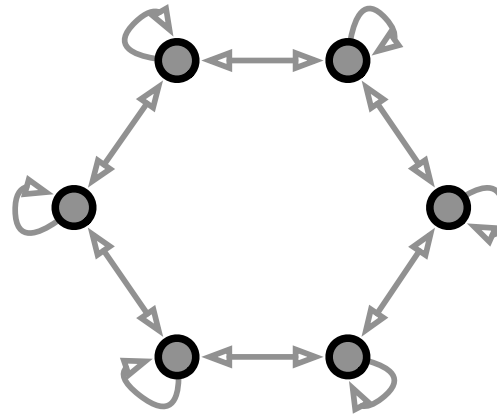
Figure 11.2: Digraph associated to the circulant matrix $A_{n,\kappa}$, for $n = 6$.

This matrix is circulant, that is, each row-vector is equal to the preceding row-vector rotated one element to the right. The associated digraph is illustrated in the Figure 11.2. From Exercise E4.3, the eigenvalues of $A_{n,\kappa}$ can be computed to be (not ordered in magnitude)

$$\lambda_i = 2\kappa \cos \frac{2\pi(i-1)}{n} + (1 - 2\kappa), \quad \text{for } i \in \{1, \ldots, n\}. \tag{11.7}$$

An illustration is given in Figure 11.3. For n even (similar results hold for n odd), plotting the eigenvalues on the segment $[-1, 1]$ shows that

$$\rho_{\text{ess}}(A_{n,\kappa}) = \max\{|\lambda_2|, |\lambda_{n/2+1}|\},$$

where

$$\lambda_2 = 2\kappa \cos \frac{2\pi}{n} + (1 - 2\kappa), \text{ and } \lambda_{n/2+1} = 1 - 4\kappa.$$

If we fix $\kappa \in \;]0, 1/2[$ and consider sufficiently large values of n, then $|\lambda_2| > |\lambda_{n/2+1}|$. In the limit of large graphs $n \to \infty$, the Taylor expansion $\cos(x) = 1 - x^2/2 + \mathcal{O}(x^4)$ leads to

$$\rho_{\text{ess}}(A_{n,\kappa}) = 1 - 4\pi^2 \kappa \frac{1}{n^2} + \mathcal{O}\left(\frac{1}{n^4}\right). \tag{11.8}$$
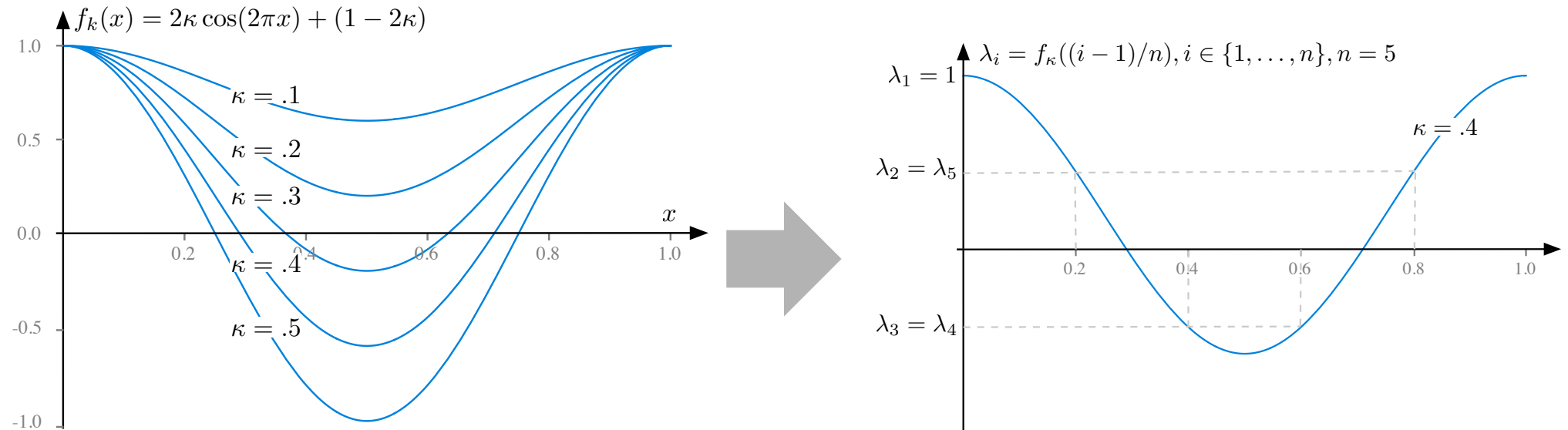
Figure 11.3: The eigenvalues of $A_{n,\kappa}$ as given in equation (11.7). The left figure includes the case of $\kappa = .5$, even if that value is strictly outside the allowed range $\kappa \in [0, .5[$.

Note that $\rho_{\text{ess}}(A_{n,\kappa}) < 1$ for any n, but the separation from $\rho_{\text{ess}}(A_{n,\kappa})$ to 1, called the *spectral gap*, shrinks with $1/n^2$.

In summary, this discussion leads to the broad statement that certain sparse large-scale graphs have slow convergence factors.

11.5   Appendix: Accelerated averaging algorithm

The averaging algorithm $x(k+1) = Ax(k)$ may converge slowly as seen in Section 11.4 due to a large $\rho_{\text{ess}}(A)$. In this section we propose a simple modification of averaging that is known to be faster. The *accelerated averaging algorithm* is defined by

$$x(k+1) = \beta A\, x(k) + (1-\beta)\, x(k-1), \qquad \text{for } k \in \mathbb{Z}_{\geq 0}, \tag{11.9}$$

where the initial conditions are $x(0) = x(-1) := x_0$, the matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, primitive, and row-stochastic, and $\beta \in \mathbb{R}$ is a parameter to be chosen.

   This iteration has some basic properties. We define the iteration matrix

$$T_\beta = \begin{bmatrix} \beta A & (1-\beta)I_n \\ I_n & \mathbb{0}_{n \times n} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}.$$

One can show that $T_\beta \mathbb{1}_{2n} = \mathbb{1}_{2n}$ for all β, and that T_β is semi-convergent if and only if $\rho_{\text{ess}}(T_\beta) < 1$. Moreover, similar to the result in (11.4) one can show that, for an appropriate value of β, the asymptotic convergence factor for this accelerated iteration is equal to $\rho_{\text{ess}}(T_\beta)$. Accordingly, in what follows, we optimize the convergence speed of the algorithm by minimizing $\rho_{\text{ess}}(T_\beta)$ with respect to β. We formally state these results and more in the following theorem.

Theorem 11.4 (Convergence and optimization of the accelerated averaging algorithm). *Consider the accelerated averaging algorithm (11.9) with $x(0) = x(-1) = x_0$, $A \in \mathbb{R}^{n \times n}$ symmetric, primitive, and row-stochastic matrix, and $\beta \in \mathbb{R}$. The following statements hold:*

(i) for all $\beta \in \mathbb{R}$, the set of fixed points of T_β is $\{\alpha \mathbb{1}_{2n} \mid \alpha \in \mathbb{R}\}$ and, if $\lim_{k \to \infty} x(k)$ exists, then it is equal to $\text{average}(x_0)\mathbb{1}_n$;

(ii) the following conditions are equivalent:

   (a) T_β is semi-convergent,
   (b) $\rho_{\text{ess}}(T_\beta) < 1$, and
   (c) $\beta \in (0, 2)$;

(iii) for $\beta \in (0, 2)$, along the accelerated averaging iteration (11.9)

$$\left\| x(k) - \text{average}(x_0)\mathbb{1}_n \right\|_2 \leq c_\varepsilon (\rho_{\text{ess}}(T_\beta) + \varepsilon)^k \left\| x(0) - \text{average}(x_0)\mathbb{1}_n \right\|_2,$$

where $\varepsilon > 0$ is an arbitrarily small constant and c_ε is a sufficiently large constant independent of x_0;

(iv) the optimal convergence rate of the accelerated averaging algorithm is

$$\min_{\beta \in (0,2)} \rho_{\text{ess}}(T_\beta) = \frac{\rho_{\text{ess}}(A)}{1 + \sqrt{1 - \rho_{\text{ess}}(A)^2}}, \tag{11.10}$$

which is obtained at

$$\beta^* = \operatorname{argmin}_{\beta \in (0,2)} \rho_{\text{ess}}(T_\beta) = \frac{2}{1 + \sqrt{1 - \rho_{\text{ess}}(A)^2}} \in (1, 2). \tag{11.11}$$

Note: A key advantage of the accelerated averaging algorithm is it is faster than standard averaging in two senses: First, it is immediate to see that $\rho_{\text{ess}}(T_{\beta^*}) = \frac{\rho_{\text{ess}}(A)}{1+\sqrt{1-\rho_{\text{ess}}(A)^2}} < \rho_{\text{ess}}(A)$. Second, Exercise E11.8 shows that performance improves also in its asymptotic order; for example, for averaging algorithms over circulant matrices, the spectral gap of order $1/n$ instead of order $1/n^2$. One important drawback of the accelerated averaging algorithm is that computation of optimal gain requires knowledge of the essential spectral radius of A.

Proof of Theorem 11.4. Regarding statement (i), we let $x^\star = \lim_{k \to \infty} x(k)$ and take the limit in both left and right hand side of the accelerated averaging algorithm (11.9) to obtain $x^* = \beta A x^* + (1 - \beta) x^*$, that is, after simple manipulations $x^\star = A x^\star$. Under the given assumptions on the matrix A and by employing the Perron–Frobenius Theorem, we obtain that $x^\star = \alpha \mathbb{1}_n$ for some $\alpha \in \mathbb{R}$. Observe also that $x(t) = \alpha \mathbb{1}_n$ is a conserved quantity for the accelerated averaging algorithm (11.9). Thus, when left-multiplying $x(t) = \alpha \mathbb{1}_n$ by $\mathbb{1}_n^\mathsf{T}$ and evaluating the result for $t = 0$, we obtain $\alpha = \operatorname{average}(x_0)$. This concludes the proof of statement (i).

Next, we prove statement (ii). We start by analyzing the matrix T_β with methods similar to those adopted for the second-order Laplacian flow in Section 8.1.1. The symmetric matrix A can be expressed as $A = U \Lambda U^\mathsf{T}$, where U is a unitary matrix and $\Lambda = \operatorname{diag}(\{\lambda_i\}_{i=1}^n)$ collects the eigenvalues of the matrix A. Because A is row-stochastic, symmetric and primitive, we set $\lambda_n = 1$ and we know $-1 < \lambda_i < 1$, for $i \in \{1, \ldots, n-1\}$. A similarity transformation with the matrix U leads us to

$$\begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}^\mathsf{T} T_\beta \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix}^\mathsf{T} \begin{bmatrix} \beta A & (1-\beta)I_n \\ I_n & 0 \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & U \end{bmatrix} = \begin{bmatrix} \beta \Lambda & (1-\beta)I_n \\ I_n & 0 \end{bmatrix}.$$

By appropriately permuting the entries of this matrix, we arrive at

$$\Gamma = \begin{bmatrix} \Gamma_1 & 0 & \dots & 0 \\ 0 & \Gamma_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \Gamma_n \end{bmatrix}, \qquad \text{where } \Gamma_i = \begin{bmatrix} \beta\,\lambda_i & 1-\beta \\ 1 & 0 \end{bmatrix}, \quad i \in \{1,\dots,n\}.$$

Note that, after the similarity transformation via the matrix U and the permutation (which is itself a similarity transformation), the spectra of Γ and T_β remain identical. We can, hence, analyze the matrix Γ to investigate the convergence rates. For a given index $i \in \{1,\dots,n\}$, the eigenvalues of Γ_i are the roots of

$$\mu_i^2 - (\beta\,\lambda_i)\mu_i + \beta - 1 = 0, \tag{11.12}$$

which are given by

$$\mu_{1,2;i} = \frac{\beta\,\lambda_i \pm \sqrt{\beta^2\,\lambda_i^2 - 4\beta + 4}}{2}. \tag{11.13}$$

For the system to converge to steady-state consensus, all eigenvalues $\mu_{1,2;i}$, $i \in \{1,\dots,n\}$, should lie within the unit disc, with only one eigenvalue on the unit circle. For Γ_n with $\lambda_n = 1$, we note that the eigenvalues are $\{1, \beta - 1\}$. Therefore, a necessary convergence condition for $\beta \in \mathbb{R}$ is

$$-1 < \beta - 1 < 1 \quad \text{or} \quad 0 < \beta < 2. \tag{11.14}$$

For the other block matrices Γ_i, $i \in \{1,\dots,n-1\}$, the eigenvalues are given by equation (11.13) and we note that: the sum of the roots is $\mu_{1;i} + \mu_{2;i} = \beta\,\lambda_i$, and the product of the roots is $\mu_{1;i} \cdot \mu_{2;i} = \beta - 1$. We consider the following cases:

a) Assume Γ_i has real-valued roots: For the roots to lie within the unit circle, we require $|\mu_{1;i}| < 1$, $|\mu_{2;i}| < 1$, and $\mu_{1;i}^2 + \mu_{2;i}^2 < 2$ for all $i \in \{1,\dots,n-1\}$. Regarding the latter:

$$\mu_{1;i}^i + \mu_{2;i}^2 = (\mu_{1;i} + \mu_{2;i})^2 - 2 \cdot \mu_{1;i} \cdot \mu_{2;i} < 2$$
$$\iff \quad \beta^2\lambda_i^2 - 2\beta + 2 < 2$$
$$\iff \quad \beta^2 - 2\beta < 0 \qquad\qquad \text{(for arbitrary } |\lambda_i| < 1)$$
$$\iff \quad \beta(\beta - 2) < 0 \quad \text{or} \quad \beta \in (0,2). \tag{11.15}$$

We now verify $|\mu_{1;i}| < 1$, $|\mu_{2;i}| < 1$. For Γ_i, $i \in \{1, \ldots, n-1\}$, with $|\lambda_i| < 1$, it can be calculated explicitly that $|\mu_{1;i}| < 1$, $|\mu_{2;i}| < 1$ if $\beta \in (0, 2)$.

b) Assume Γ_i has complex conjugate roots: As the coefficients of equation (11.12) are all real (β is real and λ_i is real as the matrix A is symmetric), the complex-conjugate roots have the same magnitude. We require the magnitudes to be strictly less than 1:

$$|\mu_{1;i}| = |\mu_{2;i}| = \sqrt{\beta - 1} < 1 \quad \implies \quad 0 < (\beta - 1) < 1 \text{ or } \beta \in (0, 2). \tag{11.16}$$

Equations (11.14), (11.15), and (11.16) together imply that the iteration converges for values of $\beta \in (0, 2)$. This concludes the proof of statement (ii).

Regarding statement (iii), it is an immediate consequence of Exercise E5.10 and some ad-hoc bounds. We leave it to the reader to fill out the details.

Finally, we prove statement (iv). In order to minimize the modulus of the eigenvalues of Γ_i, we choose β such that the discriminant in the expression (11.13) becomes zero:

$$\beta^2 \lambda_i^2 - 4\beta + 4 = 0. \tag{11.17}$$

Let us keep the index $i \in \{1, \ldots, n-1\}$ fixed. Two possible values of β arise from equation (11.17):

$$\beta \in \left\{ \frac{2}{1 + \sqrt{1 - \lambda_i^2}}, \frac{2}{1 - \sqrt{1 - \lambda_i^2}} \right\},$$

Because the second root may lead to a value of β outside the existence interval $(0, 2)$, we restrict ourselves to the optimal selection (for the index i) of the gain β as

$$\beta = \frac{2}{1 + \sqrt{1 - \lambda_i^2}}.$$

Among all choices of the gain β for different $i \in \{1, \ldots, n-1\}$, we note that

$$\beta^\star = 2/(1 + \sqrt{1 - \rho_{\text{ess}}(A)^2}),$$

as in equation (11.11), is the optimal choice to minimize the maximum magnitude of $|\mu_{1,2;i}|$ for $i \in \{1, \ldots, n-1\}$. Furthermore, since $1 > \rho_{\text{ess}}(A) \geq 0$, we have $2 > \beta^\star \geq 1$, and thus the magnitudes of all eigenvalues of Γ is strictly less than 1, except for the the

eigenvalue at 1. The magnitudes of the other eigenvalues of Γ for $\beta = \beta^\star$ are

$$\Big\{ \underbrace{1,\ |\beta^\star - 1|}_{\Gamma_n} \Big\}, \underbrace{\Big\{ |\sqrt{\beta^\star - 1}|,\ |\sqrt{\beta^\star - 1}| \Big\}}_{\Gamma_{n-1}},$$

$$\underbrace{\Big\{ |\mu_{1;n-2}(\beta^\star)|,\ |\mu_{2;n-2}(\beta^\star)| \Big\}}_{\Gamma_{n-2}}, \ \ldots,\ \underbrace{\Big\{ |\mu_{1;1}(\beta^\star)|,\ |\mu_{2;1}(\beta^\star)| \Big\}}_{\Gamma_1} \Big\}.$$

Furthermore, it can be verified that for $\beta = \beta^\star$ we have identical magnitudes $|\mu_{1;i}(\beta^\star)| = |\mu_{2;i}(\beta^\star)| = \sqrt{\beta^\star - 1}$ for all $i \in \{1, \ldots, n-2\}$. Finally, note that $\sqrt{\beta^\star - 1} \geq |\beta^\star - 1| = \beta^\star - 1$ so that

$$\rho_{\text{ess}}(T_{\beta^*}) = \rho_{\text{ess}}(\Gamma) = \sqrt{\beta^\star - 1} = \frac{\rho_{\text{ess}}(A)}{1 + \sqrt{1 - \rho_{\text{ess}}(A)^2}} < \rho_{\text{ess}}(A).$$

∎

11.6 Appendix: Design of fastest distributed averaging

We are interested in optimization problems of the form:

$$\text{minimize } r_{\text{asym}}(A) \text{ or } r_{\text{step}}(A)$$
$$\text{subject to } A \text{ compatible with a digraph } G, \text{ doubly-stochastic and primitive}$$

where A is compatible with G if its only non-zero entries correspond to the edges E of the graph. In other words, if $E_{ij} = \mathbb{e}_i \mathbb{e}_j^\mathsf{T}$ is the matrix with entry (i, j) equal to one and all other entries equal to zero, then $A = \sum_{(i,j) \in E} a_{ij} E_{ij}$ for arbitrary weights $a_{ij} \in \mathbb{R}$. We refer to such problems as fastest distributed averaging (FDAs) problems.

   Note: In what follows, we remove the constraint $A \geq 0$ to widen the set of matrices of interest. Accordingly, we remove the constraint of A being primitive. Convergence to average consensus is guaranteed by (1) achieving convergence factors less than 1, (2) subject to row-sums and column-sums equal to 1.

Problem 11.5 (Asymmetric FDA with asymptotic convergence factor).

$$\textit{minimize } \rho\big(A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big)$$
$$\textit{subject to } A = \sum_{(i,j) \in E} a_{ij} E_{ij}, \quad A\mathbb{1}_n = \mathbb{1}_n, \quad \mathbb{1}_n^\mathsf{T} A = \mathbb{1}_n^\mathsf{T}$$

The asymmetric FDA is a hard optimization problem. Even though the constraints are linear, the objective function, i.e., the spectral radius of a matrix, is not convex (and, additionally, not even Lipschitz continuous).

Problem 11.6 (Asymmetric FDA with per-step convergence factor).

$$\textit{minimize } \big\|A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n\big\|_2$$
$$\textit{subject to } A = \sum_{(i,j) \in E} a_{ij} E_{ij}, \quad A\mathbb{1}_n = \mathbb{1}_n, \quad \mathbb{1}_n^\mathsf{T} A = \mathbb{1}_n^\mathsf{T}$$

Problem 11.7 (Symmetric FDA problem).

$$\text{minimize } \rho\left(A - \mathbb{1}_n\mathbb{1}_n^\mathsf{T}/n\right)$$
$$\text{subject to } A = \sum_{(i,j)\in E} a_{ij}E_{ij}, \;\; A = A^\mathsf{T}, \;\; A\mathbb{1}_n = \mathbb{1}_n$$

Recall here that $A = A^\mathsf{T}$ *implies* $\rho(A) = \|A\|_2$.

Both Problems 11.6 and 11.7 are convex and can be rewritten as so-called *semidefinite programs* (SDPs); see (Xiao and Boyd, 2004). An SDP is an optimization problem where (1) the variable is a positive semidefinite matrix, (2) the objective function is linear, and (3) the constraints are affine equations. SDPs can be efficiently solved by software tools such as CVX; see (Grant and Boyd, 2014).

11.7   Historical notes and further reading

The main ideas in Sections 11.1 and 11.2 are standard in the literature on row-stochastic matrices and Markov chains. Example recent treatments in taken in the control literature include (Olshevsky and Tsitsiklis, 2009; Garin and Schenato, 2010; Fagnani, 2014).

Recent work has focused on achieving finite-time or linear-time average consensus; we here mention only the works by (Cortés, 2006; Wang and Xiao, 2010; Olshevsky, 2017).

The cumulative quadratic disagreement in Section 11.3 is taken from (Carli et al., 2009). Theorem 11.3 may be extended to the setting of normal matrices, as opposed to symmetric, as illustrated in (Carli et al., 2009); it is not known how to compute the cumulative quadratic disagreement for arbitrary doubly-stochastic primitive matrices.

Regarding Section 11.4, for more results on the study of circulant matrices and on the elegant settings of Cayley graphs we refer to (Davis, 1979; Carli et al., 2008b).

The accelerated consensus algorithm (11.9) is rooted in momentum methods for optimization (Polyak, 1964), and it has been applied to averaging algorithms for example in (Muthukrishnan et al., 1998; Bof et al., 2016).

11.8   Exercises

E11.1   **Induced norm of deflated doubly stochastic, primitive matrices with positive diagonal.** Assume A is doubly stochastic, primitive and has a strictly-positive diagonal. Show that

$$r_{\text{step}}(A) = \|A - \mathbb{1}_n \mathbb{1}_n^{\mathsf{T}}/n\|_2 < 1.$$

E11.2   **Spectrum of deflated doubly-stochastic, primitive, and symmetric matrices.** Consider a matrix A doubly stochastic, primitive and symmetric. Let $1 = \lambda_1 \geq \ldots \geq \lambda_n$ denote its real eigenvalue with corresponding orthonormal eigenvectors $v_1, \ldots, v_n$. Show that the matrix $A - \mathbb{1}_n \mathbb{1}_n^\mathsf{T}/n$ has eigenvalues $0, \lambda_2 \geq \ldots \geq \lambda_n$ with eigenvectors $v_1, \ldots, v_n$.

E11.3   **Spectral gap of regular cycle graphs.** A *k-regular cycle graph* is an undirected cycle graph with n-nodes each connected to itself and its $2k$ nearest neighbors with a uniform weight equal to $1/(2k+1)$. The associated doubly-stochastic adjacency matrix $A_{n,k}$ is a circulant matrix with first row given by

$$A_{n,k}(1,:) = \begin{bmatrix} \frac{1}{2k+1} & \cdots & \frac{1}{2k+1} & 0 & \cdots & 0 & \frac{1}{2k+1} & \cdots & \frac{1}{2k+1} \end{bmatrix}.$$

Using the results in Exercise E4.3, compute

(i)   the eigenvalues of $A_{n,k}$ as a function of n and k;
(ii) the limit of the spectral gap for fixed k as $n \to \infty$; and
(iii) the limit of the spectral gap for $2k = n - 1$ as $n \to \infty$.

E11.4 $\mathcal{H}_2$ **performance of balanced averaging in continuous time (Young et al., 2010).** Consider the continuous-time averaging dynamics with disturbance

$$\dot{x}(t) = -Lx(t) + w(t),$$

where $L = L^\mathsf{T}$ is the Laplacian matrix of an undirected and connected graph and $w(t)$ is an exogenous disturbance input signal. Pick a matrix $Q \in \mathbb{R}^{p \times n}$ satisfying $Q\mathbb{1}_n = \mathbb{0}_p$ and define the output signal $y(t) = Qx(t) \in \mathbb{R}^p$ as the solution from zero initial conditions $x(0) = \mathbb{0}_n$. Let $\Sigma_{L,Q}$ denote the input-output system from w to y and define its $\mathcal{H}_2$ *norm* by

$$\|\Sigma_{L,Q}\|_{\mathcal{H}_2}^2 = \operatorname{trace}\left(\int_0^\infty H(t)^\mathsf{T} H(t)dt\right), \tag{E11.1}$$

where $H(t) = Qe^{-Lt}$ is the so-called *impulse response matrix*. Show that

(i) $\|\Sigma_{L,Q}\|_{\mathcal{H}_2} = \sqrt{\operatorname{trace}(P)}$, where P is the solution to the *Lyapunov equality*

$$LP + PL = Q^\mathsf{T}Q; \tag{E11.2}$$

(ii) $\|\Sigma_{L,Q}\|_{\mathcal{H}_2} = \sqrt{\operatorname{trace}\left(L^\dagger Q^\mathsf{T}Q\right)/2}$, where $L^\dagger$ is the pseudoinverse of L; and

(iii) defining *short-range* and *long-range output matrices* Q_{sr} and Q_{lr} by $Q_{\mathrm{sr}}^\mathsf{T}Q_{\mathrm{sr}} = L$ and $Q_{\mathrm{lr}}^\mathsf{T}Q_{\mathrm{lr}} = I_n - \frac{1}{n}\mathbb{1}_n\mathbb{1}_n^\mathsf{T}$, respectively, we have

$$\|\Sigma_{L,Q}\|_{\mathcal{H}_2}^2 = \begin{cases} n - 1, & \text{for } Q = Q_{\mathrm{sr}}, \\ \displaystyle\sum_{i=2}^n \frac{1}{\lambda_i(L)}, & \text{for } Q = Q_{\mathrm{lr}}. \end{cases}$$

Hint: The $\mathcal{H}_2$ norm has several interesting interpretations, including the total output signal energy in response to a unit impulse input or the root mean square of the output signal in response to a white noise input with identity covariance. You may find useful Theorem 7.4 and Exercise E6.10.

E11.5   **Convergence rate for the Laplacian flow.** Consider a weight-balanced, strongly connected digraph G with self-loops, degree matrices $D_{\text{out}} = D_{\text{in}} = I_n$, doubly-stochastic adjacency matrix A, and Laplacian matrix L. Consider the associated Laplacian flow

$$\dot{x}(t) = -Lx(t).$$

For $x_{\text{ave}} := \frac{1_n^\mathsf{T} x(0)}{n}$, define the disagreement vector by $\delta(t) = x(t) - x_{\text{ave}} 1_n$.

  (i) Show that the average $t \mapsto \frac{1_n^\mathsf{T} x(t)}{n}$ is conserved and that, consequently, $1_n^\mathsf{T} \delta(t) = 0$ for all $t \geq 0$.
  (ii) Derive the matrix E describing the disagreeement dynamics

$$\dot{\delta}(t) = E\delta(t).$$

 (iii) Describe the spectrum $\operatorname{spec}(E)$ of E as a function of the spectrum $\operatorname{spec}(A)$. Show that $\operatorname{spec}(E)$ has a simple eigenvalue at $\lambda = 0$ with corresponding normalized eigenvector $v_1 := 1_n/\sqrt{n}$.
 (iv) The Jordan form J of E can be described as follows

$$E = P \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_m \end{bmatrix} P^{-1} =: \begin{bmatrix} c_1 & \tilde{c} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \tilde{J} \end{bmatrix} \begin{bmatrix} r_1 \\ \tilde{R} \end{bmatrix},$$

   where c_1 is the first column of P and r_1 is the first row of P^{-1}. Show that

$$\delta(t) = \tilde{c}\exp(\tilde{J}t)\tilde{R}\delta(0).$$

  (v) Use statements (iii) and (iv) to show that, for all $\varepsilon > 0$, there exists $c_\varepsilon > 0$ satisfying

$$\|\delta(t)\| \leq c_\varepsilon (e^\mu + \varepsilon)^t \|\delta(0)\|,$$

   where $\mu = \max\{\Re(\lambda) - 1 \mid \lambda \in \operatorname{spec}(A)\backslash\{1\}\} < 0$. Show that, if $A = A^\mathsf{T}$, then $\mu \leq \rho_{\text{ess}}(A) - 1$.

   Hint: Use arguments similar to those in Exercise E5.10 and in the proof of Theorem 7.4.

E11.6   **Convergence factors in digraphs with equal out-degree.** Consider the unweighted digraphs in the figure below with their associated discrete-time averaging systems $x(t+1) = A_1 x(t)$ and $x(t+1) = A_2 x(t)$. For which digraph is the worst-case discrete-time consensus protocol (i.e., the evolution starting from the worst-case initial condition) guaranteed to converge faster? Assign to each edge the same weight equal to $\frac{1}{3}$.
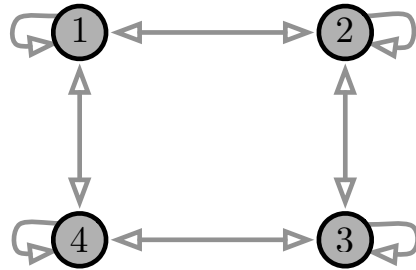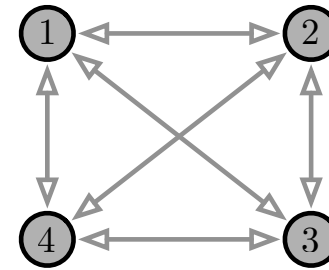
(a) Digraph 1   (b) Digraph 2

E11.7   **Convergence estimates.** Consider a discrete-time averaging system with 4 agents, state variable $x \in \mathbb{R}^4$, dynamics $x(k+1) = Ax(k)$, and averaging matrix $A = \sum_{i=1}^3 \alpha_i v_i v_i^{\mathsf{T}} \in \mathbb{R}^{4 \times 4}$ with

$$\alpha_1 = 1, \alpha_2 = \frac{1}{2}, \alpha_3 = \frac{1}{4}, \quad v_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}, v_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

(i) Verify A is row-stochastic, symmetric and primitive.

(ii) Suppose $x(0) = [0, 8, 2, 2]^{\mathsf{T}}$. It is possible that $x(3) = [4, 3, 2, 3]^{\mathsf{T}}$?

E11.8   **Scalability of accelerated consensus.**

(i) Prove the following series expansion around $x = 0$:

$$f(x) = \frac{1 - x}{1 + \sqrt{1 - (1 - x)^2}} = 1 - \sqrt{2}\sqrt{x} + o(x).$$

Next, consider a sequence of row-stochastic matrices $\{A_n \in \mathbb{R}^{n \times n}\}_{n \in \mathbb{N}}$ of increasing dimension, and the corresponding accelerated consensus algorithms with sequence of optimal iteration matrices $\{T_{\beta^*, n} \in \mathbb{R}^{2n \times 2n}\}_{n \in \mathbb{N}}$.

(ii) Prove that, if $\rho_{\text{ess}}(A_n) = 1 - g(n)$ with $g(n) = o(n)$ as $n \to \infty$, then the following series expansion holds as $n \to \infty$:

$$\rho_{\text{ess}}(T_{\beta^*, n}) = 1 - \sqrt{2}\sqrt{g(n)} + o(g(n)).$$

(iii) Show that, for circulant matrices $\{A_n\}_n$ with spectral radius given in equation (11.8) in Section 11.4, there exists a constant c such that the accelerated consensus algorithm satisfies

$$\rho_{\text{ess}}(T_{\beta^*, n}) = 1 - c\frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right).$$

Chapter 12

Time-varying Averaging Algorithms

In this chapter we discuss time-varying averaging systems, that is, systems in which the row-stochastic matrix is a function of time. We provide sufficient conditions on the sequence of digraphs associated to the sequence of row-stochastic matrices for consensus to be achieved. We focus mainly on the discrete-time setting, but present the main result also for continuous-time systems.

It is well known that, for time-varying systems, the analysis of eigenvalues is not appropriate anymore. In the following example, two matrices with spectral radius equal to $1/2$ are multiplied to obtain a spectral radius larger than 1:

$$\begin{bmatrix} \frac{1}{2} & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} & 0 \\ 0 & 0 \end{bmatrix}.$$

This example explains how it is not possible to predict the convergence of arbitrary products of matrices, just based on their spectral radii. Convergence proofs in this chapter will be based upon ergodicity coefficients and contraction inequalities.

12.1 Examples and models of time-varying discrete-time algorithms

In time-varying or time-varying algorithms the averaging row-stochastic matrix is not constant throughout time, but instead changes values and, possibly, switches among a finite number of values. Here are examples of discrete-time averaging algorithms with switching matrices.

Example 12.1 (Shared Communication Channel). *We consider a shared communication digraph $G_{\text{shared-comm}}$ whereby, at each communication round, only one node can transmit to all its out-neighbors over a common bus and every receiving node will implement a single averaging step. For example, if agent j receives the message from agent i, then agent j will implement:*

$$x_j^+ := \frac{1}{2}(x_i + x_j).$$ (12.1)

Each node is allocated a communication slot in a periodic deterministic fashion, e.g., in a round-robin scheduling, *where the n agents are numbered and, for each i, agent i talks only at times $i, n+i, 2n+i, \ldots, kn+i$ for $k \in \mathbb{Z}_{\geq 0}$. For example, in Figure* 12.1 *we illustrate the communication digraph and in Figure* 12.2 *the resulting round-robin communication protocol.*
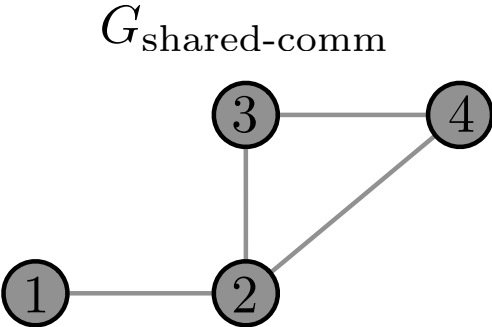
Figure 12.1: Example communication digraph
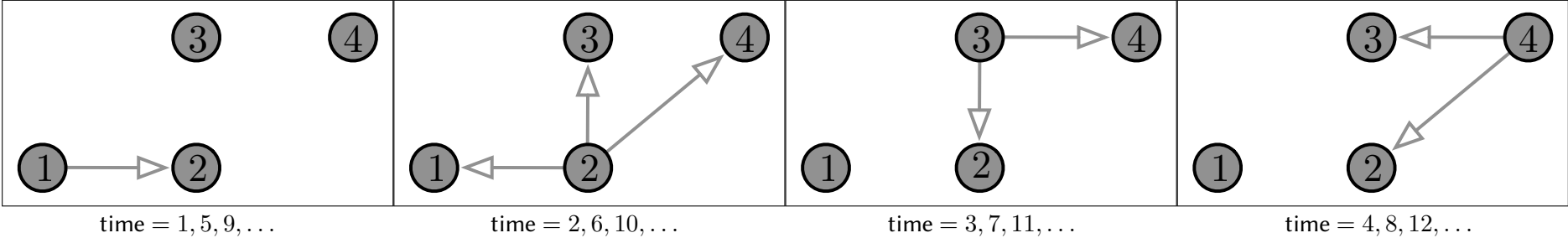
Figure 12.2: Round-robin communication protocol.

Formally, let A_i denote the averaging matrix corresponding to the transmission by agent i to its out-neighbors. With round robin scheduling, we have

$$x(n + 1) = A_n A_{n-1} \ldots A_1 x(1).$$

$\bullet$

Example 12.2 (Asynchronous Execution). *Imagine each node has a different clock, so that there is no common time schedule. Suppose that messages are safely delivered even if transmitting and receiving agents are not synchronized. Each time an agent wakes up, the available information from its neighbors varies. At an iteration instant for agent i, assuming agent i has new messages/information from agents $i_1, \ldots, i_m$, agent i will implement:*

$$x_i^+ := \frac{1}{m+1} x_i + \frac{1}{m+1}(x_{i_1} + \cdots + x_{i_m}).$$

Given arbitrary clocks, one can consider the set of times at which one of the n agents performs an iteration. Then the system is a discrete-time averaging algorithm. It is possible to carefully characterize all possible sequences of events (who transmitted to agent i when it wakes up). ●

12.2 Models of time-varying averaging algorithms

Consider a sequence of row-stochastic matrices $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$, or equivalently a time-varying row-stochastic matrix $k \mapsto A(k)$. The associated *time-varying averaging algorithm* is the discrete-time dynamical system

$$x(k+1) = A(k)x(k), \quad k \in \mathbb{Z}_{\geq 0}. \tag{12.2}$$

Let $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ be the sequence of weighted digraphs associated to $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$.

Note that $(1, \mathbb{1}_n)$ is an eigenpair for each matrix $A(k)$. Hence, all points in the consensus set $\{\alpha \mathbb{1}_n \mid \alpha \in \mathbb{R}\}$ are equilibria for the algorithm. We aim to provide conditions under which each solution converges to consensus.

We start with a useful definition, for two digraphs $G = (V, E)$ and $G' = (V', E')$, *union* of G and G' is defined by

$$G \cup G' = (V \cup V', E \cup E').$$

In what follows, we will need to compute only the union of digraphs with the same set of nodes; in that case, the graph union is essentially defined by the union of the edge sets. Some useful properties of the product of multiple row-stochastic matrices and of the unions of multiple digraphs are presented in Exercise E12.1.

12.3   Convergence over time-varying graphs connected at all times

Let us first consider the case where each $A(k)$ is symmetric and induces an undirected digraph $G(k)$ with possible self-loops. (Recall that a digraph is undirected if (v, u) is an edge if and only if (u, v) is an edge.)

Theorem 12.3 (Convergence under connectivity at all times). *Let $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ be a sequence of symmetric and doubly-stochastic matrices with associated digraphs $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ so that*

AC1) *each non-zero edge weight $a_{ij}(k)$, including the self-loops weights $a_{ii}(k)$, is larger than a constant $\varepsilon > 0$; and*

AC2) *each digraph $G(k)$ is strongly connected and aperiodic.*

Then the solution to $x(k + 1) = A(k)x(k)$ converges exponentially fast to $\mathrm{average}\big(x(0)\big)\mathbb{1}_n$.

Note: In Assumption (AC2) strong connectivity is equivalent to connectivity of the undirected digraph $G(k)$ regarded as a graph (by removing any possible self-loop).

Note: Assumption (AC1) prevents the weights from becoming arbitrarily close to zero as $k \to \infty$ and, as we show below, ensures that $\rho_{\text{ess}}(A(k))$ is upper bounded by a number strictly lower than 1 at every time $k \in \mathbb{Z}_{\geq 0}$. To gain some intuition into what can go wrong, consider a sequence of symmetric and doubly-stochastic averaging matrices $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ with entries given by

$$A(k) = \begin{bmatrix} 1 - \exp(-1/(k+1)^\alpha) & \exp(-1/(k+1)^\alpha) \\ \exp(-1/(k+1)^\alpha) & 1 - \exp(-1/(k+1)^\alpha) \end{bmatrix}$$

for $k \in \mathbb{Z}_{\geq 0}$ and exponent $\alpha \geq 1$. These matrices fail to satisfy Assumption (AC1). For any $\alpha \geq 1$ and for k, we know the $\rho_{\text{ess}}(A(k) < 1$. For any $\alpha \geq 1$ and for $k \to \infty$, this matrix converges to $A_\infty = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ with spectrum $\text{spec}(A_\infty) = \{-1, +1\}$ and essential spectral radius $\rho_{\text{ess}}(A_\infty) = 1$. One can show that,

(i) for $\alpha = 1$, the convergence of $A(k)$ to A_∞ is so slow that $\{x(k)\}_k$ converges to $\text{average}(x(0))\mathbb{1}_n$,

(ii) for $\alpha > 1$, the convergence of $A(k)$ to A_∞ is so fast that $\{x(k)\}_k$ oscillates indefinitely.[1]

---

[1] A simplified version of this example is the scalar iteration $x(k+1) = \exp(-1/(k+1)^\alpha)x(k)$ whose solution satisfies $\log(x(k)) = -\sum_{\kappa=0}^{k-1} \frac{1}{(\kappa+1)^\alpha} + \log(x_0)$. For $\alpha = 1$, $\lim_{k \to \infty} \log(x(k))$ diverges to $-\infty$, and $\lim_{k \to \infty} x(k)$ converges to zero. Instead, for $\alpha > 1$, $\lim_{k \to \infty} \log(x(k))$ exists finite, and thus $\lim_{k \to \infty} x(k)$ does not converge to zero.

Proof of Theorem 12.3. First, we reason as follows. At fixed n, there exist only a finite number of possible connected unweighted graphs and, for each given graph, the set of matrices with edge weights in the interval $[\varepsilon, 1]$ is compact. It is known that the following maps are continuous: the function from a matrix to its eigenvalues, the function from a complex number to its magnitude, and the function from $n - 1$ non-negative numbers to their maximum. Hence, by composition, the essential spectral radius ρ_{ess} is a continuous function of the matrix entries defined over a compact set and, therefore, it attains its maximum value. Because each digraph is strongly connected and aperiodic, each matrix is primitive. Because the essential spectral radius of each possible matrix is strictly less than 1, so is its maximum value. In summary, we now know that, under assumptions (AC1) and (AC2), there exists a $c \in [0, 1[$ so that $\rho_{\mathrm{ess}}(A(k)) \leq c < 1$ for all $k \in \mathbb{Z}_{\geq 0}$.

Second, From Section 5.3, we recall the notion of the disagreement vector $\delta(k) = x(k) - \mathrm{average}(x(0))\mathbb{1}_n$ and quadratic disagreement function $V(\delta) = \|\delta\|_2^2$. We also recall

$$\tau_2(A(k)) = \max_{\|y\|_2=1, y \perp \mathbb{1}_n} \|A(k)y\|_2 = \rho_{\mathrm{ess}}(A(k)) \leq c < 1. \tag{12.3}$$

Combining these two sets of ideas, it is immediate to compute

$$V(\delta(k+1)) = V(A(k)\delta(k)) = \|A(k)\delta(k)\|_2^2 \leq \rho_{\mathrm{ess}}(A(k))^2\|\delta(k)\|_2^2 \leq c^2 V(\delta(k)).$$

It follows that $V(\delta(k)) \leq c^{2k}V(\delta(0))$ or $\|\delta(k)\|_2 \leq c^k\|\delta(0)\|_2$, that is, $\delta(k)$ converges to zero exponentially fast. Equivalently, as $k \to \infty$, $x(k)$ converges exponentially fast to $\mathrm{average}(x(0))\mathbb{1}_n$. ∎

This proof is based on a positive "energy function" that decreases along the system's evolutions (we postpone a careful discussion of Lyapunov theory to Chapter 15). The same quadratic function is useful also for sequences of primitive row-stochastic matrices $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ with a common dominant left eigenvector, see Exercise E12.5. More general cases require a different type (not quadratic) of "decreasing energy" functions.

12.4   Convergence over time-varying digraphs connected over time

We are now ready to state the main result in this chapter.

Theorem 12.4 (Consensus for time-varying algorithms). *Let $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ be a sequence of row-stochastic matrices with associated digraphs $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$. Assume that*

(A1) *each digraph $G(k)$ has a self-loop at each node;*

(A2) *each non-zero edge weight $a_{ij}(k)$, including the self-loops weights $a_{ii}(k)$, is larger than a constant $\varepsilon > 0$; and*

(A3) *there exists a duration $\delta \in \mathbb{N}$ such that, for all times $k \in \mathbb{Z}_{\geq 0}$, the union digraph $G(k) \cup \ldots \cup G(k + \delta - 1)$ contains a globally reachable node.*

Then

(i)   *there exists a non-negative vector $w \in \mathbb{R}^n$ normalized to $w_1 + \cdots + w_n = 1$ such that $\lim_{k \to \infty} A(k) \cdot A(k-1) \cdot \ldots \cdot A(0) = \mathbb{1}_n w^\mathsf{T}$;*

(ii) *the solution to $x(k+1) = A(k)x(k)$ converges exponentially fast to $(w^\mathsf{T} x(0))\mathbb{1}_n$;*

(iii) *if additionally each matrix in the sequence is doubly-stochastic, then $w = \frac{1}{n}\mathbb{1}_n$ so that*

$$\lim_{k \to \infty} x(k) = \operatorname{average}(x(0))\mathbb{1}_n.$$

Note: In a sequence with property (A2), edges can appear and disappear, but the weight of each edge (that appears an infinite number of times) does not go to zero as $k \to \infty$.

Note: This result is analogous to the time-invariant result that we saw in Chapter 5. The existence of a globally reachable node is the connectivity requirement in both cases.

Note: Assumption (A3) is a *uniform* connectivity requirement, that is, any interval of length δ must have the connectivity property. In equivalent words, the connectivity property holds for any contiguous interval of duration δ.

Example 12.5 (Shared communication channel with round robin scheduling). *Consider the shared communication channel model with round-robin scheduling. Assume the algorithm is implemented over a communication graph $G_{\text{shared-comm}}$ that is strongly connected.*

Consider now the assumptions in Theorem 12.4. Assumption (A1) is satisfied because in equation (12.1) the self-loop weight is equal to $1/2$. Similarly, Assumption (A2) is satisfied because the edge weight is equal to $1/2$. Finally, Assumption (A3) is satisfied with duration

δ selected equal to n, because after n rounds each node has transmitted precisely once and so all edges of the communication graph $G_{\text{shared-comm}}$ are present in the union graph. Therefore, the algorithm converges to consensus. However, the algorithm does not converge to average consensus since it is false that the averaging matrices are doubly-stochastic.

Note: round robin is not necessarily the only scheduling protocol with convergence guarantees. Indeed, consensus is achieved so long as each node is guaranteed a transmission slot once every bounded period of time. $\bullet$

Next, we provide a second theorem on convergence over time-varying averaging systems, whereby we assume the matrix to be symmetric and the corresponding graphs to be connected over time.

Theorem 12.6 (Consensus for symmetric time-varying algorithms). *Let $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ be a sequence of symmetric row-stochastic matrices with associated graphs $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$. Let the matrix sequence $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ satisfy Assumptions* (A1) *and* (A2) *in Theorem* 12.4 *as well as*

(A4) *for all $k \in \mathbb{Z}_{\geq 0}$, the graph $\cup_{\tau \geq k} G(\tau)$ is connected.*

Then

(i) $\lim_{k \to \infty} A(k) \cdot A(k-1) \cdot \ldots \cdot A(0) = \frac{1}{n} \mathbb{1}_n \mathbb{1}_n^\mathsf{T}$;
(ii) each solution to $x(k+1) = A(k)x(k)$ converges exponentially fast to $\mathrm{average}\big(x(0)\big)\mathbb{1}_n$.

Note: this result is analogous to the time-invariant result that we saw in Chapter 5. For symmetric row-stochastic matrices and undirected graphs, the connectivity of an appropriate graph is the requirement in both cases.

Note: Assumption (A3) in Theorem 12.4 requires the existence of a finite time-interval of duration δ so that the union graph $\cup_{k \leq \tau \leq k+\delta-1} G(\tau)$ contains a globally reachable node for all times $k \geq 0$. This assumption is weakened in the symmetric case in Theorem 12.6 to Assumption (A4) requiring that the union graph $\cup_{\tau \geq k} G(\tau)$ is connected for all times $k \geq 0$.

Finally, we conclude this section with an instructive example.

Example 12.7 (Uniform connectivity is required for non-symmetric matrices). *We have learned that, for asymmetric matrices, a uniform connectivity property* (A3) *is required, whereas for symmetric matrices, uniform connectivity is not required (see* (A4)*). Here is a counter-example from (Hendrickx, 2008, Chapter 9) showing that Assumption* (A3) *cannot be relaxed for asymmetric graphs. Initialize a group of $n = 3$ agents to*

$$x_1 < -1, \quad x_2 < -1, \quad x_3 > +1.$$

Step 1: Perform $x_1^+ := (x_1 + x_3)/2$, $x_2^+ := x_2$, $x_3^+ := x_3$ a number of times δ_1 until

$$x_1 > +1, \quad x_2 < -1, \quad x_3 > +1.$$

Step 2: Perform $x_1^+ := x_1$, $x_2^+ := x_2$, $x_3^+ := (x_2 + x_3)/2$ a number of times δ_2 until

$$x_1 > +1, \quad x_2 < -1, \quad x_3 < -1.$$

Step 3: Perform $x_1^+ := x_1$, $x_2^+ := (x_1 + x_2)/2$, $x_3^+ := x_3$ a number of times δ_3 until

$$x_1 > +1, \quad x_2 > +1, \quad x_3 < -1.$$

And repeat this process.

| Step 1 | | Step 2 | | Step 3 | | union |
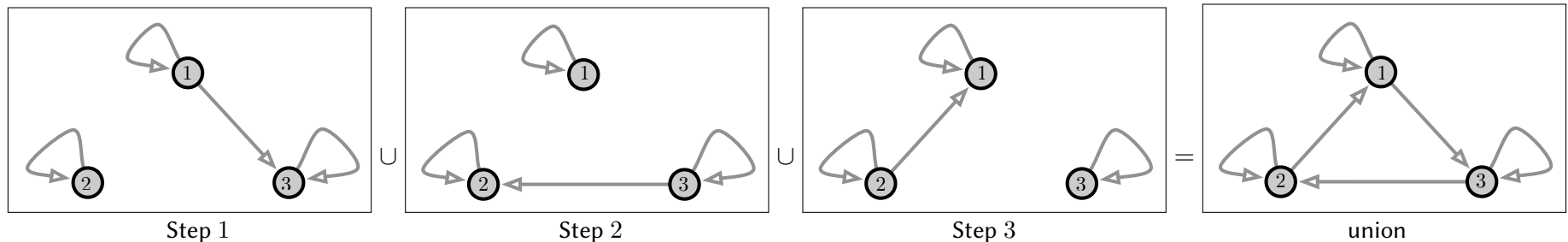
By design, on steps $1, 4, 7, \ldots$, the variable x_1 is changed to become larger than $+1$ by computing averages with $x_3 > +1$. However, note that, every time this happens, the variable $x_3 > +1$ is increasingly smaller and closer to $+1$. Hence, the durations of steps $1, 4, 7, \ldots$

increase: $\delta_1 < \delta_4 < \delta_7 < \ldots$, since more updates of x_1 are required for x_1 to become larger than $+1$. Indeed, one can formally show the following:

(i) *The agents do not converge to consensus so that one of the assumptions of Theorem 12.4 must be violated.*

(ii) *It is easy to see that* (A1) *and* (A2) *are satisfied.*

(iii) *Regarding connectivity, note that, for all $k \in \mathbb{Z}_{\geq 0}$, the digraph $\cup_{\tau \geq k} G(\tau)$ contains a globally reachable node. However, this property is not quite equivalent to Assumption* (A3).

(iv) *Assumption* (A3) *in Theorem 12.4 must be violated: there does not exist a duration $\delta \in \mathbb{N}$ such that, for all $k \in \mathbb{Z}_{\geq 0}$, the digraph $G(k) \cup \ldots \cup G(k + \delta - 1)$ contains a globally reachable node.*

(v) *Indeed, one can show that $\lim_{k \to \infty} \delta_k = \infty$ since, as we keep iterating steps 1, 2 and 3, their duration grows unbounded.*　　●

12.5   Appendix: Proofs

12.5.1   Connectivity over time

Before presenting the convergence to consensus proof for time-varying averaging systems, we provide one more useful result. This result allows us to manipulate our assumption of connectivity over time.

Lemma 12.8 (Global reachability over time). *Given a sequence of digraphs* $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ *such that each digraph* $G(k)$ *has a self-loop at each node, the following two properties are equivalent:*

(i) *there exists a duration* $\delta \in \mathbb{N}$ *such that, for all times* $k \in \mathbb{Z}_{\geq 0}$, *the union digraph* $G(k) \cup \ldots \cup G(k+\delta-1)$ *contains a directed spanning tree;*

(ii) *there exists a duration* $\Delta \in \mathbb{N}$ *such that, for all times* $k \in \mathbb{Z}_{\geq 0}$, *there exists a node* $j = j(k)$ *that reaches all nodes* $i \in \{1, \ldots, n\}$ *over the interval* $\{k, k+\Delta-1\}$ *in the following sense: there exists a sequence of nodes* $\{j, h_1, \ldots, h_{\Delta-1}, i\}$ *such that* (j, h_1) *is an edge at time* k, (h_1, h_2) *is an edge at time* $k+1$, ..., $(h_{\Delta-2}, h_{\Delta-1})$ *is an edge at time* $k+\Delta-2$, *and* $(h_{\Delta-1}, i)$ *is an edge at time* $k+\Delta-1$;

or, equivalently, for the reverse digraph,

(iii) *there exists a duration* $\delta \in \mathbb{N}$ *such that, for all times* $k \in \mathbb{Z}_{\geq 0}$, *the union digraph* $G(k) \cup \ldots \cup G(k+\delta-1)$ *contains a globally reachable node;*

(iv) *there exists a duration* $\Delta \in \mathbb{N}$ *such that, for all times* $k \in \mathbb{Z}_{\geq 0}$, *there exists a node* j *reachable from all nodes* $i \in \{1, \ldots, n\}$ *over the interval* $\{k, k+\Delta-1\}$ *in the following sense: there exists a sequence of nodes* $\{j, h_1, \ldots, h_{\Delta-1}, i\}$ *such that* (h_1, j) *is an edge at time* k, (h_2, h_1) *is an edge at time* $k+1$, ..., $(h_{\Delta-1}, h_{\Delta-2})$ *is an edge at time* $k+\Delta-2$, *and* $(i, h_{\Delta-1})$ *is an edge at time* $k+\Delta-1$.

Note: It is sometimes easy to see if a sequence of digraphs satisfies properties (i) and (iii). Property (iv) is directly useful in the analysis later in the chapter. Regarding the proof of the lemma, it is easy to check that (ii) implies (i) and that (iv) implies (iii) with $\delta = \Delta$. The converse is left as Exercise E12.3.

12.5.2 Proof of Theorem 12.4: the max-min function is exponentially decreasing

We are finally ready to prove Theorem 12.4. We start by noting that Assumptions (A1) and (A3) imply property Lemma 12.8(iv) about the existence of a duration Δ with certain properties. Next, without loss of generality, we assume that at some time $h\Delta$, for some $h \in \mathbb{N}$, the solution $x(h\Delta)$ is not equal to a multiple of $\mathbb{1}_n$ and, therefore, satisfies $V_{\text{max-min}}(x(h\Delta)) > 0$. Clearly,

$$x((h+1)\Delta) = A((h+1)\Delta - 1) \ldots A(h\Delta + 1) \cdot A(h\Delta)\, x(h\Delta)$$
$$=: \mathcal{A}x(h\Delta).$$

By Assumption (A3), we know that there exists a node j reachable from all nodes i over the interval $\{h\Delta, (h+1)\Delta - 1\}$ in the following sense: there exists a sequence of nodes $\{j, h_1, \ldots, h_{\Delta-1}, i\}$ such that all following edges exist in the sequence of digraphs: (h_1, j) at time $h\Delta$, (h_2, h_1) at time $h\Delta + 1$, ..., $(i, h_{\Delta-1})$ at time $(h+1)\Delta - 1$. Therefore, Assumption (A2) implies

$$a_{h_1,j}(h\Delta) \geq \varepsilon, \;\; a_{h_2,h_1}(h\Delta + 1) \geq \varepsilon, \;\; \ldots, \;\; a_{i,h_{\Delta-1}}((h+1)\Delta - 1) \geq \varepsilon,$$

and therefore their product satisfies

$$a_{i,h_{\Delta-1}}((h+1)\Delta - 1) \cdot a_{h_{\Delta-1},h_{\Delta-2}}((h+1)\Delta - 2) \ldots a_{h_2,h_1}(h\Delta + 1) \cdot a_{h_1,j}(h\Delta) \geq \varepsilon^\Delta.$$

Remarkably, this product is one term in the (i, j) entry of the row-stochastic matrix $\mathcal{A} = A((h+1)\Delta - 1) \ldots A(h\Delta)$. In summary, Assumption (A3) implies that there exists a node j such that, for all i, $\mathcal{A}_{ij} \geq \varepsilon^\Delta$ or, in other words, the row-stochastic matrix $\mathcal{A}$ has a positive column lower bounded by $\varepsilon^\Delta \mathbb{1}_n$.

We now invoke Lemma 5.5 from Section 5.3 to obtain that the row-stochastic matrix $\mathcal{A}$ is scrambling with $\tau_1(\mathcal{A}) \leq 1 - \varepsilon^\Delta$ and that the max-min disagreement function decreases according to

$$V_{\text{max-min}}\big(x((h+1)\Delta)\big) \leq (1 - \varepsilon^\Delta)V_{\text{max-min}}\big(x(h\Delta)\big).$$

This inequality proves exponential convergence of the cost function $k \mapsto V_{\text{max-min}}(x(k))$ to zero and, together with the positive definiteness property of the $V_{\text{max-min}}$ function, convergence of $x(k)$ to a multiple of $\mathbb{1}_n$. We leave the other statements in Theorem 12.4 to the reader and refer to (Moreau, 2005; Hendrickx, 2008) for further details.

12.6   Time-varying algorithms in continuous-time

We now briefly consider the continuous-time linear time-varying system

$$\dot{x}(t) = -L(t)x(t).$$

We associate a time-varying graph $G(t)$ (without self loops) to the time-varying Laplacian $L(t)$ in the usual manner.

For example, in Chapter 7, we discussed how the heading in some flocking models is described by the continuous-time Laplacian flow:

$$\dot{\theta} = -L\theta,$$

where each θ is the heading of a bird, and where L is the Laplacian of an appropriate weighted digraph G: each bird is a node and each directed edge (i, j) has weight $1/d_{\text{out}}(i)$. We discussed also the need to consider time-varying graphs: birds average their heading only with other birds within sensing range, but this sensing relationship may change with time.

Recall that the solution to a continuous-time time-varying system can be given in terms of the state transition matrix:

$$x(t) = \Phi(t, 0)x(0),$$

We refer to (Hespanha, 2009) for the proper definition and study of the state transition matrix.

Theorem 12.9 (Consensus for time-varying algorithms in continuous time). *Let $t \mapsto A(t)$ be a time-varying adjacency matrix with associated time-varying digraph $t \mapsto G(t)$, $t \in \mathbb{R}_{\geq 0}$. Assume*

(A1) *each non-zero edge weight $a_{ij}(t)$ is larger than a constant $\varepsilon > 0$,*

(A2) *there exists a duration $T > 0$ such that, for all $t \in \mathbb{R}_{\geq 0}$, the digraph associated to the adjacency matrix*

$$\int_t^{t+T} L(\tau)d\tau$$

contains a globally reachable node.

Then

(i) *there exists a non-negative $w \in \mathbb{R}^n$ normalized to $w_1 + \cdots + w_n = 1$ such that the state transition matrix $\Phi(t, 0)$ associated to $-L(t)$ satisfies $\lim_{t \to \infty} \Phi(t, 0) = \mathbb{1}_n w^\mathsf{T}$,*

(ii) *the solution to $\dot{x}(t) = -L(t)x(t)$ converges exponentially fast to $(w^\mathsf{T} x(0))\mathbb{1}_n$,*

(iii) *if additionally, the $\mathbb{1}_n^\mathsf{T} L(t) = \mathbb{0}_n^\mathsf{T}$ for almost all times t (that is, the digraph is weight-balanced at all times, except a set of measure zero), then $w = \frac{1}{n}\mathbb{1}_n$ so that*

$$\lim_{t \to \infty} x(t) = \text{average}(x(0))\mathbb{1}_n.$$

12.7 Historical notes and further reading

For historical notes on ergodicity coefficients we refer to Chapter 5.

The main result in this chapter, namely Theorem 12.4, appeared in the control literature in Moreau (2005). Note that Theorem 12.4 provides only sufficient condition for consensus in time-varying averaging systems. For results on necessary and sufficient conditions we refer the reader to the recent works (Blondel and Olshevsky, 2014; Xia and Cao, 2014) and references therein.

In the context of time-varying averaging systems, other relevant references on first and second order, discrete and continuous time systems include (Tsitsiklis, 1984; Tsitsiklis et al., 1986; Hong et al., 2006, 2007; Cao et al., 2008; Carli et al., 2008b).

For references on time-varying continuous-time averaging systems we refer to (Moreau, 2004; Lin et al., 2007; Hendrickx and Tsitsiklis, 2013).

12.8   Exercises

E12.1   **On the product of stochastic matrices (Jadbabaie et al., 2003).** For $k \geq 2$, consider non-negative $n \times n$ matrices $A_1, A_2, \ldots, A_k$ with positive diagonal entries. Let $a_{\min}$ (resp. $a_{\max}$) be the smallest (resp. largest) diagonal entry of $A_1, A_2, \ldots, A_k$ and let $G_1, \ldots, G_k$ be the digraphs associated with $A_1, \ldots, A_k$. Show that

   (i) $A_1 A_2 \ldots A_k \geq \left(\dfrac{a_{\min}^2}{2 a_{\max}} \right)^{k-1} (A_1 + A_2 + \cdots + A_k)$, and

   (ii) if the digraph $G_1 \cup \ldots \cup G_k$ is strongly connected, then the matrix $A_1 \ldots A_k$ is irreducible.

   Hint: Set $A_i = a_{\min} I_n + B_i$ for a non-negative B_i, and show statement (i) by induction on k.

E12.2   **Products of primitive matrices with positive diagonal.** Let A and $A_1, A_2, \ldots, A_{n-1}$ be primitive $n \times n$ matrices with positive diagonal entries. Let $x \in \mathbb{R}^n_{\geq 0}$ be a non-negative vector with at least one zero entry. Show that

   (i) the number of positive entries of Ax is strictly larger than the number of positive entries of x, and
   (ii) $A_1 A_2 \ldots A_{n-1} > 0$.

E12.3 **A simple proof.** Prove Lemma 12.8.
Hint: You will want to use Exercise E3.3.

E12.4   **Alternative sufficient condition.** As in Theorem 12.4, let $\{A(k)\}_{k\in\mathbb{Z}_{\geq 0}}$ be a sequence of row-stochastic matrices with associated digraphs $\{G(k)\}_{k\in\mathbb{Z}_{\geq 0}}$. Prove that the same asymptotic properties in Theorem 12.4 hold true under the following Assumption (A5), instead of Assumptions (A1), (A2), and (A3):

(A5) there exists a node j such that, for all times $k \in \mathbb{Z}_{\geq 0}$, each edge weight $a_{ij}(k)$, $i \in \{1, \ldots, n\}$, is larger than a constant $\varepsilon > 0$.

In other words, Assumption (A5) requires that all digraphs $G(k)$ contain all edges $a_{ij}(k)$, $i \in \{1, \ldots, n\}$, and that all these edges have weights larger than a strictly positive constant.

Hint: Modify the proof of Theorem 12.4.

E12.5 **Convergence over digraphs strongly-connected at all times.** Consider a sequence of row-stochastic matrices $\{A(k)\}_{k\in\mathbb{Z}_{\geq 0}}$ with associated digraphs $\{G(k)\}_{k\in\mathbb{Z}_{\geq 0}}$ so that

(A1) each non-zero edge weight $a_{ij}(k)$, including the self-loops weights $a_{ii}(k)$, is larger than a constant $\varepsilon > 0$;

(A2) each digraph $G(k)$ is strongly connected and aperiodic point-wise in time; and

(A3) there is a positive vector $w \in \mathbb{R}^n$ satisfying $\mathbb{1}_n^\mathsf{T} w = 1$ and $w^\mathsf{T} A(k) = w^\mathsf{T}$ for all $k \in \mathbb{Z}_{\geq 0}$.

Without relying on Theorem 12.4, show that

(i) the function $\delta \mapsto V(\delta) = \delta^\mathsf{T} \operatorname{diag}(w)\delta$ satisfies $V(A(k)\delta) < V(\delta)$ for all $k \in \mathbb{Z}_{\geq 0}$ and $\delta \neq \mathbb{0}_n$, and

(ii) the solution to $x(k+1) = A(k)x(k)$ satisfies to $\lim_{k\to\infty} x(k) = (w^\mathsf{T} x(0))\mathbb{1}_n$.

Hint: To establish (ii), adopt the following version of the Lyapunov Theorem: Let $\mathbb{0}_n$ be an equilibrium for the smooth discrete-time system $x(k+1) = f(x(k))$. Suppose there exists a continuous function $V: \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ satisfying $V(\mathbb{0}_n) = 0$, $V(x) > 0$ and $V(f(x)) < V(x)$ for all $x \neq \mathbb{0}_n$. Then $\mathbb{0}_n$ is asymptotically stable (which also means that every trajectory converges to $\mathbb{0}_n$). A comprehensive discussions of Lyapunov theory is postponed to Chapter 15.

Chapter 13

Randomized Averaging Algorithms

In this chapter we discuss averaging algorithms defined by sequences of random stochastic matrices. In other words, we imagine that at each discrete instant, the averaging matrix is selected randomly according to some stochastic model. We refer to such algorithms as randomized averaging algorithms. Randomized averaging algorithms are well behaved and easy to study in the sense that much information can be learned simply from the expectation of the averaging matrix.

13.1   Examples of randomized averaging algorithms

Consider the following models of randomized averaging algorithms.

Uniform Symmetric Gossip. Given an undirected graph G, at each iteration, select uniformly likely one of the graph edges, say agents i and j talk, and they both perform $(1/2, 1/2)$ averaging, that is:

$$x_i(k + 1) = x_j(k + 1) := \frac{1}{2}\big(x_i(k) + x_j(k)\big).$$

Packet Loss in Communication Network. Given a strongly connected and aperiodic digraph, at each communication round, packets travel over directed edges and, with some likelihood, each edge may drop the packet. (If information is not received, then the receiving node can either do no update whatsoever, or adjust its averaging weights to compensate for the packet loss).

Broadcast Wireless Communication. Given a digraph, at each communication round, a randomly-selected node transmits to all its out-neighbors. (Here we imagine that simultaneous transmissions are prohibited by wireless interference.)

Opinion Dynamics with Stochastic Interactions and Prominent Agents. Given an undirected graph and a probability $0 < p < 1$, at each iteration, select uniformly likely one of the graph edges and perform: with probability p both agents perform the $(1/2, 1/2)$ update, and with probability $(1 - p)$ only one agent performs the update and the "prominent agent" does not.

Note that, in the second, third and fourth example models, the row-stochastic matrices at each iteration are not symmetric in general, even if the original digraph was undirected.

13.2   A brief review of probability theory

We briefly review a few basic concepts from probability theory and refer the reader for example to (Breiman, 1992).

- Loosely speaking, a *random variable* $X\colon \Omega \to E$ is a measurable function from the *set of possible outcomes* Ω to some set E which is typically a subset of $\mathbb{R}$.

- The *probability* of an event (i.e., a subset of possible outcomes) is the measure of the likelihood that the event will occur. An event occurs *almost surely* if it occurs with probability equal to 1.

- The random variable X is called *discrete* if its image is finite or countably infinite. In this case, X is described by a *probability mass function* assigning a probability to each value in the image of X.

  Specifically, if X takes value in $\{x_1, \ldots, x_M\} \subset \mathbb{R}$, then the probability mass function $p\colon \{x_1, \ldots, x_M\} \to [0,1]$ satisfies $p_X(x_i) \geq 0$ and $\sum_{i=1}^{n} p_X(x_i) = 1$, and determines the probability of X being equal to x_i by $\mathbb{P}[X = x_i] = p_X(x_i)$.

- The random variable X is called *continuous* if its image is uncountably infinite. If X is an absolutely continuous function, X is described by a *probability density function* assigning a probability to intervals in the image of X.

  Specifically, if X takes value in $\mathbb{R}$, then the probability density function $f_X\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ satisfies $f(x) \geq 0$ and $\int_{\mathbb{R}} f(x)dx = 1$, and determines the probability of X taking value in the interval $[a, b]$ by $\mathbb{P}[a \leq X \leq b] = \int_a^b f(x)dx$.

- The *expected value* of a discrete variable is $\mathbb{E}[X] = \sum_{i=1}^{M} x_i p_X(x_i)$.

  The *expected value* of a continuous variable is $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.

- A (finite or infinite) sequence of random variables is *independent and identically distributed (i.i.d.)* if each random variable has the same probability mass/distribution as the others and all are mutually independent.

13.3   Randomized averaging algorithms

In this section we consider random sequences of row stochastic sequences. Accordingly, let $A(k)$ be the row-stochastic averaging matrix occurring randomly at time k and $G(k)$ be its associated graph. We then consider the *randomized averaging algorithm*

$$x(k + 1) = A(k)x(k).$$

We are now ready to present the main result of this chapter.

Theorem 13.1 (Consensus for randomized algorithms). *Let $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ be a sequence of random row-stochastic matrices with associated digraphs $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$. Assume*

(A1) *the sequence of variables $\{A(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ is i.i.d.,*

(A2) *at each time k, the random matrix $A(k)$ has a strictly positive diagonal so that each digraph in the sequence $\{G(k)\}_{k \in \mathbb{Z}_{\geq 0}}$ has a self-loop at each node, and*

(A3) *the digraph associated to the expected matrix $\mathbb{E}[A(k)]$, for any k, has a globally reachable node.*

Then the following statements hold almost surely:

(i) *there exists a random non-negative vector $w \in \mathbb{R}^n$ with $w_1 + \cdots + w_n = 1$ such that*

$$\lim_{k \to \infty} A(k) \cdot A(k-1) \cdot \ldots \cdot A(0) = \mathbb{1}_n w^\mathsf{T} \quad \text{almost surely,}$$

(ii) *as $k \to \infty$, each solution $x(k)$ of $x(k+1) = A(k)x(k)$ satisfies*

$$\lim_{k \to \infty} x(k) = \big(w^\mathsf{T} x(0)\big) \mathbb{1}_n \quad \text{almost surely,}$$

(iii) *if additionally each random matrix is doubly-stochastic, then $w = \frac{1}{n}\mathbb{1}_n$ so that*

$$\lim_{k \to \infty} x(k) = \text{average}\big(x(0)\big) \mathbb{1}_n.$$

Note: if each random matrix is doubly-stochastic, then $\mathbb{E}[A(k)]$ is doubly-stochastic. The converse is easily seen to be false.
Note: Assumption (A1) is restrictive and more general conditions are sufficient; see the discussion below in Section 13.4.

13.3.1   Additional results on uniform symmetric gossip algorithms

Recall: given undirected graph G, at each iteration, select uniformly likely one of the graph edges, say agents i and j talk, and they both perform $(1/2, 1/2)$ averaging, that is:

$$x_i(k+1) = x_j(k+1) := \frac{1}{2}\big(x_i(k) + x_j(k)\big).$$

Corollary 13.2 (Convergence for uniform symmetric gossip). *If the graph G is connected, then each solution to the uniform symmetric gossip converges to average consensus with probability 1.*

Proof based on Theorem 13.1. The corollary can be established by verifying that Assumptions (A1)–(A3) in Theorem 13.1 are satisfied. Regarding (A3), note that the graph associated to the expected averaging matrix is G.   ■

We provide also an alternative elegant proof.

Proof based on Theorem 12.6. For any time $k_0 \geq 0$ and any edge (i, j), consider the event "the edge (i, j) is not selected for update at any time larger than k_0." Since the probability that (i, j) is not selected at any time k is $1 - 1/m$, where m is the number of edges, the probability that (i, j) is not selected at any times after k_0 is

$$\lim_{k \to \infty} \left(1 - \frac{1}{m}\right)^{k-k_0} = 0.$$

With this fact one can verify that all assumptions in Theorem 12.6 are satisfied by the random sequence of matrices almost surely. Hence, almost sure convergence follows. Finally, since each matrix is doubly stochastic, $\mathrm{average}(x(k))$ is preserved, and the solution converges to $\mathrm{average}(x(0))\mathbb{1}_n$. ∎

13.3.2 Additional results on the mean-square convergence factor

Given a sequence of stochastic averaging matrices $\{A(k)\}_{k\in\mathbb{Z}_{\geq 0}}$ and corresponding solutions $x(k)$ to $x(k+1) = A(k)x(k)$, we define the *mean-square convergence factor* by

$$r_{\text{mean-square}}\big(\{A(k)\}_{k\in\mathbb{Z}_{\geq 0}}\big) = \sup_{x(0)\neq x_{\text{final}}} \limsup_{k\to\infty} \left(\mathbb{E}\left[\|x(k) - \text{average}(x(k))\mathbb{1}_n\|_2^2\right] \right)^{1/k}.$$

We now present upper and lower bounds for the mean-square convergence factor.

Theorem 13.3 (Upper and lower bounds on the mean-square convergence factor). *Under the same assumptions as in Theorem 13.1, the mean-square convergence factor satisfies*

$$\rho_{\text{ess}}\big(\mathbb{E}[A(k)]\big)^2 \leq r_{\text{mean-square}} \leq \rho\Big(\mathbb{E}\left[A(k)^{\mathsf{T}}(I_n - \mathbb{1}_n\mathbb{1}_n^{\mathsf{T}}/n)A(k)\right] \Big).$$

13.4 Historical notes and further reading

In this chapter we present results from (Fagnani and Zampieri, 2008; Tahbaz-Salehi and Jadbabaie, 2008; Garin and Schenato, 2010) that build on classic references such as (Chatterjee and Seneta, 1977; Cogburn, 1984). Specifically, references for the main Theorem 13.1 are (Tahbaz-Salehi and Jadbabaie, 2008) and (Fagnani and Zampieri, 2008). Note that Assumption (A1) is restrictive and more general conditions are sufficient. For example, Tahbaz-Salehi and Jadbabaie (2010) treat the case of a sequence of row-stochastic matrices generated by an ergodic and stationary random process. Related analysis and modeling results are presented in (Hatano and Mesbahi, 2005; Bajović et al., 2013; Matei et al., 2013; Touri and Nedić, 2014; Ravazzi et al., 2015).

For a comprehensive analysis of the mean-square convergence factor we refer to (Fagnani and Zampieri, 2008, Proposition 4.4).

Frasca and Hendrickx (2013) provide an upper bound on the mean square deviation of the consensus value from the initial average and show that, when the network size grows, the deviation tends to zero.

A detailed analysis of the uniform symmetric gossip model is given by Boyd et al. (2006). A detailed analysis of the model with stochastic interactions and prominent agents is given by (Acemoglu and Ozdaglar, 2011); see also (Acemoglu et al., 2013).

In this book we will not discuss averaging algorithms in the presence of quantization effects, we refer the reader instead to (Kashyap et al., 2007; Nedić et al., 2009; Frasca et al., 2009). Similarly, regarding averaging in the presence of noise, we refer to (Xiao et al., 2007; Bamieh et al., 2012; Lovisari et al., 2013; Jadbabaie and Olshevsky, 2019). Finally, regarding averaging in the presence of delays, we refer to (Olfati-Saber and Murray, 2004; Hu and Hong, 2007; Lin and Jia, 2008).

13.5   Table of asymptotic behaviors for averaging systems

Dynamics	Assumptions & Asymptotic Behavior	References
discrete-time: $x(k+1) = Ax(k)$, A row-stochastic adjacency matrix of digraph G	G has a globally reachable node subgraph of globally reachable nodes is aperiodic $\Longrightarrow$ $\lim_{k\to\infty} x(k) = (w^\mathsf{T}x(0))\mathbb{1}_n$, where $w \geq 0$, $w^\mathsf{T}A = w^\mathsf{T}$, and $\mathbb{1}_n^\mathsf{T}w = 1$	Thm 5.1
continuous-time: $\dot{x}(t) = -Lx(t)$, L Laplacian matrix of digraph G	G has a globally reachable node $\Longrightarrow$ $\lim_{t\to\infty} x(t) = (w^\mathsf{T}x(0))\mathbb{1}_n$, where $w \geq 0$, $w^\mathsf{T}L = \mathbb{0}_n^\mathsf{T}$, and $\mathbb{1}_n^\mathsf{T}w = 1$	Thm 7.4
time-varying discrete-time: $x(k+1) = A(k)x(k)$, $A(k)$ row-stochastic adjacency matrix of digraph $G(k)$, $k \in \mathbb{Z}_{\geq 0}$	(i) at each time k, $G(k)$ has self-loop at each node, (ii) each $a_{ij}(k) > 0$ is larger than $\varepsilon > 0$, (iii) there exists duration δ s.t., for all time k, $G(k) \cup \ldots \cup G(k+\delta-1)$ has a globally reachable node $\Longrightarrow$ $\lim_{k\to\infty} x(k) = (w^\mathsf{T}x(0))\mathbb{1}_n$, where $w \geq 0$, $\mathbb{1}_n^\mathsf{T}w = 1$	Thm 12.4
time-varying symmetric discrete-time: $x(k+1) = A(k)x(k)$, $A(k)$ symmetric stochastic adjacency of $G(k)$, $k \in \mathbb{Z}_{\geq 0}$	(i) at each time k, $G(k)$ has self-loop at each node, (ii) each $a_{ij}(k) > 0$ is larger than $\varepsilon > 0$, (iii) for all time k, $\cup_{\tau \geq k} G(\tau)$ is connected $\Longrightarrow$ $\lim_{k\to\infty} x(k) = \text{average}(x(0))\mathbb{1}_n$	Thm 12.6
time-varying continuous-time: $\dot{x}(t) = -L(t)x(t)$, $L(t)$ Laplacian matrix of digraph $G(t)$, $t \in \mathbb{R}_{\geq 0}$	(i) each $a_{ij}(t) > 0$ is larger than $\varepsilon > 0$, (ii) there exists duration T s.t., for all time t, digraph associated to $\int_t^{t+T} L(\tau)d\tau$ has a globally reachable node $\Longrightarrow$ $\lim_{t\to\infty} x(t) = (w^\mathsf{T}x(0))\mathbb{1}_n$, where $w \geq 0$, $\mathbb{1}_n^\mathsf{T}w = 1$	Thm 12.9
randomized discrete-time: $x(k+1) = A(k)x(k)$, $A(k)$ random row-stochastic adjacency matrix of digraph $G(k)$, $k \in \mathbb{Z}_{\geq 0}$	(i) $\{A(k)\}_{k\in\mathbb{Z}_{\geq 0}}$ is i.i.d., (ii) each matrix has strictly positive diagonal, (iii) digraph associated to $\mathbb{E}[A(k)]$ has a globally reachable node, $\Longrightarrow$ $\lim_{k\to\infty} x(k) = (w^\mathsf{T}x(0))\mathbb{1}_n$ almost surely	Thm 13.1

Part III

Nonlinear Systems

Chapter 14

Motivating Problems and Systems

In this chapter we begin our study of nonlinear network systems by introducing some example models and problems. Although the models presented are simple and their mathematical analysis is elementary, these models provide the appropriate notation, concepts, and intuition required to consider more realistic and complex models.

14.1   Lotka-Volterra population models

The Lotka-Volterra population models are one the simplest and most widely adopted frameworks for modeling the dynamics of interacting populations in mathematical ecology. These equations were originally developed in (Lotka, 1920; Volterra, 1928). In what follows we introduce various single-species and multi-species model of population dynamics. We start with single-species models. We let $x(t)$ denote the population number or its density at time t. The ratio $\dot{x}/x$ is the average contribution of an individual to the growth of the population.

Single-species constant growth model   In a simplest model, one may assume $\dot{x}/x$ is equal to a constant *growth rate r*. This assumption however leads to exponential growth or decay $x(t) = x(0)\,e^{rt}$ depending upon whether r is positive or negative. Of course, exponential growth may be reasonable only for short periods of time and violates a reasonable assumption of *bounded resources* for large times.

Single-species logistic growth model   In large populations it is natural to assume that resources would diminish with the growing size of the population. In a very simple model, one may assume $\dot{x}/x = r(1 - x/\kappa)$, where $r > 0$ is the intrinsic growth rate and $\kappa > 0$ is called the *carrying capacity*. This assumption leads to the so-called *logistic system*

$$\dot{x}(t) = rx(t)\big(1 - x(t)/\kappa\big). \tag{14.1}$$

This dynamical system has the following behavior:

 (i) there are two equilibrium points 0 and κ,

(ii) the solution is

$$x(t) = \frac{\kappa x(0)\,e^{rt}}{\kappa + x(0)(e^{rt} - 1)},$$

(iii) all solutions with $0 < x(0) < \kappa$ are monotonically increasing and converge asymptotically to κ,

(iv) all solutions with $\kappa < x(0)$ are monotonically decreasing and converge asymptotically to κ.

The reader is invited to show these facts and related ones in Exercise E 14.1. The evolution of the logistic equation from multiple initial values is illustrated in Figure 14.1.

$$x(t) = \frac{\kappa x(0)\mathrm{e}^{rt}}{\kappa + x(0)(\mathrm{e}^{rt} - 1)}$$
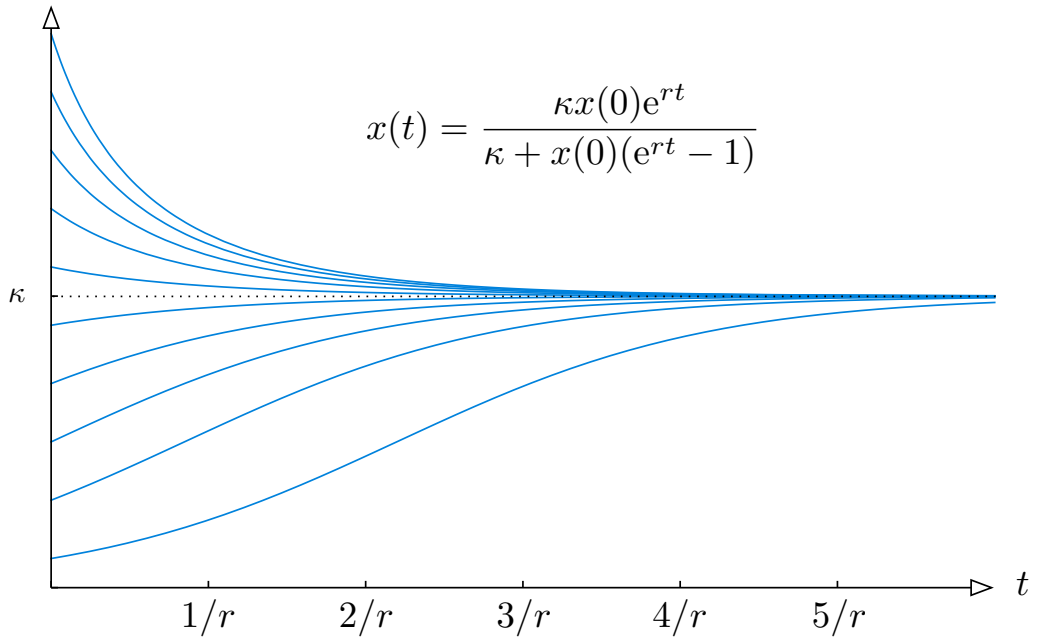
Figure 14.1: Solutions to the logistic equations from 10 initial conditions

Multi-species Lotka-Volterra model with signed interactions Finally, we consider the case of $n \geq 2$ interacting species. We assume logistic growth model for each species with an additional term due to the interaction with the other species. Specifically, we write the growth rate for species $i \in \{1, \ldots, n\}$,

$$\frac{\dot{x}_i}{x_i} = r_i + a_{ii}x_i + \sum_{j=1, j \neq i}^{n} a_{ij}x_j, \tag{14.2}$$

where the first two terms are the logistic equation (so that a_{ii} is typically negative because of bounded resources and the carrying capacity is $\kappa_i = -r_i/a_{ii}$), and the third term is the combined effect of the pairwise interactions with all other species. The vector r is called the *intrinsic growth rate*, the matrix $A = [a_{ij}]$ is called the *interaction matrix*, and the ordinary differential equations (14.2) are called the *Lotka-Volterra model* for $n \geq 2$ interacting species. For $x \in \mathbb{R}_{\geq 0}^n$, this model is written in vector form as

$$\dot{x} = \operatorname{diag}(x)(Ax + r) =: f_{\mathsf{LV}}(x). \tag{14.3}$$

(a) Common clownfish (*Amphiprion ocellaris*) near magnificent sea anemones (*Heteractis magnifica*) on the Great Barrier Reef, Australia. Clownfish and anemones provide an example of ecological mutualism in that each species benefits from the activity of the other. Public domain image from Wikipedia.

(b) The Canadian lynx (*Lynx canadensis*) is a major predator of the snowshoe hare (*Lepus americanus*). Historical records of animals captures indicate that the lynx and hare numbers rise and fall periodically; see (Odum, 1959). Public domain image from Rudolfo's Usenet Animal Pictures Gallery (no longer in existence).

(c) Subadult male lion (*Panthera Leo*) and spotted hyena (*Crocuta Crocuta*) compete for the same resources in the Maasai Mara National Reserve in Narok County, Kenya. Picture "Hyänen und Löwe im Morgenlicht" by lubye134, licensed under Creative Commons Attribution 2.0 Generic (BY 2.0).

Figure 14.2: Mutualism, predation and competition in population dynamics

As illustrated in Figure 14.2, for any two species i and j, the sign of a_{ij} and a_{ji} in the interaction matrix A is determined by which of the following three possible types of interaction is being modeled:

(+, +) = mutualism: for $a_{ij} > 0$ and $a_{ji} > 0$, the two species are in symbiosis and cooperation. The presence of species i has a positive effect on the growth of species j and vice versa.

(+,-) = predation: for $a_{ij} > 0$ and $a_{ji} < 0$, the species are in a predator-prey or host-parasite relationship. In other words, the presence of a prey (or host) species j favors the growth of the predator (or parasite) species i, wheres the presence of the predator species has a negative effect on the growth of the prey.

(-,-) = competition: for $a_{ij} < 0$ and $a_{ji} < 0$, the two species compete for a common resources of sorts and have therefore a negative effect on each other.

Note: the typical availability of bounded resources suggests it is ecologically meaningful to assume that the interaction matrix A is Hurwitz and that, to model the setting in which species live in isolation, the diagonal entries a_{ii} are negative.

Scientific questions of interest include:

(i) Does the Lotka-Volterra system have equilibrium points? Are they stable?

(ii) How does the presence of mutualism, predation, and/or competition affect the dynamic behavior?

(iii) Does the model predict extinction or periodic evolution of species?

14.2   Kuramoto coupled-oscillator models

In this section we introduce network of coupled oscillators and, in particular, phase-coupled oscillators. We start with two simple definitions. Given a connected, weighted, and undirected graph $G = (\{1, \ldots, n\}, E, A)$ and angles $\theta_1, \ldots, \theta_n$ associated to each node in the network, define the *coupled oscillators model* by

$$\dot{\theta}_i = \omega_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}. \tag{14.4}$$

A special case of this model is due to (Kuramoto, 1975); the *Kuramoto coupled oscillators model* is characterized by a complete homogeneous graph, i.e., a graph with identical edge weights $a_{ij} = K/n$ for all $i, j \in \{1, \ldots, n\}$ and for some *coupling strength K*. The Kuramoto model is

$$\dot{\theta}_i = \omega_i - \frac{K}{n} \sum_{j=1}^{n} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}. \tag{14.5}$$

Note: for $n = 2$, adopting the notation $\omega = \omega_1 - \omega_2$ and $a = a_{12} + a_{21}$, the coupled oscillator model can be written as a one-dimensional system in the difference variable $\theta = \theta_1 - \theta_2$ as:

$$\dot{\theta} = \omega - a \sin(\theta). \tag{14.6}$$

Coupled oscillator models arise naturally in many circumstances; in what follows we present three examples taken from (Dörfler and Bullo, 2014).

Example #1: A spring network on a ring   We start by studying a system of n dynamic particles constrained to rotate around a unit-radius circle and assumed to possibly overlap without ever colliding. Each particle is subject to (1) a non-conservative torque τ_i, (2) a linear damping torque, and (3) a total elastic torque. This system is illustrated in Figure 14.3.
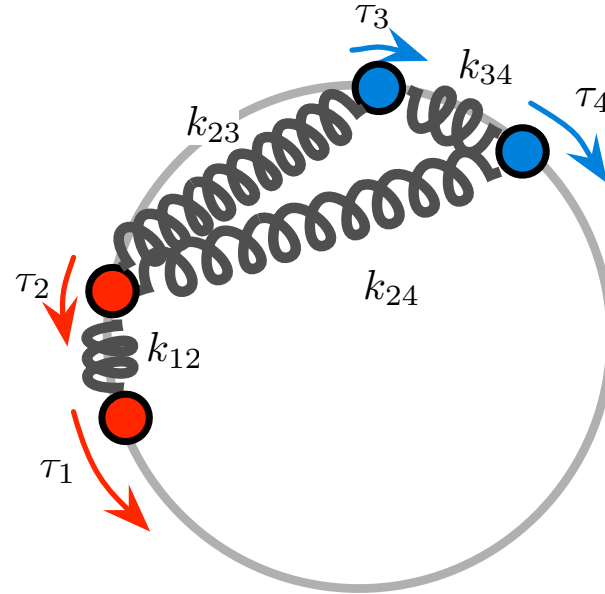
Figure 14.3: Mechanical analog of a coupled oscillator network

We assume that pairs of interacting particles i and j are coupled through elastic springs with stiffness $k_{ij} > 0$; we set $k_{ij} = 0$ if the particles are not interconnected. The elastic energy stored by the spring between particles at angles θ_i and θ_j is

$$\mathsf{U}_{ij}(\theta_i, \theta_j) = \frac{k_{ij}}{2}\text{distance}^2 = \frac{k_{ij}}{2}\big((\cos\theta_i - \cos\theta_j)^2 + (\sin\theta_i - \sin\theta_j)^2\big)$$
$$= k_{ij}\big(1 - \cos(\theta_i)\cos(\theta_j) - \sin(\theta_i)\sin(\theta_j)\big) = k_{ij}\big(1 - \cos(\theta_i - \theta_j)\big),$$

so that the elastic torque on particle i is

$$\mathsf{T}_i(\theta_i, \theta_j) = -\frac{\partial}{\partial\theta_i}\mathsf{U}_{ij}(\theta_i, \theta_j) = -k_{ij}\sin(\theta_i - \theta_j).$$

Newton's Law applied to this rotating system implies that this spring network obeys the dynamics

$$m_i \ddot{\theta}_i + d_i \dot{\theta}_i = \tau_i - \sum_{j=1}^{n} k_{ij} \sin(\theta_i - \theta_j),$$

where m_i and d_i are inertia and damping coefficients. In the limit of small masses m_i and uniformly-high viscous damping $d = d_i$, that is, $m_i/d \approx 0$, the model simplifies to the coupled oscillator network (14.4)

$$\dot{\theta}_i = \omega_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}.$$

with natural rotation frequencies $\omega_i = \tau_i/d$ and with coupling strengths $a_{ij} = k_{ij}/d$.

Example #2: The structure-preserving power network model   As second example we consider an AC power network, visualized in Figure 14.4, with n buses including generators and load buses. We present two simplified models for this network, a static power-balance model and a dynamic continuous-time model.

(a) Line diagram

(b) Equivalent graph representation
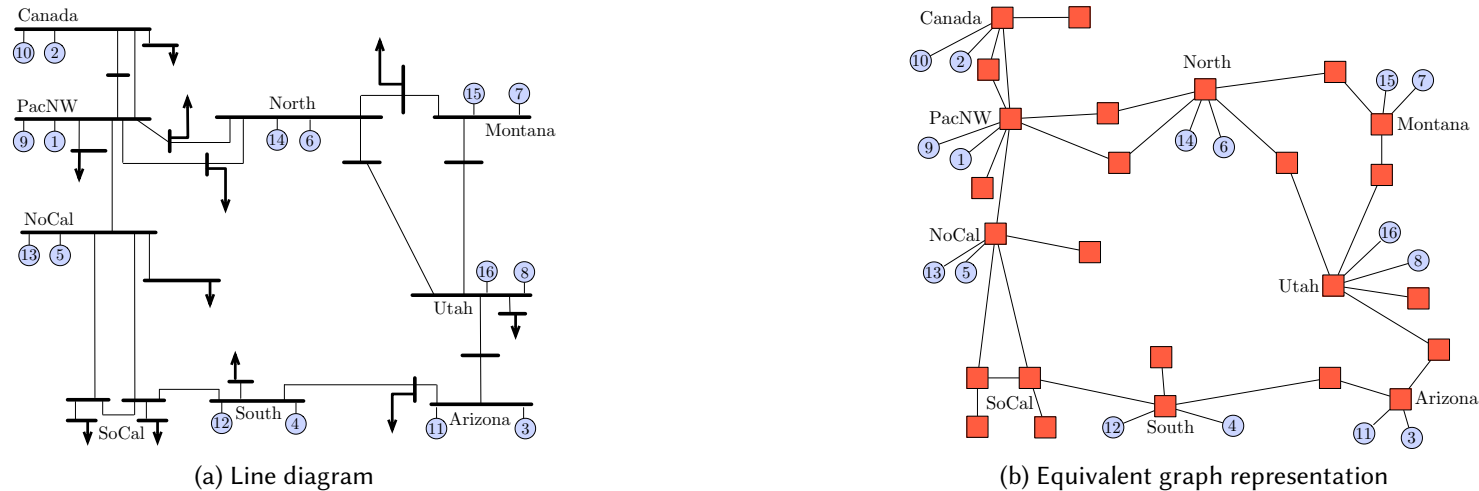
Figure 14.4: A simplified aggregated model with 16 generators and 25 load busses of the Western North American power grid, ofter referred to as the Western Interconnect. This model is often studied in the context of inter-area oscillations (Trudnowski et al., 1991). In the equivalent graph representation, generators are represented by light disks and load buses by dark boxes.

The transmission network is described by an admittance matrix $Y \in \mathbb{C}^{n \times n}$ that is symmetric and sparse with line impedances $Z_{ij} = Z_{ji}$ for each branch $\{i, j\} \in E$. The network admittance matrix is sparse matrix with nonzero off-diagonal entries $Y_{ij} = -1/Z_{ij}$ for each branch $\{i, j\} \in E$; the diagonal elements $Y_{ii} = -\sum_{j=1, j \neq i}^{n} Y_{ij}$ assure zero row-sums.

The static model is described by the following two concepts. Firstly, according to Kirchhoff's current law, the current injection at node i is balanced by the current flows from adjacent nodes:

$$I_i = \sum_{j=1}^{n} \frac{1}{Z_{ij}} (V_i - V_j) = \sum_{j=1}^{n} Y_{ij} V_j \,.$$

Here, I_i and V_i are the *phasor representations* of the nodal current injections and nodal voltages, so that, for example, $V_i = |V_i| \, e^{i\theta_i}$

corresponds to the signal $|V_i|\cos(\omega_0 t + \theta_i)$. (Recall $i = \sqrt{-1}$.) The complex power injection $S_i = V_i \cdot \overline{I}_i$ (where $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$) then satisfies the power balance equation

$$S_i = V_i \cdot \sum_{j=1}^{n} \overline{Y}_{ij}\overline{V}_j = \sum_{j=1}^{n} \overline{Y}_{ij}|V_i||V_j|e^{i(\theta_i - \theta_j)} \, .$$

Secondly, for a lossless network the real part of the power balance equations at each node is

$$\underbrace{P_i}_{\text{active power injection}} = \sum_{j=1}^{n} \underbrace{a_{ij} \cdot \sin(\theta_i - \theta_j)}_{\text{active power flow from } i \text{ to } j} \, , \quad i \in \{1, \ldots, n\}, \tag{14.7}$$

where $a_{ij} = |V_i||V_j||Y_{ij}|$ denotes the maximum power transfer over the transmission line $\{i, j\}$, and $P_i = \Re(S_i)$ is the active power injection into the network at node i, which is positive for generators and negative for loads. The systems of equations (14.7) are the *active power flow equations* at balance; see Figure 14.5.

(a) The active power satisfies a conservation law at each node: $P_i = \sum_{j=1}^{n} f_{ij}$

(b) Along each edge, the amount of active power transfered is proportional to the power angle: $f_{ij} = a_{ij}\sin(\theta_i - \theta_j)$
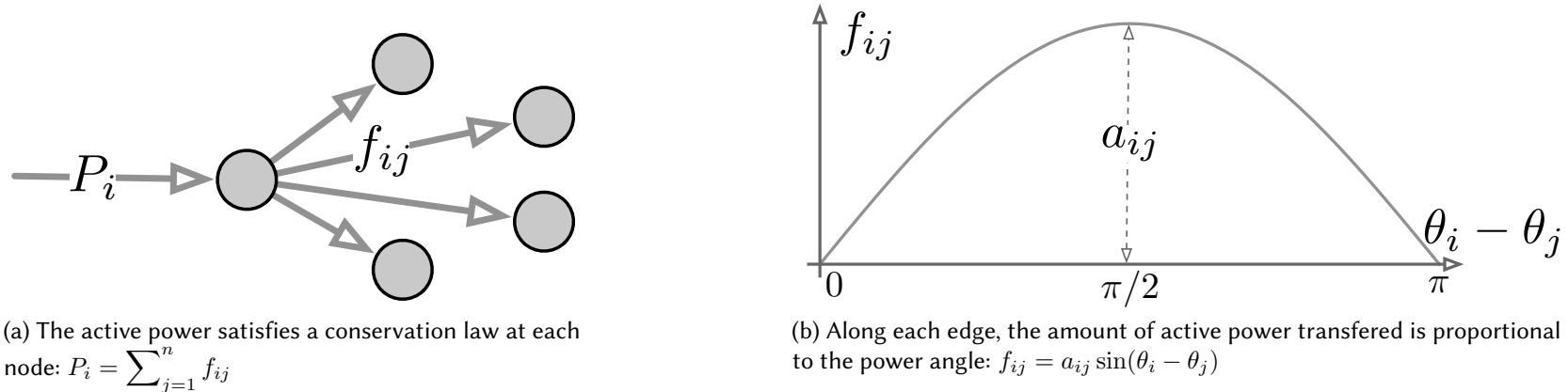
Figure 14.5: Interpretation of the active power flow equations (14.7).

Next, we discuss a simplified dynamic model. Many appropriate dynamic models have been proposed for each network node: zeroth order (for so-called constant power loads), first-order models (for so-called frequency-dependent loads and inverter-based generators), and second and higher order for generators; see (Bergen and Hill, 1981). For extreme simplicity, we here assume that every node is described by a first-order integrator with the following intuition: node i speeds up (i.e., θ_i increases) when the power balance at node i is positive, and slows down (i.e., θ_i decreases) when the power balance at node i is negative. This assumption leads immediately to the coupled-oscillators model (14.4) written as:

$$\dot{\theta}_i = P_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j). \tag{14.8}$$

The systems of equations (14.8) are a first-order simplified version of the so-called coupled swing equations; see (Bergen and Hill, 1981). A more realistic model of power network necessarily include higher-order dynamics for the generators, uncertain load models, mixed resistive-inductive lines, and the modeling of reactive power.

Example #3: Flocking, schooling, and vehicle coordination   As third example, we consider a set of n kinematic particles in the plane $\mathbb{R}^2$, which we identify with the complex plane $\mathbb{C}$. Each particle $i \in \{1, \ldots, n\}$ is characterized by its position $r_i \in \mathbb{C}$, its heading angle $\theta_i \in \mathbb{S}^1$, and a steering control law $u_i(r, \theta)$ depending on the position and heading of itself and other vehicles, see Figure 14.6.(a). For simplicity, we assume that all particles have unit speed. The particle kinematics are then given by

$$\dot{r}_i = e^{i\theta_i},$$
$$\dot{\theta}_i = u_i(r, \theta),$$

(14.9)

for $i \in \{1, \ldots, n\}$. If no control is applied, then particle i travels in a straight line with orientation $\theta_i(0)$, and if $u_i = \omega_i \in \mathbb{R}$ is a nonzero constant, then particle i traverses a circle with radius $1/|\omega_i|$.

The interaction among the particles is modeled by a graph $G = (\{1, \ldots, n\}, E, A)$ determined by communication and sensing patterns. As shown by Vicsek et al. (1995), motion patterns emerge if the controllers use only relative phase information between neighboring particles. As we will discuss later, we may adopt potential gradient control strategies (i.e., a negative gradient flow) to coordinate the relative heading angles $\theta_i(t) - \theta_j(t)$. As shown in Example #1, an intuitive extension of the quadratic elastic spring potential to the circle is the function $\mathsf{U}_{ij} \colon \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{R}$ defined by

$$\mathsf{U}_{ij}(\theta_i, \theta_j) = a_{ij}(1 - \cos(\theta_i - \theta_j)),$$

for each edge $\{i, j\}$ of the graph. Note that the potential $\mathsf{U}_{ij}(\theta_i, \theta_j)$ achieves its unique minimum value if the heading angles θ_i and θ_j are synchronized and its unique maximum when θ_i and θ_j are out of phase by an angle π.

These considerations motivate the *affine gradient control law*

$$\dot{\theta}_i = \omega_0 - K \frac{\partial}{\partial \theta_i} \sum_{\{i,j\} \in E} \mathsf{U}_{ij}(\theta_i - \theta_j) = \omega_0 - K \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j), \quad i \in \{1, \ldots, n\}.$$

(14.10)

to synchronize the heading angles of the particles for $K > 0$ (gradient descent), respectively, to disperse the heading angles for $K < 0$ (gradient ascent). The term ω_0 can induce additional rotations (for $\omega_0 \neq 0$) or translations (for $\omega_0 = 0$). A few representative trajectories are illustrated in Figure 14.6.

The controlled phase dynamics (14.10) give rise to elegant and useful coordination patterns that mimic animal flocking behavior (Leonard et al., 2012) and fish schools. Inspired by these biological phenomena, scientists have studied the controlled phase

dynamics (14.10) and their variations in the context of tracking and formation controllers in swarms of autonomous vehicles (Paley et al., 2007).
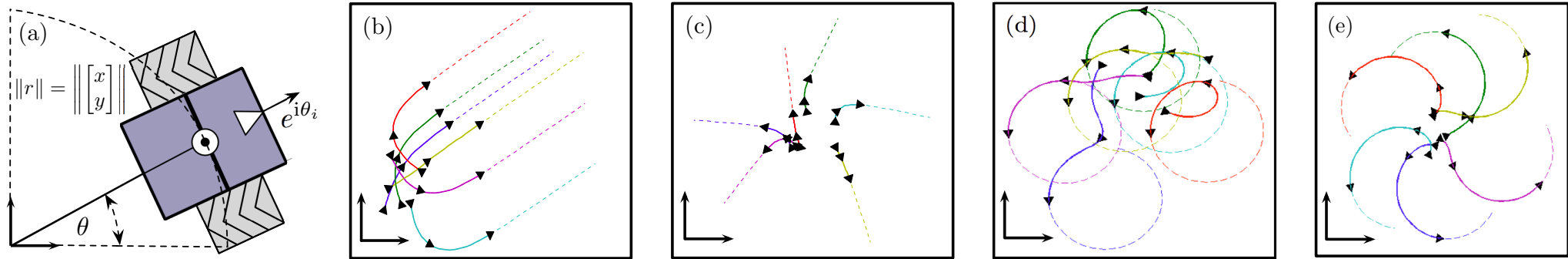
Figure 14.6: Figure (a) illustrates the particle kinematics (14.9). Figures (b)-(e) illustrate the controlled dynamics (14.9)-(14.10) with $n = 6$ particles, a complete interaction graph, and identical and constant natural frequencies: $\omega_0(t) = 0$ in figures (b) and (c) and $\omega_0(t) = 1$ in figures (d) and (e). The values of K are $K = +1$ in figures (b) and (d) and $K = -1$ in figure (c) and (e). The arrows depict the orientation, the dashed curves show the long-term position dynamics, and the solid curves show the initial transient position dynamics. As illustrated, the resulting motion displays synchronized or dispersed heading angles for $K = \pm 1$, and translational motion for $\omega_0 = 0$, respectively circular motion for $\omega_0 = 1$. Image reprinted from (Dörfler and Bullo, 2014) with permission from Elsevier.

Scientific questions of interest for coupled oscillator model include:

 (i) When do the oscillators asymptotically achieve frequency synchronization, that is, when do they asymptotically reach an equal velocity?

 (ii) When do they reach phase synchronization?

(iii) Are frequency (or phase) synchronized solutions stable and attractive in some sense?

14.3   Exercises

E14.1   **Logistic ordinary differential equation.** Given a growth rate $r > 0$ and a carrying capacity $\kappa > 0$, consider the logistic equation (14.1) defined by

$$\dot{x} = rx(1 - x/\kappa),$$

with initial condition $x(0) \in \mathbb{R}_{\geq 0}$. Show that

(i) there are two equilibrium points 0 and κ,
(ii) the solution is

$$x(t) = \frac{\kappa x(0)\, e^{rt}}{\kappa + x(0)(e^{rt} - 1)}, \tag{E14.1}$$

and it takes value in $\mathbb{R}_{\geq 0}$,

(iii) all solutions with $0 < x(0) < \kappa$ are monotonically increasing and converge asymptotically to κ,
(iv) all solutions with $\kappa < x(0)$ are monotonically decreasing and converge asymptotically to κ, and
(v) if $x(0) < \kappa/2$, then the solution $x(t)$ has an inflection point when $x(t) = \kappa/2$.

E14.2   **Simulating coupled oscillators.** Simulate in your favorite programming language and software package the coupled Kuramoto oscillators in equation (14.5). Set $n = 10$, define a vector $\omega \in \mathbb{R}^{10}$ with entries deterministically uniformly-spaced between -1 and 1. Select random initial phases.

(i) Simulate the resulting differential equations for $K = 10$ and $K = 0.1$.
(ii) Find the approximate value of K at which the qualitative behavior of the system changes from asynchrony to synchrony.

Turn in your code, a few printouts (as few as possible), and your written responses.

Chapter 15

Stability Theory for Dynamical Systems

In this chapter we provide a brief self-contained review of stability theory for nonlinear dynamical systems. We review the key ideas and theorems in stability theory, including the Lyapunov Stability Criteria and the Krasovskiĭ-LaSalle Invariance Principle. We then apply these theoretical tools to a number of example systems, including linear and linearized systems, negative gradient systems, continuous-time averaging dynamics (i.e., the Laplacian flow) and positive linear systems described by Metzler matrices.

This chapter is not meant to provide a comprehensive treatment, e.g., we leave out matters of existence and uniqueness of solutions and we do not include proofs. Section 15.10 below provides numerous references for further reading. We start the chapter by introducing a running example with three prototypical dynamical systems.

Example 15.1 (Gradient and mechanical systems). *We start by introducing a differentiable function $V : \mathbb{R} \to \mathbb{R}$; for example see Figure 15.1. Based on V and on two positive coefficients m and d, we define three instructive and prototypical dynamical systems:*
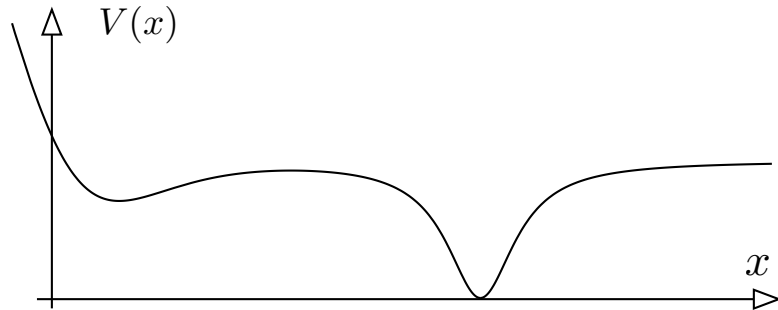
Figure 15.1: A differentiable function V playing the role of a potential energy function (i.e., a function describing the potential energy stored) in a negative gradient system, a conservative mechanical systems or a dissipative mechanical systems. Specifically, $V(x) = -x\,\mathrm{e}^{-x}/(1 + \mathrm{e}^{-x}) + (x - 10)^2/(1 + (x - 10)^2)$.

$$\text{negative gradient system:} \qquad \dot{x} = -\frac{\partial V}{\partial x}(x), \tag{15.1}$$

$$\text{conservative mechanical system:} \qquad m\ddot{x} = -\frac{\partial V}{\partial x}(x), \tag{15.2}$$

$$\text{dissipative mechanical system:} \qquad m\ddot{x} = -\frac{\partial V}{\partial x}(x) - d\dot{x}. \tag{15.3}$$

In the study of physical systems, the parameter m is an inertia, d is a damping coefficient, and the function V is the potential energy function, describing the potential energy stored in the system.

These example are also know as a (first order, second order, or second order dissipative) particle on an energy landscape, or the "rolling ball on a hill" examples. According to Newton's law, the correct physical systems are models (15.2) and (15.3), but we will also see interesting examples of first-order negative gradient systems (15.1).   •

15.1   On sets and functions

Before proceeding we review some basic general properties of sets and functions. First, we recall that a set $W \subset \mathbb{R}^n$ is *bounded* if there exists a constant K that each $w \in W$ satisfies $\|w\| \leq K$, *closed* if it contains its boundary (or, equivalently, if it contains all its limit points), and *compact* if it is bounded and closed.

Second, given a differentiable function $V : \mathbb{R}^n \to \mathbb{R}$, a *critical point* of V is a point $x^* \in \mathbb{R}^n$ satisfying

$$\frac{\partial V}{\partial x}(x^*) = \mathbb{0}_n.$$

A critical point x^* is a *local minimum point* (resp. *local strict minimum point*) of V if there exists a distance $\varepsilon > 0$ such that $V(x^*) \leq V(x)$ (resp. $V(x^*) < V(x)$) for all $x \neq x^*$ within distance ε of x^*. The point x^* is a global minimum if $V(x^*) < V(x)$ for all $x \neq x^*$. Local and global maximum points are defined similarly.

Given a constant $\ell \in \mathbb{R}$, we define the *ℓ-level set* of V and the *ℓ-sublevel set* of V by

$$V^{-1}(\ell) = \{y \in \mathbb{R}^n \mid V(y) = \ell\}, \quad \text{and} \quad V_{\leq}^{-1}(\ell) = \{y \in \mathbb{R}^n \mid V(y) \leq \ell\}.$$

These notions are illustrated in Figure 15.2.
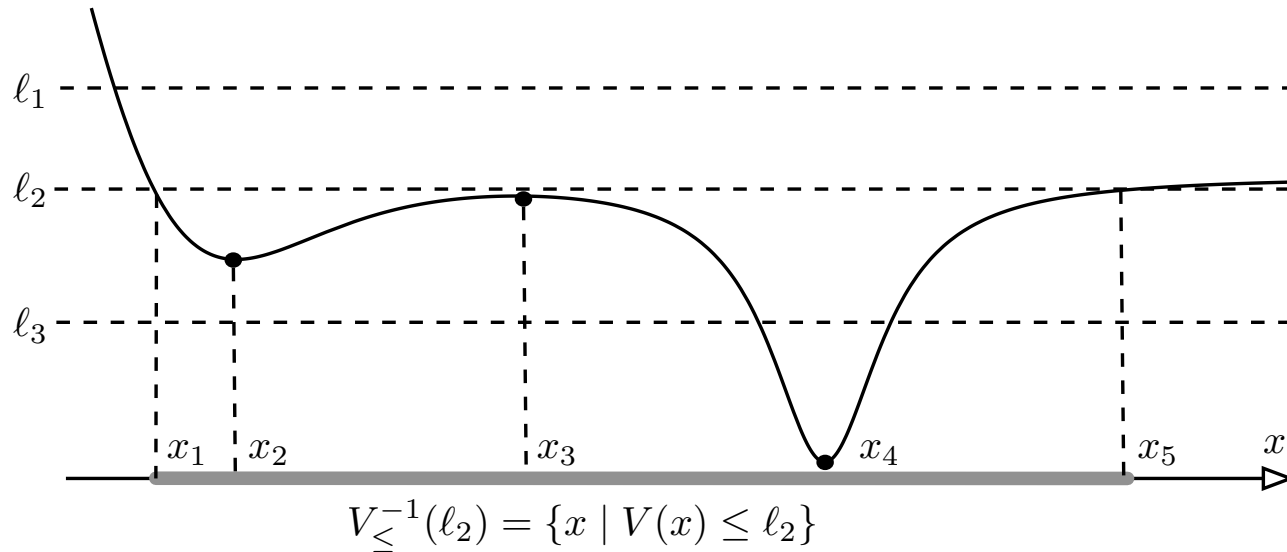
Figure 15.2: A differentiable function, its sublevel set and its critical points. The sublevel set $V_{\leq}^{-1}(\ell_1) = \{x \mid V(x) \leq \ell_1\}$ is unbounded. The sublevel set $V_{\leq}^{-1}(\ell_2) = [x_1, x_5]$ is compact and contains three critical points (x_2 and x_4 are local minima and x_3 is a local maximum). Finally, the sublevel set $V_{\leq}^{-1}(\ell_3)$ is compact and contains a single critical point, the global minimum x_4.

Third, given a point $x_0 \in \mathbb{R}^n$, a function $V : \mathbb{R}^n \to \mathbb{R}$ is

(i) *locally positive-definite* (resp. *positive-semidefinite*) *about* x_0 if $V(x_0) = 0$ and if there exists a neighborhood U of x_0 such that $V(x) > 0$ (resp. $V(x) \geq 0$) for all $x \in U \setminus \{x_0\}$,

(ii) *globally positive-definite about* x_0 if $V(x_0) = 0$ and $V(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{x_0\}$, and

(iii) *locally* (resp. *globally*) *negative-definite* if $-V$ is *locally* (resp. *globally*) *positive-definite*; and *negative-semidefinite* if $-V$ is *positive-semidefinite*.

Note: Assume a differentiable V is locally positive-definite about x_0. Pick $\alpha > V(x_0)$. One can show that the sublevel set $V_{\leq}^{-1}(\alpha)$ contains a neighborhood of x_0. Indeed, in Figure 15.2, V is locally positive-definite about x_4 and $V_{\leq}^{-1}(\ell_2)$ and $V_{\leq}^{-1}(\ell_3)$ are both compact intervals containing x_4.

Fourth and finally, a non-negative continuous function $V \colon X \to \mathbb{R}_{\geq 0}$ is

(i) *radially unbounded* if $X = \mathbb{R}^n$ and $V(x) \to \infty$ along any trajectory such that $\|x\| \to \infty$, that is, any sequence $\{x_n\}_{n \in \mathbb{N}}$ with the property that $\lim_{n \to \infty} \|x_n\| = \infty$ satisfies $\lim_{n \to \infty} V(x_n) = \infty$, and

(ii) *proper* if, for all $\ell \in \mathbb{R}$, the ℓ-sublevel set of V is compact.

We illustrate these concepts in Figure 15.3 and state a useful equivalence without proof.

Lemma 15.2. *A continuous function $V \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is proper if and only if it is radially unbounded.*

$$V(x) = -x \frac{e^{-x}}{1 + e^{-x}} + \frac{(x-10)^2}{1 + (x-10)^2}$$

$$V_{\text{log-lin}}(x) = x - 1 - \log(x)$$
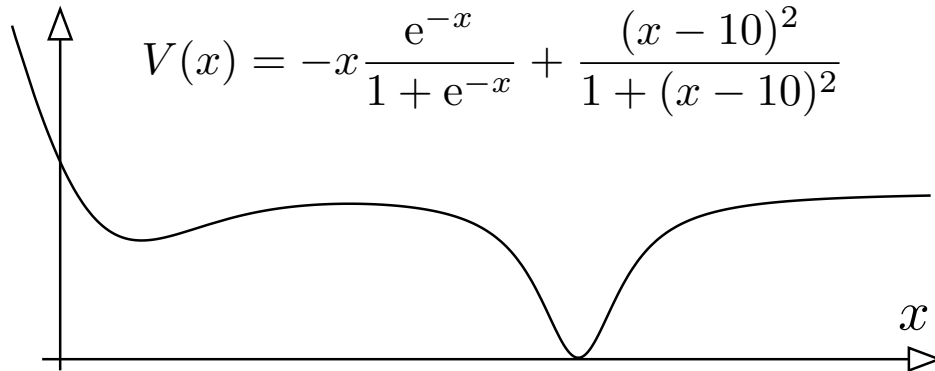
(a) This function $V \colon \mathbb{R} \to \mathbb{R}$ is not radially unbounded because $\lim_{x \to +\infty} V(x) = 1$.
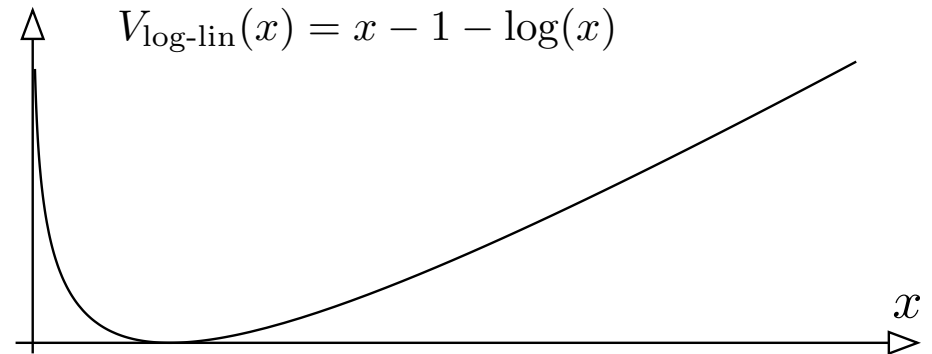
(b) The function $V_{\text{log-lin}} \mathbb{R}_{>0} \to \mathbb{R}$ is proper on $X = \mathbb{R}_{>0}$ since each sublevel set is a compact interval.

Figure 15.3: Example proper and not proper functions

15.2   Dynamical systems and stability notions

Dynamical systems

A *(continuous-time) dynamical system* is a pair (X, f) where X, called the *state space*, is a subset of $\mathbb{R}^n$ and f, called the *vector field*, is a map from X to $\mathbb{R}^n$. Given an initial state $x_0 \in X$, the *solution* (also called *trajectory* or *evolution*) of the dynamical system is a curve $t \mapsto x(t) \in X$ satisfying the differential equation

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0.$$

A dynamical system (X, f) is *linear* if $x \mapsto f(x) = Ax$ for some square matrix A.

Typically, the map f is assumed to have some continuity properties so that the solution exists and is unique for at least small times. Moreover, some of our examples are defined on closed submanifolds of $\mathbb{R}^n$ (e.g., the Lotka-Volterra model (14.3) is defined over the positive orthant $\mathbb{R}^n_{\geq 0}$), and the coupled oscillator model (14.4) is defined over the set of n angles) and additional assumptions are required to ensure that the solution exists for all times in X. We do not discuss these topics in great detail here, we simply assume the systems admit solutions inside X for all time, and refer to the references in Section 15.10 below.

Equilibrium points and their stability

An *equilibrium point* for the dynamical systems (X, f) is a point $x^* \in X$ such that $f(x^*) = \mathbb{0}_n$. If the initial state is $x(0) = x^*$, then the solution exists unique for all time and is constant: $x(t) = x^*$ for all $t \in \mathbb{R}_{\geq 0}$.

An equilibrium point x^* for the dynamical system (X, f) is

(i) *stable* (or *Lyapunov stable*) if, for each $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ so that if $\|x(0) - x^*\| < \delta$, then $\|x(t) - x^*\| < \varepsilon$ for all $t \geq 0$,

(ii) *unstable* if it is not stable, and

(iii) *locally asymptotically stable* if it is stable and if there exists $\delta > 0$ so that $\lim_{t \to \infty} x(t) = x^*$ for all trajectories satisfying $\|x(0) - x^*\| < \delta$.

These three concepts are illustrated in Figure 15.4.
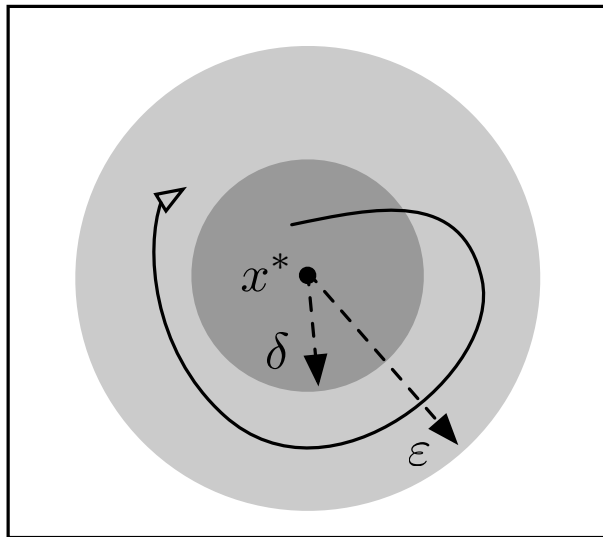
(a) Stable equilibrium: for all ε, each solution starting inside a sufficiently small δ-disk remains inside the ε-disk.
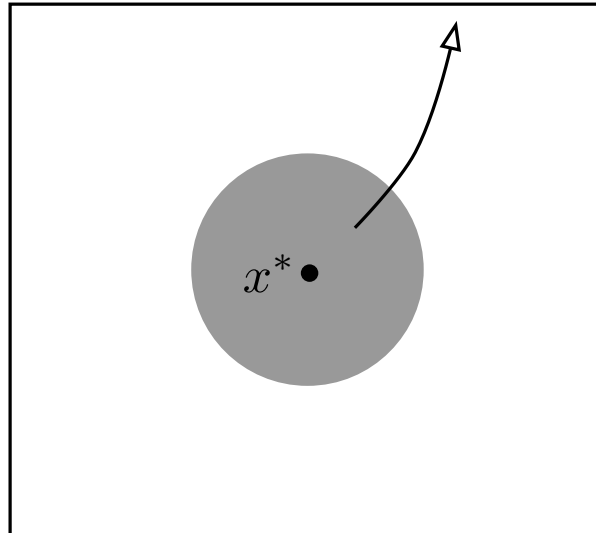
(b) Unstable equilibrium: no matter how small δ is, at least one solution starting inside the δ-disk diverges.
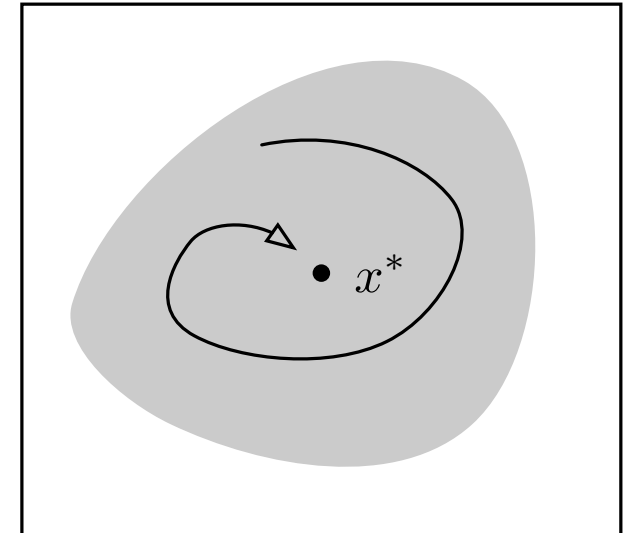
(c) Asymptotically stable equilibrium: solutions starting in a sufficiently small δ-disk converge asymptotically to the equilibrium.

Figure 15.4: Illustrations of a stable, an unstable and an asymptotically stable equilibrium.

These first three notions are local in nature. To characterize global properties of a dynamical system (X, f), we introduce the following notions. Given a locally asymptotically stable equilibrium point x^*,

(i) the set of initial conditions $x_0 \in X$ whose corresponding solution $x(t)$ converges to x^* is called the *region of attraction* of x^*,

(ii) x^* is said to be *globally asymptotically stable* if its region of attraction is the whole space X, and

(iii) x^* is said to be *globally* (respectively, *locally*) *exponentially stable* if it is globally (respectively, locally) asymptotically stable and there exist positive constants c_1 and c_2 such that all trajectories starting in the region of attraction satisfy

$$\|x(t) - x^*\| \leq c_1 \|x(0) - x^*\| \, \mathrm{e}^{-c_2 t} .$$

Example 15.3 (Gradient and mechanical systems: Example 15.1 continued). *It is instructive to report some numerical simulations of the three dynamical systems and state some conjectures about their equilibria and stability properties. These conjectures will be established in the next section.*

●
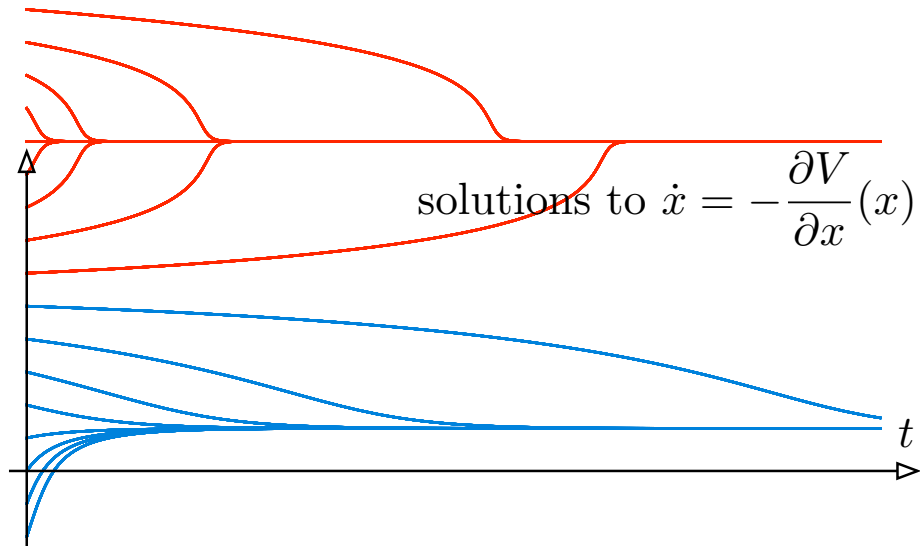
(a) Conjecture: each solution converges to one of the two local minima.
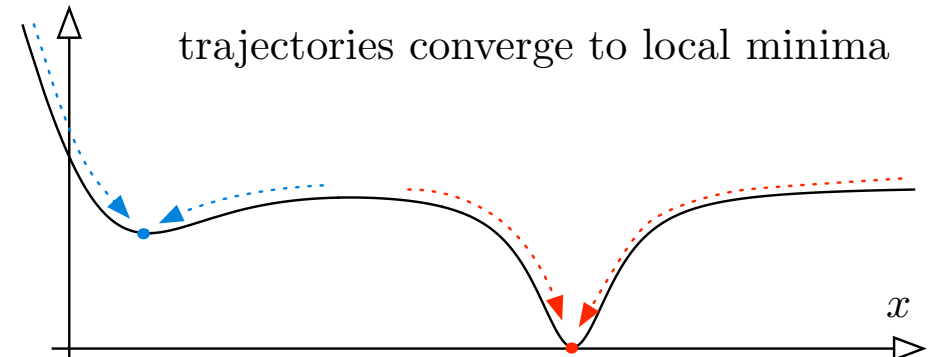
(b) Sketch of the motion on the potential energy surface.
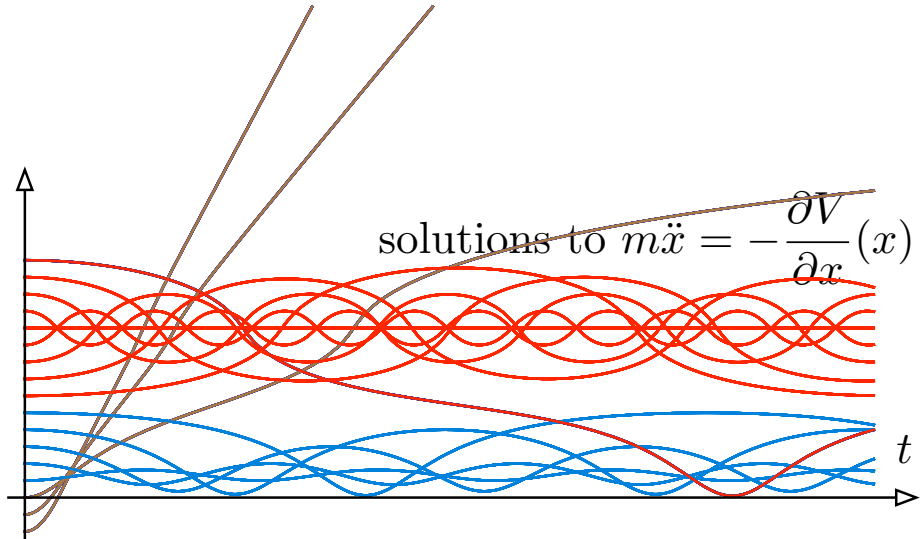
(c) Conjecture: each solution oscillates around a local minimum or diverge.
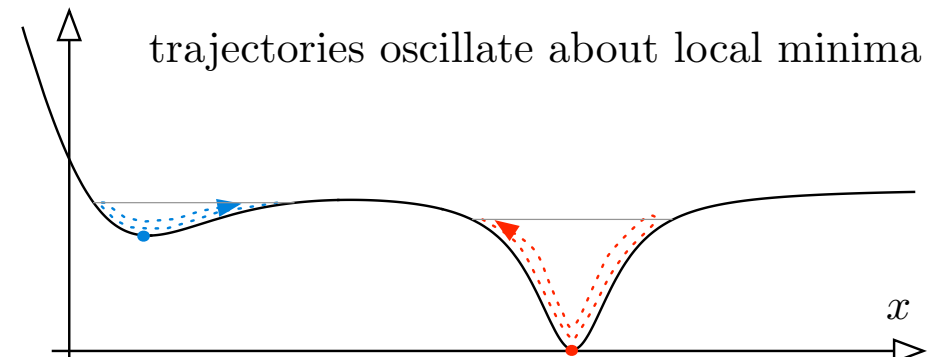
(d) Sketch of the motion on the potential energy surface.

15.3   The Lyapunov Stability Criteria

We are now ready to provide a critical tool in the study of the stability and convergence properties of a dynamical system. Roughly speaking, Lyapunov's idea is to use the concept of an energy function with a local/global minimum that is non-increasing along the system's solution.

Before proceeding, we require one final useful notion. The *Lie derivative* (also called the *directional derivative*) of a differentiable function $V \colon \mathbb{R}^n \to \mathbb{R}$ with respect to a vector field $f \colon \mathbb{R}^n \to \mathbb{R}^n$ is the function $\mathscr{L}_f V \colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$\mathscr{L}_f V(x) = \frac{\partial V}{\partial x}(x)f(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}(x)f_i(x). \tag{15.4}$$

Along the flow of a dynamical system (X, f), we have

$$\frac{d}{dt}V(x(t)) = \dot{V}(x(t)) = \mathscr{L}_f V(x(t)). \tag{15.5}$$

With this notation we note that $V \colon \mathbb{R}^n \to \mathbb{R}$ is non-increasing along every trajectory $x \colon \mathbb{R}_{\geq 0} \to X$ of (X, f) if

$$\dot{V}(x(t)) = \mathscr{L}_f V(x(t)) \leq 0,$$

or, equivalently, if each point $x \in X$ satisfies $\mathscr{L}_f V(x) \leq 0$. Because of this last inequality, when the vector field f is clear from the context, it is customary to adopt a slight abuse of notation and write $\dot{V}(x) = \mathscr{L}_f V(x)$.

We are now ready to present the main result of this section.

Theorem 15.4 (Lyapunov Stability Criteria). *Consider a dynamical system* $(\mathbb{R}^n, f)$ *with differentiable vector field* f *and with an equilibrium point* $x^* \in \mathbb{R}^n$. *The equilibrium point* x^* *is*

stable *if there exists a continuously-differentiable function* $V : \mathbb{R}^n \to \mathbb{R}$*, called a* weak Lyapunov function*, satisfying*

(L1) V *is locally positive-definite about* x^*,

(L2) $\mathscr{L}_f V$ *is locally negative-semidefinite about* x^*;

locally asymptotically stable *if there exists a continuously-differentiable function* $V : \mathbb{R}^n \to \mathbb{R}$*, called a* local Lyapunov function, *satisfying Assumption* (L1) *and*

(L3) $\mathscr{L}_f V$ *is locally negative-definite about* x^*;

globally asymptotically stable *if there exists a continuously-differentiable function* $V : \mathbb{R}^n \to \mathbb{R}$*, called a* global Lyapunov function, *satisfying*

(L4) V *is globally positive-definite about* x^*,

(L5) $\mathscr{L}_f V$ *is globally negative-definite about* x^*,

(L6) V *is proper.*

Note the immediate implications: (L4) $\implies$ (L1) and (L5) $\implies$ (L3) $\implies$ (L2).

Note: Theorem 15.4 assumes the existence of a Lyapunov function with certain properties, but does not provide constructive methods to design or compute one. In what follows we will see that Lyapunov functions can be designed for certain classes of systems. But, in general, the design of Lyapunov function is challenging. A common procedure is based on trial-and-error: one selects a so-called *candidate Lyapunov function* and verifies which, if any, of the properties (L1)–(L6) is satisfied.

Example 15.5 (Gradient and mechanical systems: Example 15.3 continued). *We now apply the Lyapunov Stability Criteria in Theorem 15.4 to the example dynamical systems in Example 15.1. Based on the properties of the function V in Figure 15.2 with local minimum points x_2 and x_4, we establish most of the conjectures from Example 15.3. Note that the vector fields and the Lyapunov functions we adopt in what follows are all continuously differentiable.*

Negative gradient systems: For the dynamics $\dot{x} = -\partial V/\partial x$, we select the function $V(x) - V(x_2)$ as candidate Lyapunov function about x_2. We compute

$$\dot{V}(x) = -\|\partial V/\partial x\|^2 \leq 0.$$

Note that $V - V(x_2)$ is locally positive definite about x_2 (Assumption (L1)) and $\dot{V}$ is locally negative definite about x_2 (Assumption (L3)); hence $V - V(x_2)$ is a local Lyapunov function for the equilibrium point x_2. An identical argument applies to x_4. Hence, both local minima x_2 and x_4 are locally asymptotically stable;

Conservative and dissipative mechanical systems: Given an inertia coefficient $m > 0$ and a damping coefficient $d \geq 0$, we write the conservative and the dissipative mechanical systems in first order form as:

$$\dot{x} = v, \quad m\dot{v} = -dv - \frac{\partial V}{\partial x}(x),$$

where $(x, v) \in \mathbb{R}^2$ are the position and velocity coordinates. As candidate Lyapunov function about the equilibrium point $(x_2, 0)$, we consider the mechanical energy $\mathsf{E} \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$ given by the sum of kinetic and potential energy:

$$\mathsf{E}(x, v) = \frac{1}{2}mv^2 + V(x).$$

We compute its derivative along trajectories of the considered mechanical system as follows:

$$\dot{\mathsf{E}}(x, v) = mv\dot{v} + \frac{\partial V}{\partial x}(x)\dot{x} = v\left(-dv - \frac{\partial V}{\partial x}(x)\right) + \frac{\partial V}{\partial x}(x)v = -dv^2 \leq 0.$$

This calculation, and x_2 being a local minimum of V, together establish that, for $d \geq 0$, the function $\mathsf{E} - V(x_2)$ is locally positive definite about x_2 (Assumption (L1)) and $\dot{\mathsf{E}}$ is locally negative semidefinite about $(x_2, 0)$ (Assumption (L2)). Hence, the function $\mathsf{E} - V(x_2)$ is a weak Lyapunov function for the equilibrium point $(x_2, 0)$ and, therefore, the point $(x_2, 0)$ is stable for both the conservative and the dissipative mechanical system. An identical argument applies to the point $(x_4, 0)$.

Note that we obtain the correct properties, i.e., consistent with the simulations in the previous exercise, for negative gradient system and for the conservative mechanical system. But more work is required to show that the local minima are locally asymptotically stable for the dissipative mechanical system. ●
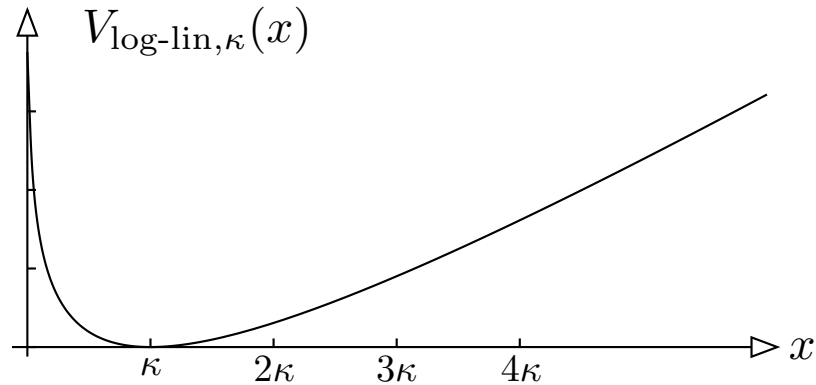
Figure 15.6: The function $V_{\text{log-lin},\kappa}(x) = x - \kappa - \kappa \log(x/\kappa)$, with $\kappa > 0$.

Example 15.6 (The logistic equation). *As second example, we consider the logistic equation* (14.1):

$$\dot{x}(t) = rx(t)\left(1 - \frac{x(t)}{\kappa}\right) =: f_{\text{logistic}}(x),$$

with growth rate r and carrying capacity κ. We neglect the possible initial condition $x(0) = 0$ (with subsequent equilibrium solution $x(t) = 0$ for all $t \geq 0$) and restrict out attention to solutions in $X = \mathbb{R}_{>0}$.

For $\kappa > 0$, define the logarithmic-linear function $V_{\text{log-lin},\kappa} : \mathbb{R}_{>0} \to \mathbb{R}$, illustrated in Figure 15.6, by

$$V_{\text{log-lin},\kappa}(x) = x - \kappa - \kappa \log\left(\frac{x}{\kappa}\right).$$

In Exercise E15.1 we ask the reader to verify that

(i) $V_{\text{log-lin},\kappa}$ is continuously differentiable with $\frac{d}{dx}V_{\text{log-lin},\kappa}(x) = (x - \kappa)/x$,
(ii) $V_{\text{log-lin},\kappa}(x) \geq 0$ for all $x > 0$ and $V_{\text{log-lin},\kappa}(x) = 0$ if and only if $x = \kappa$, and
(iii) $\lim_{x \to 0^+} V_{\text{log-lin},\kappa}(x) = \lim_{x \to \infty} V_{\text{log-lin},\kappa}(x) = +\infty$.

Next we compute

$$\mathscr{L}_{f_{\text{logistic}}} V_{\text{log-lin},\kappa}(x) = \frac{x - \kappa}{x} \cdot rx\left(1 - \frac{x}{\kappa}\right) = -\frac{r}{\kappa}(x - \kappa)^2.$$

In summary, we have established that f_{logistic} is a differentiable vector field, $x^* = \kappa$ is an equilibrium point, $V_{\text{log-lin},\kappa}$ is globally positive definite about κ, $\mathcal{L}_{f_{\text{logistic}}} V_{\text{log-lin},\kappa}$ is globally negative definite about κ, and $V_{\text{log-lin},\kappa}$ is proper. Hence, $V_{\text{log-lin},\kappa}$ is a global Lyapunov function and $x^* = \kappa$ is globally asymptotically stable. (This result is consistent with the behavior characterized in Exercise E14.1.) •

15.4 The Krasovskiĭ-LaSalle Invariance Principle

While the Lyapunov Stability Criteria are very useful, it is sometimes difficult to find a Lyapunov function with a negative-definite Lie derivative. To overcome this obstacle, in this section we introduce a powerful tool for the convergence analysis, namely the Krasovskiĭ-LaSalle Invariance Principle.

Before stating the main result, we introduce two useful concepts:

(i) A curve $t \mapsto x(t)$ *approaches* a set $S \subset \mathbb{R}^n$ as $t \to +\infty$ if the distance[1] from $x(t)$ to the set S converges to 0 as $t \to +\infty$.

If the set S consists of a single point s and $t \mapsto x(t)$ approaches S, then $t \mapsto x(t)$ converges to s in the usual sense: $\lim_{t \to +\infty} x(t) = s$. If the set S consists of multiple disconnected components and $t \mapsto x(t)$ approaches S, then $t \mapsto x(t)$ must approach one of the disconnected components of S. Specifically, if the set S is composed of a finite number of points, then $t \mapsto x(t)$ must converge to one of the points.

(ii) Given a dynamical system (X, f), a set $W \subset X$ is *invariant* (or *f-invariant*) if each solution starting in W remains in W, that is, if $x(0) \in W$ implies $x(t) \in W$ for all $t \geq 0$.

For example, any sublevel set of a function is invariant for the corresponding negative gradient flow.

We are now ready to present the main result of this section.

---

[1] Here we define the distance from a point y to a set Z to be $\inf_{z \in Z} \|y - z\|$.

Theorem 15.7 (Krasovskiĭ-LaSalle Invariance Principle). *For a dynamical system (X, f) with differentiable f, assume that*

KL1) *all trajectories of (X, f) are bounded,*

KL2) *there exists a closed invariant set $W \subset X$, and*

KL3) *there exists a continuously-differentiable function $V : W \to \mathbb{R}$ satisfying $\mathscr{L}_f V(x) \le 0$ for all $x \in W$.*

Then for each solution $t \mapsto x(t)$ starting in W there exists $c \in \mathbb{R}$ such that x converges to the largest invariant set contained in

$$\left\{ x \in W \mid \mathscr{L}_f V(x) = 0 \right\} \cap V^{-1}(c).$$

Note: if the closed invariant set $W \subset X$ in Assumption (KL2) is also bounded, then Assumption (KL1) is automatically satisfied.

Note: unlike in the Lyapunov Stability Criteria, the Krasovskiĭ-LaSalle Invariance Principle does not require the function V to be locally positive definite and establishes certain asymptotic convergence properties without requiring the Lie derivative of V to be locally negative definite.

Note: in some examples it is sufficient for one's purposes to show that $x(t) \to \left\{ x \in W \mid \mathscr{L}_f V(x) = 0 \right\}$. In other cases, however, one really needs to analyze the largest invariant set inside $\left\{ x \in W \mid \mathscr{L}_f V(x) = 0 \right\}$.

Note: If the largest invariant set is the union of multiple disjoint non-empty sets, then the solution to the negative gradient flow must converge to one of these disjoint sets.

Example 15.8 (Gradient and mechanical systems: Example 15.5 continued). *We continue the analysis of the example dynamical systems in Examples 15.1 and 15.5. Specifically, we sharpen here our results about the dissipative mechanical system about a local minimum point x_2 (or x_4) based on the Krasovskiĭ-LaSalle Invariance Principle.*

First, we note that the assumptions of the Krasovskiĭ-LaSalle Invariance Principle in Theorem 15.7 are satisfied:

(i) *the function* E *and the vector field (the right-hand side of the mechanical system) are continuously differentiable;*

(ii) *the derivative* Ė *is locally negative semidefinite; and*

(iii) *for any initial condition $(x_0, v_0) \in \mathbb{R}^2$ sufficiently close to $(x_2, 0)$ the sublevel set $\{(x, v) \in \mathbb{R}^2 \mid \mathsf{E}(x, v) \leq \mathsf{E}(x_0, v_0)\}$ is compact due to the local positive definiteness of V at x_2.*

It follows that $(x(t), v(t))$ converges to largest invariant set contained in

$$C = \{(x, v) \in \mathbb{R}^2 \mid \mathsf{E}(x, v) \leq \mathsf{E}(x_0, v_0), v = 0\} = \{(x, 0) \in \mathbb{R}^2 \mid \mathsf{E}(x, 0) \leq \mathsf{E}(x_0, v_0)\}.$$

A subset of C is invariant if any trajectory initiating in the subset remains in it. But this is only true if the starting position $\bar{x}$ satisfies $\frac{\partial}{\partial x} V(\bar{x}) = 0$, because otherwise the resulting trajectory would experience a strictly non-zero $\dot{v}(0)$ and hence leave C. In other words, the largest invariant set inside C is $\{(x, 0) \in \mathbb{R}^2 \mid \mathsf{E}(x, 0) \leq \mathsf{E}(x_0, v_0), \frac{\partial}{\partial x} V(x) = 0\}$. But the local minimum point x_2 is the unique critical point in the sublevel set and, therefore,

$$\lim_{t \to +\infty} (x(t), v(t)) = (x_2, 0).$$

$\bullet$

15.5   Application #1: Linear and linearized systems

It is interesting to study the convergence properties of a linear system. Recall that a symmetric matrix is positive definite if all its eigenvalues are strictly positive.

Theorem 15.9 (Convergence of linear systems). *For a matrix $A \in \mathbb{R}^{n \times n}$, the following properties are equivalent:*

   (i) *each solution to the differential equation $\dot{x} = Ax$ satisfies $\lim_{t \to +\infty} x(t) = \mathbb{0}_n$,*

  (ii) *A is Hurwitz, i.e., all the eigenvalues of A have strictly-negative real parts, and*

 (iii) *for every positive-definite matrix Q, there exists a unique solution positive-definite matrix P to the so-called* Lyapunov matrix equation:

$$A^{\mathsf{T}}P + PA = -Q.$$

   Note: one can show that statement (iii) implies statement (i) using the Lyapunov Stability Criteria with quadratic Lyapunov function $V(x) = x^{\mathsf{T}}Px$, whose Lie derivative along the systems solutions is $\dot{V} = x^{\mathsf{T}}(A^{\mathsf{T}}P + PA)x = -x^{\mathsf{T}}Qx \leq 0$.

Next, we show a very useful way to apply linear stability methods to analyze the local stability of a nonlinear system.

The *linearization at the equilibrium point x^** of the dynamical system (X, f) is the linear dynamical system defined by the differential equation $\dot{y} = Ay$, where

$$A = \frac{\partial f}{\partial x}(x^*).$$

Theorem 15.10 (Convergence of nonlinear systems via linearization). *Consider a dynamical system (X, f) with an equilibrium point x^*, with twice differentiable vector field f, and with linearization A at x^*. The following statements hold:*

(i) the equilibrium point x^* is locally exponentially stable if all the eigenvalues of A have strictly-negative real parts; and

(ii) the equilibrium point x^* is unstable if at least one eigenvalue of A has strictly-positive real part.

Example 15.11 (Two coupled oscillators). *For $\theta \in \mathbb{R}$, consider the dynamical system (14.6) arising from two coupled oscillators:*

$$\dot{\theta} = f(\theta) = \omega - \sin(\theta).$$

If $\omega \in [0, 1[$, then there are two equilibrium points inside the range $\theta \in [0, 2\pi[$:

$$\theta_1^* = \arcsin(\omega) \in [0, \pi/2[, \qquad and \qquad \theta_2^* = \pi - \arcsin(\omega) \in \,]\pi/2, +\pi].$$

Moreover, for $\theta \in \mathbb{R}$, the 2π-periodic set of equilibria are $\{\theta_1^* + 2k\pi \mid k \in \mathbb{Z}\}$ and $\{\theta_2^* + 2k\pi \mid k \in \mathbb{Z}\}$. The linearization matrix $A(\theta_i^*) = \frac{\partial f}{\partial \theta}(\theta_i^*) = -\cos(\theta_i^*)$ for $i \in \{1, 2\}$ shows that θ_1^* is locally exponentially stable and θ_2^* is unstable. •

Example 15.12 (A third order scalar system). *Pick a scalar c and, for $x \in \mathbb{R}$, consider the dynamical system*

$$\dot{x} = f(x) = c \cdot x^3.$$

The linearization at the equilibrium $x^* = 0$ is indefinite: $A(x^*) = 0$. Thus, Theorem 15.10 offers no conclusions other than the equilibrium cannot be exponentially stable. On the other hand, the Krasovskiĭ-LaSalle Invariance Principle shows that for $c < 0$ every trajectory converges to $x^* = 0$. Here, a non-increasing and differentiable function is given by $V(x) = x^2$ with Lie derivative $\mathscr{L}_f V(x) = -2cx^4 \leq 0$. Since $V(x(t))$ is non-increasing along the solution to the dynamical system, a compact invariant set is then readily given by any sublevel set $\{x \mid V(x) \leq \ell\}$ for $\ell \geq 0$. •

15.6 Application #2: Positive linear systems and Metzler matrices

In this short section we study the positive linear system $\dot{x} = Mx$, $x \in \mathbb{R}^n_{\geq 0}$, with equilibrium point $\mathbb{0}_n$, and with matrix M being Metzler matrix.

The following corollary is based upon Theorem 10.14 and illustrates how each of the conditions (iv), (v), and (vi) corresponds to a Lyapunov function of a specific form for a Hurwitz Metzler system.

Corollary 15.13 (Lyapunov functions for positive linear systems). *Let M be a Hurwitz Metzler matrix. The positive linear system $\dot{x} = Mx$, $x \in \mathbb{R}^n_{\geq 0}$, with equilibrium point $\mathbb{0}_n$, admits the following global Lyapunov functions:*

$$V_1(x) = \max_{i \in \{1, \dots, n\}} x_i / \xi_i, \qquad \textit{for } \xi > \mathbb{0}_n \textit{ satisfying } M\xi < \mathbb{0}_n,$$
$$V_2(x) = \eta^\mathsf{T} x, \qquad \textit{for } \eta > \mathbb{0}_n \textit{ satisfying } \eta^\mathsf{T} M < \mathbb{0}_n, \textit{ and}$$
$$V_3(x) = x^\mathsf{T} P x, \qquad \textit{for a diagonal matrix } P \succ 0 \textit{ satisfying } M^\mathsf{T} P + P M \prec 0.$$

We illustrate the level sets of these three global Lyapunov functions in Figure 15.7.

(a) Rectangular level set for global Lyapunov function V_1

(b) Linear level set for global Lyapunov function V_2

(c) Quadratic level set for global Lyapunov function V_3
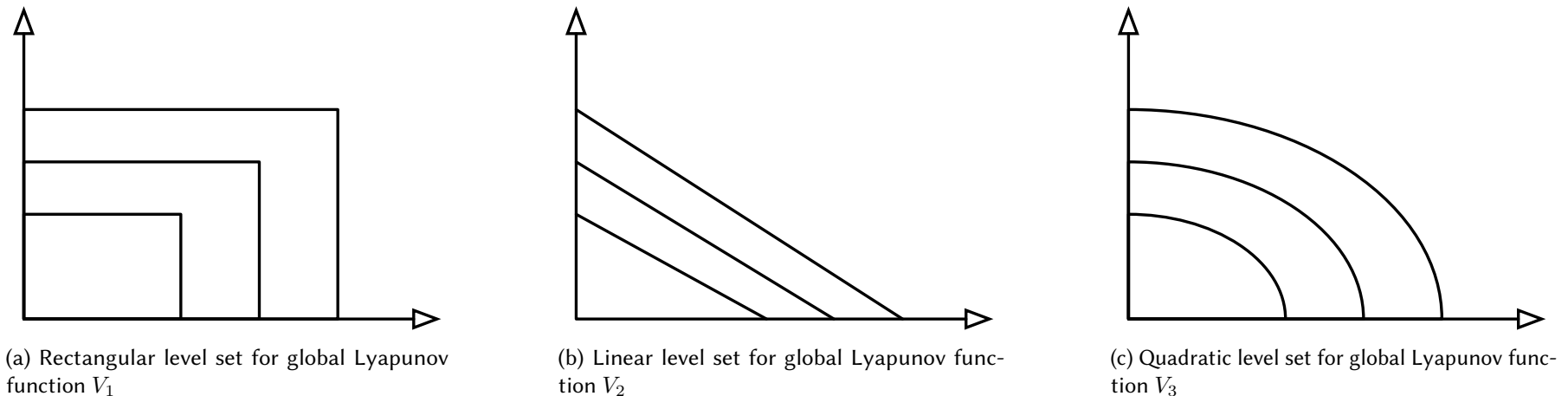
Figure 15.7: Level sets of global Lyapunov functions for Hurwitz positive linear systems, as established in Corollary 15.13

15.7 Application #3: Negative gradient systems

We now summarize and extend the analysis given in Example 15.3 of the stability properties of negative gradient systems. Recall for convenience that, given a differentiable function $V \colon \mathbb{R}^n \to \mathbb{R}$, the *negative gradient flow* defined by V is the dynamical system

$$\dot{x}(t) = -\frac{\partial V}{\partial x}(x(t)). \tag{15.6}$$

We start by noting that, as in the Exercise, the Lie derivative of V along the negative gradient flow is

$$\mathscr{L}_{-\frac{\partial V}{\partial x}} V(x) = -\left\| \frac{\partial V}{\partial x}(x) \right\|^2 \le 0,$$

and that, therefore, each sublevel set $V_{\le}^{-1}(\ell)$, for $\ell \in \mathbb{R}$ is invariant (provided it is non-empty).

Given a twice differentiable function $V \colon \mathbb{R}^n \to \mathbb{R}$ and a point $x \in \mathbb{R}^n$, the *Hessian matrix of V*, denoted by $\operatorname{Hess} V(x) \in \mathbb{R}^{n \times n}$, is the symmetric matrix of second order partial derivatives at x: $(\operatorname{Hess} V)_{ij}(x) = \partial^2 V / \partial x_i \partial x_j(s)$. Given a critical point x^* of V, if the Hessian matrix $\operatorname{Hess} V(x^*)$ is positive definite, then x^* is an isolated local minimum point of V. The converse is not true; as a counterexample, consider the function $V(x) = x^4$ and the critical point $x^* = 0$.

Theorem 15.14 (Convergence of negative gradient flow). *Let $V : \mathbb{R}^n \to \mathbb{R}$ be twice-differentiable and assume its sublevel set $V_{\leq}^{-1}(\ell) = \{x \in \mathbb{R}^n \mid V(x) \leq \ell\}$ is compact for some $\ell \in \mathbb{R}$. Then the negative gradient flow (15.6) has the following properties:*

(i) each solution $t \mapsto x(t)$ starting in $V_{\leq}^{-1}(\ell)$ satisfies $\lim_{t \to +\infty} V(x(t)) = c$, for some $c \leq \ell$, and approaches the set of critical points of V:

$$\left\{ x \in \mathbb{R}^n \mid \frac{\partial V}{\partial x}(x) = \mathbb{0}_n \right\},$$

(ii) each local minimum point x^* is locally asymptotically stable and it is locally exponentially stable if and only if $\operatorname{Hess} V(x^*)$ is positive definite,

(iii) a critical point x^* is unstable if at least one eigenvalue of $\operatorname{Hess} V(x^*)$ is strictly negative,

(iv) if the function V is analytic, then every solution starting in a compact sublevel set has finite length (as a curve in $\mathbb{R}^n$) and converges to a single equilibrium point.

Proof. To show statement (i), we verify that the assumptions of the Krasovskiĭ-LaSalle Invariance Principle are satisfied as follows. First, as set W we adopt the sublevel set $V_{\leq}^{-1}(\ell)$ which is compact by assumption and is invariant. Second we know the Lie derivative of V along the vector field is non-positive. Statement (i) is now an immediate consequence of the Krasovskiĭ-LaSalle Invariance Principle.

The statements (ii) and (iii) follow from observing that the linearization of the negative gradient system at the equilibrium x^* is the negative Hessian matrix evaluated at x^* and from applying Theorem 15.10.

Regarding statement (iv), we refer to the original source (Łojasiewicz, 1984) and to the review in (Absil et al., 2005, Section 2). ∎

Note: If the function V has isolated critical points, then the negative gradient flow evolving in a compact set must converge to a single critical point.

15.8   Application #4: Continuous-time averaging systems and Laplacian matrices

In this section we revisit the continuous-time averaging system, i.e., the Laplacian flow,

$$\dot{x} = -Lx.$$

As in Section 12.5, we define the *max-min function* $V_{\text{max-min}} \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ by

$$V_{\text{max-min}}(x) = \max_{i \in \{1,\ldots,n\}} x_i - \min_{i \in \{1,\ldots,n\}} x_i,$$

and that $V_{\text{max-min}}(x) \geq 0$, and $V_{\text{max-min}}(x) = 0$ if an only if $x = \alpha \mathbb{1}_n$ for some $\alpha \in \mathbb{R}$.

Lemma 15.15 (The max-min function along the Laplacian flow). *Let $L \in \mathbb{R}^{n \times n}$ be the Laplacian matrix of a weighted digraph G. Let $x(t)$ be the solution to the Laplacian flow $\dot{x} = -Lx$. Then*
(i) $t \mapsto V_{\text{max-min}}(x(t))$ is non-increasing,
(ii) if G has a globally reachable node, then, for some $\alpha \in \mathbb{R}$,

$$\lim_{t \to \infty} V_{\text{max-min}}(x(t)) = 0 \quad \text{and} \quad \lim_{t \to \infty} x(t) = \alpha \mathbb{1}_n.$$

Numerous proofs for these results are possible (e.g., statement (ii) is established in Theorem 7.4). A second approach is to use the properties of the row-stochastic matrix $\exp(-Lt)$, $t \in \mathbb{R}_{\geq 0}$, as established in Theorem 7.2.

Here we pursue a strategy based on adopting $V_{\text{max-min}}$ as a weak Lyapunov function and, because $V_{\text{max-min}}$ is not continuously-differentiable, applying an appropriate generalization of the Krasovskiĭ-LaSalle Invariance Principle in Theorem 15.7. For our purposes here, it suffices to present the following concepts.

Definition 15.16. *The upper right Dini derivative and upper left Dini derivative of a continuous function $f :]a, b[\to \mathbb{R}$ at a point $t \in]a, b[$ are defined by, respectively,*

$$D^+ f(t) = \limsup_{\Delta t > 0, \Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}, \quad \text{and} \quad D^- f(t) = \limsup_{\Delta t < 0, \Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}.$$

Recall that the *limit superior* of a real sequence $\{a_n\}_{n \in \mathbb{N}}$ is defined by

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} \sup_{m \geq n} a_m.$$

Similarly, the limit superior of a function $f :]0, \varepsilon[\to \mathbb{R}$ is defined by

$$\limsup_{h \to 0, h > 0} f(h) = \inf_{h > 0} \sup_{x \in]0, \varepsilon[} f(x). \tag{15.7}$$

Since the sup operator is always well defined (possibly equal to $+\infty$), so are the Dini derivatives.

Lemma 15.17 (Properties of the upper Dini derivatives). *Given a continuous function $f :]a, b[\to \mathbb{R}$,*

(i) if f is differentiable at $t \in]a, b[$, then $D^+ f(t) = D^- f(t) = \frac{d}{dt} f(t)$ is the usual derivative of f at t, and

(ii) if $D^+ f(t) \leq 0$ or $D^- f(t) \leq 0$ for all $t \in]a, b[$, then f is non-increasing on $]a, b[$.

Morever, given differentiable functions $f_1, \ldots, f_m :]a, b[\to \mathbb{R}$, the max function $f_{\max}(t) = \max\{f_i(t) \mid i \in \{1, \ldots, m\}\}$ satisfies

(iii) $D^+ f_{\max}(t) = \max \left\{ \dfrac{d}{dt} f_i(t) \;\middle|\; i \in \mathrm{argmax}(f_{\max}(t)) \right\}$,

$\qquad D^- f_{\max}(t) = \min \left\{ \dfrac{d}{dt} f_i(t) \;\middle|\; i \in \mathrm{argmax}(f_{\max}(t)) \right\}$, *and*

(iv) if $D^+ f_{\max}(t) \leq 0$ for all $t \in]a, b[$, then $f_{\max}$ is non-increasing on $]a, b[$.

Note: statement (i) follows from the definition of derivative of a differentiable function. Statement (ii) is a consequence of Theorem 1.14 in (Giorgi and Komlósi, 1992), to which we refer for all proofs. Statement (iii) is known as Danskin's Lemma.

Proof of Lemma 15.15. Define the quantities $x_{\max}(t) = \max(x(t))$ and $x_{\min}(t) = \min(x(t))$ as well as $\operatorname{argmax}(x(t)) = \{i \in \{1, \ldots, n\} \mid x_i(t) = x_{\max}(t)\}$ and $\operatorname{argmin}(x(t)) = \{i \in \{1, \ldots, n\} \mid x_i(t) = x_{\min}(t)\}$. Along the Laplacian flow $\dot{x}_i = \sum_{j=1}^{n} a_{ij}(x_j - x_i)$, Lemma 15.17(iii) (Danskin's Lemma) implies

$$D^+ V_{\text{max-min}}(x(t)) = \max\{\dot{x}_i(t) \mid i \in \operatorname{argmax}(x(t))\} - \min\{\dot{x}_i(t) \mid i \in \operatorname{argmin}(x(t))\}$$

$$= \max\Big\{ \sum_{j=1}^{n} a_{ij}(x_j - x_{\max}) \mid i \in \operatorname{argmax}(x(t))\Big\}$$

$$- \min\Big\{ \sum_{j=1}^{n} a_{ij}(x_j - x_{\min}) \mid i \in \operatorname{argmin}(x(t))\Big\},$$

where we have used $-\min(x) = \max(-x)$. Because $x_j - x_{\max} \leq 0$ and $x_j - x_{\min} \geq 0$ for all $j \in \{1, \ldots, n\}$, we have established that $D^+ V_{\text{max-min}}(x(t))$ is the sum of two non-positive terms. This property, combined with Lemma 15.17(iv), implies that $t \mapsto V_{\text{max-min}}(x(t))$ is non-increasing, thereby completing the proof of statement (i).

To establish statement (ii) we invoke a generalized version of the Krasovskiĭ-LaSalle Invariance Principle 15.7. First, we note that statement (i) implies that any solution is bounded inside $[x_{\min}(0), x_{\max}(0)]^n$; this is a sufficient property (in lieu of the compactness of the set W). Second, we know the continuous function $V_{\text{max-min}}$ along the Laplacian flow is non-increasing (in lieu of the same property for a Lie derivative of a continuously-differentiable function). Therefore, we now know that there exists c such that the solution starting from $x(0)$ converges to the largest invariant set C contained in

$$\big\{x \in [x_{\min}(0), x_{\max}(0)]^n \mid D^+ V_{\text{max-min}}(x)\big|_{\dot{x}=-Lx} = 0\big\} \cap V_{\text{max-min}}^{-1}(c).$$

Because $V_{\text{max-min}}$ is non-negative, we know $c \geq 0$. We now assume by absurd that $c > 0$, we let $y(t)$ be a trajectory originating in C, and we aim to show that $V_{\text{max-min}}(y(t))$ decreases along time (which is a contradiction because C is invariant).

Let k be a globally reachable node. Let i (resp. j) be an arbitrary index in $\operatorname{argmax}(y(0))$ (resp. $\operatorname{argmin}(y(0))$) so that $y_i(0) - y_j(0) = c > 0$. Without loss of generality we assume $y_k(0) < y_i(0)$. (Otherwise it would need to be $y_k(0) > y_j(0)$ and we would proceed similarly.) Recall we know $\dot{y}_i(0) \leq 0$. We now note that, if $\dot{y}_i(t) = 0$ for all $t \in (0, \varepsilon)$ for a positive ε, then the equation $\dot{y}_i = \sum_j a_{ij}(y_j - y_i)$ and the property $y_i(0) = \max y(0)$ together imply that $y_j(t) = y_i(t)$ for all $t \in (0, \varepsilon)$ and for all j such that

$a_{ij} > 0$. Iterating this argument along the directed walk from i to k, we get the contradiction that $y_k(t) = y_i(t)$ for all $t \in (0, \varepsilon)$. Therefore, we know that $\dot{y}_i(t) < 0$ for small times. Because i is an arbitrary index in $\mathrm{argmax}(y(0))$, we have proved that $t \mapsto \max y(t)$ is strictly decreasing for small times. This establishes that C is not invariant if $c > 0$ and completes the proof of statement (ii). ∎

15.9   Application #5: Interconnected stable systems

We consider the *interconnection of n dynamical systems*

$$\dot{x}_i = f_i(t, x_i, x_{-i}), \quad \text{for } i \in \{1, \dots, n\}, \tag{15.8}$$

where $x_i \in \mathbb{R}^{N_i}$, $N = \sum_{i=1}^{n} N_i$, and $x_{-i} \in \mathbb{R}^{N-N_i}$. We assume the origin is an equilibrium, that is,

$$f_i(t, \mathbb{0}_{N_i}, \mathbb{0}_{N-N_i}) = \mathbb{0}_{N_i}, \quad \text{for all } t. \tag{15.9}$$

We assume that there exist positive definite and differentible functions $V_i(t, x_i)$ an positive definite functions $\phi_i(x_i)$, for $i \in \{1, \dots, n\}$, such that, for all t, $x_i \in \mathbb{R}^{N_i}$ and $x_{-i} \in \mathbb{R}^{N-N_i}$,

(A1) $\dfrac{\partial V_i}{\partial t}(t, x_i) + \dfrac{\partial V_i}{\partial x_i} f_i(t, x_i, \mathbb{0}_{N-N_i}) \leq -c_i \phi_i(x_i)^2$, for some $c_i > 0$,

(A2) $\left\| \dfrac{\partial V_i}{\partial x_i}(t, x_i) \right\| \leq \phi_i(x_i)$ and $\| f_i(t, x_i, x_{-i}) - f_i(t, x_i, \mathbb{0}_{N-N_i}) \| \leq \displaystyle\sum_{j=1, j \neq i}^{n} \gamma_{ij} \phi_j(x_j)$ for some $\gamma_{ij} \geq 0$.

Note that Assumption (A1) is a Lie derivative assumption and Assumption (A2) contains two Lipschitz boundedness conditions. Next, define the *gain matrix*

$$\Gamma = \begin{bmatrix} -c_1 & \cdots & \gamma_{1n} \\ \vdots & & \vdots \\ \gamma_{n1} & \cdots & -c_n \end{bmatrix}.$$

Theorem 15.18 (Stability of interconnected stable system). *Consider the interconnected system* (15.8) *satisfying the equilibrium condition* (15.9), *as well as Assumptions* (A1) *and* (A2)*. If the Metzler gain matrix* Γ *is Hurwitz, then*

(i) $V(t, x) = \sum_{i=1}^{n} p_i V_i(t, x_i)$ *is a local Lyapunov function for any* $\mathrm{diag}(p) \succ 0$ *satisfying* $\mathrm{diag}(p)\Gamma + \Gamma^{\mathsf{T}} \mathrm{diag}(p) \prec 0,$

(ii) *the origin is locally asymptotically stable, and*

(iii) *if additionally each* V_i *is proper, then* V *is a global Lyapunov function and the origin is globally asymptotically stable.*

Proof. Regarding statement (i), we start by noting that such diagonal matrices exist because Γ is Hurwitz Metzler and Theorem 10.14(vi) applies. Define $\phi(x) = \begin{bmatrix} \phi_1(x_1) & \dots & \phi_n(x_n) \end{bmatrix}^\mathsf{T}$ and compute the Lie derivative of the candidate Lyapunov function $V(t,x) = \sum_{i=1}^n p_i V_i(t,x_i)$ as follows:

$$
\begin{aligned}
\frac{d}{dt}V(t,x) &= \sum_{i=1}^n p_i \left(\frac{\partial V_i}{\partial t}(t,x_i) + \frac{\partial V_i}{\partial x_i} f_i(t,x_i,x_{-i}) \right) \\
&= \sum_{i=1}^n p_i \left(\frac{\partial V_i}{\partial t}(t,x_i) + \frac{\partial V_i}{\partial x_i} f_i(t,x_i,\mathbb{0}_{N-N_i}) \right) \\
&\quad + \sum_{i=1}^n p_i \frac{\partial V_i}{\partial x_i} \left(f_i(t,x_i,x_{-i}) - f_i(t,x_i,\mathbb{0}_{N-N_i}) \right) \\
&\leq -\sum_{i=1}^n p_i c_i \phi_i(x_i)^2 + \sum_{i=1}^n p_i \left\| \frac{\partial V_i}{\partial x_i} \right\| \left\| f_i(t,x_i,x_{-i}) - f_i(t,x_i,\mathbb{0}_{N-N_i}) \right\| \quad (15.10) \\
&\leq -\sum_{i=1}^n p_i c_i \phi_i(x_i)^2 + \sum_{i=1}^n \sum_{j=1,j\neq i}^n p_i \gamma_{ij} \phi_i(x_i)\phi_j(x_j) \quad (15.11) \\
&= \tfrac{1}{2}\phi(x)^\mathsf{T} \left(\Gamma^\mathsf{T} P + P\Gamma \right) \phi(x) < 0,
\end{aligned}
$$

for all x such that $\phi(x) \neq \mathbb{0}_n$. In this derivation, inequality (15.10) follows from Assumption (A1) and inequality (15.11) follows from Assumption (A2). Statement (i)-(iii) now follow directly from the Lyapunov Stability Criteria in Theorem 15.4. ∎

Remark 15.19 (Conservativeness of stability test). *Assumptions* (A1)-(A2) *are conservative because any interconnection is treated as a disturbance and its effect is upper bounded in a conservative way. For example, pick* $\varepsilon > 0$ *and define*

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\varepsilon & 1 \\ -1 & -\varepsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =: A_\varepsilon \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{15.12}$$

The matrix A_ε *has eigenvalues* $-\varepsilon \pm i$ *and so it is Hurwitz for all* $\varepsilon > 0$. *Treating the system as the interconnection of two stable subsystems leads to the gain matrix* $\begin{bmatrix} -\varepsilon & 1 \\ 1 & -\varepsilon \end{bmatrix}$ *with eigenvalues* $-\varepsilon \pm 1$, *which fails to be Hurwitz whenever* $\varepsilon < 1$.   ●

15.10   Historical notes and further reading

Classic historical works on stability properties of physical systems include (Lagrange, 1788; Maxwell, 1868; Thomson and Tait, 1867). Modern stability theory started with the work by Lyapunov (1892), who proposed the key ideas towards a general treatment of stability notions and tests for nonlinear dynamical systems. Lyapunov's ideas were extended by Barbashin and Krasovskiĭ (1952); Krasovskiĭ (1963) and LaSalle (1960, 1968, 1976) through their work on invariance principles. Other influential works include (Chetaev, 1961; Hahn, 1967).

For comprehensive treatments, we refer the reader to the numerous excellent texts in this area, e.g., including the classic control texts (Sontag, 1998; Khalil, 2002; Vidyasagar, 2002), the classic dynamical systems texts (Hirsch and Smale, 1974; Arnol'd, 1992; Guckenheimer and Holmes, 1990), and the more recent works (Haddad and Chellaboina, 2008; Goebel et al., 2012; Blanchini and Miani, 2015).

This chapter has treated systems evolving in continuous time. Naturally, it is possible to develop a Lyapunov theory for discrete-time systems, even though remarkably there are only few references on this topic; see (LaSalle, 1976, Chapter 1). For discrete-time Lyapunov functions for averaging systems we refer to Section 5.3.

Our treatment of Metzler matrices in Section 15.6 is standard, a recent discussion is given by (Rantzer, 2015).

We refer to (Clarke et al., 1998; Cortés, 2008) for a comprehensive review of stability theory for nonsmooth systems and Lyapunov functions. Properties of the Dini derivatives are reviewed by Giorgi and Komlósi (1992). The usefulness of Dini derivatives in continuous-time averaging systems is highlighted for example by Lin et al. (2007); see also (Danskin, 1966) for Danskin's Lemma.

The treatment of interconnected stable systems in Section 15.9 originates in the method of vector Lyapunov functions developed by (Bellman, 1962; Matrosov, 1962). Classic references include are (Šiljak, 1978, 1991; Lakshmikantham et al., 1991).

15.11 Exercises

E15.1   **The logarithmic-linear function.** For $\kappa > 0$, define the function $V_{\text{log-lin},\kappa} : \mathbb{R}_{>0} \to \mathbb{R}$ by

$$V_{\text{log-lin},\kappa}(x) = x - \kappa - \kappa \log\left(\frac{x}{\kappa}\right).$$

Show that

(i) $V_{\text{log-lin},\kappa}$ is continuously differentiable and $\frac{d}{dx} V_{\text{log-lin},\kappa}(x) = (x - \kappa)/x$,

(ii) $V_{\text{log-lin},\kappa}(x) \geq 0$ for all $x > 0$ and $V_{\text{log-lin},\kappa}(x) = 0$ if and only if $x = \kappa$, and

(iii) $\lim_{x \to 0^+} V_{\text{log-lin},\kappa}(x) = \lim_{x \to \infty} V_{\text{log-lin},\kappa}(x) = +\infty$.

E15.2   **Grönwall-Bellman Comparison Lemma.** Given continuous functions of time $t \mapsto a(t) \in \mathbb{R}$ and $t \mapsto gamma(t) \in \mathbb{R}$, assume the differentiable signal $t \mapsto x(t)$ satisfies the differential inequality

$$\dot{z}(t) \le a(t)z(t) + \gamma(t).$$

Show that, for all $0 \le s \le t \in \mathbb{R}_{\ge 0}$, with the shorthand $A(s,t) = \int_s^t a(\tau)d\tau$,

$$z(t) \le e^{A(0,t)} z(0) + \int_0^t e^{A(\tau,t)} \gamma(\tau)d\tau. \tag{E15.1}$$

In other words, $z(t)$ is upper bounded by the solution to the corresponding differential equality.

E15.3   **The negative gradient flow of a strictly convex function.** Recall that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *convex* if $f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$ for all $x \neq y$ in $\mathbb{R}^n$ and for all $\alpha, \beta \geq 0$ satisfying $\alpha + \beta = 1$. A function is *strictly convex* if the previous inequality holds strictly.

   Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be strictly convex and twice differentiable. Show global convergence of the associated negative gradient flow, $\dot{x} = -\frac{\partial}{\partial x} f(x)$, to the global minimizer x^* of f using the Lyapunov function candidate $V(x) = (x - x^*)^\mathsf{T}(x - x^*)$ and the Krasovskiĭ-LaSalle Invariance Principle in Theorem 15.7.

   Hint: Use the *global underestimate property* of a strictly convex function stated as follows: $f(y) - f(x) > \frac{\partial}{\partial x} f(x)(y - x)$ for all distinct x and y in the domain of f.

E15.4   **Region of attraction for an example nonlinear systems.** Consider the nonlinear system

$$\dot{x}_1 = -2x_1 - 2x_2 - 4x_1^3 x_2^2,$$
$$\dot{x}_2 = -2x_1 - 2x_2 - 2x_1^4 x_2.$$

Is the origin locally asymptotically stable? What is the region of attraction?

E15.5   **A useful corollary by Barbashin and Krasovskiĭ (1952).** Consider a dynamical system $(\mathbb{R}^n, f)$ with differentiable vector field f and with an equilibrium point $x^* \in \mathbb{R}^n$.

Assume the continuously-differentiable $V: \mathbb{R}^n \to \mathbb{R}$ is a weak Lyapunov function, but not a local Lyapunov function (as defined in Theorem 15.4). In other words, assume V is locally positive-definite about x^* (Assumption (L1)) and $\mathscr{L}_f V$ is locally negative-semidefinite about x^* (Assumption (L2)), but $\mathscr{L}_f V$ is not locally negative-definite about x^* (Assumption (L3)). Then Lyapunov Theorem 15.4 implies that x^* is stable but not necessarily locally asymptotically stable.

Now, assume:

(L7) $\{x^*\}$ is the only positively invariant set in $\{x \in W \mid \mathscr{L}_f V(x) = 0\}$, where W be a neighborhood of x^* on which V is positive-definite and $\mathscr{L}_f V$ is negative-semidefinite.

Prove that Assumptions (L1), (L2) and (L7) imply the equilibrium point x^* is locally asymptotically stable.

E15.6 **Limit sets of dynamical systems.** Consider the following nonlinear dynamical system

$$\dot{x}_1 = 4x_1^2 x_2 - f_1(x_1)(x_1^2 + 2x_2^2 - 4),$$ (E15.2a)
$$\dot{x}_2 = -2x_1^3 - f_2(x_2)(x_1^2 + 2x_2^2 - 4),$$ (E15.2b)

where the differentiable functions $f_1(x)$, $f_2(x)$ have the same sign as their arguments, i.e., $x_i f_i(x_i) > 0$ if $x_i \neq 0$, $f_i(0) = 0$, and $f_i'(0) > 0$. This vector field exhibit some very unconventional limit sets. In what follows you will investigate this vector field and show that each trajectory converge to an equilibrium, but that none of the equilibria is Lyapunov stable.

(i) Show that $\mathcal{E} = \{x \in \mathbb{R}^2 \mid x_1^2 + 2x_2^2 = 4\}$ is an invariant set. Calculate the equilibria on the set $\mathcal{E}$.
(ii) Show that all trajectories converge either to the invariant set $\mathcal{E}$ or to the origin $(0, 0)$.
(iii) Determine the largest invariant set inside $\mathcal{E}$, such that all trajectories originating in $\mathcal{E}$ converge to that set.
(iv) Show that the origin $(0, 0)$ and all equilibria on $\mathcal{E}$ are unstable, i.e., not stable in the sense of Lyapunov. Sketch the vector field.

E15.7   **An invariant triangle.** Consider the dynamical system

$$\dot{x}_1 = -x_2 + x_1 x_2,$$

$$\dot{x}_2 = x_1 + \frac{1}{2}(x_1^2 - x_2^2).$$

Show that

(i)   the equilibrium points $A = (-2, 0)$, $B = (1, -\sqrt{3})$, and $C = (1, \sqrt{3})$ are unstable (note that A, B, and C are not the only equilibrium points);
(ii) the triangle $\mathcal{D}$ defined by three points A, B, C is positively invariant, that is, trajectories starting in $\mathcal{D}$ do not leave $\mathcal{D}$; and
      Hint: Show that trajectory cannot leave $\mathcal{D}$ through any of the line segments AB, BC, and CA.
(iii) the energy function

$$V(x_1, x_2) = -\frac{1}{2}(x_1^2 + x_2^2) + \frac{1}{2}\left(x_1 x_2^2 - \frac{1}{3}x_1^3\right)$$

is conserved along the trajectories of the dynamical system.

E15.8   **The continuous-time Hopfield neural network (Hopfield, 1982).** Consider the additive RC model of a neuron:

$$C_i \dot{x}_i + \frac{1}{R_i} x_i = I_i + \sum_{j=1}^{n} w_{ij} y_j, \tag{E15.3}$$

where the neuron internal voltage is x_i and the neuron output voltage is

$$y_i = g(x_i). \tag{E15.4}$$

Here (i) w_{ij} are conductances of the neural network; we assume the network is undirected and weighted, and (ii) as monotonically-increasing activation function g, we adopt $g(x) = \tanh(x/2)$ and note its monotonically-increasing inverse $g^{-1}(y) = -\log((1-y)/(1+y))$.
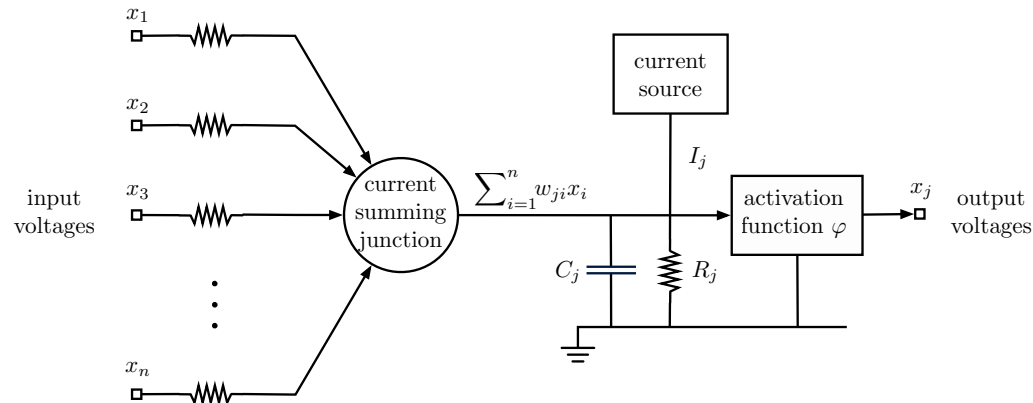
Figure E15.1: Additive model of neuron with current summing junction.

Define the *energy function* of the network by

$$E(x_1, \ldots, x_n) = -\frac{1}{2} \sum_{i,j=1}^{n} w_{ij} y_i y_j + \sum_{i=1}^{n} \left(\frac{1}{R_i} \int_0^{y_i} g^{-1}(v) dv - I_i x_i \right), \tag{E15.5}$$

and perform the following steps:

(i) discuss the energy landscape with numerous minima,
(ii) characterize as best as possible the equilibria of the system,
(iii) will the trajectories of this system either converge to one of the equilibria, or oscillate or behave in some other way?

Note: *(Haykin, 2008, Chapter 13) reviews numerous neurodynamics results; e.g., see Section 13.7 on the Hopfield network. A nice discussion can also be found at* `http://www.scholarpedia.org/article/Hopfield_network`

E15.9 **Persidskii models (Persidskii, 1969; Kaszkurewicz and Bhaya, 1993).** Given a matrix $A \in \mathbb{R}^{n \times n}$ and a diagonal continuous map $\Phi \colon \mathbb{R}^n \to \mathbb{R}^n$, define the *Persidskii dynamical system*

$$\dot{x} = A\Phi(x) =: f_{\text{Per}}(x). \qquad (\text{E}15.6)$$

By diagonal map we mean that $\Phi(x) = \begin{bmatrix} \phi_1(x_1) & \cdots & \phi_n(x_n) \end{bmatrix}^{\mathsf{T}}$. Assume there exists a positive vector $p \in \mathbb{R}^n_{>0}$ such that A satisfies the stability LMI

$$\operatorname{diag}(p)A + A^{\mathsf{T}}\operatorname{diag}(p) \prec 0.$$

Such a matrix A is said to be *diagonally stable*. Define the function $V \colon \mathbb{R}^n \to \mathbb{R}$ by

$$V(x) = \sum_{i=1}^n p_i \int_0^{x_i} \phi_i(z)dz. \qquad (\text{E}15.7)$$

Show that

(i) V is continuously differentiable and satisfies
$$\mathscr{L}_{f_{\text{Per}}}V(x) = \tfrac{1}{2}\Phi(x)^{\mathsf{T}}\big(\operatorname{diag}(p)A + A^{\mathsf{T}}\operatorname{diag}(p)\big)\Phi(x); \qquad (\text{E}15.8)$$

(ii) if $\phi_i(0) = 0$ and $\phi_i(y)y > 0$ for each $i \in \{1, \ldots, n\}$ and for all $y \in \mathbb{R} \setminus \{0\}$, then V is globally positive-definite and $\mathscr{L}_{f_{\text{Per}}}V(x)$ is globally negative-definite about $\mathbb{0}_n$. The equilibrium $\mathbb{0}_n$ is therefore locally asymptotically stable for f_{Per}, and

(iii) if additionally $\int_0^{x_i} \phi_i(z)dz \to \infty$ as $|x_i| \to \infty$ for all i, then V is proper and $\mathbb{0}_n$ is globally asymptotically stable for f_{Per}.

Chapter 16

Lotka-Volterra Population Dynamics

In this chapter we study the behavior of the Lotka-Volterra population model, that was introduced in Section 14.1. First we illustrate the behavior of the 2-dimensional model via simple phase portraits. Then, using Lyapunov stability theory from Chapter 14 we provide sufficient conditions for the general n-dimensional model to have a globally asymptotically stable point. As a special case, we study the case of cooperative models.

Recall that the Lotka-Volterra vector field for $n \geq 2$ interacting species, as given in equation (14.3), is

$$\dot{x} = \mathrm{diag}(x)\big(Ax + r\big) =: f_{\mathsf{LV}}(x), \tag{16.1}$$

where the matrix $A = [a_{ij}]$ is called the interaction matrix, and the vector r is called the intrinsic growth rate. In components, $\dot{x}_i = x_i \sum_{j=1}^{n}(a_{ij}x_j + r_i)$.

16.1   Two-species model and analysis

In this section we consider the *two-species Lotka-Volterra system*

$$
\begin{aligned}
\dot{x}_1 &= x_1(r_1 + a_{11}x_1 + a_{12}x_2), \\
\dot{x}_2 &= x_2(r_2 + a_{21}x_1 + a_{22}x_2),
\end{aligned}
\tag{16.2}
$$

with parameters (r_1, r_2) and $(a_{11}, a_{12}, a_{21}, a_{22})$. It is possible to fully characterize the dynamics behavior of this system as a function of the six scalar parameters. As explained in Section 14.1, to model bounded resources, our standing assumptions are:

$$
r_i > 0, \quad \text{and } a_{ii} < 0, \text{ for } i \in \{1, 2\}.
$$

We study various cases depending upon the sign of a_{12} and a_{21}.

To study the phase portrait of this two-dimensional system, we establish the following details:

(i) along the axis $x_2 = 0$, there exists a unique non-trivial equilibrium point $x_1^* = -r_1/a_{11}$;

(ii) similarly, along the axis $x_1 = 0$, there exists a unique non-trivial equilibrium point $x_2^* = -r_2/a_{22}$;

(iii) the x_1-*null-line* is the set of points (x_1, x_2) where $\dot{x}_1 = 0$, that is, the line in the (x_1, x_2) plane defined by $r_1 + a_{11}x_1 + a_{12}x_2 = 0$;

(iv) similarly, the x_2-*null-line* is the (x_1, x_2) plane defined by $r_2 + a_{21}x_1 + a_{22}x_2 = 0$.

Clearly, the x_1-null-line (respectively the x_2-null-line) passes through the equilibrium point x_1^* (respectively x_2^*).

In what follows we study the cases of mutualistic interactions and competitive interactions. We refer to Exercise E16.2 for a specially-interesting case of predator-prey interactions.

16.1.1   Mutualism

Here we assume inter-species mutualism, that is, we assume both inter-species coefficients a_{12} and a_{21} are positive. We identify two distinct parameter ranges corresponding to distinct dynamic behavior and illustrate them in Figure 16.1.

Lemma 16.1 (Two-species mutualism). *Consider the two-species Lotka-Volterra dynamical system* (16.2) *with parameters* (r_1, r_2) *and* $(a_{11}, a_{12}, a_{21}, a_{22})$. *Assume the interaction is mutualistic, i.e., assume* $a_{12} > 0$ *and* $a_{21} > 0$. *The following statements hold:*

Case I: if $a_{12}a_{21} < a_{11}a_{22}$, *then there exists a unique positive equilibrium point* (x_1^*, x_2^*), *solution to*

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \end{bmatrix},$$

and all trajectories starting in $\mathbb{R}_{>0}^2$ *converge to it;*

Case II: otherwise, if $a_{12}a_{21} > a_{11}a_{22}$, *then there exists no positive equilibrium point and all trajectories starting in* $\mathbb{R}_{>0}^2$ *diverge.*
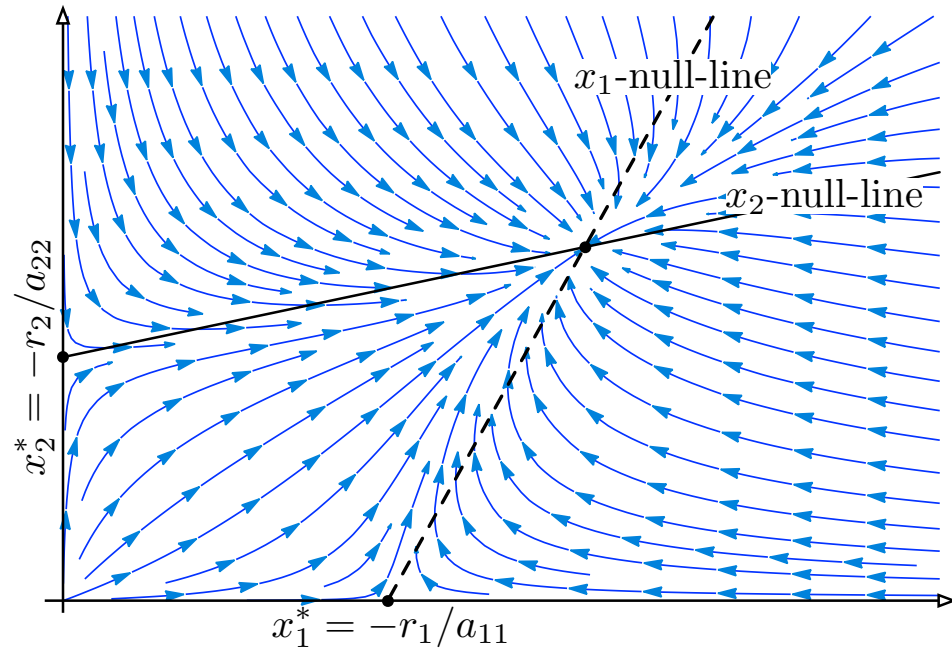
Case I: $a_{12} > 0$, $a_{21} > 0$, $a_{12}a_{21} < a_{11}a_{22}$. There exists a unique positive equilibrium point. All trajectories starting in $\mathbb{R}^2_{>0}$ converge to the equilibrium point.
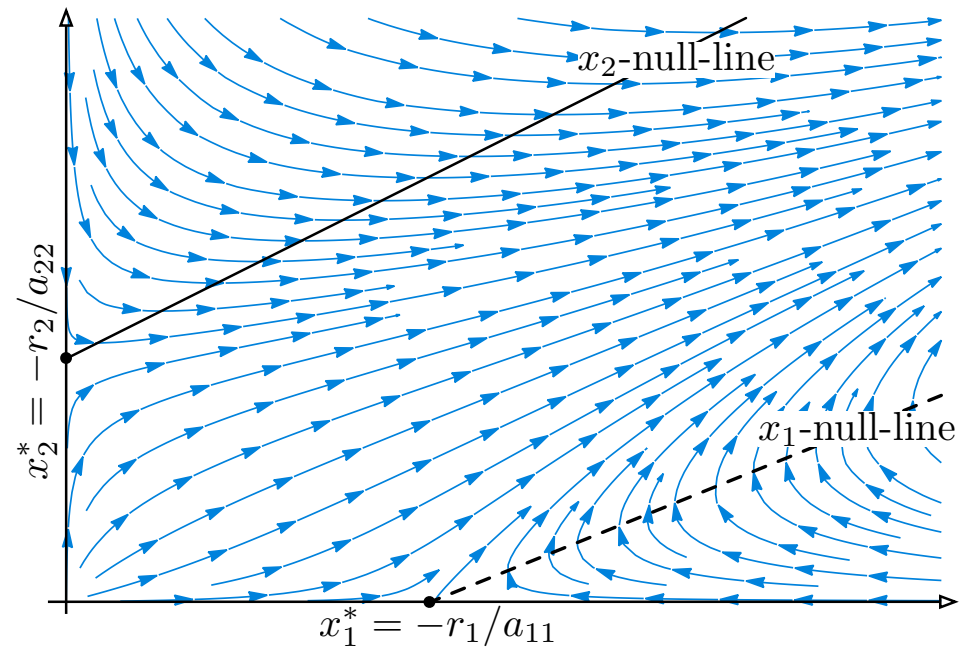
Case II: $a_{12} > 0$, $a_{21} > 0$, $a_{12}a_{21} > a_{11}a_{22}$. There exists no positive equilibrium point. All trajectories starting in $\mathbb{R}^2_{>0}$ diverge.

Figure 16.1: Two possible cases of mutualism in the two-species Lotka-Volterra system

16.1.2   Competition

Here we assume inter-species competition, that is, we assume both inter-species coefficients a_{12} and a_{21} are negative. We identify four (two sets of two) distinct parameter ranges corresponding to distinct dynamic behavior and illustrate them in Figures 16.2 and 16.3.

Lemma 16.2 (Two-species competition with a positive equilibrium). *Consider the two-species Lotka-Volterra system* (16.2) *with parameters* (r_1, r_2) *and* $(a_{11}, a_{12}, a_{21}, a_{22})$. *Assume the interaction is competitive, i.e., assume* $a_{12} < 0$ *and* $a_{21} < 0$. *The following statements hold:*

Case III: *if* $r_2/|a_{22}| < r_1/|a_{12}|$ *and* $r_1/|a_{11}| < r_2/|a_{21}|$, *then there exists a unique positive equilibrium, which attracts all trajectories starting in* $\mathbb{R}^2_{>0}$;

Case IV: *if* $r_1/|a_{12}| < r_2/|a_{22}|$ *and* $r_2/|a_{21}| < r_1/|a_{11}|$, *then the equilibrium in* $\mathbb{R}^2_{>0}$ *is unstable; all trajectories (except the equilibrium solution) converge either to the equilibrium* $(-r_1/a_{11}, 0)$ *or to the equilibrium* $(0, -r_2/a_{22})$.

As for Case I, for Cases III and IV, it is easy to compute the unique positive equilibrium point (x_1^*, x_2^*) as the solution to
$$
\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.
$$
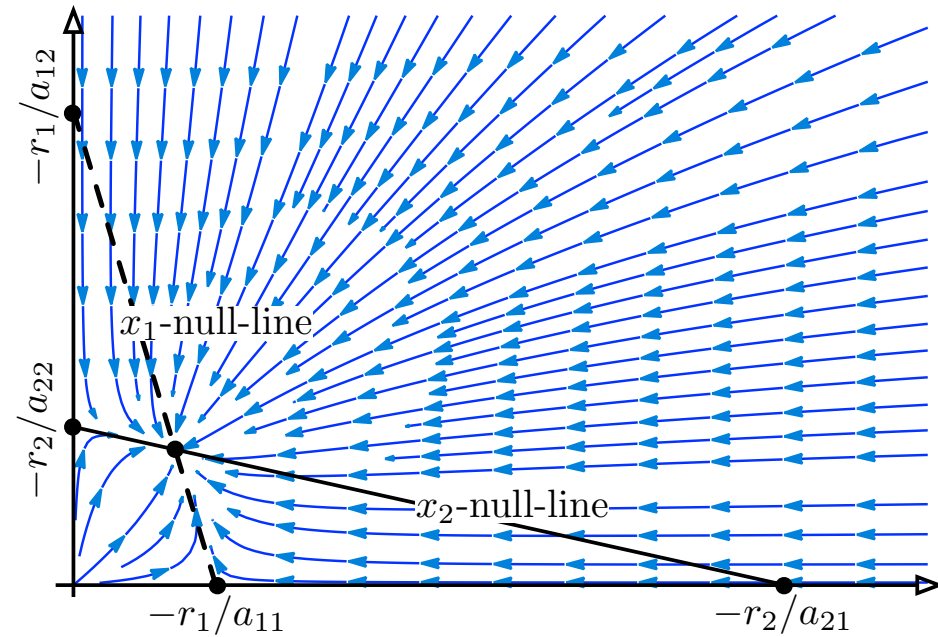
Case III: $a_{12} < 0$, $a_{21} < 0$, $r_2/|a_{22}| < r_1/|a_{12}|$, and $r_1/|a_{11}| < r_2/|a_{21}|$. There exists a unique positive equilibrium, which attracts all trajectories starting in $\mathbb{R}^2_{>0}$.
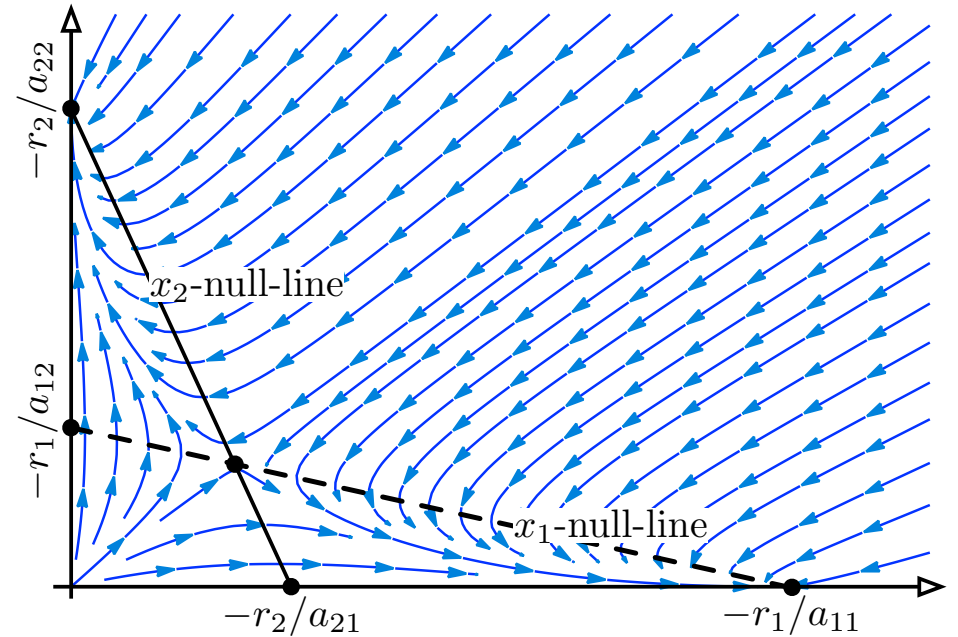
Case IV: $a_{12} < 0$, $a_{21} < 0$, $r_1/|a_{12}| < r_2/|a_{22}|$, and $r_2/|a_{21}| < r_1/|a_{11}|$. The equilibrium in $\mathbb{R}^2_{>0}$ is unstable; all trajectories (except the equilibrium solution) converge either to the equilibrium $(-r_1/a_{11}, 0)$ or to the equilibrium $(0, -r_2/a_{22})$.

Figure 16.2: Two competition cases with an equilibrium in the two-species Lotka-Volterra system

Lemma 16.3 (Two-species competition without positive equilibria). *Consider the two-species Lotka-Volterra system* (16.2) *with parameters* (r_1, r_2) *and* $(a_{11}, a_{12}, a_{21}, a_{22})$. *Assume the interaction is competitive, i.e., assume* $a_{12} < 0$ *and* $a_{21} < 0$. *The following statements hold:*

Case V: *if* $r_2/|a_{22}| < r_1/|a_{12}|$ *and* $r_2/|a_{21}| < r_1/|a_{11}|$, *then there exists no equilibrium in* $\mathbb{R}^2_{>0}$ *and all trajectories starting in* $\mathbb{R}^2_{>0}$ *converge to the equilibrium* $(-r_1/a_{11}, 0)$;

Case VI: *if* $r_1/|a_{12}| < r_2/|a_{22}|$ *and* $r_1/|a_{11}| < r_2/|a_{21}|$, *then there exists no equilibrium in* $\mathbb{R}^2_{>0}$ *and all trajectories starting in* $\mathbb{R}^2_{>0}$ *converge to the equilibrium* $(0, -r_2/a_{22})$.
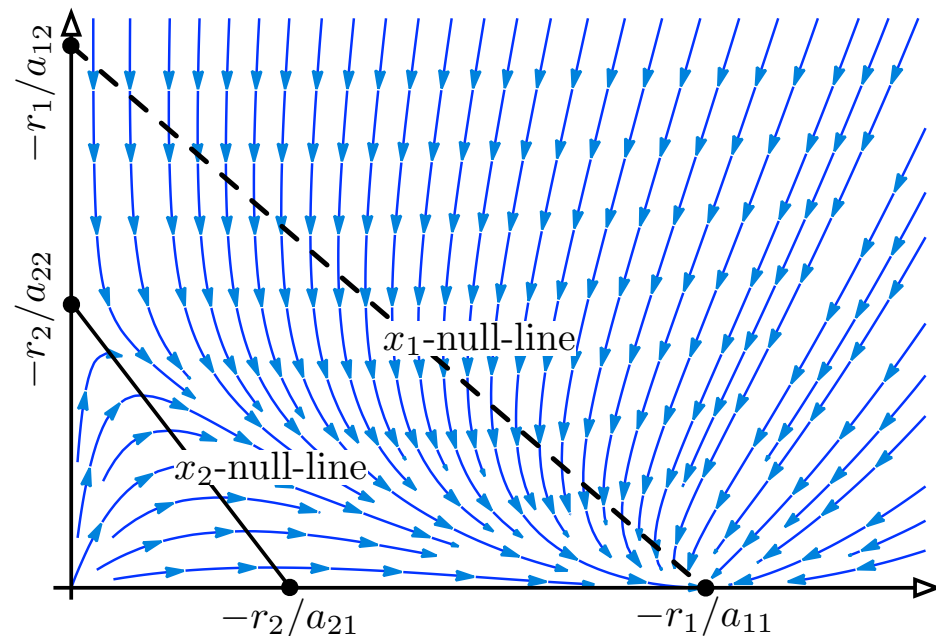
Case V: $a_{12} < 0$, $a_{21} < 0$, $r_2/|a_{22}| < r_1/|a_{12}|$, and $r_2/|a_{21}| < r_1/|a_{11}|$. There exists no equilibrium in $\mathbb{R}^2_{>0}$. All trajectories starting in $\mathbb{R}^2_{>0}$ converge to the equilibrium $(-r_1/a_{11}, 0)$.
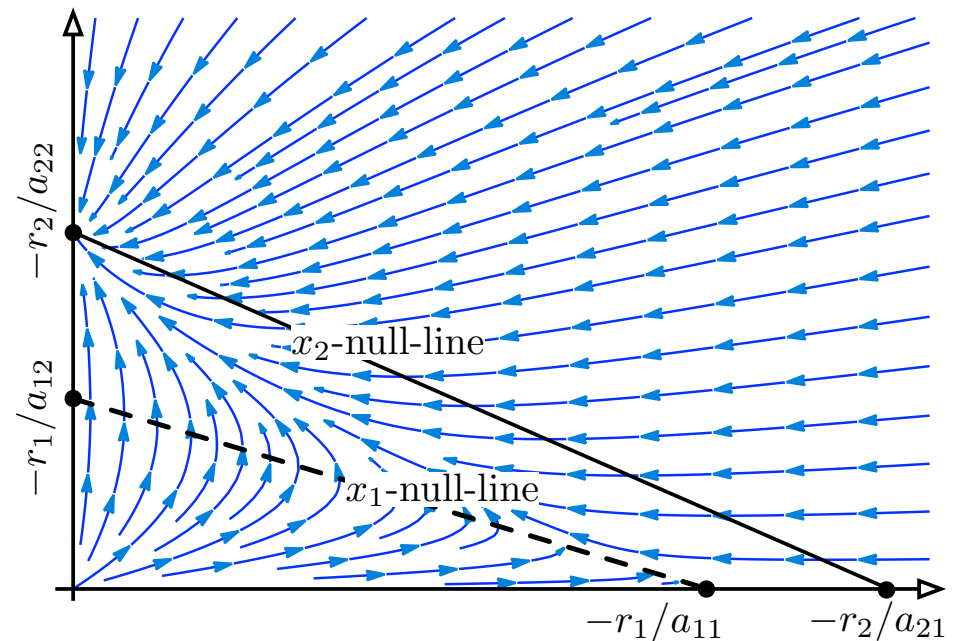
Case VI: $a_{12} < 0$, $a_{21} < 0$, $r_1/|a_{12}| < r_2/|a_{22}|$, and $r_1/|a_{11}| < r_2/|a_{21}|$. There exists no equilibrium in $\mathbb{R}^2_{>0}$. All trajectories starting in $\mathbb{R}^2_{>0}$ converge to the equilibrium $(0, -r_2/a_{22})$.

Figure 16.3: Two competition cases without equilibria in the two-species Lotka-Volterra system

16.2   General results for Lotka-Volterra models

We have seen some variety of behavior in the 2-species Lotka-Volterra model (16.2). Much richer dynamical behavior is possible in the n-species Lotka-Volterra model (14.3), including persistence, extinction, equilibria, periodic orbits, and chaotic evolution. In what follows we focus on sufficient conditions for the existence and stability of equilibrium points.

Lemma 16.4 (Lotka-Volterra is a positive system). *For $n \geq 2$, the Lotka-Volterra system (16.1) is a positive system, i.e., $x(0) \geq 0$ implies $x(t) \geq 0$ for all subsequent t. Moreover, if $x_i(0) = 0$, then $x_i(t) = 0$ for all subsequent t.*

Therefore, without loss of generality, we can assume that all initial conditions are positive vectors in $\mathbb{R}_{>0}^n$. In other words, if a locally-asymptotically stable positive equilibrium exists, the best we can hope for is to establish that its region of attraction is $\mathbb{R}_{>0}^n$. We are now ready to state the main result of this section.

Theorem 16.5 (Sufficient conditions for global asymptotic stability). *For the Lotka-Volterra system* (16.1) *with interaction matrix A and intrinsic growth rate r, assume*

(A1) *A is diagonally stable, i.e., there exists a positive vector $p \in \mathbb{R}^n_{>0}$ such that $\operatorname{diag}(p)A + A^\mathsf{T}\operatorname{diag}(p)$ is negative definite, and*

(A2) *the unique equilibrium point $x^* = -A^{-1}r$ is positive.*

Then x^* is globally asymptotically stable on $\mathbb{R}^n_{>0}$.

Proof. Note that A diagonally stable implies A Hurwitz and invertible. For $\kappa > 0$, recall the *logarithmic-linear function* $V_{\text{log-lin},\kappa}\colon \mathbb{R}_{>0} \to \mathbb{R}$ illustrated in Figure 15.6 and defined by

$$V_{\text{log-lin},\kappa}(x) = x - \kappa - \kappa \log\left(\frac{x}{\kappa}\right).$$

Assumption (A2) allows us to define $V\colon \mathbb{R}^n_{>0} \to \mathbb{R}_{\geq 0}$ by

$$V(x) = \sum_{i=1}^{n} p_i V_{\text{log-lin},x_i^*}(x_i) = \sum_{i=1}^{n} p_i\big(x_i - x_i^* - x_i^* \log(x_i/x_i^*)\big).$$

From Exercise E15.1 we know that the function $V_{\text{log-lin},\kappa}$ is continuously differentiable, takes non-negative values and satisfies $V_{\text{log-lin},\kappa}(x_i) = 0$ if and only if $x_i = \kappa$. Moreover, this function is unbounded in the limits as $x_i \to \infty$ and $x_i \to 0^+$. Therefore, V is globally positive-definite about x^* and proper.

Next, we compute the Lie derivative of V along the flow of the Lotka-Volterra vector field $f_{\text{LV}}(x) = \operatorname{diag}(x)(Ax + r)$. First, compute $\frac{d}{dx_i}V_{\text{log-lin},x_i^*}(x_i) = (x_i - x_i^*)/x_i$, so that

$$\mathscr{L}_{f_{\text{LV}}}V(x) = \sum_{i=1}^{n} p_i \frac{x_i - x_i^*}{x_i}(f_{\text{LV}}(x))_i.$$

Because A is invertible and $x^* = -A^{-1}r$, we write $Ax + r = A(x - x^*)$ and obtain

$$
\begin{aligned}
\mathscr{L}_{f_{\text{LV}}}V(x) &= \sum_{i=1}^{n} p_i(x_i - x_i^*)(A(x - x^*))_i \\
&= (x - x^*)^{\mathsf{T}} A^{\mathsf{T}} \operatorname{diag}(p)(x - x^*) \\
&= \frac{1}{2}(x - x^*)^{\mathsf{T}}(A^{\mathsf{T}} \operatorname{diag}(p) + \operatorname{diag}(p)A)(x - x^*).
\end{aligned}
$$

where we use the equality $y^{\mathsf{T}}By = y^{\mathsf{T}}(B + B^{\mathsf{T}})y/2$ for all $y \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n \times n}$. Assumption (A1) now implies that $\mathscr{L}_{f_{\text{LV}}}V(x) \leq 0$ with equality if and only if $x = x^*$. Therefore, $\mathscr{L}_{f_{\text{LV}}}V$ is globally negative-definite about x^*. According to the Lyapunov Stability Criteria in Theorem 15.4, x^* is globally asymptotically stable on $\mathbb{R}^n_{>0}$. ∎

Note: Assumption (A2) is not critical and, via a more complex treatment, a more general theorem can be obtained. For example, under the diagonal stability Assumption (A1), (Takeuchi, 1996, Theorem 3.2.1) shows the existence of a unique non-negative and globally stable equilibrium point for each $r \in \mathbb{R}^n$; this existence and uniqueness result is established via a linear complementarity problem.

16.3   Cooperative Lotka-Volterra models

In this section we focus on the case of Lotka-Volterra systems with only mutualistic interactions. In other words, we consider systems whose interaction terms satisfy $a_{ij} \geq 0$ for all i and j. For such systems, whenever $i \neq j$ we know

$$\frac{\partial}{\partial x_j}(f_{\mathsf{LV}})_i(x) = x_i a_{ij} \geq 0,$$

so that the Jacobian matrix of such systems is Metzler everywhere in $\mathbb{R}_{\geq 0}$. Such systems are called *cooperative*.

   We recall from Section 10.2 the properties of Metzler matrices. For example the Perron–Frobenius Theorem 10.2 for Metzler matrices establishes the existence of a dominant eigenvalue. Metzler matrices have so much structure that we are able to provide the following fairly comprehensive characterization: (1) Metzler matrices with a positive dominant eigenvalue have unbounded solutions of the Lotka-Volterra model (see Lemma 16.6 below), and (2) Metzler matrices with a negative dominant eigenvalue (and positive intrinsic growth rate) have a globally asymptotically-stable equilibrium point (see Theorem 16.7 below).

We start with a sufficient condition for unbounded evolutions.

Lemma 16.6 (Unbounded evolutions for unstable Metzler matrices). *Consider the Lotka-Volterra system* (16.1) *with interaction matrix A and intrinsic growth rate r. If A is a Metzler matrix with a positive dominant eigenvalue, then*

(i) there exist solutions that diverge in finite time starting from $\mathbb{R}_{>0}$, and

(ii) if $r > 0$, then all solutions starting from $\mathbb{R}_{>0}$ diverge in finite time.

Proof. Let $\lambda > 0$ and $w \geq \mathbb{0}_n$ with $\mathbb{1}_n^\mathsf{T} w = 1$ be the dominant eigenvalue and left eigenvector of A, whose existence and properties are established by the Perron–Frobenius Theorem 10.2 for Metzler matrices. Define $W \colon \mathbb{R}_{>0}^n \to \mathbb{R}_{>0}$ as the following weighted geometric average:

$$W(x) = \Pi_{i=1}^n x_i^{w_i}.$$

Along the flow of the Lotka-Volterra system, simple calculations show

$$\frac{\partial W(x)}{\partial x_i} = w_i \frac{1}{x_i} W(x)$$

$$\implies \quad \frac{\mathscr{L}_{f_{\mathsf{LV}}} W(x)}{W(x)} = \sum_{i=1}^n w_i \frac{1}{x_i} (f_{\mathsf{LV}}(x))_i = w^\mathsf{T}(Ax + r) = w^\mathsf{T}(\lambda x + r).$$

Generalizing the classic inequality $(a + b)/2 \geq (ab)^{1/2}$ for any $a, b \in \mathbb{R}_{>0}$, we recall from (Lohwater, 1982) the *weighted arithmetic-geometric mean inequality*: $w^\mathsf{T} x \geq \Pi_{i=1}^n x_i^{w_i}$ for any $x \in \mathbb{R}_{>0}^n$ and $w \in \mathbb{R}_{>0}^n$ with $\mathbb{1}_n^\mathsf{T} w = 1$. Additionally, we note that the inequality holds also for non-negative vectors w. Therefore, we have

$$\frac{\mathscr{L}_{f_{\mathsf{LV}}} W(x)}{W(x)} = w^\mathsf{T}(\lambda x + r) \geq \lambda \Pi_{i=1}^n x_i^{w_i} + w^\mathsf{T} r = \lambda W(x) + w^\mathsf{T} r,$$

so that

$$\mathscr{L}_{f_{\mathsf{LV}}} W(x) \geq W(x)(\lambda W(x) + w^\mathsf{T} r).$$

This inequality implies that, for any $x(0)$ such that $W(x(0)) > -w^\mathsf{T} r/\lambda$, the function $t \mapsto W(x(t))$ and, therefore at least one of the entries of the state $x(t)$, goes to infinity in finite time. This concludes the proof of statement (i).

Statement (ii) follows by noting that $r > 0$ implies $W(x(0)) > -w^\mathsf{T} r/\lambda$ for all $x(0) \in \mathbb{R}_{>0}^n$. ∎

Note: this lemma is true for any interaction matrix A that has a positive left eigenvector with positive eigenvalue.

We next provide a sufficient condition for global convergence to a unique equation* point.

Theorem 16.7 (Global convergence for cooperative Lotka-Volterra systems). *For the Lotka-Volterra system* (16.1) *with interaction matrix A and intrinsic growth rate r, assume*

(A3) *the interaction matrix A is Metzler and Hurwitz, and*

(A4) *the intrinsic growth rate is positive, $r > 0$.*

Then there exists a unique interior equilibrium point x^* and x^* is globally attractive on $\mathbb{R}_{>0}^n$.

Proof. We leave it to the reader to verify that, based on Assumptions (A3) and (A4), the Assumptions (A1) and (A2) of Theorem 16.5 are satisfied so that its consequences hold. ∎

Note: In (Baigent, 2010, Chapter 4), Theorem 16.7 is established via the Lyapunov function $V(x) = \max_{i \in \{1,\dots,n\}} \frac{|x_i - x_i^*|}{\xi_i}$, where x^* is the equilibrium point and $\xi = (\xi_1, \dots, \xi_n)$ is the positive vector with respect to which the Metzler Hurwitz matrix A has negative weighted row sums, as in the Metzler Hurwitz Theorem 10.14(iv).

16.4 Historical notes and further reading

The Lotka-Volterra population models are one the simplest and most widely adopted frameworks for modeling the dynamics of interacting populations in mathematical ecology. These equations were originally developed in (Lotka, 1920; Volterra, 1928).

An early reference for the analysis of the 2-species model is (Goh, 1976). Early references for the key stability result in Theorem 16.5 are (Takeuchi et al., 1978; Goh, 1979).

Textbook treatment include (Goh, 1980; Takeuchi, 1996; Baigent, 2010). For a more complete treatment of the n-special model, we refer the interested reader to (Takeuchi, 1996; Baigent, 2010). For example, Baigent (2010) discusses conservative Lotka-Volterra models (Hamiltonian structure and existence of periodic orbits), competitive and monotone models.

We refer to the texts (Hofbauer and Sigmund, 1998; Sandholm, 2010) for comprehensive discussions about the connection with between Lotka-Volterra models and evolutionary game dynamics.

16.5 Exercises

E16.1 **Proofs for 2-species behavior.** Provide proofs for Lemmas 16.1, 16.2, and 16.3.

E16.2 **The 2-dimensional Lotka-Volterra predator/prey dynamics.** In this exercise we study a 2-dimensional predator/prey model. We specialize the general Lotka-Volterra population model to the following set of equations:

$$\dot{x}(t) = \alpha x(t) - \beta x(t)y(t),$$
$$\dot{y}(t) = -\gamma y(t) + \delta x(t)y(t),$$

(E16.1)

where x is the non-negative number of preys, y is the non-negative number of predators individuals, and α, β, and γ are fixed positive systems parameters.

(i) Compute the unique non-zero equilibrium point (x^*, y^*) of the system.
(ii) Determine, if possible, the stability properties of the equilibrium points $(0,0)$ and (x^*, y^*) via linearization (Theorem 15.10).
(iii) Define the function $V(x,y) = -\delta x - \beta y + \gamma \ln(x) + \alpha \ln(y)$ and note its level sets as illustrated in Figure (E16.1).

(a) Compute the Lie derivative of $V(x,y)$ with respect to the Lotka-Volterra vector field.
(b) What can you say about the stability properties of (x^*, y^*)?
(c) Sketch the trajectories of the system for some initial conditions in the x-y positive orthant.
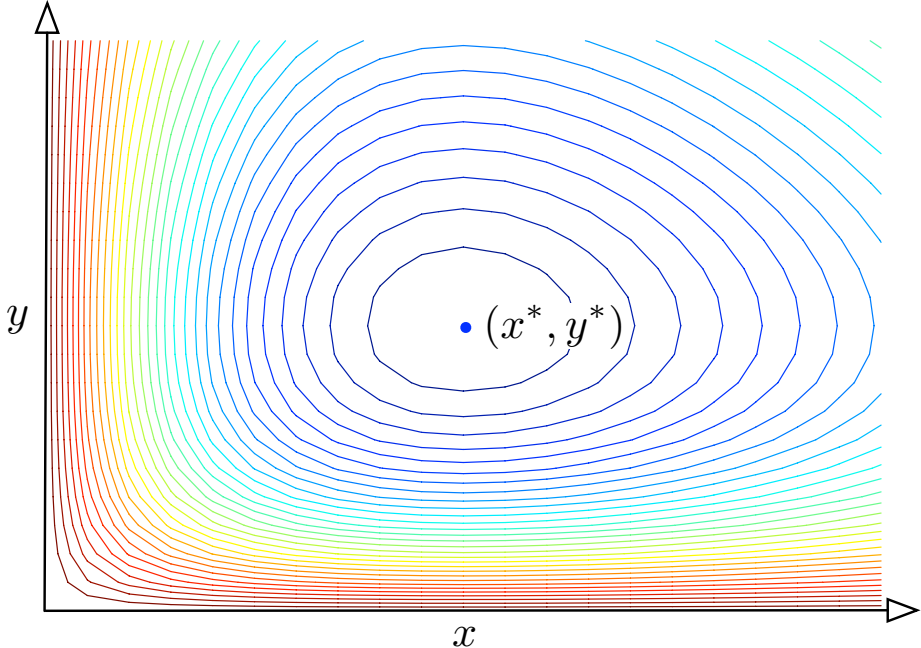
Figure E16.1: Level sets of the function $V(x, y)$ for unit parameter values

E16.3   **Unbounded evolutions caused by an unstable dominant eigenvalue.** State and prove a version of Lemma for matrices that are not necessarily Metzler, but that have a dominant eigenvalue and eigenvector (in a certain sense to be specified).

Chapter 17

Networks of Kuramoto Coupled Oscillators

In this chapter we continue our discussion about coupled-oscillator models and their behavior. Starting from the basic models discussed in Section 14.2, we here focus on characterizing synchronization and other dynamic phenomena.

Recall the two main models. Given an undirected weighted graph with adjacency matrix A and with n nodes, and given frequencies $\omega \in \mathbb{R}^n$, the coupled-oscillators model (14.4) is

$$\dot{\theta}_i = \omega_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}. \tag{17.1}$$

Moreover, given an undirected unweighted graph with n nodes, frequencies $\omega \in \mathbb{R}^n$, and coupling constant K, the Kuramoto model (14.5), is

$$\dot{\theta}_i = \omega_i - \frac{K}{n} \sum_{j=1}^{n} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}. \tag{17.2}$$

17.1 Preliminary notation and analysis

17.1.1   The geometry of the circle and the torus

Parametrization   The unit circle is $\mathbb{S}^1$. The torus $\mathbb{T}^n$ is the set consisting of n-copies of the circle. We parametrize the circle $\mathbb{S}^1$ by assuming (i) angles are measured counterclockwise, (ii) the 0 angle is the intersection of the unit circle with the positive horizontal axis, and (iii) angles take value in $[-\pi, \pi[$.

Geodesic distance   The *clockwise arc length from θ_i to θ_j* is the length of the clockwise arc from θ_i to θ_j. The counterclockwise arc length is defined analogously. The *geodesic distance between θ_i and θ_j* is the minimum between clockwise and counterclockwise arc lengths and is denoted by $|\theta_i - \theta_j|$. In the parametrization:

$$\mathrm{dist}_{\mathrm{cc}}(\theta_1, \theta_2) = \mathrm{mod}((\theta_2 - \theta_1), 2\pi), \qquad \mathrm{dist}_{\mathrm{c}}(\theta_1, \theta_2) = \mathrm{mod}((\theta_1 - \theta_2), 2\pi)$$
$$|\theta_1 - \theta_2| = \min\{\mathrm{dist}_{\mathrm{c}}(\theta_1, \theta_2), \mathrm{dist}_{\mathrm{cc}}(\theta_1, \theta_2)\}.$$

Arc subset and cohesive subset of the n-torus   Let G be an undirected weighted connected graph and let $\gamma \in [0, \pi)$.

(i) The *arc subset* $\overline{\Gamma}_{\text{arc}}(\gamma) \subset \mathbb{T}^n$ is the set of $(\theta_1, \dots, \theta_n) \in \mathbb{T}^n$ such that there exists an arc of length γ in $\mathbb{S}^1$ containing all angles $\theta_1, \dots, \theta_n$. The set $\Gamma(\gamma)$ is the interior of $\overline{\Gamma}_{\text{arc}}(\gamma)$;

(ii) The *cohesive subset* $\Delta^G(\gamma) \subseteq \mathbb{T}^n$ is

$$\Delta^G(\gamma) = \{\theta \in \mathbb{T}^n \mid |\theta_i - \theta_j| \leq \gamma, \quad \text{for all edges } (i, j)\}.$$

Note:

(i) For example, $\theta \in \overline{\Gamma}_{\text{arc}}(\pi)$ implies all angles $\theta_1, \dots, \theta_n$ belong to a closed half circle.

(ii) Clearly, $\overline{\Gamma}_{\text{arc}}(\gamma) \subset \Delta^G(\gamma)$ for any graph G. The converse is not true in general. For example, $\{\theta \in \mathbb{T}^n \mid |\theta_i - \theta_j| \leq \pi \text{ for all } i, j\}$ is equal to the entire $\mathbb{T}^n$. A weak converse statement is studied in Exercise E17.1.

(iii) If $\theta = (\theta_1, \dots, \theta_n) \in \Gamma_{\text{arc}}(\pi)$, then $\text{average}(\theta)$ is well posed. (The average of n angles is ill-posed in general, e.g., there is no reasonable definition of the average of two diametrically-opposed points.)

Rotations   Given the angle $\alpha \in [-\pi, \pi[$, the rotation of the n-tuple $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{T}^n$ by α, denoted by $\mathrm{rot}_\alpha(\theta)$, is the counterclockwise rotation of each entry $(\theta_1, \ldots, \theta_n)$ by α. For $\theta = \in \mathbb{T}^n$, we also define its *rotation set* to be

$$[\theta] = \{\mathrm{rot}_\alpha(\theta) \in \mathbb{T}^n \mid \alpha \in [-\pi, \pi[\}.$$

The coupled oscillator model (17.1) is *invariant* under rotations, that is, given a solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ to the coupled oscillator model, a rotation of $\mathrm{rot}_\alpha(\theta(t))$ by any angle α is again a solution.

17.1.2 Synchronization notions

Consider the following notions of synchronization for a solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$:

Frequency synchrony: A solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ is *frequency synchronized* if $\dot{\theta}_i(t) = \dot{\theta}_j(t)$ for all time t and for all i and j.

Phase synchrony: A solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ is *phase synchronized* if $\theta_i(t) = \theta_j(t)$ for all time t and for all i and j.

Phase cohesiveness: A solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ is *phase cohesive* with respect to $\gamma > 0$ if one of the following conditions holds for all time t:

(i) $\theta(t) \in \Gamma_{\mathrm{arc}}(\gamma)$; or

(ii) $\theta(t) \in \Delta^G(\gamma)$, for a graph of interest G.

Asymptotic notions: We will also talk about solutions that *asymptotically achieve* certain synchronization properties. For example, a solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ *achieves phase synchronization* if $\lim_{t \to \infty} |\theta_i(t) - \theta_j(t)| = 0$. Analogous definitions can be given for asymptotic frequency synchronization and asymptotic phase cohesiveness.

Finally, notice that phase synchrony is the extreme case of all phase cohesiveness notions with $\gamma = 0$.

17.1.3   Preliminary results

We have the following result on the synchronization frequency.

Lemma 17.1 (Synchronization frequency). *Consider the coupled oscillator model* (17.1) *with frequencies* $\omega \in \mathbb{R}^n$ *defined over a connected weighted undirected graph. If a solution achieves frequency synchronization, then it does so with a constant synchronization frequency equal to*

$$\omega_{\mathsf{sync}} \triangleq \frac{1}{n} \sum_{i=1}^{n} \omega_i = \mathrm{average}(\omega).$$

Proof. This fact is obtained by summing all equations (17.1) for $i \in \{1, \dots, n\}$.   ∎

Lemma 17.1 implies that, by expressing each angle with respect to a rotating frame with frequency ω_{sync} and by replacing ω_i by $\omega_i - \omega_{\mathsf{sync}}$, we obtain $\omega_{\mathsf{sync}} = 0$ or, equivalently, $\omega \in \mathbb{1}_n^{\perp}$. In this rotating frame a frequency-synchronized solution is an equilibrium. Due to the rotational invariance of the coupled oscillator model (17.1), it follows that if $\theta^* \in \mathbb{T}^n$ is an equilibrium point, then every point in the rotation set

$$[\theta^*] = \{\theta \in \mathbb{T}^n \mid \mathrm{rot}_\alpha(\theta^*), \alpha \in [-\pi, \pi[\}$$

is also an equilibrium. We refer to $[\theta^*]$ as an *equilibrium set*.

We have the following important result on local stability properties of equilibria.

Lemma 17.2 (Linearization and frequency synchronization). *Consider the coupled oscillator model* (17.1) *with frequencies* $\omega \in \mathbb{1}_n^{\perp}$ *defined over a connected weighted undirected graph with incidence matrix* B. *The following statements hold:*

(i) (Jacobian:) the Jacobian of the coupled oscillator model at $\theta \in \mathbb{T}^n$ *is*

$$J(\theta) = -B \operatorname{diag}(\{a_{ij} \cos(\theta_i - \theta_j)\}_{\{i,j\} \in E}) B^{\mathsf{T}};$$

(ii) (local stability:) if there exists an equilibrium $\theta^* \in \Delta^G(\gamma), \gamma < \pi/2$, *then*

   (a) $-J(\theta^*)$ *is a Laplacian matrix; and*

   (b) the equilibrium set $[\theta^*]$ *is locally exponentially stable;*

(iii) (frequency synchronization:) if a solution $\theta(t)$ *is phase cohesive in the sense that* $\theta(t) \in \Delta^G(\gamma), \gamma < \pi/2$, *for all* $t \geq 0$, *then there exists a phase cohesive equilibrium* $\theta^* \in \Delta^G(\gamma)$ *and* $\theta(t)$ *achieves exponential frequency synchronization converging to* $[\theta^*]$.

Proof. We start with statements (i) and (ii)a. Given $\theta \in \mathbb{T}^n$, we define the undirected graph $G_{\text{cosine}}(\theta)$ with the same nodes and edges as G and with edge weights $a_{ij} \cos(\theta_i - \theta_j)$. Next, we compute

$$\frac{\partial}{\partial \theta_i}\left(\omega_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j)\right) = -\sum_{j=1}^{n} a_{ij} \cos(\theta_i - \theta_j),$$

$$\frac{\partial}{\partial \theta_j}\left(\omega_i - \sum_{k=1}^{n} a_{ik} \sin(\theta_i - \theta_k)\right) = a_{ij} \cos(\theta_i - \theta_j).$$

Therefore, the Jacobian is equal to minus the Laplacian matrix of the (possibly negatively weighted) graph $G_{\text{cosine}}(\theta)$ and statement (i) follows from Lemma 9.1. Regarding statement (ii)a, if $|\theta_i^* - \theta_j^*| < \pi/2$ for all $\{i, j\} \in E$, then $\cos(\theta_i^* - \theta_j^*) > 0$ for all $\{i, j\} \in E$, so that $G_{\text{cosine}}(\theta)$ has strictly non-negative weights and all usual properties of Laplacian matrices hold.

To prove statement (ii)b notice that $J(\theta^*)$ is negative semidefinite with the nullspace $\mathbb{1}_n$ arising from the rotational symmetry. All other eigenvectors are orthogonal to $\mathbb{1}_n$ and have negative eigenvalues. We now restrict our analysis to the orthogonal complement of $\mathbb{1}_n$: we define a coordinate transformation matrix $Q \in \mathbb{R}^{(n-1) \times n}$ with orthonormal rows orthogonal to $\mathbb{1}_n$,

$$Q\mathbb{1}_n = \mathbb{0}_{n-1} \quad \text{and} \quad QQ^{\mathsf{T}} = I_{n-1},$$

and we note that $QJ(\theta^*)Q^{\mathsf{T}}$ has negative eigenvalues. Therefore, in the original coordinates, the zero eigenspace $\mathbb{1}_n$ is exponentially stable. Theorem 15.10 implies that the equilibrium set $[\theta^*]$ is locally exponentially stable.

Regarding statement (iii), define $x_i(t) = \dot{\theta}_i(t)$. Then $\dot{x}(t) = J(\theta(t))x(t)$ is a time-varying averaging system. The associated undirected graph has time-varying yet strictly positive weights $a_{ij} \cos(\theta_i(t) - \theta_j(t)) \geq a_{ij} \cos(\gamma) > 0$ for each $\{i, j\} \in E$. Hence, the weighted graph is connected for each $t \geq 0$. From the analysis of time-varying averaging systems in Theorem 12.9, the exponential convergence of $x(t)$ to $\text{average}(x(0))\mathbb{1}_n$ follows. Equivalently, the frequencies synchronize. By continuity, the limiting value of $\theta(t)$ must be an equilibrium. ∎

17.1.4   The order parameter and the mean field model

An alternative synchronization measure (besides phase cohesiveness) is the magnitude of the *order parameter*

$$re^{\mathrm{i}\psi} = \frac{1}{n}\sum_{j=1}^{n} e^{\mathrm{i}\theta_j}\,. \tag{17.3}$$

The order parameter (17.3) is the centroid of all oscillators represented as points on the unit circle in $\mathbb{C}^1$. The magnitude r of the order parameter is a synchronization measure:

- if the oscillators are phase-synchronized, then $r = 1$;
- if the oscillators are spaced equally on the unit circle, then $r = 0$; and
- for $r \in\,]0, 1[$ and oscillators contained in a semi-circle, the associated configuration of oscillators satisfy a certain level of phase cohesiveness; see Exercise E17.2.

By means of the order parameter $re^{\mathrm{i}\psi}$ the all-to-all Kuramoto model (17.2) can be rewritten in the insightful form

$$\dot{\theta}_i = \omega_i - Kr\sin(\theta_i - \psi)\,, \quad i \in \{1, \dots, n\}\,. \tag{17.4}$$

(We ask the reader to establish this identity in Exercise E17.3.) Equation (17.4) gives the intuition that the oscillators synchronize because of their coupling to a *mean field* represented by the order parameter $re^{\mathrm{i}\psi}$, which itself is a function of $\theta(t)$. Intuitively, for small coupling strength K each oscillator rotates with its distinct natural frequency ω_i, whereas for large coupling strength K all angles $\theta_i(t)$ will entrain to the mean field $re^{\mathrm{i}\psi}$, and the oscillators synchronize. The transition from incoherence to synchrony occurs at a critical threshold value of the coupling strength, denoted by K_{critical}.

17.2   Synchronization of identical oscillators

We start our discussion with the following insightful lemma.

Lemma 17.3. *Consider the coupled oscillator model* (17.1). *If $\omega_i \neq \omega_j$ for some distinct $i, j \in \{1, \ldots, n\}$, then the oscillators cannot achieve phase synchronization.*

Proof. We prove the lemma by contradiction. Assume that all oscillators are in phase synchrony $\theta_i(t) = \theta_j(t)$ for $t \geq 0$ and $i, j \in \{1, \ldots, n\}$. Then equating the dynamics, $\dot{\theta}_i(t) = \dot{\theta}_j(t)$, implies that $\omega_i = \omega_j$.   ■

We now consider oscillators with identical natural frequencies, $\omega_i = \omega \in \mathbb{R}$ for all $i \in \{1, \ldots, n\}$. By working in a rotating frame with frequency ω, we have $\omega = 0$. Thus, we consider the model

$$\dot{\theta}_i = -\sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j), \qquad i \in \{1, \ldots, n\}. \tag{17.5}$$

Notice that phase synchronization is an equilibrium of the this model. Conversely, phase synchronization cannot be an equilibrium of the original coupled oscillator model (17.1) if $\omega_i \neq \omega_j$.

17.2.1 An averaging-based approach

Let us first analyze the coupled oscillator model (17.5) with initial conditions restricted to an open semi-circle, $\theta(0) \in \Gamma_{\text{arc}}(\gamma)$ for some $\gamma \in [0, \pi[$. In this case, the oscillators remain in a semi-circle at least for small times $t > 0$ and the two coordinate transformations

$$x_i(t) = \tan(\theta_i(t)) \quad \text{(with } x_i \in \mathbb{R}\text{)}, \quad \text{and} \quad y_i(t) = \theta_i(t) \quad \text{(with } y_i \in \mathbb{R}\text{)}$$

are well-defined and bijective (at least for small times).

In the x_i-coordinates, the coupled oscillator model reads as the time-varying continuous-time averaging system

$$\dot{x}_i(t) = -\sum_{j=1}^{n} b_{ij}(t)(x_i(t) - x_j(t)), \tag{17.6}$$

where $b_{ij}(t) = a_{ij}\sqrt{(1 + x_i(t)^2)/(1 + x_j(t)^2)}$ and $b_{ij}(t) \geq a_{ij}\cos(\gamma/2)$; see Exercise E17.7 for a derivation. Similarly, in the y_i-coordinates, the coupled oscillator model reads as

$$\dot{y}_i(t) = -\sum_{j=1}^{n} c_{ij}(t)(y_i(t) - y_j(t)), \tag{17.7}$$

where $c_{ij}(t) = a_{ij}\operatorname{sinc}(y_i(t) - y_j(t))$ and $c_{ij}(t) \geq a_{ij}\operatorname{sinc}(\gamma)$. Notice that both averaging formulations (17.6) and (17.7) are well-defined as long as the the oscillators remain in a semi-circle $\Gamma_{\text{arc}}(\gamma)$ for some $\gamma \in [0, \pi[$.

Theorem 17.4 (Phase cohesiveness and synchronization in open semicircle). *Consider the coupled oscillator model* (17.5) *with identical frequencies defined over a connected weighted undirected graph with Laplacian matrix L. Then*

(i) (phase cohesiveness:) for each $\gamma \in [0, \pi[$ each solution originating in $\Gamma_{\mathrm{arc}}(\gamma)$ remains in $\Gamma_{\mathrm{arc}}(\gamma)$ for all times;

(ii) (asymptotic phase synchronization:) each trajectory originating in $\Gamma_{\mathrm{arc}}(\gamma)$ for $\gamma \in [0, \pi[$ achieves exponential phase synchronization, that is,

$$\|\theta(t) - \mathrm{average}(\theta(0))\mathbb{1}_n\|_2 \leq \|\theta(0) - \mathrm{average}(\theta(0))\mathbb{1}_n\|_2 e^{\lambda_{\mathrm{ps}} t}, \tag{17.8}$$

where $\lambda_{\mathrm{ps}} = -\lambda_2(L)\cos(\gamma/2)$.

Proof. Consider the averaging formulations (17.6) and (17.7) with initial conditions $\theta(0) \in \Gamma_{\mathrm{arc}}(\gamma)$ for some $\gamma \in [0, \pi[$. By continuity, for small positive times $t > 0$, the oscillators remain in a semi-circle, the time-varying weights $b_{ij}(t) \geq a_{ij}(\cos(\gamma/2)$ and $c_{ij}(t) \geq a_{ij}\mathrm{sinc}(\gamma)$ are strictly positive for each $\{i, j\} \in E$, the associated time-dependent graph is connected. As one establishes in the proof of Theorem 12.9, the max-min function $V_{\mathrm{max\text{-}min}}$, defined in equation (5.14), evaluated along the solutions to the time-varying consensus systems (17.6) and (17.7) are strictly decreasing for until consensus is reached.

Thus, the oscillators remain in $\Gamma_{\mathrm{arc}}(\gamma)$ phase synchronization exponentially fast. Since the graph is undirected, we can also conclude convergence to the average phase. Finally, the explicit convergence estimate (17.8) follows, for example, by analyzing (17.6) with the disagreement Lyapunov function and using $b_{ij}(t) \geq a_{ij}\cos(\gamma/2)$. ∎

17.2.2   The potential landscape, convergence and phase synchronization

The consensus analysis in Theorem 17.4 leads to a powerful result but is inherently restricted to a semi-circle. To overcome this limitation, we use potential functions as an analysis tool. Inspired by Examples #1 and #3 in Section 14.2, define the potential function $\mathsf{U}\colon \mathbb{T}^n \to \mathbb{R}$ by

$$\mathsf{U}(\theta) = \sum_{\{i,j\}\in E} a_{ij}\big(1 - \cos(\theta_i - \theta_j)\big)\,. \tag{17.9}$$

Then the coupled oscillator model (17.1) (with all $\omega_i = 0$) is identical to the negative gradient flow

$$\dot{\theta} = -\frac{\partial \mathsf{U}(\theta)}{\partial \theta}\,. \tag{17.10}$$

Among the many critical points of the potential function U in equation (17.9), each point in the set of phase-synchronized angles is a global minimum of U. This fact can be easily seen since each summand in (17.9) is bounded in $[0, 2a_{ij}]$ and the lower bound is reached only if neighboring oscillators are phase-synchronized.

Theorem 17.5 (Phase synchronization). *Consider the coupled oscillator model* (17.5) *with identical frequencies defined over a connected weighted undirected graph. Then*

(i) *(global convergence:) for all initial conditions* $\theta(0) \in \mathbb{T}^n$, *the phases* $\theta_i(t)$ *converge to the set of critical points* $\{\theta \in \mathbb{T}^n \mid \partial \mathsf{U}(\theta)/\partial \theta = \mathbb{0}_n\}$; *and*

(ii) *(local stability:) phase synchronization is a locally exponentially stable equilibrium set.*

Proof. Since the coupled oscillator model (17.1) is a negative gradient flow, we can apply Theorem 15.14. Note that U is analytic and the state space is the compact manifold $\mathbb{T}^n$. Specifically, statement (i) is statement Theorem 15.14(i).

   Statement (ii) follows from the Jacobian result in Lemma 17.2 and Theorem 15.10. ∎

Theorem 17.5 together with Theorem 17.4 gives a fairly complete picture of the local convergence and phase synchronization properties of the coupled oscillator model (17.5). Regarding global properties, a stronger result can be made in case of an all-to-all homogeneous coupling graph, that is, for the Kuramoto model (17.2).

Corollary 17.6 (Almost global phase synchronization for the Kuramoto model). *Consider the Kuramoto model* (17.2) *with identical natural frequencies $\omega_i = \omega_j$ for all $i, j \in \{1, \ldots, n\}$. Then for almost all initial conditions in $\mathbb{T}^n$, the oscillators achieve phase synchronization.*

Proof. For identical natural frequencies, the Kuramoto model (17.2) can be put in rotating coordinates so that $\omega_i = 0$ for all $i \in \{1, \ldots, n\}$; see Section 17.2. The Kuramoto model reads in the order-parameter formulation (17.4) as

$$\dot{\theta}_i = -Kr \sin(\theta_i - \psi), \quad i \in \{1, \ldots, n\}. \tag{17.11}$$

The associated potential function reads as (see Exercise E17.5)

$$\mathsf{U}(\theta) = \sum_{\{i,j\} \in E} a_{ij}\big(1 - \cos(\theta_i - \theta_j)\big) = \frac{Kn}{2}(1 - r^2), \tag{17.12}$$

and its unique global minimum is obtained for $r = 1$, that is, in the phase-synchronized state. By Theorem 17.5, all angles converge to the set of equilibria which are from (17.11) either (i) $r = 0$, (ii) $r > 0$ and in-phase with the order parameter $\theta_i = \psi$, or (iii) $r > 0$ and out-of-phase with the order parameter $\theta_i = \psi + k\pi$ for $k \in \mathbb{Z} \setminus \{0\}$ for all $i \in \{1, \ldots, n\}$. In the latter case, any infinitesimal deviation from an out-of-phase equilibrium causes the potential (17.12) to decrease, that is, the out-of-phase equilibria are unstable. Likewise, the equilibria with $r = 0$ correspond to the global maxima of the potential (17.12), and any infinitesimal deviation from these equilibria causes the potential (17.12) to decrease. It follows that, from almost all initial conditions[1], the oscillators converge to phase-synchronized equilibria $\theta_i = \psi$ for all $i \in \{1, \ldots, n\}$. ∎

---

[1]To be precise further analysis is needed. A linearization of the Kuramoto model (17.11) at the unstable out-of-phase equilibria yields that these are exponentially unstable. The region of attraction (the so-called stable manifold) of such exponentially unstable equilibria is known to be a zero measure set (Potrie and Monzón, 2009, Proposition 4.1).

17.2.3   Phase balancing

Applications in neuroscience, vehicle coordination, and central pattern generators for robotic locomotion motivate the study of coherent behaviors with synchronized frequencies where the phases are not synchronized, but rather dispersed in appropriate patterns. While the phase-synchronized state can be characterized by the order parameter r achieving its maximal (unit) magnitude, we say that a solution $\theta\colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ to the coupled oscillator model (17.1) achieves *phase balancing* if all phases θ_i asymptotically converge to the *splay set*

$$\left\{ \theta \in \mathbb{T}^n \mid r(\theta) = \Big| \sum\nolimits_{j=1}^{n} e^{\mathrm{i}\theta_j}/n \Big| = 0 \right\},$$

that is, asymptotically the oscillators are uniformly distributed over the unit circle $\mathbb{S}^1$ so that their centroid converges to the origin.

For a complete homogeneous graph with coupling strength $a_{ij} = K/n$, i.e., for the Kuramoto model (17.2), we have a remarkable identity between the magnitude of the order parameter r and the potential function $\mathsf{U}(\theta)$

$$\mathsf{U}(\theta) = \frac{Kn}{2}\left(1 - r^2\right). \tag{17.13}$$

(We ask the reader to establish this identity in Exercise E17.5.) For the complete graph, the correspondence (17.13) shows that the global minimum of the potential function $\mathsf{U}(\theta) = 0$ (for $r = 1$) corresponds to phase-synchronization and the global maximum $\mathsf{U}(\theta) = Kn/2$ (for $r = 0$) corresponds to phase balancing. This motivates the following gradient ascent dynamics to reach phase balancing:

$$\dot{\theta} = +\frac{\partial \mathsf{U}(\theta)}{\partial \theta}, \qquad \text{or, equivalently,} \quad \dot{\theta}_i = \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j). \tag{17.14}$$

Theorem 17.7 (Phase balancing). *Consider the coupled oscillator model (17.14) with a connected, undirected, and weighted graph. Then*

(i) *(global convergence:) for all initial conditions $\theta(0) \in \mathbb{T}^n$, the phases $\theta_i(t)$ converge to the set of critical points $\{\theta \in \mathbb{T}^n \mid \partial \mathsf{U}(\theta)/\partial \theta = \mathbb{0}_n\}$; and*

(ii) *(local stability:) for a complete graph with uniform weights $a_{ij} = K/n$, phase balancing is the global maximizer of the potential function (17.13) and is a locally asymptotically stable equilibrium set.*

Proof. The proof statement (i) is analogous to the proof of statement (i) in Theorem 17.5.

To prove statement (ii), notice that, for a complete graph, the phase balanced set characterized by $r = 0$ achieves the global maximum of the potential $\mathsf{U}(\theta) = \frac{Kn}{2}\left(1 - r^2\right)$. By Theorem 15.14, local maxima of the potential are locally asymptotically stable for the gradient ascent dynamics (17.14). ∎

17.3   Synchronization of heterogeneous oscillators

In this section we analyze non-identical oscillators with $\omega_i \neq \omega_j$. As shown in Lemma 17.3, these oscillator networks cannot achieve phase synchronization. On the other hand frequency synchronization with a certain degree of phase cohesiveness can be achieved provided that the natural frequencies satisfy certain bounds relative to the network coupling. We start off with the following necessary conditions.

Lemma 17.8 (Necessary condition for synchronization). *Consider the coupled oscillator model* (17.1) *with graph with adjacency matrix A, frequencies $\omega \in \mathbb{1}_n^{\perp}$, and nodal degree $d_i = \sum_{j=1}^{n} a_{ij}$ for each node $i \in \{1, \dots, n\}$. If there exists a frequency-synchronized solution satisfying the phase cohesiveness $|\theta_i - \theta_j| \leq \gamma$ for all $\{i, j\} \in E$ and for some $\gamma \in [0, \pi/2]$, then*

(i) (absolute bound:) for each node $i \in \{1, \dots, n\}$,

$$d_i \sin(\gamma) \geq |\omega_i|, \tag{17.15}$$

(ii) (incremental bound:) for distinct $i, j \in \{1, \dots, n\}$,

$$(d_i + d_j) \sin(\gamma) \geq |\omega_i - \omega_j|. \tag{17.16}$$

Proof. Statement (i) follows directly from the fact that synchronized solutions must satisfy the equilibrium equation $\dot{\theta}_i = 0$. Since the sinusoidal interaction terms in equation (17.1) are upper bounded by the nodal degree $d_i = \sum_{j=1}^{n} a_{ij}$, condition (17.15) is necessary for the existence of an equilibrium.

Statement (ii) follows from the fact that frequency-synchronized solutions must satisfy $\dot{\theta}_i - \dot{\theta}_j = 0$. By analogous arguments, we arrive at the necessary condition (17.16). ∎

17.3.1   Synchronization of heterogeneous oscillators over complete homogeneous graphs

We now consider the Kuramoto model over a complete homogeneous graph in equation (17.2). As discussed in Subsection 17.1.4, the Kuramoto model synchronizes provided that the coupling gain K is larger than some critical value K_{critical}. The necessary condition (17.16) delivers a lower bound for K_{critical} given by

$$K \geq \frac{n}{2(n-1)} \left(\max_i \omega_i - \min_i \omega_i \right).$$

Here we evaluated the left-hand side of (17.16) for $a_{ij} = K/n$, for the maximum $\gamma = \pi/2$, and for all distinct $i, j \in \{1, \dots, n\}$. Perhaps surprisingly, the lower necessary bound (17.3.1) is a factor $1/2$ away from the upper sufficient bound.

Theorem 17.9 (Synchronization test for all-to-all Kuramoto model). *Consider the Kuramoto model* (17.2) *with natural frequencies* $\omega \in \mathbb{1}_n^\perp$ *and coupling strength* K. *Assume*

$$K > K_{\text{critical}} \triangleq \max_i \omega_i - \min_i \omega_i, \tag{17.17}$$

and define the arc lengths $\gamma_{\min} \in [0, \pi/2[$ *and* $\gamma_{\max} \in]\pi/2, \pi]$ *as the unique solutions to* $\sin(\gamma_{\min}) = \sin(\gamma_{\max}) = K_{\text{critical}}/K$.
  The following statements hold:

 (i) (phase cohesiveness:) each solution starting in $\Gamma_{\text{arc}}(\gamma)$, *for* $\gamma \in [\gamma_{\min}, \gamma_{\max}]$, *remains in* $\Gamma_{\text{arc}}(\gamma)$ *for all times;*
 (ii) (asymptotic phase cohesiveness:) each solution starting in $\Gamma_{\text{arc}}(\gamma_{\max})$ *asymptotically reaches the set* $\overline{\Gamma}_{\text{arc}}(\gamma_{\min})$; *and*
(iii) (asymptotic frequency synchronization:) each solution starting in $\Gamma_{\text{arc}}(\gamma_{\max})$ *achieves frequency synchronization.*

  Moreover, the following converse statement is true: Given an interval $[\omega_{\min}, \omega_{\max}]$, *the coupling strength* K *satisfies* $K > \omega_{\max} - \omega_{\min}$ *if, for all frequencies* ω *supported on* $[\omega_{\min}, \omega_{\max}]$ *and for the arc length* $\gamma_{\max}$ *computed as above, the set* $\Gamma_{\text{arc}}(\gamma_{\max})$ *is positively invariant.*

We illustrate the definitions of $\gamma_{\min}$, $\gamma_{\max}$, and $\Gamma_{\text{arc}}(\gamma)$, for $\gamma \in [\gamma_{\min}, \gamma_{\max}]$ in Figure 17.1.
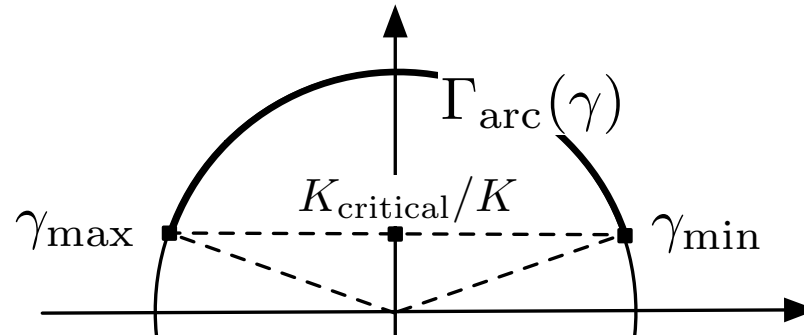
Figure 17.1: Illustrating the definitions of $\gamma_{\min}$, $\gamma_{\max}$, and $\Gamma_{\text{arc}}(\gamma)$, for $\gamma \in [\gamma_{\min}, \gamma_{\max}]$.

Proof. We start with statement (i). Define the function $W : \Gamma_{\text{arc}}(\pi) \to [0, \pi[$ by

$$W(\psi) = \max\{|\psi_i - \psi_j| \mid i, j \in \{1, \ldots, n\}\}.$$

The arc containing all angles ψ has two boundary points: a counterclockwise maximum and a counterclockwise minimum. If $U_{\text{max}}(\psi)$ (resp. $U_{\text{min}}(\psi)$) denotes the set indices of the angles $\psi_1, \ldots, \psi_n$ that are equal to the counterclockwise maximum (resp. the counterclockwise minimum), then

$$W(\psi) = |\psi_{m'} - \psi_{k'}|, \quad \text{for all } m' \in U_{\text{max}}(\psi) \text{ and } k' \in U_{\text{min}}(\psi).$$

We now assume $\theta(0) \in \overline{\Gamma}_{\text{arc}}(\gamma)$, for $\gamma \in [\gamma_{\text{min}}, \gamma_{\text{max}}]$, and aim to show that $\theta(t) \in \overline{\Gamma}_{\text{arc}}(\gamma)$ for all times $t > 0$. By continuity, $\overline{\Gamma}_{\text{arc}}(\gamma)$ is positively invariant if and only if $W(\theta(t))$ does not increase at any time t such that $W(\theta(t)) = \gamma$.

In the next equation we compute the maximum possible amount of infinitesimal increase of $t \mapsto W(\theta(t))$ along system (17.2). Based on the notion of upper Dini derivative and the treatment in Section 15.8, we compute

$$D^+ W(\theta(t)) := \limsup_{\Delta t \to 0^+} \frac{W(\theta(t + \Delta t)) - W(\theta(t))}{\Delta t} = \dot{\theta}_m(t) - \dot{\theta}_k(t),$$

where the indices $m \in U_{\max}(\theta(t))$ and $k \in U_{\min}(\theta(t))$ have the property that $\dot{\theta}_m(t) = \max\{\dot{\theta}_{m'}(t) \mid m' \in U_{\max}(\theta(t))\}$ and $\dot{\theta}_k(t) = \min\{\dot{\theta}_{k'}(t) \mid k' \in U_{\min}(\theta(t))\}$. In components

$$D^+ W(\theta(t)) = \omega_m - \omega_k - \frac{K}{n} \sum_{j=1}^n \left(\sin(\theta_m(t) - \theta_j(t)) + \sin(\theta_j(t) - \theta_k(t)) \right).$$

The trigonometric identity $\sin(x) + \sin(y) = 2\sin(\frac{x+y}{2})\cos(\frac{x-y}{2})$ leads to

$$D^+ W(\theta(t)) = \omega_m - \omega_k$$

$$- \frac{K}{n} \sum_{i=1}^n \left(2 \sin\left(\frac{\theta_m(t) - \theta_k(t)}{2} \right) \cos\left(\frac{\theta_m(t) - \theta_i(t)}{2} - \frac{\theta_i(t) - \theta_k(t)}{2} \right) \right).$$

Measuring angles counterclockwise and modulo 2π, the equality $W(\theta(t)) = \gamma$ implies $\theta_m(t) - \theta_k(t) = \gamma$, $\theta_m(t) - \theta_i(t) \in [0, \gamma]$, and $\theta_i(t) - \theta_k(t) \in [0, \gamma]$. Moreover,

$$\min_\theta \cos\left(\frac{\theta_m - \theta_i}{2} - \frac{\theta_i - \theta_k}{2} \right) = \cos\left(\max_\theta \left| \frac{\theta_m - \theta_i}{2} - \frac{\theta_i - \theta_k}{2} \right| \right) = \cos(\gamma/2),$$

so that

$$D^+ W(\theta(t)) \leq \omega_m - \omega_k - \frac{K}{n} \sum_{i=1}^n \left(2 \sin\left(\frac{\gamma}{2} \right) \cos\left(\frac{\gamma}{2} \right) \right).$$

Applying the reverse identity $2\sin(x)\cos(y) = \sin(x - y) + \sin(x + y)$, we obtain

$$D^+ W(\theta(t)) \leq \omega_m - \omega_k - \frac{K}{n} \sum_{i=1}^n \sin(\gamma) \leq (\max_i \omega_i - \min_i \omega_i) - K \sin(\gamma).$$

Hence, the $W(\theta(t))$ does not increase at all t such that $W(\theta(t)) = \gamma$ if $K \sin(\gamma) \geq K_{\text{critical}} = \max_i \omega_i - \min_i \omega_i$.

Given the structure of the level sets of $\gamma \mapsto K \sin(\gamma)$, there exists an open interval of arc lengths $\gamma \in [0, \pi]$ satisfying $K \sin(\gamma) \geq \max_i \omega_i - \min_i \omega_i$ if and only if equation (17.17) is true with the strict equality sign at $\gamma^* = \pi/2$, that is, if $K > K_{\text{critical}}$. Additionally, if $K > K_{\text{critical}}$, there exists a unique $\gamma_{\text{min}} \in [0, \pi/2[$ and a unique $\gamma_{\text{max}} \in]\pi/2, \pi]$ that satisfy equation (17.17) with the equality sign. In summary, for every $\gamma \in [\gamma_{\text{min}}, \gamma_{\text{max}}]$, if $W(\theta(t)) = \gamma$, then the arc length $W(\theta(t))$ is non-increasing. This concludes the proof of statement (i).

Moreover, pick $\varepsilon \ll \gamma_{\text{max}} - \gamma_{\text{min}}$. For all $\gamma \in [\gamma_{\text{min}} + \varepsilon, \gamma_{\text{max}} - \varepsilon]$, there exists a positive $\delta(\varepsilon)$ with the property that, if $W(\theta(t)) = \gamma$, then $D^+ W(\theta(t)) \leq -\delta(\varepsilon)$. Hence, each solution $\theta \colon \mathbb{R}_{\geq 0} \to \mathbb{T}^n$ starting in $\Gamma_{\text{arc}}(\gamma_{\text{max}} - \varepsilon)$ must satisfy $W(\theta(t)) \leq \gamma_{\text{min}} - \varepsilon$ after time at most $(\gamma_{\text{max}} - \gamma_{\text{min}})/\delta(\varepsilon)$. This proves statement (ii).

Regarding statement (iii), we just proved that for every $\theta(0) \in \Gamma_{\text{arc}}(\gamma_{\text{max}})$ and for all $\gamma \in]\gamma_{\text{min}}, \gamma_{\text{max}}]$ there exists a finite time $T \geq 0$ such that $\theta(t) \in \overline{\Gamma}_{\text{arc}}(\gamma)$ for all $t \geq T$ and for some $\gamma < \pi/2$. It follows that $|\theta_i(t) - \theta_j(t)| \leq \gamma < \pi/2$ for all $\{i, j\} \in E$ and for all $t \geq T$. We now invoke Lemma 17.2(iii) to conclude the proof of statement (iii).

The converse statement can be established by noticing that all of the above inequalities and estimates are exact for a bipolar distribution of natural frequencies $\omega_i \in \{\underline{\omega}, \overline{\omega}\}$ for all $i \in \{1, \ldots, n\}$. We refer the reader for these details to the full proof in (Dörfler and Bullo, 2011). ∎

17.3.2 Synchronization of heterogeneous oscillators over weighted undirected graphs

We here adopt the following notation:

$$\|\omega\|_{2,\text{ pairs}} = \sqrt{\frac{1}{2}\sum_{i,j=1}^{n}(\omega_i - \omega_j)^2}, \qquad \text{and} \qquad \|\theta\|_{2,\text{ pairs}} = \sqrt{\frac{1}{2}\sum_{i,j=1}^{n}|\theta_i - \theta_j|^2}.$$

Theorem 17.10 (Synchronization test). *Consider the coupled oscillator model* (17.1) *with frequencies* $\omega \in \mathbb{1}_n^{\perp}$ *defined over a connected weighted undirected graph with Laplacian matrix* L. *Assume*

$$\lambda_2(L) > \lambda_{\text{critical}} \triangleq \|\omega\|_{2,\text{ pairs}}, \tag{17.18}$$

and define $\gamma_{\max} \in]\pi/2, \pi]$ *and* $\gamma_{\min} \in [0, \pi/2[$ *as the solutions to* $(\pi/2) \cdot \text{sinc}(\gamma_{\max}) = \sin(\gamma_{\min}) = \lambda_{\text{critical}}/\lambda_2(L)$. *The following statements hold:*

(i) *(phase cohesiveness:) each solution starting in* $\{\theta \in \Gamma_{\text{arc}}(\pi) \mid \|\theta\|_{2,\text{ pairs}} \leq \gamma\}$, *for* $\gamma \in [\gamma_{\min}, \gamma_{\max}]$, *remains in* $\{\theta \in \Gamma_{\text{arc}}(\pi) \mid \|\theta\|_{2,\text{ pairs}} \leq \gamma\}$ *for all times,*

(ii) *(asymptotic phase cohesiveness:) each solution starting in* $\{\theta \in \Gamma_{\text{arc}}(\pi) \mid \|\theta\|_{2,\text{ pairs}} < \gamma_{\max}\}$ *asymptotically reaches the set* $\{\theta \in \Gamma_{\text{arc}}(\pi) \mid \|\theta\|_{2,\text{ pairs}} \leq \gamma_{\min}\}$; *and*

(iii) *(asymptotic frequency synchronization:) each solution starting in* $\{\theta \in \Gamma_{\text{arc}}(\pi) \mid \|\theta\|_{2,\text{ pairs}} < \gamma_{\max}\}$ *achieves frequency synchronization.*

The proof of Theorem 17.10 follows the reasoning of the proof of Theorem 17.9 using the quadratic Lyapunov function $\|\theta\|_{2,\text{ pairs}}^2$. The full proof is in (Dörfler and Bullo, 2012, Appendix B).

17.4   Historical notes and further reading

The scientific interest in synchronization of coupled oscillators can be traced back to the work by Huygens (1673) on "an odd kind of sympathy" between coupled pendulum clocks. The model of coupled oscillator which we study was originally proposed by Winfree (1967). For complete interaction graphs, this model is nowadays known as the Kuramoto model due to the work by Kuramoto (1975, 1984). A detailed historical account is given by Strogatz (2000).

The Kuramoto model and its variations appear in the study of biological synchronization phenomena such as pacemaker cells in the heart (Michaels et al., 1987), circadian rhythms (Liu et al., 1997), neuroscience (Varela et al., 2001; Brown et al., 2003; Crook et al., 1997), metabolic synchrony in yeast cell populations (Ghosh et al., 1971), flashing fireflies (Buck, 1988), chirping crickets (Walker, 1969), and rhythmic applause (Néda et al., 2000), among others. The Kuramoto model also appears in physics and chemistry in modeling and analysis of spin glass models (Daido, 1992; Jongen et al., 2001), flavor evolutions of neutrinos (Pantaleone, 1998), and in the analysis of chemical oscillations (Kiss et al., 2002). Some technological applications include deep brain stimulation (Tass, 2003), vehicle coordination (Paley et al., 2007; Sepulchre et al., 2007; Klein et al., 2008), semiconductor lasers (Kozyreff et al., 2000; Hoppensteadt and Izhikevich, 2000), microwave oscillators (York and Compton, 1991), clock synchronization in wireless networks (Simeone et al., 2008), and droop-controlled inverters in microgrids (Simpson-Porco et al., 2013).

Our treatment borrows ideas from (Dörfler and Bullo, 2011, 2014). Recent surveys include (Strogatz, 2000; Acebrón et al., 2005; Arenas et al., 2008; Mauroy et al., 2012; Dörfler and Bullo, 2014). We refer to (Mallada et al., 2016; Gushchin et al., 2016) for a more general treatment with odd-coupling functions and with varying coupling strengths.

17.5   Exercises

E17.1   **Phase cohesiveness and arc length.** Pick $\gamma < 2\pi/3$ and $n \geq 3$. Show the following statement: if $\theta \in \mathbb{T}^n$ satisfies $|\theta_i - \theta_j| \leq \gamma$ for all $i, j \in \{1, \ldots, n\}$, then there exists an arc of length γ containing all angles, that is, $\theta \in \overline{\Gamma}_{\mathrm{arc}}(\gamma)$.

E17.2 **Order parameter and arc length.** Given $n \geq 2$ and $\theta \in \mathbb{T}^n$, the *shortest arc length* $\gamma(\theta)$ is the length of the shortest arc containing all angles, i.e., the smallest $\gamma(\theta)$ such that $\theta \in \overline{\Gamma}_{\mathrm{arc}}(\gamma(\theta))$. Given $\theta \in \mathbb{T}^n$, the *order parameter* is the centroid of $(\theta_1, \ldots, \theta_n)$ understood as points on the unit circle in the complex plane $\mathbb{C}$:

$$r(\theta)\, e^{\psi(\theta)\mathrm{i}} := \frac{1}{n} \sum_{j=1}^{n} e^{\theta_j \mathrm{i}} \ .$$

where recall $\mathrm{i} = \sqrt{-1}$. Show that

(i) if $\gamma(\theta) \in [0, \pi]$, then $r(\theta) \in [\cos(\gamma(\theta)/2), 1]$.

The order parameter magnitude r is known to measure synchronization. Show the following statements:

(iii) if all oscillators are phase-synchronized, then $r = 1$, and
(iv) if all oscillators are spaced equally on the unit circle (the so-called *splay state*), then $r = 0$.

E17.3   **Order parameter and mean-field dynamics.** Show that the Kuramoto model (17.2) is equivalent to the so-called mean-field model (17.4) with the order parameter r defined in (17.3).

E17.4   **Multiplicity of equilibria in the Kuramoto model.** A common misconception in the literature is that the Kuramoto model has a unique equilibrium set in the
phase cohesive set $\{\theta \in \mathbb{T}^n \mid |\theta_i - \theta_j| < \pi/2$ for all $\{i, j\} \in E\}$. Consider now the example of a Kuramoto oscillator network defined over a symmetric cycle graph
with identical unit weights and zero natural frequencies. The equilibria are determined by

$$0 = \sin(\theta_i - \theta_{i-1}) + \sin(\theta_i - \theta_{i+1}),$$

where $i \in \{1, \ldots, n\}$ and all indices are evaluated modulo n. Show that for $n > 4$ there are at least two disjoint equilibrium sets in the phase cohesive set
$\{\theta \in \mathbb{T}^n \mid |\theta_i - \theta_j| < \pi/2$ for all $\{i, j\} \in E\}$.

E17.5   **Potential and order parameter.** Recall $U(\theta) = \sum_{\{i,j\}\in E} a_{ij}\big(1 - \cos(\theta_i - \theta_j)\big)$. Prove $U(\theta) = \frac{Kn}{2}(1 - r^2)$ for a complete homogeneous graph with coupling strength $a_{ij} = K/n$.

E17.6   **Analysis of the two-node case.** Present a complete analysis of a system of two coupled oscillators:

$$\dot{\theta}_1 = \omega_1 - a_{12}\sin(\theta_1 - \theta_2),$$
$$\dot{\theta}_2 = \omega_2 - a_{21}\sin(\theta_2 - \theta_1),$$

where $a_{12} = a_{21}$ and $\omega_1 + \omega_2 = 0$. When do equilibria exist? What are their stability properties and their basins of attraction?

E17.7   **Averaging analysis of coupled oscillators in a semi-circle.** Consider the coupled oscillator model (17.5) with $\theta \in \Gamma_{\mathrm{arc}}(\gamma)$ for some $\gamma < \pi$. Show that the coordinate transformations $x_i = \tan(\theta_i)$, with $x_i \in \mathbb{R}$, gives the averaging system (17.6) with $b_{ij} \geq a_{ij} \cos(\gamma/2)$.

E17.8   **Phase synchronization in spring network.** Consider the spring network from Example #1 in Section 14.2 with identical oscillators, no external torques, and a connected, undirected, and weighted graph:

$$m_i \ddot{\theta}_i + d_i \dot{\theta}_i + \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j) = 0, \quad i \in \{1, \ldots, n\}.$$

Prove the phase synchronization result (in Theorem 17.5) for this spring network.

E17.9   **Synchronization on acyclic graphs.** For frequencies $\sum_{i=1}^{n} \omega_i = 0$, consider the coupled oscillator model

$$\dot{\theta}_i = -\sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j).$$

Assume the adjacency matrix A with elements $a_{ij} = a_{ji} \in \{0,1\}$ is associated to an undirected, connected, and acyclic graph. Show that the following statements are equivalent:

(i) there exists a locally stable frequency-synchronized solution in the set $\{\theta \in \mathbb{T}^n \mid |\theta_i - \theta_j| < \pi/2 \text{ for all } \{i,j\} \in E\}$,
(ii) $\left\| B^\mathsf{T} L^\dagger \omega \right\|_\infty < 1$, where B and L are the network incidence and Laplacian matrices.

Hint: Follow the derivation in Appendix 10.5.2.

E17.10   **Distributed averaging-based integral control for coupled oscillators.** Consider a set of n controllable coupled oscillators governed by the second-order dynamics

$$\dot{\theta}_i = \omega_i,$$ (E17.1a)

$$m_i \dot{\omega}_i = -d_i \omega_i - \sum_{j=1}^{n} a_{ij} \sin(\theta_i - \theta_j) + u_i,$$ (E17.1b)

where $i \in \{1, \ldots, n\}$ is the index set, each oscillator has the state $(\theta_i, \omega_i) \in \mathbb{T}^1 \times \mathbb{R}$, $u_i \in \mathbb{R}$ is a control input to oscillator i, and $m_i > 0$ and $d_i > 0$ are the inertia and damping coefficients. The oscillators are coupled through an undirected, connected, and weighted graph $G = (V, E, A)$ with node set $V = \{1, \ldots, n\}$, edge set $E \subset V \times V$, and adjacency matrix $A = A^{\mathsf{T}} \in \mathbb{R}^{n \times n}$. To reject disturbances affecting the oscillators, consider the distributed averaging-based integral control law (see Exercise E6.18)

$$u_i = -q_i,$$ (E17.2a)

$$\dot{q}_i = \omega_i - \sum_{j=1}^{n} b_{ij}(q_i - q_j),$$ (E17.2b)

where $q_i \in \mathbb{R}$ is a controller state for each agent $i \in \{1, \ldots, n\}$, and the matrix B with elements b_{ij} is the adjacency matrix of an undirected and connected graph. Your tasks are as follows:

(i)   characterize the set of equilibria $(\theta^\star, \omega^\star, q^\star)$ of the closed-loop system (E17.1)-(E17.2),
(ii) show that all trajectories converge to the set of equilibria, and
(iii) show that the phase synchronization set $\{\theta \in \mathbb{T}^n \mid \theta_i = \theta_j \text{ for all } i, j \in \{1, \ldots, n\}\}$ together with $\omega = q = \mathbb{0}_n$ is an equilibrium and that it is locally asymptotically stable.

Bibliography

R. P. Abelson. Mathematical models of the distribution of attitudes under controversy. In N. Frederiksen and H. Gulliksen, editors, *Contributions to Mathematical Psychology*, volume 14, pages 142–160. Holt, Rinehart, & Winston, 1964. ISBN 0030430100.

P.-A. Absil, R. Mahony, and B. Andrews. Convergence of the iterates of descent methods for analytic cost functions. *SIAM Journal on Control and Optimization*, 6(2):531–547, 2005. doi.

J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort, and R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77(1):137–185, 2005. doi.

D. Acemoglu and A. Ozdaglar. Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49, 2011. doi.

D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operation Research*, 38(1):1–27, 2013. doi.

R. P. Agaev and P. Y. Chebotarev. The matrix of maximum out forests of a digraph and its applications. *Automation and Remote Control*, 61(9):1424–1450, 2000. URL https://arxiv.org/pdf/math/0602059.

M. Arcak. Passivity as a design tool for group coordination. *IEEE Transactions on Automatic Control*, 52(8):1380–1390, 2007. doi.

M. Arcak, C. Meissen, and A. Packard. *Networks of Dissipative Systems: Compositional Certification of Stability, Performance, and Safety*. Springer, 2016. ISBN 978-3-319-29928-0. doi.

A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Physics Reports*, 469(3):93–153, 2008. doi.

V. I. Arnol'd. *Ordinary Differential Equations*. Springer, 1992. ISBN 3-540-54813-0. Translation of the third Russian edition by R. Cooke.

B. Avramovic, P. V. Kokotovíc, J. R. Winkelman, and J. H. Chow. Area decomposition for electromechanical models of power systems. *Automatica*, 16(6):637–648, 1980. doi.

H. Bai, M. Arcak, and J. Wen. *Cooperative Control Design*. Springer, 2011. ISBN 1461429072.

S. Baigent. Lotka-Volterra Dynamics — An Introduction. Unpublished Lecture Notes, University of College, London, Mar. 2010. URL http://www.ltcc.ac.uk/media/london-taught-course-centre/documents/Bio-Mathematics-(APPLIED).pdf.

D. Bajović, J. Xavier, J. M. F. Moura, and B. Sinopoli. Consensus and products of random stochastic matrices: Exact rate for convergence in probability. *IEEE Transactions on Signal Processing*, 61(10):2557–2571, 2013. doi.

C. Balbuena, D. Ferrero, X. Marcote, and I. Pelayo. Algebraic properties of a digraph and its line digraph. *Journal of Interconnection Networks*, 04(04):377–393, 2003. doi.

B. Bamieh, M. R. Jovanovic, P. Mitra, and S. Patterson. Coherence in large-scale networks: Dimension-dependent limitations of local feedback. *IEEE Transactions on Automatic Control*, 57(9):2235–2249, 2012. doi.

R. B. Bapat. Moore-Penrose inverse of the incidence matrix of a tree. *Linear and Multilinear Algebra*, 42(2):159–167, 1997. doi.

E. A. Barbashin and N. N. Krasovskiĭ. On global stability of motion. *Doklady Akademii Nauk SSSR*, 86(3):453–456, 1952. (In Russian).

P. Barooah and J. P. Hespanha. Estimation on graphs from relative measurements. *IEEE Control Systems*, 27(4):57–74, 2007. doi.

P. Barooah and J. P. Hespanha. Estimation from relative measurements: Electrical analogy and large graphs. *IEEE Transactions on Signal Processing*, 56(6):2181–2193, 2008. doi.

D. Bauso and G. Notarstefano. Distributed n-player approachability and consensus in coalitional games. *IEEE Transactions on Automatic Control*, 60(11):3107–3112, 2015. doi.

A. Bavelas. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, 22:725–730, 1950. **doi**.

R. Bellman. Vector Lyapunov functions. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 1(1):32–34, 1962. **doi**.

F. Benezit, V. Blondel, P. Thiran, J. Tsitsiklis, and M. Vetterli. Weighted gossip: Distributed averaging using non-doubly stochastic matrices. In *IEEE International Symposium on Information Theory*, pages 1753–1757, June 2010. **doi**.

M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005. **doi**.

A. R. Bergen and D. J. Hill. A structure preserving model for power system stability analysis. *IEEE Transactions on Power Apparatus and Systems*, 100(1):25–35, 1981. **doi**.

A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. SIAM, 1994. ISBN 978-0898713213.

D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997. ISBN 1886529019.

J. Bierkens. Non-reversible Metropolis-Hastings. *Statistics and Computing*, 26(6):1213–1228, 2016. **doi**.

N. Biggs. *Algebraic Graph Theory*. Cambridge University Press, 2 edition, 1994. ISBN 0521458978.

N. Biggs. Algebraic potential theory on graphs. *Bulletin of the London Mathematical Society*, 29(6):641–682, 1997. **doi**.

L. J. Billera and P. Diaconis. A geometric interpretation of the Metropolis-Hastings algorithm. *Statistical Science*, 16(4):335–339, 2001. **doi**.

D. Bindel, J. Kleinberg, and S. Oren. How bad is forming your own opinion? *Games and Economic Behavior*, 92:248–265, 2015. **doi**.

F. Blanchini and S. Miani. *Set-Theoretic Methods in Control*. Springer, 2015. ISBN 9783319179322.

F. Blanchini, P. Colaneri, and M. E. Valcher. Co-positive Lyapunov functions for the stabilization of positive switched systems. *IEEE Transactions on Automatic Control*, 57(12):3038–3050, 2012. **doi**.

V. D. Blondel and A. Olshevsky. How to decide consensus? A combinatorial necessary and sufficient condition and a proof that consensus is decidable but NP-hard. *SIAM Journal on Control and Optimization*, 52(5):2707–2726, 2014. .

S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424 (4-5):175–308, 2006. .

N. Bof, R. Carli, and L. Schenato. On the performance of consensus based versus Lagrangian based algorithms for quadratic cost functions. In *European Control Conference*, pages 160–165, Aalborg, Denmark, June 2016. .

B. Bollobás. *Modern Graph Theory*. Springer, 1998. ISBN 0387984887.

S. Bolognani, S. Del Favero, L. Schenato, and D. Varagnolo. Consensus-based distributed sensor calibration and least-square parameter identification in WSNs. *International Journal of Robust and Nonlinear Control*, 20(2):176–193, 2010. .

P. Bonacich. Technique for analyzing overlapping memberships. *Sociological Methodology*, 4:176–185, 1972a. .

P. Bonacich. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*, 2(1): 113–120, 1972b. .

S. P. Borgatti and M. G. Everett. A graph-theoretic perspective on centrality. *Social Networks*, 28(4):466–484, 2006. .

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994. ISBN 089871334X.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6): 2508–2530, 2006. .

U. Brandes. Centrality: concepts and methods. Slides, May 2006. URL http://vw.indiana.edu/netsci06. The International Workshop/School and Conference on Network Science.

U. Brandes and T. Erlebach. *Network Analysis: Methodological Foundations*. Springer, 2005. ISBN 3540249796.

L. Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. SIAM, 1992. ISBN 0-89871-296-3. Corrected reprint of the 1968 original.

S. Brin and L. Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks*, 30:107–117, 1998. **doi**.

E. Brown, P. Holmes, and J. Moehlis. Globally coupled oscillator networks. In E. Kaplan, J. E. Marsden, and K. R. Sreenivasan, editors, *Perspectives and Problems in Nonlinear Science: A Celebratory Volume in Honor of Larry Sirovich*, pages 183–215. Springer, 2003. **doi**.

R. A. Brualdi and H. J. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, 1991. ISBN 0521322650.

A. M. Bruckstein, N. Cohen, and A. Efrat. Ants, crickets, and frogs in cyclic pursuit. Technical Report CIS 9105, Center for Intelligent Systems, Technion, Haifa, Israel, July 1991. URL `http://www.cs.technion.ac.il/tech-reports`.

J. Buck. Synchronous rhythmic flashing of fireflies. II. *Quarterly Review of Biology*, 63(3):265–289, 1988. **doi**.

F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Princeton University Press, 2009. ISBN 978-0-691-14195-4. URL `http://www.coordinationbook.info`.

Z. Burda, J. Duda, J. M. Luck, and B. Waclaw. Localization of the maximal entropy random walk. *Physical Review Letters*, 102:160602, 2009. **doi**.

M. Cao, A. S. Morse, and B. D. O. Anderson. Agreeing asynchronously. *IEEE Transactions on Automatic Control*, 53(8):1826–1838, 2008. **doi**.

R. Carli and S. Zampieri. Network clock synchronization based on the second-order linear consensus algorithm. *IEEE Transactions on Automatic Control*, 59(2):409–422, 2014. **doi**.

R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. A PI consensus controller for networked clocks synchronization. In *IFAC World Congress*, volume 41, pages 10289–10294, 2008a. **doi**.

R. Carli, F. Fagnani, A. Speranzon, and S. Zampieri. Communication constraints in the average consensus problem. *Automatica*, 44 (3):671–684, 2008b. **doi**.

R. Carli, F. Garin, and S. Zampieri. Quadratic indices for the analysis of consensus algorithms. In *IEEE Information Theory and Applications Workshop*, pages 96–104, San Diego, USA, Feb. 2009. 🔴.

C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2):591–646, 2009. 🔴.

H. Caswell. *Matrix Population Models*. Sinauer Associates, 2 edition, 2006. ISBN 087893121X.

A. Cayley. On the Theory of Analytic Forms Called Trees. *Philosophical Magazine*, 13:19–30, 1857.

A. G. Chandrasekhar, H. Larreguy, and J. P. Xandri. Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, 88(1):1–32, 2020. 🔴.

N. D. Charkes, P. T. M. Jr, and C. Philips. Studies of skeletal tracer kinetics. I. Digital-computer solution of a five-compartment model of [18f] fluoride kinetics in humans. *Journal of Nuclear Medicine*, 19(12):1301–1309, 1978.

S. Chatterjee and E. Seneta. Towards consensus: Some convergence theorems on repeated averaging. *Journal of Applied Probability*, 14(1):89–97, 1977. 🔴.

G. Chen, X. Wang, and X. Li. *Fundamentals of Complex Networks: Models, Structures and Dynamics*. John Wiley & Sons, 2015. ISBN 978-1118718117. 🔴.

A. Cherukuri, B. Gharesifard, and J. Cortes. Saddle-point dynamics: Conditions for asymptotic stability of saddle points. *SIAM Journal on Control and Optimization*, 55(1):486–511, 2017. 🔴.

N. G. Chetaev. *The Stability of Motion*. Pergamon, 1961. Translation from Russian by M. Nadler.

G. E. Cho and C. D. Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335(1):137–150, 2001. 🔴.

N. Chopra and M. W. Spong. On exponential synchronization of Kuramoto oscillators. *IEEE Transactions on Automatic Control*, 54(2):353–357, 2009. 🔴.

J. H. Chow, editor. *Time-Scale Modeling of Dynamic Networks with Applications to Power Systems*. Lecture Notes in Control and Information Sciences. Springer, 1982. ISBN 978-3-540-12106-0.

J. H. Chow and P. Kokotović. Time scale modeling of sparse dynamic networks. *IEEE Transactions on Automatic Control*, 30(8): 714–722, 1985. doi.

J. H. Chow, J. Cullum, and R. A. Willoughby. A sparsity-based technique for identifying slow-coherent areas in large power systems. *IEEE Transactions on Power Apparatus and Systems*, 103(3):463–473, 1984. doi.

F. H. Clarke, Y. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998. ISBN 0387983368.

R. Cogburn. The ergodic theory of Markov chains in random environments. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 66(1):109–128, 1984. doi.

M. B. Cohen, A. Madry, D. Tsipras, and A. Vladu. Matrix scaling and balancing via box constrained newton's method and interior point methods. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, Oct. 2017. doi.

G. Como, K. Savla, D. Acemoglu, M. A. Dahleh, and E. Frazzoli. Robust distributed routing in dynamical networks – Part I: Locally responsive policies and weak resilience. *IEEE Transactions on Automatic Control*, 58(2):317–332, 2013. doi.

S. Coogan and M. Arcak. A compartmental model for traffic networks and its dynamical behavior. *IEEE Transactions on Automatic Control*, 60(10):2698–2703, 2015. doi.

J. Cortés. Finite-time convergent gradient flows with applications to network consensus. *Automatica*, 42(11):1993–2000, 2006. doi.

J. Cortés. Discontinuous dynamical systems. *IEEE Control Systems*, 28(3):36–73, 2008. doi.

E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*. Springer, 2014. ISBN 978-3-319-06619-6.

S. M. Crook, G. B. Ermentrout, M. C. Vanier, and J. M. Bower. The role of axonal delay in the synchronization of networks of coupled cortical oscillators. *Journal of Computational Neuroscience*, 4(2):161–172, 1997. doi.

D. Cvetković and S. K. Simić. Towards a spectral theory of graphs based on the signless Laplacian, I. *Publications de l'Institut Mathematique*, 85(99):19–33, 2009. **doi**.

H. Daido. Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions. *Physical Review Letters*, 68(7):1073–1076, 1992. **doi**.

J. M. Danskin. The theory of max-min, with applications. *SIAM Journal on Applied Mathematics*, 14(4):641–664, 1966. **doi**.

K. C. Das and P. Kumar. Some new bounds on the spectral radius of graphs. *Discrete Mathematics*, 281(1):149–161, 2004. **doi**.

S. Dashkovskiy, H. Ito, and F. R. Wirth. On a small gain theorem for ISS networks in dissipative Lyapunov form. *European Journal of Control*, 4:1–9, 2011. **doi**.

P. J. Davis. *Circulant Matrices*. John Wiley & Sons, 1979. ISBN 0-471-05771-1.

T. A. Davis and Y. Hu. The University of Florida sparse matrix collection. *ACM Transactions on Mathematical Software*, 38(1):1–25, 2011. **doi**.

M. H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974. **doi**.

P. M. DeMarzo, D. Vayanos, and J. Zwiebel. Persuasion bias, social influence, and unidimensional opinions. *Quarterly Journal of Economics*, 118(3):909–968, 2003. **doi**.

P. V. den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1):29–48, 2002. **doi**.

M. Desai and V. Rao. A characterization of the smallest eigenvalue of a graph. *Journal of Graph Theory*, 18(2):181–194, 1994. **doi**.

R. Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2000. ISBN 3642142788.

R. L. Dobrushin. Central limit theorem for nonstationary Markov chains. I. *Theory of Probability & Its Applications*, 1(1):65–80, 1956. **doi**.

J.-G. Dong and L. Qiu. Complex Laplacians and applications in multi-agent systems, 2014. URL https://arxiv.org/pdf/1406.1862.

F. Dörfler and F. Bullo. On the critical coupling for Kuramoto oscillators. *SIAM Journal on Applied Dynamical Systems*, 10(3):1070–1099, 2011. doi.

F. Dörfler and F. Bullo. Exploring synchronization in complex oscillator networks, 2012. URL http://arxiv.org/pdf/1209.1335. Extended version including proofs.

F. Dörfler and F. Bullo. Kron reduction of graphs with applications to electrical networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1):150–163, 2013. doi.

F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: A survey. *Automatica*, 50(6):1539–1564, 2014. doi.

F. Dörfler, J. W. Simpson-Porco, and F. Bullo. Electrical networks and algebraic graph theory: Models, properties, and applications. *Proceedings of the IEEE*, 106(5):977–1005, 2018. doi.

P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. Mathematical Association of America, 1984. ISBN 0883850249. URL https://math.dartmouth.edu/~doyle/docs/walks/walks.pdf.

X. Duan, S. Jafarpour, and F. Bullo. Graph-theoretic stability conditions for Metzler matrices and monotone systems. *SIAM Journal on Control and Optimization*, 59(5):3447–3471, 2021. doi.

C. L. DuBois. UCI Network Data Repository, 2008. URL http://networkdata.ics.uci.edu.

D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010. ISBN 0521195330.

B. C. Eaves, A. J. Hoffman, U. G. Rothblum, and H. Schneider. Line-sum-symmetric scalings of square nonnegative matrices. In *Mathematical Programming Essays in Honor of George B. Dantzig Part II*, pages 124–141. Springer, 1985. doi.

Y. Ebihara, D. Peaucelle, and D. Arzelier. Analysis and synthesis of interconnected positive systems. *IEEE Transactions on Automatic Control*, 62(2):652–667, 2017. doi.

L. Euler. Solutio Problematis ad Geometriam Situs Pertinentis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 8: 128–140, 1741. Also in Opera Omnia (1), Vol. 7, 1-10.

F. Fagnani. Consensus dynamics over networks. Lecture notes for Winter School on Complex Networks, INRIA. Downloaded on 12/23/2016, Jan. 2014. URL `http://www-sop.inria.fr/members/Giovanni.Neglia/complexnetworks14`.

F. Fagnani and P. Frasca. *Introduction to Averaging Dynamics over Networks*. Springer, 2017. ISBN 978-3-319-68022-4. .

F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *IEEE Journal on Selected Areas in Communications*, 26(4):634–649, 2008. .

L. Farina and S. Rinaldi. *Positive Linear Systems: Theory and Applications*. John Wiley & Sons, 2000. ISBN 0471384569.

M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973. URL `http://dml.cz/dmlcz/101168`.

D. Fife. Which linear compartmental systems contain traps? *Mathematical Biosciences*, 14(3):311–315, 1972. .

S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, 2010. .

D. M. Foster and J. A. Jacquez. Multiple zeros for eigenvalues and the multiplicity of traps of a linear compartmental system. *Mathematical Biosciences*, 26(1):89–97, 1975. .

L. R. Foulds. *Graph Theory Applications*. Springer, 1995. ISBN 0387975993.

B. A. Francis and M. Maggiore. *Flocking and Rendezvous in Distributed Robotics*. Springer, 2016. ISBN 978-3-319-24727-4.

P. Frasca and J. M. Hendrickx. On the mean square error of randomized averaging algorithms. *Automatica*, 49(8):2496–2501, 2013. .

P. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus on networks with quantized communication. *International Journal of Robust and Nonlinear Control*, 19(16):1787–1816, 2009. .

L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977. .

J. R. P. French Jr. A formal theory of social power. *Psychological Review*, 63(3):181–194, 1956. **doi**.

N. E. Friedkin. Theoretical foundations for centrality measures. *American Journal of Sociology*, 96(6):1478–1504, 1991. **doi**.

N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. In S. R. Thye, E. J. Lawler, M. W. Macy, and H. A. Walker, editors, *Advances in Group Processes*, volume 16, pages 1–29. Emerald Group Publishing Limited, 1999. ISBN 0762304529.

N. E. Friedkin and E. C. Johnsen. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Cambridge University Press, 2011. ISBN 9781107002463.

N. E. Friedkin and E. C. Johnsen. Two steps to obfuscation. *Social Networks*, 39:12–13, 2014. **doi**.

N. E. Friedkin, P. Jia, and F. Bullo. A theory of the evolution of social power: Natural trajectories of interpersonal influence systems along issue sequences. *Sociological Science*, 3:444–472, 2016. **doi**.

F. G. Frobenius. Über matrizen aus nicht negativen Elementen. 1912. **doi**. Königliche Gesellschaft der Wissenschaften.

T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11): 1129–1164, 1991. **doi**.

P. A. Fuhrmann and U. Helmke. *The Mathematics of Networks of Linear Systems*. Springer, 2015. ISBN 3319166468.

F. R. Gantmacher. *The Theory of Matrices*, volume 1 and 2. Chelsea, New York, 1959. ISBN 0-8218-1376-5 and 0-8218-2664-6. Translation of German edition by K. A. Hirsch.

C. Gao, J. Cortés, and F. Bullo. Notes on averaging over acyclic digraphs and discrete coverage control. *Automatica*, 44(8):2120–2127, 2008. **doi**.

F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. In A. Bemporad, M. Heemels, and M. Johansson, editors, *Networked Control Systems*, pages 75–107. Springer, 2010. **doi**.

M. George, S. Jafarpour, and F. Bullo. Markov chains with maximum entropy for robotic surveillance. *IEEE Transactions on Automatic Control*, 64(4):1566–1580, 2019. **doi**.

A. K. Ghosh, B. Chance, and E. K. Pye. Metabolic coupling and synchronization of NADH oscillations in yeast cell populations. *Archives of Biochemistry and Biophysics*, 145(1):319–331, 1971. doi.

G. Giorgi and S. Komlósi. Dini derivatives in optimization — Part I. *Rivista di Matematica Per Le Scienze Economiche e Sociali*, 15(1): 3–30, 1992. doi.

D. Gleich. Spectral Graph Partitioning and the Laplacian with Matlab, Jan. 2006. URL `https://www.cs.purdue.edu/homes/dgleich/demos/matlab/spectral/spectral.html`. (Last retrieved on Dec 1, 2021).

D. F. Gleich. PageRank beyond the Web. *SIAM Review*, 57(3):321–363, 2015. doi.

C. D. Godsil and G. F. Royle. *Algebraic Graph Theory*. Springer, 2001. ISBN 0387952411.

R. Goebel, R. G. Sanfelice, and A. R. Teel. *Hybrid Dynamical Systems: Modeling, Stability, and Robustness*. Princeton University Press, 2012. ISBN 9780691153896.

B. S. Goh. Global stability in two species interactions. *Journal of Mathematical Biology*, 3(3-4):313–318, 1976. doi.

B. S. Goh. Stability in models of mutualism. *American Naturalist*, pages 261–275, 1979. doi.

B.-S. Goh. *Management and Analysis of Biological Populations*. Elsevier, 1980. ISBN 978-0-444-41793-0.

M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1, Mar. 2014. URL `http://cvxr.com/cvx`.

J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer, 1990. ISBN 0387908196.

A. Gushchin, E. Mallada, and A. Tang. Phase-coupled oscillators with plastic coupling: Synchronization and stability. *IEEE Transactions on Network Science and Engineering*, 3(4):240–256, 2016. doi.

I. Gutman and W. Xiao. Generalized inverse of the Laplacian matrix and some applications. *Bulletin (Académie Serbe des Sciences et des Arts. Classe des sciences mathématiques et naturelles. Sciences mathématiques)*, 129(29):15–23, 2004. URL `http://emis.ams.org/journals/BSANU/29/2.html`.

H. Habibagahi and J. Quirk. Hicksian stability and Walras' law. *Review of Economic Studies*, 40(2):249–258, 1973. [doi].

W. H. Haddad, V. Chellaboina, and Q. Hui. *Nonnegative and Compartmental Dynamical Systems*. Princeton University Press, 2010. ISBN 0691144117.

W. M. Haddad and V. Chellaboina. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton University Press, 2008. ISBN 9780691133294.

W. Hahn. *Stability of Motion*. Springer, 1967. ISBN 978-3-642-50085-5.

F. Harary. A criterion for unanimity in French's theory of social power. In D. Cartwright, editor, *Studies in Social Power*, pages 168–182. University of Michigan, 1959. ISBN 0879442301. URL http://psycnet.apa.org/psycinfo/1960-06701-006.

F. Harary. A graph theoretic approach to matrix inversion by partitioning. *Numerische Mathematik*, 4(1):128–135, 1962. [doi].

F. Harary. *Graph Theory*. Addison-Wesley, 1969. [doi].

D. J. Hartfiel. *Markov Set-Chains*. Springer, 1998. ISBN 978-3-540-68711-5. [doi].

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. [doi].

T. Hatanaka, Y. Igarashi, M. Fujita, and M. W. Spong. Passivity-based pose synchronization in three dimensions. *IEEE Transactions on Automatic Control*, 57(2):360–375, 2012. [doi].

Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11):1867–1872, 2005. [doi].

S. Haykin. *Neural Networks and Learning Machines*. Prentice Hall, 3 edition, 2008. ISBN 0131471392.

R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulations. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002. URL http://jasss.soc.surrey.ac.uk/5/3/2.html.

J. M. Hendrickx. *Graphs and Networks for the Analysis of Autonomous Agent Systems*. PhD thesis, Departement d'Ingenierie Mathematique, Université Catholique de Louvain, Belgium, Feb. 2008.

J. M. Hendrickx and J. N. Tsitsiklis. Convergence of type-symmetric and cut-balanced consensus seeking systems. *IEEE Transactions on Automatic Control*, 58(1):214–218, 2013. doi.

J. P. Hespanha. *Linear Systems Theory*. Princeton University Press, 2009. ISBN 0691140219.

H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653, 2000. doi.

M. W. Hirsch and S. Smale. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, 1974. ISBN 0123495504.

J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998. ISBN 052162570X.

L. Hogben, editor. *Handbook of Linear Algebra*. Chapman and Hall/CRC, 2 edition, 2013. ISBN 1466507284.

Y. Hong, J. Hu, and L. Gao. Tracking control for multi-agent consensus with an active leader and variable topology. *Automatica*, 42 (7):1177–1182, 2006. doi.

Y. Hong, L. Gao, D. Cheng, and J. Hu. Lyapunov-based approach to multiagent systems with switching jointly connected interconnection. *IEEE Transactions on Automatic Control*, 52(5):943–948, 2007. doi.

J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. doi.

F. C. Hoppensteadt and E. M. Izhikevich. Synchronization of laser oscillators, associative memory, and optical neurocomputing. *Physical Review E*, 62(3):4010–4013, 2000. doi.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985. ISBN 0521386322.

R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994. ISBN 0521467136.

J. Hu and Y. Hong. Leader-following coordination of multi-agent systems with coupling time delays. *Physica A: Statistical Mechanics and its Applications*, 374(2):853–863, 2007. doi.

Y. Hu. Efficient, high-quality force-directed graph drawing. *Mathematica Journal*, 10(1):37–71, 2005.

C. Huygens. *Horologium Oscillatorium*. Paris, France, 1673.

M. Idel. A review of matrix scaling and Sinkhorn's normal form for matrices and positive maps, 2016. arXiv preprint arXiv:1609.06349.

I. C. F. Ipsen and T. M. Selee. Ergodicity coefficients defined by vector norms. *SIAM Journal on Matrix Analysis and Applications*, 32 (1):153–200, 2011. **doi**.

H. Ishii and R. Tempo. Distributed randomized algorithms for the PageRank computation. *IEEE Transactions on Automatic Control*, 55(9):1987–2002, 2010. **doi**.

H. Ishii and R. Tempo. The PageRank problem, multiagent consensus, and web aggregation: A systems and control viewpoint. *IEEE Control Systems*, 34(3):34–53, 2014. **doi**.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2010. ISBN 0691148201.

J. A. Jacquez and C. P. Simon. Qualitative theory of compartmental systems. *SIAM Review*, 35(1):43–79, 1993. **doi**.

A. Jadbabaie and A. Olshevsky. Scaling laws for consensus protocols subject to noise. *IEEE Transactions on Automatic Control*, 64(4): 1389–1402, 2019. **doi**.

A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003. **doi**.

S. Jafarpour and F. Bullo. Synchronization of Kuramoto oscillators via cutset projections. *IEEE Transactions on Automatic Control*, 64 (7):2830–2844, 2019. **doi**.

G. Jongen, J. Anemüller, D. Bollé, A. C. C. Coolen, and C. Perez-Vicente. Coupled dynamics of fast spins and slow exchange interactions in the XY spin glass. *Journal of Physics A: Mathematical and General*, 34(19):3957–3984, 2001. **doi**.

T. Kaczorek. *Positive 1D and 2D Systems*. Springer, 2001. **doi**.

A. Kashyap, T. Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007. **doi**.

E. Kaszkurewicz and A. Bhaya. Robust stability and diagonal liapunov functions. *SIAM Journal on Matrix Analysis and Applications*, 14(2):508–520, 1993. doi.

L. Katz. A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43, 1953. doi.

T. Kavitha, C. Liebchen, K. Mehlhorn, D. Michail, R. Rizzi, T. Ueckerdt, and K. A. Zweig. Cycle bases in graphs characterization, algorithms, complexity, and applications. *Computer Science Review*, 3(4):199–243, 2009. doi.

D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *IEEE Symposium on Foundations of Computer Science*, pages 482–491, Washington, DC, Oct. 2003. doi.

H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002. ISBN 0130673897.

G. Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Annalen der Physik und Chemie*, 148(12):497–508, 1847. doi.

I. Z. Kiss, Y. Zhai, and J. L. Hudson. Emerging coherence in a population of chemical oscillators. *Science*, 296(5573):1676–1678, 2002. doi.

M. S. Klamkin and D. J. Newman. Cyclic pursuit or "the three bugs problem". *American Mathematical Monthly*, 78(6):631–639, 1971. doi.

D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993. doi.

D. J. Klein, P. Lee, K. A. Morgansen, and T. Javidi. Integration of communication and control using discrete time Kuramoto models for multivehicle coordination over broadcast networks. *IEEE Journal on Selected Areas in Communications*, 26(4):695–705, 2008. doi.

J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999. doi.

G. Kozyreff, A. G. Vladimirov, and P. Mandel. Global coupling with time delay in an array of semiconductor lasers. *Physical Review Letters*, 85(18):3809–3812, 2000. doi.

D. Krackhardt. Cognitive social structures. *Social Networks*, 9(2):109–134, 1987. doi.

N. N. Krasovskiĭ. *Stability of Motion. Applications of Lyapunov's Second Method to Differential Systems and Equations with Delay*. Stanford University Press, 1963. Translation of the 1959 edition in Russian by J. L. Brenner.

Y. Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In H. Araki, editor, *Int. Symposium on Mathematical Problems in Theoretical Physics*, volume 39 of *Lecture Notes in Physics*, pages 420–422. Springer, 1975. ISBN 978-3-540-07174-7. doi.

Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer, 1984. ISBN 0387133224.

J. L. Lagrange. *Mécanique Analytique*. Chez la Veuve Desaint, Paris, 1788.

V. Lakshmikantham, V. M. Matrosov, and S. Sivasundaram. *Vector Lyapunov Functions and Stability Analysis of Nonlinear Systems*. Kluwer, 1991. ISBN 0792311523.

P. Lancaster and H. K. Farahat. Norms on direct sums and tensor products. *Mathematics of Computation*, 26(118):401–414, 1972. doi.

J. P. LaSalle. Some extensions of Liapunov's second method. *IRE Transactions on Circuit Theory*, CT-7:520–527, 1960. doi.

J. P. LaSalle. Stability theory for ordinary differential equations. *Journal of Differential Equations*, 4:57–65, 1968. doi.

J. P. LaSalle. *The Stability of Dynamical Systems*. SIAM, 1976. ISBN 9780898710229. doi.

A. J. Laub. *Matrix Analysis for Scientists and Engineers*. SIAM, 2005. ISBN 0898717906.

N. E. Leonard, T. Shen, B. Nabet, L. Scardovi, I. D. Couzin, and S. A. Levin. Decision versus compromise for animal groups in motion. *Proceedings of the National Academy of Sciences*, 109(1):227–232, 2012. doi.

P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 3(3):183–212, 1945. doi.

Z. Li and Z. Duan. *Cooperative Control of Multi-Agent Systems: A Consensus Region Approach*. CRC Press, 2014. ISBN 1466569948.

Z. Li, Z. Duan, G. Chen, and L. Huang. Consensus of multiagent systems and synchronization of complex networks: A unified viewpoint. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57(1):213–224, 2010. doi.

P. Lin and Y. Jia. Average consensus in networks of multi-agents with both switching topology and coupling time-delay. *Physica A: Statistical Mechanics and its Applications*, 387(1):303–313, 2008.

Z. Lin, B. Francis, and M. Maggiore. Necessary and sufficient graphical conditions for formation control of unicycles. *IEEE Transactions on Automatic Control*, 50(1):121–127, 2005.

Z. Lin, B. Francis, and M. Maggiore. State agreement for continuous-time coupled nonlinear systems. *SIAM Journal on Control and Optimization*, 46(1):288–307, 2007.

Z. Lin, W. Ding, G. Yan, C. Yu, and A. Giua. Leader–follower formation via complex Laplacian. *Automatica*, 49(6):1900–1906, 2013.

C. Liu, D. R. Weaver, S. H. Strogatz, and S. M. Reppert. Cellular construction of a circadian clock: Period determination in the suprachiasmatic nuclei. *Cell*, 91(6):855–860, 1997.

J. Liu, S. Mou, A. S. Morse, B. D. O. Anderson, and C. Yu. Deterministic gossiping. *Proceedings of the IEEE*, 99(9):1505–1524, 2011.

A. J. Lohwater. Introduction to Inequalities. Unpublished Lecture Notes, reproduced with permission of Marjorie Lohwater, 1982. URL http://www.mediafire.com/?1mw1tkgozzu.

S. Łojasiewicz. Sur les trajectoires du gradient d'une fonction analytique. *Seminari di Geometria 1982-1983*, pages 115–117, 1984. Istituto di Geometria, Dipartimento di Matematica, Università di Bologna, Italy.

A. J. Lotka. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences*, 6(7): 410–415, 1920.

L. Lovász. Random walks on graphs: A survey. In T. S. D. Miklós, V. T. Sós, editor, *Combinatorics: Paul Erdős is Eighty*, volume 2, pages 353–398. János Bolyai Mathematical Society, 1993. ISBN 9638022744.

L. Lovász. *Graphs and Geometry*. American Mathematical Society, 2019. ISBN 9781470450878.

E. Lovisari, F. Garin, and S. Zampieri. Resistance-based performance analysis of the consensus algorithm over geometric graphs. *SIAM Journal on Control and Optimization*, 51(5):3918–3945, 2013.

D. G. Luenberger. *Introduction to Dynamic Systems: Theory, Models, and Applications*. John Wiley & Sons, 1979. ISBN 0471025941.

D. G. Luenberger and Y. Ye. *Linear and Nonlinear Programming*. Springer, 4 edition, 2008. ISBN 9780387745022.

A. M. Lyapunov. *Obščaya zadača ob ustoĭčivosti dviženiya*. Fakul'teta i Khar'kovskogo Matematicheskogo Obshchestva, Kharkov, 1892. Translation: (Lyapunov, 1992).

A. M. Lyapunov. *The General Problem of the Stability of Motion*. Taylor & Francis, 1992. Translation from Russian by A. T. Fuller.

Q. Ma, Y.-Y. Liu, and A. Olshevsky. Optimal lockdown for pandemic control, 2022. URL http://arxiv.org/abs/2010.12923.

N. M. Maia de Abreu. Old and new results on algebraic connectivity of graphs. *Linear Algebra and its Applications*, 423(1):53–73, 2007. doi.

E. Mallada, X. Meng, M. Hack, L. Zhang, and A. Tang. Skewless network clock synchronization without discontinuity: Convergence and performance. *IEEE/ACM Transactions on Networking*, 23(5):1619–1633, 2015. doi.

E. Mallada, R. A. Freeman, and A. K. Tang. Distributed synchronization of heterogeneous oscillators on networks with arbitrary topology. *IEEE Transactions on Control of Network Systems*, 3(1):1–12, 2016. doi.

J. R. Marden, G. Arslan, and J. S. Shamma. Joint strategy fictitious play with inertia for potential games. *IEEE Transactions on Automatic Control*, 54(2):208–220, 2009. doi.

A. A. Markov. Extensions of the law of large numbers to dependent quantities. *Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete*, 15, 1906. (in Russian).

J. A. Marshall, M. E. Broucke, and B. A. Francis. Formations of vehicles in cyclic pursuit. *IEEE Transactions on Automatic Control*, 49(11):1963–1974, 2004. doi.

R. Marsli and F. J. Hall. Some properties of ergodicity coefficients with applications in spectral graph theory. *Linear and Multilinear Algebra*, 0(0):1–21, 2020. doi.

S. Martínez, J. Cortés, and F. Bullo. Motion coordination with distributed information. *IEEE Control Systems*, 27(4):75–88, 2007. doi.

I. Matei, J. S. Baras, and C. Somarakis. Convergence results for the linear consensus problem under Markovian random graphs. *SIAM Journal on Control and Optimization*, 51(2):1574–1591, 2013. **doi**.

V. M. Matrosov. On the theory of stability of motion. *Journal of Applied Mathematics and Mechanics*, 26(6):1506–1522, 1962. **doi**.

A. Mauroy, P. Sacré, and R. J. Sepulchre. Kick synchronization versus diffusive synchronization. In *IEEE Conf. on Decision and Control*, pages 7171–7183, Maui, HI, USA, Dec. 2012. **doi**.

J. C. Maxwell. On governors. *Proceedings of the Royal Society. London. Series A. Mathematical and Physical Sciences*, 16:270–283, 1868. **doi**.

W. Mei, S. Mohagheghi, S. Zampieri, and F. Bullo. On the dynamics of deterministic epidemic propagation over networks. *Annual Reviews in Control*, 44:116–128, 2017. **doi**.

Z. Meng, W. Xia, K. H. Johansson, and S. Hirche. Stability of positive switched linear systems: Weak excitation and robustness to time-varying delay. *IEEE Transactions on Automatic Control*, 62(1):399–405, 2017. **doi**.

R. Merris. Laplacian matrices of a graph: A survey. *Linear Algebra and its Applications*, 197:143–176, 1994. **doi**.

M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010. ISBN 9781400835355.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953. **doi**.

C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2001. ISBN 0898714540.

D. C. Michaels, E. P. Matyas, and J. Jalife. Mechanisms of sinoatrial pacemaker synchronization: A new hypothesis. *Circulation Research*, 61(5):704–714, 1987. **doi**.

A. N. Michel and R. K. Miller. *Qualitative Analysis of Large Scale Dynamical Systems*. Academic Press, 1977. ISBN 0-12-493850-7.

H. Minc. *Nonnegative Matrices*. John Wiley & Sons, 1988. ISBN 0471839663.

A. MirTabatabaei and F. Bullo. Opinion dynamics in heterogeneous networks: Convergence conjectures and theorems. *SIAM Journal on Control and Optimization*, 50(5):2763–2785, 2012. doi.

B. Mohar. The Laplacian spectrum of graphs. In Y. Alavi, G. Chartrand, O. R. Oellermann, and A. J. Schwenk, editors, *Graph Theory, Combinatorics, and Applications*, volume 2, pages 871–898. John Wiley & Sons, 1991. ISBN 0471532452. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.96.2577`.

J. M. Montenbruck, G. S. Schmidt, G. S. Seyboth, and F. Allgöwer. On the necessity of diffusive couplings in linear synchronization problems with quadratic cost. *IEEE Transactions on Automatic Control*, 60(11):3029–3034, 2015. doi.

L. Moreau. Stability of continuous-time distributed consensus algorithms. In *IEEE Conf. on Decision and Control*, pages 3998–4003, Nassau, Bahamas, Dec. 2004. doi.

L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005. doi.

R. M. Murray. Recent research in cooperative control of multivehicle systems. *ASME Journal of Dynamic Systems, Measurement, and Control*, 129(5):571–583, 2007. doi.

S. Muthukrishnan, B. Ghosh, and M. H. Schultz. First-and second-order diffusive methods for rapid, coarse, distributed load balancing. *Theory of Computing Systems*, 31(4):331–354, 1998. doi.

K. S. Narendra and R. Shorten. Hurwitz stability of Metzler matrices. *IEEE Transactions on Automatic Control*, 55(6):1484–1487, 2010. doi.

Z. Néda, E. Ravasz, T. Vicsek, Y. Brechet, and A.-L. Barabási. Physics of the rhythmic applause. *Physical Review E*, 61(6):6987–6992, 2000. doi.

A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009. doi.

Y. Nesterov. Efficiency of coordinate descent methods on huge-scale optimization problems. *SIAM Journal on Optimization*, 22(2): 341–362, 2012. **doi**.

M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003. **doi**.

M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010. ISBN 0199206651.

V. Nikiforov. Some inequalities for the largest eigenvalue of a graph. *Combinatorics, Probability and Computing*, 11(2):179–189, 2002. **doi**.

I. Noy-Meir. Desert ecosystems. I. Environment and producers. *Annual Review of Ecology and Systematics*, pages 25–51, 1973. **doi**.

E. P. Odum. *Fundamentals of Ecology*. Saunders Company, 1959.

R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004. **doi**.

R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma. Belief consensus and distributed hypothesis testing in sensor networks. In P. J. Antsaklis and P. Tabuada, editors, *Network Embedded Sensing and Control. (Proceedings of NESC'05 Worskhop)*, Lecture Notes in Control and Information Sciences, pages 169–182. Springer, 2006. ISBN 3540327940. **doi**.

A. Olshevsky. Linear time average consensus and distributed optimization on fixed graphs. *SIAM Journal on Control and Optimization*, 55(6):3990–4014, 2017. **doi**.

A. Olshevsky and J. N. Tsitsiklis. Convergence speed in distributed consensus and averaging. *SIAM Journal on Control and Optimization*, 48(1):33–55, 2009. **doi**.

E. E. Osborne. On pre-conditioning of matrices. *Journal of the ACM*, 7(4):338–345, 1960. **doi**.

G. F. Oster and C. A. Desoer. Tellegen's theorem and thermodynamic inequalities. *Journal of Theoretical Biology*, 32(2):219–241, 1971. **doi**.

R. W. Owens. An algorithm to solve the Frobenius problem. *Mathematics Magazine*, 76(4):264–275, 2003. **doi**.

L. Page. Method for node ranking in a linked database, Sept. 2001. URL https://www.google.com/patents/US6285999. US Patent 6,285,999.

D. A. Paley, N. E. Leonard, R. Sepulchre, D. Grunbaum, and J. K. Parrish. Oscillator models and collective motion. *IEEE Control Systems*, 27(4):89–105, 2007. doi.

J. Pantaleone. Stability of incoherence in an isotropic gas of oscillating neutrinos. *Physical Review D*, 58(7):073002, 1998. doi.

O. Perron. Zur Theorie der Matrices. *Mathematische Annalen*, 64(2):248–263, 1907. doi.

S. K. Persidskii. Concerning problem of absolute stability. *Avtomat. i Telemekh.*, (12):5–11, 1969. (Automation and Remote Control, 1970, 30:12, 1889-1895).

G. Piovan, I. Shames, B. Fidan, F. Bullo, and B. D. O. Anderson. On frame and orientation localization for relative sensing networks. *Automatica*, 49(1):206–213, 2013. doi.

M. Pirani and S. Sundaram. On the smallest eigenvalue of grounded Laplacian matrices. *IEEE Transactions on Automatic Control*, 61 (2):509–514, 2016. doi.

B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964. doi.

V. H. Poor. *An Introduction to Signal Detection and Estimation*. Springer, 2 edition, 1998. ISBN 0387941738.

M. A. Porter and J. P. Gleeson. *Dynamical Systems on Networks: A Tutorial*. Springer, 2016. ISBN 978-3-319-26641-1. doi.

M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009. URL http://www.ams.org/notices/200909/rtx090901082p.pdf.

R. Potrie and P. Monzón. Local implications of almost global stability. *Dynamical Systems*, 24(1):109–115, 2009. doi.

A. V. Proskurnikov and R. Tempo. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43:65–79, 2017. doi.

A. Rantzer. Scalable control of positive systems. *European Journal of Control*, 24:72–80, 2015. doi.

B. S. Y. Rao and H. F. Durrant-Whyte. A decentralized Bayesian algorithm for identification of tracked targets. *IEEE Transactions on Systems, Man & Cybernetics*, 23(6):1683–1698, 1993. doi.

C. Ravazzi, P. Frasca, R. Tempo, and H. Ishii. Ergodic randomized algorithms and dynamics over networks. *IEEE Transactions on Control of Network Systems*, 2(1):78–87, 2015. doi.

N. Reff. Spectral properties of complex unit gain graphs. *Linear Algebra and its Applications*, 436(9):3165–3176, 2012. doi.

W. Ren. On consensus algorithms for double-integrator dynamics. *IEEE Transactions on Automatic Control*, 53(6):1503–1509, 2008a. doi.

W. Ren. Synchronization of coupled harmonic oscillators with local interaction. *Automatica*, 44:3196–3200, 2008b. doi.

W. Ren and W. Atkins. Second-order consensus protocols in multiple vehicle systems with local interactions. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, pages 15–18, San Francisco, CA, USA, Aug. 2005. doi.

W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, 2005. doi.

W. Ren and R. W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*. Communications and Control Engineering. Springer, 2008. ISBN 978-1-84800-014-8.

W. Ren, R. W. Beard, and E. M. Atkins. Information consensus in multivehicle cooperative control. *IEEE Control Systems*, 27(2):71–82, 2007. doi.

C. W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, 21(4):25–34, 1987. doi.

L. A. Rossman. EPANET 2m, water distribution system modeling software. Technical report, US Environmental Protection Agency, Water Supply and Water Resources Division, 2000.

G. Sabidussi. Graph multiplication. *Mathematische Zeitschrift*, 72:446–457, 1960. [doi].

V. R. Saksena, J. O'Reilly, and P. V. Kokotovíc. Singular perturbations and time-scale methods in control theory: Survey 1976-1983. *Automatica*, 20(3):273–293, 1984. [doi].

S. F. Sampson. *Crisis in a Cloister*. PhD thesis, Department of Sociology, Cornell University, 1969.

W. H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2010. ISBN 0262195879.

L. Schenato and F. Fiorentin. Average TimeSynch: A consensus-based protocol for clock synchronization in wireless sensor networks. *Automatica*, 47(9):1878–1886, 2011. [doi].

M. H. Schneider and S. A. Zenios. A comparative study of algorithms for matrix balancing. *Operations Research*, 38(3):439–455, 1990. [doi].

E. Seneta. On the historical development of the theory of finite inhomogeneous Markov chains. *Mathematical Proceedings of the Cambridge Philosophical Society*, 74(3):507–513, 1973. [doi].

E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 2 edition, 1981. ISBN 0387297650.

E. Seneta. Perturbation of the stationary distribution measured by ergodicity coefficients. *Advances in Applied Probability*, 20(1): 228–230, 1988. [doi].

R. Sepulchre, D. A. Paley, and N. E. Leonard. Stabilization of planar collective motion: All-to-all communication. *IEEE Transactions on Automatic Control*, 52(5):811–824, 2007. [doi].

R. Shorten and K. S. Narendra. On a theorem of Redheffer concerning diagonal stability. *Linear Algebra and its Applications*, 431(12): 2317–2329, 2009. [doi].

D. D. Šiljak. *Large-Scale Dynamic Systems Stability & Structure*. North-Holland, 1978. ISBN 0486462854.

D. D. Šiljak. *Decentralized Control of Complex Systems*. Academic Press, 1991. ISBN 0-12-643430-1.

J. R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, 84(501):460–467, 2000. doi.

O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. H. Strogatz. Distributed synchronization in wireless networks. *IEEE Signal Processing Magazine*, 25(5):81–97, 2008. doi.

J. W. Simpson-Porco, F. Dörfler, and F. Bullo. Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica*, 49(9):2603–2611, 2013. doi.

R. Sinkhorn and P. Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2): 343–348, 1967. doi.

F. Sivrikaya and B. Yener. Time synchronization in sensor networks: A survey. *IEEE Network*, 18(4):45–50, 2004. doi.

K. D. Smith and F. Bullo. Convex optimization of the basic reproduction number. *IEEE Transactions on Automatic Control*, Oct. 2021. URL https://arxiv.org/abs/2109.07643.

S. L. Smith, M. E. Broucke, and B. A. Francis. A hierarchical cyclic pursuit scheme for vehicle networks. *Automatica*, 41(6):1045–1053, 2005. doi.

E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer, 2 edition, 1998. ISBN 0387984895.

E. D. Sontag. Monotone and near-monotone biochemical networks. *Systems and Synthetic Biology*, 1(2):59–87, 2007. doi.

D. Spielman. Graphs, vectors, and matrices. *Bulletin of the American Mathematical Society*, 54(1):45–61, 2017. doi.

M. W. Spong and N. Chopra. Synchronization of networked Lagrangian systems. In *Lagrangian and Hamiltonian Methods for Nonlinear Control 2006*, volume 366 of *Lecture Notes in Control and Information Sciences*, pages 47–59. Springer, 2007. ISBN 978-3-540-73889-3.

G. Strang. The fundamental theorem of linear algebra. *American Mathematical Monthly*, 100(9):848–855, 1993. doi.

S. H. Strogatz. From Kuramoto to Crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20, 2000. doi.

B. Sundararaman, U. Buy, and A. D. Kshemkalyani. Clock synchronization for wireless sensor networks: a survey. *Ad Hoc Networks*, 3(3):281–323, 2005. **doi**.

A. Tahbaz-Salehi and A. Jadbabaie. A necessary and sufficient condition for consensus over random networks. *IEEE Transactions on Automatic Control*, 53(3):791–795, 2008. **doi**.

A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Transactions on Automatic Control*, 55(1):225–230, 2010. **doi**.

Y. Takeuchi. *Global Dynamical Properties of Lotka-Volterra Systems*. World Scientific Publishing, 1996. ISBN 9810224710.

Y. Takeuchi, N. Adachi, and H. Tokumaru. The stability of generalized Volterra equations. *Journal of Mathematical Analysis and Applications*, 62(3):453–473, 1978. **doi**.

P. A. Tass. A model of desynchronizing deep brain stimulation with a demand-controlled coordinated reset of neural subpopulations. *Biological Cybernetics*, 89(2):81–88, 2003. **doi**.

W. Thomson and P. G. Tait. *Treatise on Natural Philosophy*. Oxford University Press, 1867.

B. Touri and A. Nedić. Product of random stochastic matrices. *IEEE Transactions on Automatic Control*, 59(2):437–448, 2014. **doi**.

D. J. Trudnowski, J. R. Smith, T. A. Short, and D. A. Pierre. An application of Prony methods in PSS design for multimachine systems. *IEEE Transactions on Power Systems*, 6(1):118–126, 1991. **doi**.

J. N. Tsitsiklis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, Massachusetts Institute of Technology, Nov. 1984.

J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986. **doi**.

E. S. Tuna. Conditions for synchronizability in arrays of coupled linear systems. *IEEE Transactions on Automatic Control*, 54(10):2416–2420, 2012. **doi**.

A. J. van der Schaft. Modeling of physical network systems. *Systems & Control Letters*, 2015. **doi**.

A. J. van der Schaft and J. Wei. A Hamiltonian perspective on the control of dynamical distribution networks. In *Lagrangian and Hamiltonian Methods for Nonlinear Control 2012*, pages 24–29. Elsevier, 2012. **doi**.

C. F. Van Loan. The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics*, 123(1):85–100, 2000. **doi**.

H. J. van Waarde, M. K. Camlibel, and H. L. Trentelman. Comments on "On the necessity of diffusive couplings in linear synchronization problems with quadratic cost". *IEEE Transactions on Automatic Control*, 62(6):3099–3101, 2017. **doi**.

F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4):229–239, 2001. **doi**.

R. S. Varga. *Matrix Iterative Analysis*. Prentice Hall, 1962. ISBN 0135655072.

T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6-7):1226–1229, 1995. **doi**.

M. Vidyasagar. *Input-Output Analysis of Large-Scale Interconnected Systems: Decomposition, Well-Posedness and Stability*. Springer, 1981. ISBN 978-3-540-10501-5.

M. Vidyasagar. *Nonlinear Systems Analysis*. SIAM, 2002. ISBN 9780898715262. **doi**.

N. K. Vishnoi. $Lx = b$, Laplacian solvers and their algorithmic applications. *Theoretical Computer Science*, 8(1-2):1–141, 2013. **doi**.

V. Volterra. Variations and fluctuations of the number of individuals in animal species living together. *ICES Journal of Marine Science*, 3(1):3–51, 1928. **doi**.

T. J. Walker. Acoustic synchrony: Two mechanisms in the snowy tree cricket. *Science*, 166(3907):891–894, 1969. **doi**.

G. G. Walter and M. Contreras. *Compartmental Modeling with Networks*. Birkhäuser, 1999. ISBN 0817640193. **doi**.

L. Wang and F. Xiao. Finite-time consensus problems for networks of dynamic agents. *IEEE Transactions on Automatic Control*, 55(4): 950–955, 2010. doi.

A. Watton and D. W. Kydon. Analytical aspects of the N-bug problem. *American Journal of Physics*, 37(2):220–221, 1969. doi.

J. T. Wen and M. Arcak. A unifying passivity framework for network flow control. *IEEE Transactions on Automatic Control*, 49(2): 162–174, 2004. doi.

H. Wielandt. Unzerlegbare, nicht negative Matrizen. *Mathematische Zeitschrift*, 52:642–648, 1950. doi.

A. T. Winfree. Biological rhythms and the behavior of populations of coupled oscillators. *Journal of Theoretical Biology*, 16(1):15–42, 1967. doi.

J. Wolfowitz. Product of indecomposable, aperiodic, stochastic matrices. *Proceedings of American Mathematical Society*, 14(5):733–737, 1963. doi.

C. Wu, R. Pines, M. Margaliot, and J.-J. E. Slotine. Generalization of the multiplicative and additive compounds of square matrices and contraction in the Hausdorff dimension. *IEEE Transactions on Automatic Control*, 2022. doi.

C. W. Wu. *Synchronization in Complex Networks of Nonlinear Dynamical Systems*. World Scientific, 2007. ISBN 978-981-270-973-8.

C. W. Wu and L. O. Chua. Application of Kronecker products to the analysis of systems with uniform linear coupling. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 42(10):775–778, 1995. doi.

T. Xia and L. Scardovi. Output-feedback synchronizability of linear time-invariant systems. *Systems & Control Letters*, 94:152–158, 2016. doi.

W. Xia and M. Cao. Sarymsakov matrices and asynchronous implementation of distributed coordination algorithms. *IEEE Transactions on Automatic Control*, 59(8):2228–2233, 2014. doi.

W. Xia and M. Cao. Analysis and applications of spectral properties of grounded Laplacian matrices for directed networks. *Automatica*, 80:10–16, 2017. doi.

L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Systems & Control Letters*, 53:65–78, 2004. **doi**.

L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Symposium on Information Processing of Sensor Networks*, pages 63–70, Los Angeles, CA, USA, Apr. 2005. **doi**.

L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, 67(1):33–46, 2007. **doi**.

R. A. York and R. C. Compton. Quasi-optical power combining using mutually synchronized oscillator arrays. *IEEE Transactions on Microwave Theory and Techniques*, 39(6):1000–1009, 1991. **doi**.

G. F. Young, L. Scardovi, and N. E. Leonard. Robustness of noisy consensus dynamics with directed communication. In *American Control Conference*, pages 6312–6317, Baltimore, USA, 2010. **doi**.

W. Yu, G. Chen, and M. Cao. Some necessary and sufficient conditions for second-order consensus in multi-agent dynamical systems. *Automatica*, 46(6):1089–1095, 2010. **doi**.

D. Zelazo. *Graph-Theoretic Methods for the Analysis and Synthesis of Networked Dynamic Systems*. PhD thesis, University of Washington, Aug. 2009.

D. Zelazo and M. Mesbahi. Edge agreement: Graph-theoretic performance bounds and passivity analysis. *IEEE Transactions on Automatic Control*, 56(3):544–555, 2011. **doi**.

D. Zelazo, S. Schuler, and F. Allgöwer. Performance and design of cycles in consensus networks. *Systems & Control Letters*, 62(1):85–96, 2013. **doi**.

Y. Zhang and Y. P. Tian. Consentability and protocol design of multi-agent systems with stochastic switching topology. *Automatica*, 45:1195–1201, 2009. **doi**.

W. X. Zhao, H. F. Chen, and H. T. Fang. Convergence of distributed randomized PageRank algorithms. *IEEE Transactions on Automatic Control*, 58(12):3255–3259, 2013. **doi**.

J. Zhu, Y. Tian, and J. Kuang. On the general consensus protocol of multi-agent systems with double-integrator dynamics. *Linear Algebra and its Applications*, 431(5-7):701–715, 2009. [doi].