Network coarse-graining through intertwining dualities

Kiang, B.A.

B. A. Kiang

Network coarse-graining through intertwining dualities

Master's thesis

September 6th 2019

Thesis supervisors:
dr. L. Avena
prof.dr. W.T.F. den Hollander

Leiden University
Mathematical Institute

Abstract

Metastability is a phenomenon where a dynamical system can move between different states that are not its global equilibrium state. On short time scales the system can find itself equilibrized in a certain region of its state space (a local equilibrium), whereas on a long time scale it will make quick transitions between new, different regions of its state space. These local equilibria are referred to as the metastable states.

One of the uses of metastability is for model reduction. In this thesis we will restrict ourselves to Markovian processes and consider the networks associated to the transitions of the Markov chains. Instead of considering a Markov process on a very large state space, one can look at the process on a reduced state space representing these metastable states. The idea is that this coarse-grained network "mimics" the behaviour of the original network. We shall give two different mathematical definitions for metastability of Markov chains.

In most cases where metastability is studied, limiting asymptotics are wielded. One must think of taking limits of large volume or low temperature. However in the paper [1] by Avena, Castell, Gaudillière and Mélot a new framework is introduced by which to describe "metastability" *without* the use of these limits. The network of transitions of a given Markov process is coarse-grained to a state space that represents probability measures which focus on different regions of the original *finite* state space (the local equilibria). It does so through the use of intertwining dualities. We say that a $n \times n$-matrix A is intertwined with a $m \times m$-matrix C with respect to a $m \times n$-matrix B if

$$BA = CB.$$

For our discussion we are given a Markov process on a finite state space with an associated transition matrix P in order to find another Markov process on a smaller state space with transition matrix $\overline{P}$ and a matrix Λ such that

$$\Lambda P = \overline{P}\Lambda$$

where the rows of Λ are probability measures on the original state space (representing the local equilibria).

In this thesis we will explore this framework based on intertwining on a toy model consisting of three nodes that we want to reduce to a network of two nodes.

The goal is to illustrate the method in [1] in this explicit model and explore which evolutions among the local equilibria can be described; how this relates to the spectrum of the transition matrix P of the Markov chain in this model; and its implications on the mixing time.

Outline of the thesis

The goal of this thesis is to test and explore on a simple toy model a novel framework, introduced in [1], to describe the evolution of local equilibria of Markov processes on finite state spaces.

The outline of the thesis is as follows:

In **Chapter 1** we present a brief introduction to metastability theory. In particular we shall explore the development of metastability theory and discuss metastability in the context of Markov chains.

Chapter 2 contains a presentation of the framework of [1] and the main results that we use for the thesis.

Chapters 3 and 4 contain original material.

In **Chapter 3** we see how this framework applies to a toy model consisting of a network with three nodes. Even though the model is simple, the phenomenology that arises is very rich and is explored in Section 3.4.

We then continue our exploration of the framework in **Chapter 4**, where we consider a random walk on the same model, but now a walk consisting of a fixed number $T \in \mathbb{N}$ steps instead of solely one step.

Contents

Abstract . . . iii

Outline of the thesis . . . v

Contents . . . vii

1 Introduction to metastability . . . 1
- 1.1 Some informative examples . . . 2
- 1.2 Models for metastability . . . 3
- 1.3 Further development of metastability theory . . . 4
- 1.4 Metastability as model reduction . . . 5
  - 1.4.1 General framework . . . 7
  - 1.4.2 Last passage . . . 8
  - 1.4.3 Trace process . . . 8
  - 1.4.4 Remarks . . . 9
- 1.5 Relation to this thesis . . . 10

2 Evolution of local equilibria through intertwining . . . 11
- 2.1 Intertwining . . . 11
- 2.2 Coarse-grained dynamics and the evolution of local equilibria . . . 12
- 2.3 Spectral characterization of solutions to the intertwining equations . . . 13
- 2.4 Approximate solutions to the intertwining equations . . . 14

3 One-step random walk on the 3-node toy model . . . 17
- 3.1 The 3-node toy model . . . 17
- 3.2 Intertwining: characterization of solutions for the one-step distribution . . . 19
- 3.3 Testing the spectral solutions on the 3-node toy model . . . 31
  - 3.3.1 Solutions from the spectral knowledge (Theorem 2.2) . . . 31
  - 3.3.2 Testing the universal solutions of Proposition 2.1 . . . 38
- 3.4 Evolution of local equilibria for the one-step simple random walk on the 3-node model . . 41
- 3.5 Analysis of the solutions . . . 43
  - 3.5.1 Behaviour of the geometric escape times . . . 43
  - 3.5.2 Bounds on the mixing times in the context of coarse graining . . . 45

4 T-step random walk on the 3-node toy model . . . 51
- 4.1 Finding the T-step distribution . . . 52
- 4.2 Characterization of intertwining solutions for odd T-step distribution . . . 55
- 4.3 Characterization of intertwining solutions for even T-step distribution . . . 59
- 4.4 Evolution of local equilibria for the T-step simple random walk on the 3-node model . . 63
- 4.5 Testing the universal spectral solutions for the T-step distribution . . . 66

5 Conclusion . . . 69

Acknowledgements . . . 71

Bibliography . . . 73

Chapter 1

Introduction to metastability

A dynamical system tends to have an equilibrium at its state of least energy. Eventually the system will converge to this stable equilibrium state. However, it could be that the system has other quasi-stable states. On different time scales the system can move between such quasi-equilibria under influence of noisy dynamics before it finally settles in its true equilibrium.
For a very big state space it can find itself equilibrized within a certain region of that state space on very short time scales, whereas on very big time scales the system can encounter a different, new region of the space that serves as quasi-equilibrium through very fast transitions.
The phenomenon described here is known as *metastability* and these quasi-equilibria are called metastable states.

Metastability is encountered in many natural occurrences, be they physical, chemical, economical, biological or other.

- Phase transitions such as the freezing of water are examples of this in physics. Here water will still be in the metastable state of the liquid phase even though its temperature is below zero degrees Celsius, before quickly freezing to the stable state of ice.
- In chemistry two compounds that react can first remain in a metastable state before transitioning to the stable state. For example we observe this when carbon dioxide and water form bicarbonate. This reaction is very slow, so the carbon dioxide and water stay in a mixture before the application of a catalytic enzyme triggers the transition to bicarbonate.
- In computational neuroscience metastability, where brain signals can persist for a long time in metastable states that are not the equilibrium state, is used to analyze how the brain responds to random environmental cues.

We shall describe metastability from the point of view of *statistical mechanics*. Statistical mechanics uses probabilistic techniques in order to describe systems of many particles and to demonstrate the relation between concepts of the macroscopic view and the description of microscopic behaviour.
In *equilibrium statistical mechanics* the Gibbs measure is used. This is a probability measure on a configuration space and is given by Boltzmann weight factors that are based on interaction Hamiltonians. A first-order phase transition is a transition of certain internal properties in a system due to a change in external variables such as temperature or pressure. This transition as a function of these external variables is discontinuous. As this variable is changed, the system stays for a long and unpredictable time in the old phase (the metastable state) before quickly transitioning to the true equilibrium. So metastability falls under *non-equilibrium statistical mechanics*, which is occupied with the dynamical properties of the system whereas the equilibrium statistical mechanics is occupied with the static properties.

In this chapter we shall first consider a few informative examples that highlight the basic characteristics of metastability, after which we present some simple mathematical models for metastability and the further development of this subject. We follow the discussion in Chapters 1 and 2 of [2]. Then we present a mathematical description of metastability on Markov chains [18]. Finally we give an outline of this thesis.

1.1 Some informative examples

The example we consider now is the *formation of rain*.
The transition from water vapour to liquid is determined by the free energy of the system. When the free energy of the gas-phase is greater than that of the liquid-phase, the water molecules will aggregate and form a rain droplet. The effective free energy $\Delta G(r)$ of a droplet (depending on its radius r) is the sum of the interfacial energy σr^{d-1} between the two phases (where $\sigma > 0$ is this effect per unit surface) and the difference between the bulk free energies $-\delta r^d$ of the two phases (where $\delta > 0$ is this effect per unit volume):

$$\Delta G(r) = \sigma r^{d-1} - \delta r^d. \tag{1.1}$$

This function is shown below:

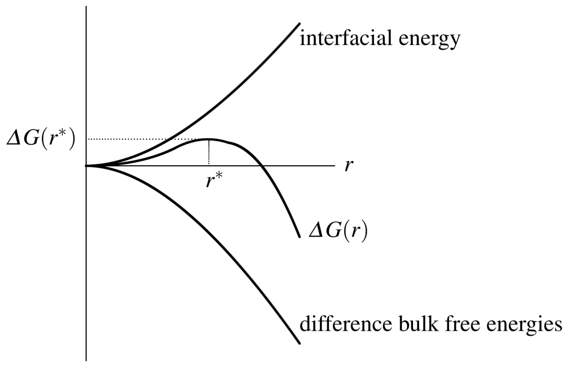


Figure 1.1: Effective free energy $\Delta G(r)$ as a function of its radius r. [2]

We can see that the effective free energy of the droplet increases with the radius r until a critical radius r^*, after which it only decreases. If a droplet of radius larger than the critical one is formed, then this radius will tend to grow. Those smaller will evaporate.
In order to grow larger than r^*, the system must temporarily violate the laws of thermodynamics and increase the free energy. These *thermal fluctuations* can produce supercritical droplets, however if r^* is very large, they will do so only rarely.
Thus we find a quick transition from the metastable gas-phase (which persists for a long time) to the stable liquid-phase after many failed attempts at the formation of a supercritical droplet.

Another example where we can see the formation of a supercritical droplet is the *kinetic Ising model* [3]. In the figure below we see a spin system that starts with all spins aligned minus and then, after the appearance of a critical droplet, transitions to the equilibrium state where all spins are aligned plus.

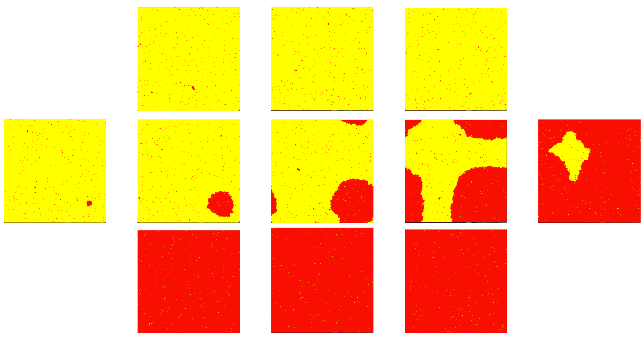


Figure 1.2: The Ising model starts out in the metastable minus phase where it stays for a long time until a supercritical droplet is formed. After that it quickly transitions to the equilibrium of all plus. [3]

The first three pictures are taken at times 471, 7482 and 13403 and represent the metastable state of the minus phase. After many unsuccessful attempts, a supercritical droplet is finally formed after a long time at time 14674. The system then quickly transitions to the stable state of the plus phase. The following pictures are taken at times 15194, 15432, 15892, 16558, 17328, 23645 and 40048 respectively.

1.2 Models for metastability

We discuss now two toy models that clearly describe metastability, namely the *Kramers model for Brownian motion in a double-well* and the *finite-state Markov process with exponentially small transition probabilities*.

One of the earliest mathematical descriptions of metastability was developed by van 't Hoff[4] in 1884 and refined by Arrhenius[5] in 1889, with the Arrhenius equation for temperature dependence of the rate constant R associated with a chemical reaction:

$$R = A \exp\left(-\frac{E}{k_B T}\right) \tag{1.2}$$

where A is the amplitude, E the activation energy of the reaction, k_B the Boltzmann constant and T the temperature.
Before the molecules can react they must achieve an energy E. The fraction of molecules that have that energy at temperature T is proportional to $\exp\left(-\frac{E}{k_B T}\right)$, so the probability of a single collision to cause a reaction is $\exp\left(-\frac{E}{k_B T}\right)$. We can see A as the average number of collisions per unit time, R as the average number of collisions per unit time that causes a reaction and $1/R$ as the average reaction time. The exponential captures the leading asymptotic behaviour.

The Arrhenius equation (1.2) turns out to be a very good formula for the average metastable crossover time of many models with stochastic dynamics in small volumes at low temperatures.
In 1940 Kramers[6] developed a toy model to mathematically verify equation (1.2) that replaces the microscopic collisions by a Brownian motion in a mesoscopic system. The model describes Brownian motion in a double-well potential using the one-dimensional diffusion equation

$$dX_t = b(X_t)dt + \sqrt{2\varepsilon} dB_t, \tag{1.3}$$

where X_t is the particle's position at time t in a drift field $b = -W'$, $W : \mathbb{R} \to \mathbb{R}$ is a double-well potential (a function with two local minima and two steep walls) and B_t is the position of the standard Brownian motion at time t.

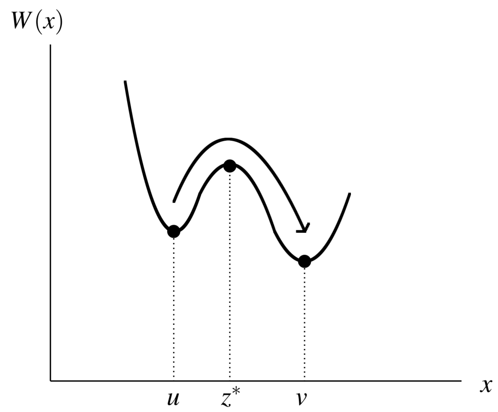


Figure 1.3: The double-well potential which has two wells at u and v. [2]

We see an example of a double-well potential above, where we have a local minimum at u, a global minimum at v and a saddle point at z^*.

The Kramers formula for average transition time from local minimum u to global minimum v through saddle point z^* is given by

$$\mathbb{E}_u[\tau_v] = [1 + o(1)] \frac{2\pi}{\sqrt{[-W''(z^*)]W''(u)}} \exp[(W(z^*) - W(u))/\varepsilon]. \tag{1.4}$$

Notice the similarity with equation (1.2), where now we have activation energy $E = W(u) - W(z^*)$, amplitude $A = \frac{2\pi}{\sqrt{[-W''(z^*)]W''(u)}}$ and inverse temperature $\frac{1}{k_B T} = \frac{1}{\varepsilon}$. Furthermore, the leading order asymptotics are exponential as we have

$$\varepsilon \ln \mathbb{E}_u[\tau_v] = [1 + o(1)](W(z^*) - W(u)) \tag{1.5}$$

as $\varepsilon \downarrow 0$.
A multi-dimensional generalisation of equation (1.4) was found by Eyring[7] and is called the Eyring-Kramers formula. We shall not present it here.

Seeing as the particle in the Kramers model spends most of the time near the two minima, the model can be simplified further to a system of two states u and v, the wells in the potential. The particle then jumps from u to v after approximately a time τ_v, the first hitting time of v starting from u. Similarly it jumps from v to u after approximately τ_u. In the limit $\varepsilon \downarrow 0$, it holds that τ_v and τ_u tend to exponentially distributed random variables. Hence we can approximate the Kramers model by a continuous-time Markov process on the state space with two states u and v. The transition rates are given by

$$\begin{aligned} c(u,v) &= \exp(-r(u,v)/\varepsilon), \quad r(u,v) = W(z^*) - W(u), \\ c(v,u) &= \exp(-r(v,u)/\varepsilon), \quad r(v,u) = W(z^*) - W(v). \end{aligned} \tag{1.6}$$

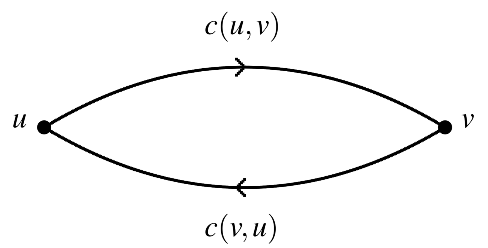


Figure 1.4: The Markov chain showing transitions between the two states u and v. [2]

The average crossover times $\mathbb{E}_u[\tau_v] = 1/c(u,v)$ and $\mathbb{E}_v[\tau_u] = 1/c(v,u)$ give us the same leading order asymptotics as in (1.5).
Of course a system can have more than two metastable states[8]. In that case we can generalize this system to a continuous-time Markov chain on a finite state space $\{m_1, ..., m_n\}$ with transition rates $c(m_i, m_j) = \exp(-r(m_i, m_j)/\varepsilon)$ for $i, j = 1, ..., n$. The theory of metastability then has the objective to find these transition rates.

1.3 Further development of metastability theory

After the Kramers model in 1940 further developments in the theory of metastability were made. These in turn led to what is known as the *pathwise approach* to metastability, proposed by Freidlin and Wentzell[8] in the 1960's and 1970's. They suggested that metastability is determined by large deviations of the random processes that govern the dynamics of the system. This theory consists of minimizing the large deviation rate function in path space in order to find the most likely path between metastable states. With that knowledge, one can ascertain the crossover time and obtain information about the system before and after the crossover. One main disadvantage of this theory however is the general difficulty with which one can discern this rate function.

A rigorous theory for metastability for particle systems was developed by Penrose and Lebowitz[9] in 1971. They identified the following three characteristics which a metastable state must possess:

1. there is only one stable state in the system,
2. the metastable state persists for a very long time,
3. the decay time from a metastable state to the stable one is much smaller than the return time from stable to metastable.

In the 1980's Davies ([10] to [14]) demonstrated that if the spectrum of the generator of a reversible Markov process contains a set of real eigenvalues that are very small and separated by a large gap from the other eigenvalues, then the system displays metastable behaviour. This *spectral approach* has as disadvantage that the assumptions on the spectrum made by Davies are generally difficult to substantiate.

In 2001 a new approach to metastability was suggested by Bovier, Eckhoff, Gayrard and Klein[15]. This *potential-theoretic approach* considers the analogy of equilibrium potentials and capacities of electric networks instead of Markov processes. In this analogy, we treat the configurations of the system as vertices in the electric network; transitions between configurations as edges; and transition probabilities as conductances of their respective edges.
In that case the hitting probability of a set of certain configurations as function of a starting configuration is viewed as the equilibrium potential where the potential is set to 1 on the target set and 0 on the vertex associated to the starting configuration. The average hitting time for metastable sets is the inverse of the capacity.
The powerful insight in this theory is that variational principles and renewal equations can be used to estimate and bound these capacities.

1.4 Metastability as model reduction

Let us now turn our attention to metastability in the context of Markov processes. Motivated by a specific example we will study how we can describe the evolution of a Markov chain by simpler dynamics, in particular one where the state space is smaller than the original yet still exhibiting the same main features of the original chain. The discussion is based on the theory presented by Landim in [18].

Define for $N \geq 1$ the set $\Lambda_N = \{1, ..., N\}^2$ and $E_{j,N}$, $0 \leq j \leq 3$, be copies of Λ_N. Then we see in the figure below that the set E_N is the union of these four squares, where each pair of neighbouring squares share exactly one point.

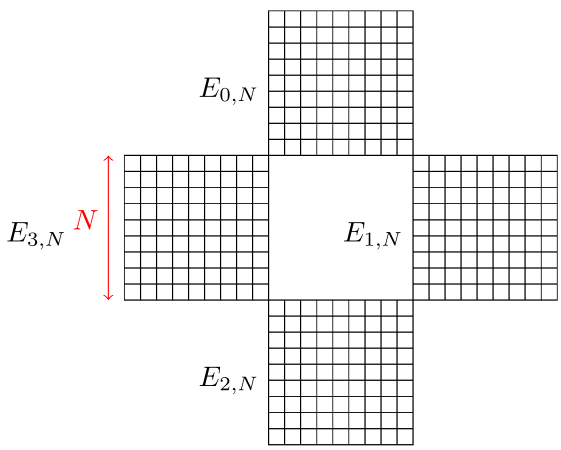


Figure 1.5: The set E_N. [18]

Let $\eta_N(t)$ be the continuous-time, irreducible Markov chain that takes values in E_N by waiting at point in E_N with mean-one exponential time and then jumping uniformly to one of the neighbours. The projection $\mathcal{Y}_N : E_N \to \{0, 1, 2, 3\}$ given by

$$\mathcal{Y}_N(\eta) = \sum_{k=0}^{3} k\chi_{E_{k,N}}(\eta),$$

where χ_A is the indicator of a set $A \subset E_N$, helps us define the evolution of the reduced model

$$Y_N(t) = \mathcal{Y}_N(\eta_N(t)).$$

It is known that the symmetric continuous-time random walk on Λ_N has mixing time of order N^2 and time that is needed to hit a point at distance N of order $\alpha_N = N^2 \log N$. Using this knowledge we can consider the time that is needed to hit one the intersection points of the $E_{j,N}$ when starting from the middle of an $E_{j,N}$. Let B be the set of these intersection points, then we denote this first hitting time as

$$H_B^N = \inf\{t \geq 0 : \eta_N(t) \in B\}.$$

Note that this hitting time is of order much larger than the mixing time of the random walk on Λ_N. This means that the chain on E_N will equilibrate in the starting $E_{j,N}$ before hitting one of the corners. After equilibration, the chain will stay for a time of order α_N in $E_{j,N}$ after which it hits an intersection point with $E_{j\pm1,N}$ (we take summation modulo 4), with probability $\frac{1}{2}$ for either of the neighbouring corners. For simplicity we take that neighbour to be $E_{j+1,N}$ and denote the corner point by ξ.
Let us now fix a sequence $(l_N)_{N\geq1}$ such that $l_N \to \infty$ and $l_N/N \to 0$. We define V_N to be the set of points in E_N that are a distance l_N or less away from ξ. After ξ is hit, $\eta_N(t)$ will move within V_N for a time of order l_N^2. As this time is much smaller than α_N, this escape time from V_N is negligible in this time scale. After escape from V_N the chain will end up in either $E_{j,N}$ or in $E_{j+1,N}$, both with probability $\frac{1}{2}$. Then it takes again a time of order α_N to hit B, meaning that the chain will equilibrate again inside the square. Thus we arrive at the initial settings and the process begins anew.

In the time scale α_N we can see that a coarse-grained process is defined by

$$\mathbf{Y}_N(t) := Y_N(t\alpha_N) = \mathcal{Y}_N(\eta_N(t\alpha_N)).$$

This is a continuous-time process with values in $\{0, 1, 2, 3\}$, where the holding rates are given by some $\lambda > 0$ and transition probabilities are $p(j, j \pm 1) = \frac{1}{2}$. Because of the memory loss that arises from the equilibration, we expect this process to converge to a Markov chain. The question that we must pose is: how exactly does $\mathbf{Y}_N(t)$ converge to a Markov chain? We see in the figure below a realization of this process. The walk stays for a time of order α_N in a state in $\{0, 1, 2, 3\}$, after which it makes very short excursions around a corner point (these excursions are represented by the black rectangles).

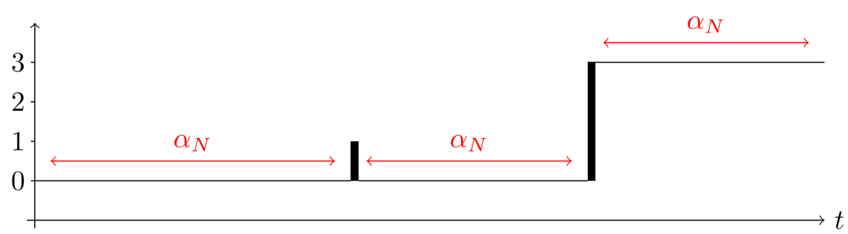


Figure 1.6: A trajectory of $\mathbf{Y}_N(t)$. [18]

These fluctuations are the reason that $\mathbf{Y}_N(t)$ does not converge to a Markov chain in any of the Skorohod topologies. In order to fix this, we shall alter the trajectories of $\mathbf{Y}_N(t)$ by removing these fluctuations.

First let us define $\mathcal{E}_N^j$ to be the set of points in $E_{j,N}$ that are at least a distance l_N away from the faces of $E_{j,N}$. This can be seen in the figure below.

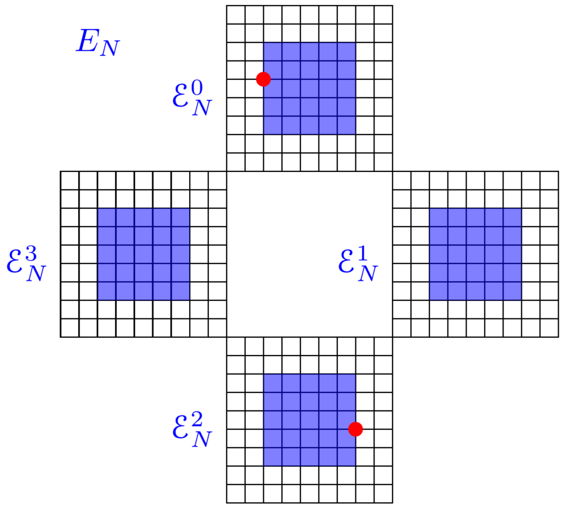


Figure 1.7: The sets $\mathcal{E}_N^j$ within each $E_{j,N}$. [18]

There are two ways in which we will alter the trajectories to obtain convergence in the Skorohod topology, namely by last passage and by the trace process. Before presenting these, we lay down a general framework.

1.4.1 General framework

Let $(E_N)_{N\geq 1}$ be a sequence of finite state spaces and $\eta_N(t)$ a continuous-time irreducible Markov chain with values in E_N. It has a generator $\mathcal{L}_N$ that acts on functions $f : E_N \to \mathbb{R}$ by

$$(\mathcal{L}_N f)(\eta) = \sum_{\xi \in E_N} R_N(n,\xi)[f(\xi) - f(\eta)];$$

and it has a unique stationary distribution π_N. For a set $A \subset E_N$, the first hitting time of A and the first return time to A are given respectively by

$$H_A := \inf\{t \geq 0 : \eta_N(t) \in A\}, \quad H_A^+ = \inf\{t \geq \tau_1 : \eta_N(t) \in A\}$$

where $\tau_1 = \inf\{t \geq 0 : \eta_N(t) \neq \eta_N(0)\}$.
Now for $n > 1$ we create a partition of E_N consisting of disjoint sets $\mathcal{E}_N^1, ..., \mathcal{E}_N^n, \Delta_N$. We also define

$$\mathcal{E}_N := \bigcup_{k=1}^{n} \mathcal{E}_N^k.$$

For $S = \{1, ..., n\}$ we define the projection that sends $\eta \in E_N$ to $S \cup \{\delta\}$ by

$$\Phi_N(\eta) := \sum_{k=1}^{n} k\chi_{\mathcal{E}_N^k}(\eta) + \delta\chi_{\Delta_N}(\eta).$$

Then

$$X_N(t) := \Phi_N(\eta_N(t))$$

is the process with values in $S \cup \{\delta\}$ that is determined by $\eta_N(t)$.

1.4.2 Last passage

In the method of last passage, we alter the trajectories through removing the fast fluctuations and only considering the last set $\mathcal{E}_N^k$ that is visited.

First we define the left limit of η_N at time t by

$$\eta_N(t-) = \lim_{s \uparrow t} \eta_N(s).$$

Then the process $X_N^V(t)$ is defined by

$$X_N^V(t) := \Phi_N(\eta_N(v_N(t)))$$

where

$$v_N(t) = \begin{cases} t & \text{if } \eta_N(t) \in \mathcal{E}_N \\ w_N(t)- & \text{otherwise,} \end{cases}$$

and

$$w_N(t) := \sup\{s \leq t : \eta_N(s) \in \mathcal{E}_N\}.$$

We see that $w_N(t)$ is the last time that η_N was in one of the $\mathcal{E}_N^k$ before time t. Thus $v_N(t)$ will always give us a time that the chain η_N is in one of the $\mathcal{E}_N^k$ and we get rid of the fast fluctuations that take place in Δ_N.
Suppose we start in $\mathcal{E}_N^k$ and then visit $\mathcal{E}_N^j$. Then in the time interval $[0, H_{\mathcal{E}_N^j})$, the process X_N^V remains constantly equal to k where $H_{\mathcal{E}_N^j}$ is of the order α_N. Since we have now got rid of the fast fluctuations, the process X_N^V in the time scale α_N converges to a Markov chain in the Skorohod topology.

We give the definition of metastability in this case:
Definition 1.1 (LP metastability)**.** The Markov chain $\eta_N(t)$ is *LP-metastable* in the time scale α_N if there exists a partition $\{\mathcal{E}_N^1, ..., \mathcal{E}_N^n, \Delta_N\}$ of E_N and a continuous-time Markov chain $\mathbf{X}(t)$ with values in $S = \{1, .., n\}$ such that

1. for all $k \in S$ and all $(\xi_N)_{N \geq 1}$ where $\xi_N \in \mathcal{E}_N^k$: starting from ξ_N, the process $\mathbf{X}_N^V(t) = X_N^V(t\alpha_N)$ converges to $\mathbf{X}(t)$ in the Skorohod topology;
2. for all $t > 0$:
$$\lim_{N \to \infty} \max_{\eta \in \mathcal{E}_N} \mathbb{E}_\eta^N \left[\int_0^t \chi_{\Delta_N}(\eta_N(s\alpha_N)) ds \right] = 0,$$
i.e. we can disregard the time in Δ_N. Here $\mathbb{E}_\eta^N$ is the expectation with respect to the probability measure $\mathbb{P}_\eta^N$ induced by $\eta_N(t)$ starting from η on the space $D([0, \infty), E_N)$ of càdlàg trajectories endowed with the Skorohod topology.

The problem with this alteration is that the process $\eta_N(v_N(t))$ is not markovian, making things hard to prove.

1.4.3 Trace process

Now we will remove the fast fluctuations by using the trace process.

We see for a non-empty, proper subset $F \subsetneq E_N$ that

$$T_F(t) := \int_0^t \chi_F(\eta_N(s)) ds$$

is the time that η_N spends in F on the interval $[0, t]$. Its generalized inverse is given by

$$S_F(t) := \sup\{s \geq 0 : T_F(s) \leq t\}.$$

As the chain is irreducible, $S_F(t)$ is almost surely finite for all $t > 0$.
Now we can define the trace of $\eta_N(t)$ on the set F to be

$$\eta_N^F(t) := \eta_N(S_F(t)).$$

It is a known result that $\eta_N^F(t)$ is an irreducible, cotinuous-time Markov chain that takes values in F. Its jump rates are

$$R_N^F(\eta, \xi) := \lambda(\eta)\mathbb{P}_\eta[H_F^+ = H_\eta], \quad \eta, \xi \in F, \quad \eta \neq \xi,$$

where for the jump rates $R_N(\eta, \xi)$ of the original chain we have holding rates $\lambda(\eta) = \sum_{\xi \in E_N} R_N(\eta, \xi)$. We can get a better understanding of these quantities from the figure below.

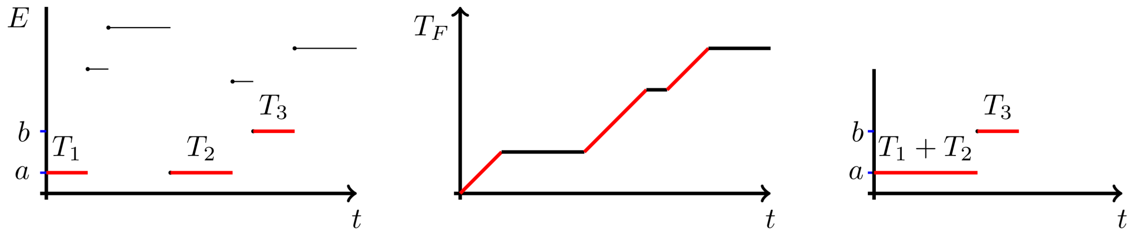


Figure 1.8: The first figure shows the trajectory of a Markov chain with values in a state space E, where we focus on the subset $F = \{a, b\}$. The second graph is $T_F(t)$. We remark that if the chain is not in F, then $T_F(t)$ is constant. Otherwise it increases linearly. The third graph shows the trace process. We can view this as the process where time is frozen when the chain hits F^C until the time when it is again in F and time continues normally. This means that we remove excursion to F^C from the trajectory and then push back what remains. [18]

For our situation we study $\eta_N^{\mathcal{E}_N}(t)$, the trace of η_N on $\mathcal{E}_N$. Defining a new projection

$$\Psi_N(\eta) = \sum_{k=1}^{n} k\chi_{\mathcal{E}_N^k}(\eta)$$

that only acts on $\mathcal{E}_N$, we can consider a new process

$$X_N^T(t) := \Psi_N(\eta^{\mathcal{E}_N}(t))$$

to give the following definition of metastability. (Note that X_N^T is the trace of X_N on S.)
Definition 1.2 (Trace metastability)**.** The Markov chain $\eta_N(t)$ is *trace-metastable* in the time scale α_N if there exists a partition $\{\mathcal{E}_N^1, ..., \mathcal{E}_N^n, \Delta_N\}$ of E_N and a continuous-time Markov chain $\mathbf{X}(t)$ with values in $S = \{1, .., n\}$ such that

1. for all $k \in S$ and all $(\xi_N)_{N\geq 1}$ where $\xi_N \in \mathcal{E}_N^k$: starting from ξ_N, the process $\mathbf{X}_N^T(t) = X_N^T(t\alpha_N)$ converges to $\mathbf{X}(t)$ in the Skorohod topology;
2. for all $t > 0$:

$$\lim_{N\to\infty} \max_{\eta \in \mathcal{E}_N} \mathbb{E}_\eta^N \left[\int_0^t \chi_{\Delta_N}(\eta_N(s\alpha_N))ds \right] = 0,$$

   i.e. we can disregard the time in Δ_N. Here $\mathbb{E}_\eta^N$ is the expectation with respect to the probability measure $\mathbb{P}_\eta^N$ induced by $\eta_N(t)$ starting from η on the space $D([0, \infty), E_N)$ of càdlàg trajectories endowed with the Skorohod topology.

1.4.4 Remarks

Remark 1. Note that the given definitions differ from our description of metastability in the earlier sections. There metastability is used to describe the transition from a metastable state to a stable one. In these definitions of metastability that would match the case where the state space consists of one transient and one absorbing state.

Remark 2. In both definitions the second condition tells us that the transition between two metastable states is very fast, as the time spent in Δ_N is negligible. This corresponds with what we discussed in earlier sections where it was known that in the metastable time-scale the transitions happen very quickly.

Remark 3. In the examples in Section 1.1 we saw that metastability happens due to a present energy barrier. If the system overcomes this barrier, it will reach a new section of the state space. In the example in this section it is not an energy barrier but a bottleneck that causes metastability.
In this thesis we shall also study a model where metastability arises due to a bottleneck.

Remark 4. The approach of Landim can be used to investigate dynamics that are represented as a Markov chain on a state space of fixed, finite size.
Transition path theory ([19] to [22]), where one studies the statistical properties of the portions of the path of a Markov process that correspond to transitions between two pre-specified subsets of the state space, is an example of this. The same goes for the *intertwining method* ([1], [3], [23], [24]), which we will study in greater detail in the next chapter.

1.5 Relation to this thesis

In the paper [1] by Avena, Castell, Gaudillière and Mélot a novel general approach is proposed by which to describe the *evolution of local equilibria of a finite Markov chain*. Different metastable states are probability measures concentrated on different parts of a possibly very large network. Then the interest lies in coarse-graining this network, or in other words to coalesce vertices of the network in order to consider a simpler Markov process on a reduced state space.
Until now we have always assumed an asymptotic regime such that we could examine metastability, be it large volume (as in [18]) or low temperature limits. Now, however, the novelty in the work of Avena et al. lies in the fact that we can in principle **describe metastable behaviour outside of these asymptotic regimes**.

The original contents of this thesis consist of working out the framework of [1] on explicit models and exploring the resulting phenomenology.

Chapter 2

Evolution of local equilibria through intertwining

In this section we will present the results of [1] that form the basis for this thesis.

Let $\mathcal{X}$ be a finite state space of size $|\mathcal{X}| = n \in \mathbb{N}$. We consider an irreducible discrete time Markov process $\{X_t, t \geq 0\}$ on $\mathcal{X}$ that is characterized by an irreducible stochastic transition matrix P. The general goal is to observe the evolution of distributions that are localized in different regions of its state space.

We will do this through intertwining dualities, the concept of which is introduced in Section 2.1. In Section 2.2 we will state the main general theorem which will explain how to make sense of the evolution of the local equilibria ("metastable states"). In Section 2.3 we will observe how all this is related to the spectrum of P. And Section 2.4 summarizes further results developed in [1] in case an analytic explicit analysis is out of reach.

2.1 Intertwining

We will now explore the tool of intertwining and discuss in which context we shall use it for our purposes.
Definition 2.1 (Intertwining)**.** Consider a Markov process $\overline{X}$ on state space $\overline{\mathcal{X}}$ of size $m \in \mathbb{N}$ with transition matrix $\overline{P}$. The Markov process X is *intertwined* with $\overline{X}$ with respect to a matrix $\Lambda : \overline{\mathcal{X}} \times \mathcal{X} \to [0, 1]$ if

$$\Lambda P = \overline{P}\Lambda. \tag{2.1}$$

Note that this is in general not a symmetric relation, unless Λ is invertible.
Furthermore we can choose for the rows of Λ to be probability measures on $\mathcal{X}$:

$$\Lambda(\overline{x}, \cdot) = \nu_{\overline{x}}.$$

In the literature intertwining has appeared in many contexts. Rogers and Pitman[16] use intertwining to state identities in laws for diffusion processes. Here they considered measures $\nu_{\overline{x}}$ with disjoint support.
Diaconis and Fills[17] examined intertwining without this restriction in order to study Markov chains. They did this in order to construct strong stationary times and to control convergence rates to equilibrium.
In these studies there was always the assumption that the size m of $\overline{\mathcal{X}}$ was much larger than or equal to the size n of $\mathcal{X}$. For our purposes we take m to be smaller than n.

In fact, for the intents of [1] our goal is:
Given $(X, P, \mathcal{X})$, find m probability measures $\nu_{\overline{x}} = \Lambda(\overline{x}, \cdot)$ (where $m < n = |\mathcal{X}|$) that are concentrated on different regions of $\mathcal{X}$ and find a transition matrix $\overline{P}$ on $\{\overline{1}, ..., \overline{m}\}$ such that

$$\Lambda P = \overline{P}\Lambda,$$

which is equivalent to

$$\nu_{\overline{x}} P = \Lambda P(\overline{x}, \cdot) = \overline{P}\Lambda(\overline{x}, \cdot) = \sum_{\overline{y} \in \overline{\mathcal{X}}} \overline{P}(\overline{x}, \overline{y}) \nu_{\overline{y}}.$$

Remark. Note that if we take all $\nu_{\overline{x}}$ to be equal to the invariant measure π of P, we see that the intertwining equation is always satisfied.
For all $x \in \mathcal{X}$ and $\overline{x} \in \overline{\mathcal{X}}$ we observe then that:

$$\begin{aligned}
\overline{P}\Pi(\overline{x}, x) &= \sum_{\overline{y} \in \overline{\mathcal{X}}} \overline{P}(\overline{x}, \overline{y}) \pi(x) \\
&= \left(\sum_{\overline{y} \in \overline{\mathcal{X}}} \overline{P}(\overline{x}, \overline{y}) \right) \pi(x) \\
&= \pi(x) \\
&= \pi P(x) \\
&= \Pi P(\overline{x}, x).
\end{aligned}$$

We will now show that if the coarse-grained process has thermalized, then so has the original.
Remark (Thermalization of coarse-grained process implies thermalization of the original process)**.** Suppose that the process X has a unique stationary distributions π, the process $\overline{X}$ has unique stationary distribution $\overline{\pi}$ and Λ is such that $\Lambda P = \overline{P}\Lambda$. Then

$$\overline{\pi}\Lambda = \pi.$$

In other words thermalization of the coarse-grained process implies thermalization of the original process.

Proof. We see that

$$\begin{aligned}
(\overline{\pi}\Lambda P)_a &= \sum_i (\overline{\pi}\Lambda)_i P_{i,a} \\
&= \sum_i \sum_j \overline{\pi}_j \Lambda_{j,i} P_{i,a} \\
&= \sum_i \sum_j \overline{\pi}_j \overline{P}_{j,i} \Lambda_{i,a} \\
&= \sum_i \overline{\pi}_i \Lambda_{i,a} \\
&= (\overline{\pi}\Lambda)_a .
\end{aligned}$$

□

2.2 Coarse-grained dynamics and the evolution of local equilibria

By finding solutions of the intertwining equations as described in the previous section, we can describe the evolution of local equilibria. In this way we can discuss metastability outside any asymptotic regime. From one local equilibrium $\nu_{\overline{x}}$ the process X evolves to another equilibrium $\nu_{\overline{y}}$ chosen according to $\overline{P}$ after a random time that is also determined by $\overline{P}$.
The following theorem (Proposition 6 of [1]) explains this in rigorous terms.

Theorem 2.1 (Evolution of local equilibria in intertwining context)**.** If X is intertwined with $\overline{X}$ with respect to measures $\nu_{\overline{x}} = \Lambda(\overline{x}, \cdot)$, then there exists a filtration $\mathcal{F}$ such that X is $\mathcal{F}$-adapted and for each $\overline{x} \in \overline{\mathcal{X}}$ there exist a stopping time $T_{\overline{x}}$ and a random variable $\overline{Y}_{\overline{x}}$ with values in $\overline{\mathcal{X}} \backslash \{\overline{x}\}$ and law $\mathbb{P}$ that satisfy

1. $T_{\overline{x}}$ is geometrically distributed with parameter $1 - \overline{P}(\overline{x}, \overline{x})$;
2. $\nu_{\overline{x}}$ is stationary until time $T_{\overline{x}}$, which means that
$$P_{\nu_{\overline{x}}}(X_t = \cdot | T_{\overline{x}} > t) = \nu_{\overline{x}}(\cdot); \tag{2.2}$$
3. for all $\overline{y} \in \overline{\mathcal{X}} \backslash \overline{x}$ we have
$$\mathbb{P}(\overline{Y}_{\overline{x}} = \overline{y}) = \frac{\overline{P}(\overline{x}, \overline{y})}{1 - \overline{P}(\overline{x}, \overline{x})}; \tag{2.3}$$
4. $$P_{\nu_{\overline{x}}}(X(T_{\overline{x}}) = \cdot | \overline{Y}_{\overline{x}} = \overline{y}) = \nu_{\overline{y}}(\cdot); \tag{2.4}$$
5. $(X(T_{\overline{x}}), \overline{Y}_{\overline{x}})$ and $T_{\overline{x}}$ are independent.

This is very alike the heuristic description of Penrose and Lebowitz in 1971 that was discussed in Section 1.3:

1. First of all since we consider an irreducible and periodic chain on a finite state space there is a unique invariant measure. This is the unique stable state of the system.
2. Secondly a metastable state $\overline{x}$ persists for a random time $T_{\overline{x}}$, which in general can be very long. This we see in (2.2).
3. The decay time from metastable state to another is much shorter. In fact the transition is immediate as we can see from (2.4).

This general theorem is interesting provided we can solve the intertwining *and* if the corresponding measures $\nu_{\overline{x}}$ are good for our purpose, in the sense that they are localized in different regions of the state space and separated as much as possible.

2.3 Spectral characterization of solutions to the intertwining equations

We are interested in finding non-degenerate solutions of the intertwining equations, i.e. where the $\nu_{\overline{x}}$ are linearly independent, and preferably such that the $\nu_{\overline{x}}$ have as little overlap as possible.
We noticed earlier that if we take all $\nu_{\overline{x}}$ to be the invariant measure, we retrieve a trivial solution. Theorem 2.2 (Lemma 10 of [1]) shows us that if non-trivial intertwining solutions $(\Lambda, \overline{P})$ exist and we have knowledge of the spectrum of P, we can find a matrix $C = (C(\overline{x}, j))_{\overline{x} \in \overline{\mathcal{X}}, j \in J}$ (where J is a subset of $\{0, 1, ..., n-1\}$ of size m) such that we can write the rows $\nu_{\overline{x}}$ of Λ as perturbations of the invariant measure by this C.

Theorem 2.2 (Spectral characterization of non-trivial intertwining equations)**.** Denote the n eigenvalues of P by
$$1 = \theta_0 > \theta_1 \geq ... \geq \theta_{n-1} \geq -1$$
and denote the set $[n] = \{0, 1, 2, ..., n-1\}$.
If Λ is of rank m such that $\Lambda P = \overline{P} \Lambda$, then there exist an orthonormal basis of eigenvectors $\{\mu_j : 0 \leq j < n\}$ of P such that
$$\mu_j P = \theta_j \mu_j$$
for all $0 \leq j < n$; a subset $J \subset [n]$ such that $0 \in J$ and $|J| = m$; and an invertible matrix

$C = (C(\overline{x}, j))_{\overline{x}\in\overline{\mathcal{X}}, j\in J}$ such that $C(\overline{x}, 0) = 1$ for all $\overline{x} \in \overline{\mathcal{X}}$ and

$$\nu_{\overline{x}} = \sum_{j\in J} C(\overline{x}, j)\mu_j \tag{2.5}$$

for all $\overline{x} \in \overline{\mathcal{X}}$ and

$$\overline{P}C(\cdot, j) = \theta_j C(\cdot, j) \tag{2.6}$$

for all $j \in J$.

Note that $\mu_0 = \pi$ and $C(\overline{x}, 0) = 1$, which means that indeed $\nu_{\overline{x}}$ adds the other eigenvectors μ_j of P to π through the perturbation matrix C.

Theorem 2.2 gives us a spectral characterization of solutions to our intertwining problem, but tells us nothing about how "good" these solutions are.
The subsequent proposition (Lemma 11 of [1]) gives us a recipe for how we can find a reversible and stochastic irreducible matrix $\overline{P}$ with a spectrum that is contained in that of P.
If we find this explicit $\overline{P}$, we can make the task of finding intertwining solutions easier as we must only find Λ instead of both Λ and $\overline{P}$. In particular the Λ's that result should be good in the sense that they are localized in different regions of the state space. However no assertions can be made about the separation between the measures.

Proposition 2.1 (Universal solutions with good properties)**.** For any

$$1 = \theta_0 > \theta_1 \geq ... \geq \theta_{m-1} \geq 0$$

there always exists a reversible and irreducible stochastic matrix $\overline{P}$ with those as eigenvalues.

We show how the matrix $\overline{P}$ in Lemma 2.2 is constructed.
For $1 \leq k \leq m$ define

$$\Sigma_k = \sum_{j<k} \theta_j. \tag{2.7}$$

Then $\overline{P}$ is given by

$$\overline{P} = \begin{pmatrix} \frac{\Sigma_1+\theta_1}{1\times 2} & \frac{\Sigma_1-\theta_1}{2\times 3} & \frac{\Sigma_1-\theta_1}{3\times 4} & \cdots & \frac{\Sigma_1-\theta_1}{(m-1)m} & \frac{\Sigma_1-\theta_1}{m} \\ \frac{\Sigma_1-\theta_1}{1\times 2} & \frac{\Sigma_2+2^2\theta_2}{2\times 3} & \frac{\Sigma_2-2\theta_2}{3\times 4} & \cdots & \frac{\Sigma_2-2\theta_2}{(m-1)m} & \frac{\Sigma_2-2\theta_2}{m} \\ \frac{\Sigma_1-\theta_1}{1\times 2} & \frac{\Sigma_2-2\theta_2}{2\times 3} & \frac{\Sigma_3+3^2\theta_3}{3\times 4} & \cdots & \frac{\Sigma_3-3\theta_3}{(m-1)m} & \frac{\Sigma_3-3\theta_3}{m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\Sigma_1-\theta_1}{1\times 2} & \frac{\Sigma_2-2\theta_2}{2\times 3} & \frac{\Sigma_3-3\theta_3}{3\times 4} & \cdots & \frac{\Sigma_{m-1}+(m-1)^2\theta_{m-1}}{(m-1)m} & \frac{\Sigma_{m-1}-(m-1)\theta_{m-1}}{m} \\ \frac{\Sigma_1-\theta_1}{1\times 2} & \frac{\Sigma_2-2\theta_2}{2\times 3} & \frac{\Sigma_3-3\theta_3}{3\times 4} & \cdots & \frac{\Sigma_{m-1}-(m-1)\theta_{m-1}}{(m-1)m} & \frac{\Sigma_m}{m} \end{pmatrix}. \tag{2.8}$$

Remark. Further we observe that in order to construct the $\overline{P}$ of Lemma 2.2, we must have full knowledge of the spectrum of P. This however, may prove too difficult in cases where we study processes on very large state spaces.
Nonetheless in [1] an algorithm based on random spanning forests was introduced to build approximate solutions to the intertwining equations. It is not the goal of this thesis to dig into this algorithmic process. Instead we shall focus on exploring analytically the solutions to a simple model.

2.4 Approximate solutions to the intertwining equations

Of course in many cases, such as when we consider immense state spaces, "good" exact solutions to the intertwining equations can be too difficult to find. In sections 5.2.1 and 5.2.2 of [3] a way of building approximate "good" solutions is presented. This method uses an algorithm and the concept of random forests to construct a measure-valued process on a smaller state space with dynamics resembling those of the original process.

First we consider a *directed and weighted graph* $G = (V, E, w)$ on state space of size $|V| = n \in \mathbb{N}$, where E is the set of directed edges and $w : V \times V \to \mathbb{R}_{>0}$ is a strictly positive weight function. A *rooted spanning forest* ϕ is a subgraph of G without cycles, where a *root* of ϕ are the vertices $x \in V$ such that for all $y \in V$ the edge (x, y) is not in ϕ. The set of roots of ϕ is denoted by $R(\phi)$ and the set of all possible rooted spanning forests by $\mathcal{F}$. We can now define what we call the *forest measure* for fixed positive $q \in \mathbb{R}_{>0}$, which is the law of the $\mathcal{F}$-valued random variable known as the *random forest* Φ_q:

$$\mathbb{P}(\Phi_q = \phi) = \frac{w(\phi) q^{|R(\phi)|}}{\sum\limits_{\psi \in \mathcal{F}} w(\psi) q^{|R(\psi)|}}.$$

Here $w(\phi) = \prod\limits_{e \in \phi} w(e)$ is the weight of $\phi \in \mathcal{F}$ and $|R(\phi)|$ is the number of roots of ϕ.

The algorithm of [3] consists of the following steps: Given an irreducible and reversible graph G on state space of size $|V| = n$ with associated Markov process X and invariant measure μ, we must

1. Choose $m \leq n$ and partition the graph into $P(G) = [A_1, ..., A_m]$.
2. Define new vertex set $\overline{V} = \{\overline{1}, ..., \overline{m}\}$.
3. Set the Λ-matrix as
$$\Lambda(\overline{x}, \cdot) = \nu_{\overline{x}}(\cdot) := \mu(\cdot|A_{\overline{x}})$$
for $\overline{x} \in \overline{V}$ and where $\mu(\cdot|A_{\overline{x}})$ is the invariant measure μ conditioned to $A_{\overline{x}} \subset V$.
4. We define $T_{q'}$ as an independent exponential random variable of parameter $q' > 0$. For $\overline{x}, \overline{y} \in \overline{V}$, the new process is given by the law
$$\overline{P}_{q'}(\overline{x}, \overline{y}) := P_{\nu_{\overline{x}}}\left[X(T_{q'}) \in A_{\overline{y}}\right].$$

The *randomization* of this deterministic algorithm is implemented by choosing

$$m = m(q) = |R(\Phi_q)|$$

and

$$P(G) = P(\Phi_q) := [A_1, ..., A_{m(q)}].$$

Some comments on this procedure:

- If the constructed $(\Lambda, \overline{P}_{q'})$ is close to a solution of the intertwining equations, then we can consider the resulting network $\overline{G}$ as a coarse-grained measure-valued description of G on time scale $T_{q'}$.
- For any $q' > 0$, because of the last step and irreducibility of the original G, the resulting $\overline{G}$ with weights given by $\overline{P}_{q'}$ is a complete graph with non-homogeneous weights. In particular $\overline{P}_{q'}$ is again irreducible and reversible.
- To see that the measures $\nu_{\overline{x}}$ have disjoint support, we refer the reader to Theorem 10 of [3].
- To see that the constructed $(\Lambda, \overline{P}_{q'})$ is close to a solution we refer the reader to Theorem 12 of [3]. Here a bound is given on
$$\mathbb{E}\left[\sum_{\overline{x}=1}^{|R(\Phi_q)|} d_{TV}(\Lambda K_{q'}(\overline{x}, \cdot), \overline{P}_{q'}\Lambda(\overline{x}, \cdot))\right]$$
where $\mathbb{E}$ is the expectation with respect to the forest measure $\mathbb{P}$. Further $K_{q'}(x, \cdot)$ is the distribution of the original process started at $x \in V$ and considered at an exponential time $T_{q'}$ of parameter $q' > 0$. The *total variation distance* function d_{TV} for two probability measures μ, ν on V is given by
$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in V} |\mu(x) - \nu(x)|.$$

Chapter 3

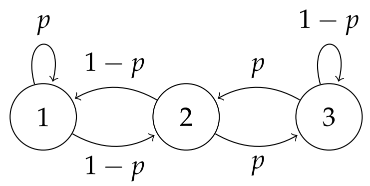
One-step random walk on the 3-node toy model

In this chapter we shall investigate the framework of [1] on an explicit model of three nodes.

- In Section 3.1 we present the model (characterized by a parameter $0 < p < 1$), calculate its invariant measure and explain the intuition behind the coarse-graining of the network.
- In Section 3.2 we solve the intertwining solutions for the one-step distribution on this model. We show that there are non-trivial solutions for all p.
- In Section 3.3 we rederive the solutions of the previous section through Theorem 2.2 and explore how "good" the universal solutions of Proposition 2.1 are.
- In Section 3.4 we investigate the evolution of local equilibria by using Theorem 2.1 and the intertwining solutions of Section 3.2.
- Finally Section 3.5 contains investigation of the phenomenology that arises from this evolution of the local equilibria and in particular find a bound on the mixing time of the original chain.

3.1 The 3-node toy model

Let G be the graph on state space $\mathcal{X} = \{1, 2, 3\}$ as in the following figure:



The transition matrix on this graph is defined using a parameter $p \in (0, 1)$ as follows:

$$P = \begin{pmatrix} p & 1-p & 0 \\ 1-p & 0 & p \\ 0 & p & 1-p \end{pmatrix}. \tag{3.1}$$

As this Markov chain is irreducible and aperiodic on a finite state space, we know that it has a unique invariant measure. To find its invariant measure $\pi = (\pi(1), \pi(2), \pi(3))$, we must solve

$$\pi P = \pi.$$

This gives us

$$\pi(1) = \pi(2) = \pi(3) = \frac{1}{3}. \tag{3.2}$$

Lemma 3.1. Furthermore for general p, the matrix P has eigenvectors

$$\mu_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \mu_1 = \begin{pmatrix} \frac{1-2p-\lambda}{p} \\ \frac{-1+p+\lambda}{p} \\ 1 \end{pmatrix} \text{ and } \mu_2 = \begin{pmatrix} \frac{1-2p+\lambda}{p} \\ \frac{-1+p-\lambda}{p} \\ 1 \end{pmatrix}$$

for eigenvalues $\theta_0 = 1$, $\theta_1 = \lambda$ and $\theta_2 = -\lambda$ respectively, where we define

$$\lambda := \sqrt{3p^2 - 3p + 1} \in [1/2, 1).$$

Proof. We find that the determinant of $\lambda I - P$ is given by

$$\begin{aligned}
\det \begin{pmatrix} \lambda - p & p-1 & 0 \\ p-1 & \lambda & -p \\ 0 & -p & \lambda - (1-p) \end{pmatrix} =& (\lambda - p)(\lambda^2 - \lambda(1-p) - p^2) + (1-p)(p-1)(\lambda - (1-p)) \\
=& \lambda^3 - \lambda^2(1-p) - p^2\lambda - p\lambda^2 + \lambda p(1-p) + p^3 - \lambda(p^2 - 2p + 1) \\
& + 3p^2 - 3p + 1 - p^3 \\
=& \lambda^3 - \lambda^2 + \lambda(-3p^2 + 3p - 1) + 3p^2 - 3p + 1.
\end{aligned}$$

This has solutions $\lambda = 1$ and $\lambda = \pm\sqrt{3p^2 - 3p + 1}$.

For the eigenvectors we must solve $P\mu = \lambda\mu$. Then we find

$$\begin{pmatrix} p\mu(1) + (1-p)\mu(2) \\ (1-p)\mu(1) + p\mu(3) \\ p\mu(2) + (1-p)\mu(3) \end{pmatrix} = \begin{pmatrix} \lambda\mu(1) \\ \lambda\mu(2) \\ \lambda\mu(3) \end{pmatrix}.$$

Thus we find

$$\mu(2) = \frac{-1 + p + \lambda}{p}\mu(3)$$

and

$$\begin{aligned}
\mu(1) =& \frac{1-p}{\lambda - p}\mu(2) \\
=& \frac{(1-p)(-1+p+\lambda)}{p(\lambda - p)}\mu(3) \\
=& \frac{-1 + p + \lambda + p - p^2 - p\lambda}{p(\lambda - p)}\mu(3) \\
=& \frac{(1 - 2p - \lambda)(\lambda - p)}{p(\lambda - p)}\mu(3) \\
=& \frac{1 - 2p - \lambda}{p}\mu(3).
\end{aligned}$$

□

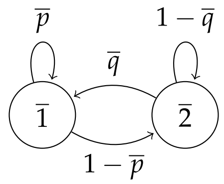
For our intertwining discussion we place the following parametric probability measures on the graph:

$$\begin{aligned}
\nu_{\bar{1}} =& (M_1, M_2, M_3), \\
\nu_{\bar{2}} =& (m_1, m_2, m_3),
\end{aligned}$$

where the M_i and m_i are in $[0, 1]$ and add up to 1 respectively. Then define the matrix Λ as

$$\Lambda = \begin{pmatrix} \nu_{\bar{1}} \\ \nu_{\bar{2}} \end{pmatrix} = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{3.3}$$

We consider a coarse-grained network on state space $\overline{\mathcal{X}} = \{\overline{1}, \overline{2}\}$. The intuition behind this is that if we take p to be large, the process on $\mathcal{X}$ will stay for a long time in either the node 1 or the pair 2-3. Thus we can collapse the nodes 2 and 3 into $\overline{2}$ that represents a measure that places most of its mass on $2, 3$ and similar for 1 with $\overline{1}$. Similarly if p is small, the process will stay for a long time in either the pair 1-2 or in 3.
We place the graph on $\overline{\mathcal{X}}$ in the following figure:



The transition matrix on this graph is given by

$$\overline{P} = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix} \tag{3.4}$$

where $\overline{p}$ and $\overline{q}$ are as of yet unknown transition probabilities in $[0, 1]$.

3.2 Intertwining: characterization of solutions for the one-step distribution

We shall now show that **there exist non-trivial solutions of the intertwining equation for all p's**. The full characterization of solutions is given in Proposition 3.1.

Proposition 3.1 (Full characterization of solutions to intertwining in the 3-node model)**.** Let P be as in (3.1), Λ as in (3.3) and $\overline{P}$ as in (3.4). The solutions $(\Lambda, \overline{P})$ to the intertwining equation $\Lambda P = \overline{P}\Lambda$ are characterized by the difference $\overline{p} - \overline{q}$, which can only be $\pm\lambda$. The solutions are identified by

- 1. If $\overline{p} - \overline{q} = +\lambda$, then either

$$0 \leq m_1 \leq \frac{1}{3} \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p} \tag{3.5}$$

or

$$0 \leq M_1 \leq \frac{1}{3} \leq m_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}; \tag{3.6}$$

2. If $\overline{p} - \overline{q} = -\lambda$, then either

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}\right\} \leq m_1 \leq \frac{1}{3} \tag{3.7}$$

$$\frac{1}{3}(1+\lambda) - \lambda m_1 \leq M_1 \leq \min\left\{\frac{-p}{1-\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\}; \tag{3.8}$$

or

$$\frac{1}{3} \leq m_1 \leq \frac{-p}{1-\lambda-p} \tag{3.9}$$

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\} \leq M_1 \leq \frac{1}{3}(1+\lambda) - \lambda m_1; \tag{3.10}$$

- $$M_2 = \frac{(1+\alpha-3p)M_1 + p}{1+\alpha} \tag{3.11}$$

  and
  $$m_2 = \frac{(1+\alpha-3p)m_1 + p}{1+\alpha}; \tag{3.12}$$

- $$M_3 = 1 - M_1 - M_2 \tag{3.13}$$

  and
  $$m_3 = 1 - m_1 - m_2; \tag{3.14}$$

and finally given $M_1 \neq m_1$,

- $$\overline{q} = \frac{p(1-p)(1-3m_1)}{(1+\alpha)(M_1 - m_1)}, \tag{3.15}$$

where $\alpha = \overline{p} - \overline{q}$.

Proof. We want to solve
$$\Lambda P = \overline{P}\Lambda.$$

In order to do this, we first find
$$\Lambda P = \begin{pmatrix} pM_1 + (1-p)M_2 & (1-p)M_1 + pM_3 & pM_2 + (1-p)M_3 \\ pm_1 + (1-p)m_2 & (1-p)m_1 + pm_3 & pm_2 + (1-p)m_3 \end{pmatrix} \tag{3.16}$$

and
$$\overline{P}\Lambda = \begin{pmatrix} \overline{p}M_1 + (1-\overline{p})m_1 & \overline{p}M_2 + (1-\overline{p})m_2 & \overline{p}M_3 + (1-\overline{p})m_3 \\ \overline{q}M_1 + (1-\overline{q})m_1 & \overline{q}M_2 + (1-\overline{q})m_2 & \overline{q}M_3 + (1-\overline{q})m_3 \end{pmatrix}. \tag{3.17}$$

The intertwining relation gives us the following constraints on Λ and $\overline{P}$:
$$pM_1 + (1-p)M_2 = \overline{p}(M_1 - m_1) + m_1 \tag{3.18}$$
$$(1-p)M_1 + pM_3 = \overline{p}(M_2 - m_2) + m_2 \tag{3.19}$$
$$pM_2 + (1-p)M_3 = \overline{p}(M_3 - m_3) + m_3 \tag{3.20}$$
$$pm_1 + (1-p)m_2 = \overline{q}(M_1 - m_1) + m_1 \tag{3.21}$$
$$(1-p)m_1 + pm_3 = \overline{q}(M_2 - m_2) + m_2 \tag{3.22}$$
$$pm_2 + (1-p)m_3 = \overline{q}(M_3 - m_3) + m_3 \tag{3.23}$$
$$M_1 + M_2 + M_3 = 1 \tag{3.24}$$
$$m_1 + m_2 + m_3 = 1 \tag{3.25}$$
$$M_1, M_2, M_3, m_1, m_2, m_3, \overline{p}, \overline{q} \in [0,1]. \tag{3.26}$$

We shall find solutions as functions of M_1 and m_1 in the following steps.

1. **Suppose that $M_1 \neq m_1$ and $M_3 \neq m_3$. Else we only find trivial solutions.**
   We see this as the equations (3.18) and (3.21) force $M_2 = m_2$ if $M_1 = m_1$ and similar for (3.20) and (3.23) if $M_3 = m_3$. Then equations (3.19) and (3.22) give us now that
   $$(1-p)M_1 + pM_3 = (1-p)m_1 + pm_3.$$
   Hence we obtain
   $$M_1 - m_1 = \frac{-p}{1-p}(M_3 - m_3).$$
   Thus $M_i = m_i$ for all i.
   From equation (3.18) and (3.19) we can then see that $M_1 = M_2 = M_3$. Plugging this in (3.24), returns the invariant measure. Note that $M_1 = M_2 = M_3 = m_1 = m_2 = m_3 = \frac{1}{3}$ clearly satisfies all constraints.

2. **If $M_2 = m_2$, then there are only non-trivial solutions for $p = \frac{1}{2}$.**
From the relation $M_1 - m_1 = \frac{-p}{1-p}(M_3 - m_3)$ and equation (3.24) we see

$$M_1 + M_2 + M_3 = \frac{-p}{1-p}(M_3 - m_3) + m_1 + m_2 + M_3 = 1.$$

Using (3.25) which states $m_1 + m_2 = 1 - m_3$, we rewrite this as

$$\frac{-p}{1-p}(M_3 - m_3) + 1 - m_3 + M_3 = 1,$$

or equivalently as

$$\frac{1-2p}{1-p}(M_3 - m_3) = 0.$$

For $p \neq \frac{1}{2}$, this gives us $M_3 = m_3$, and by the derivation above thus the invariant measure. For $p = \frac{1}{2}$ we refer the reader to Step 11.

3. **Now we use equations (3.18) to (3.230) to find a condition for $\alpha = \overline{p} - \overline{q}$.**
We combine equations (3.18)-(3.21), (3.19)-(3.22), and (3.20)-(3.23) in the following three equations respectively, where we eliminate the term m_i in each pair:

$$\begin{aligned} pM_1 + (1-p)M_2 + \overline{p}(m_1 - M_1) &= pm_1 + (1-p)m_2 + \overline{q}(m_1 - M_1) \\ (1-p)M_1 + pM_3 + \overline{p}(m_2 - M_2) &= (1-p)m_1 + pm_3 + \overline{q}(m_2 - M_2) \\ pM_2 + (1-p)M_3 + \overline{p}(m_3 - M_3) &= pm_2 + (1-p)m_3 + \overline{q}(m_3 - M_3). \end{aligned}$$

We rewrite these in the following way:

$$p(M_1 - m_1) + (1-p)(M_2 - m_2) + (\overline{p} - \overline{q})(m_1 - M_1) = 0 \tag{3.27}$$
$$(1-p)(M_1 - m_1) + p(M_3 - m_3) + (\overline{p} - \overline{q})(m_2 - M_2) = 0 \tag{3.28}$$
$$p(M_2 - m_2) + (1-p)(M_3 - m_3) + (\overline{p} - \overline{q})(m_3 - M_3) = 0. \tag{3.29}$$

We see that equations (3.27) and (3.29) give us the following relation:

$$M_2 - m_2 = \frac{\overline{p} - \overline{q} - p}{1-p}(M_1 - m_1) = \frac{\overline{p} - \overline{q} - (1-p)}{p}(M_3 - m_3). \tag{3.30}$$

Now we must check whether this satisfies (3.28) as well, so we obtain:

$$\left[\frac{(1-p)^2}{\overline{p} - \overline{q} - p} + \frac{p^2}{\overline{p} - \overline{q} - (1-p)} - (\overline{p} - \overline{q})\right](M_2 - m_2) = 0.$$

As we assumed that $M_2 \neq m_2$, this gives us a condition on $\overline{p}, \overline{q}$, namely that:

$$\frac{(1-p)^2}{\overline{p} - \overline{q} - p} + \frac{p^2}{\overline{p} - \overline{q} - (1-p)} - (\overline{p} - \overline{q}) = 0.$$

We write

$$\alpha = \overline{p} - \overline{q}$$

and compute:

$$\begin{aligned} 0 =& (1-p)^2(\alpha - (1-p)) + p^2(\alpha - p) - \alpha(\alpha - (1-p))(\alpha - p) \\ =& (1-p)^2\alpha - (1-p)^3 + p^2\alpha - p^3 - \alpha^3 + p\alpha^2 + (1-p)\alpha^2 - p(1-p)\alpha \\ =& (1 - 2p + p^2)\alpha - 1 + 3p - 3p^2 + p^3 + p^2\alpha - p^3 - \alpha^3 + p\alpha^2 + (1-p)\alpha^2 - p\alpha + p^2(\alpha) \\ =& -\alpha^3 + \alpha^2 + \alpha(3p^2 - 3p + 1) - 3p^2 + 3p - 1. \end{aligned}$$

So the condition on α becomes

$$\alpha^3 - \alpha^2 - \alpha(3p^2 - 3p + 1) + 3p^2 - 3p + 1 = 0. \tag{3.31}$$

4. **Now we find that $\alpha = \pm\lambda$.**
We do this by imposing consistency with equations (3.24) and (3.25) on the central relation (3.31) that we found. We see

$$\begin{aligned} 1 =& M_1 + M_2 + M_3 \\ =& \frac{1-p}{\overline{p}-\overline{q}-p}(M_2 - m_2) + m_1 + M_2 + \frac{p}{\overline{p}-\overline{q}-(1-p)}(M_2-m_2) + m_3 \\ =& \left[\frac{1-p}{\overline{p}-\overline{q}-p} + \frac{p}{\overline{p}-\overline{q}-(1-p)} + 1\right](M_2 - m_2) + 1. \end{aligned}$$

This returns

$$\left[\frac{1-p}{\alpha} + \frac{p}{\alpha-(1-p)} + 1\right](M_2 - m_2) = 0.$$

As we assumed that $M_2 - m_2 \neq 0$, we find another condition on $\alpha = \overline{p} - \overline{q}$ namely that

$$\begin{aligned} 0 =& (1-p)(\alpha - (1-p)) + p(\alpha - p) + (\alpha - p)(\alpha - (1-p)) \\ =& \alpha - (1-p)^2 - p^2 + \alpha^2 + p(1-p) - p\alpha - (1-p)\alpha. \end{aligned}$$

This is equivalent to

$$\begin{aligned} \alpha^2 &= (1-p)^2 + p^2 - p(1-p) \\ &= 1 - 2p + p^2 + p^2 - p + p^2 \\ &= 3p^2 - 3p + 1. \end{aligned} \tag{3.32}$$

We observe that the solution of (3.31) and (3.32) is given by

$$\alpha = \pm\sqrt{3p^2 - 3p + 1}. \tag{3.33}$$

5. **In steps 5 to 11 we consider $\alpha = +\lambda$. We start by rewriting equations (3.27) to (3.29):**

$$(\lambda - p)(M_1 - m_1) = (1-p)(M_2 - m_2) \tag{3.34}$$

$$\lambda(M_2 - m_2) = (1-p)(M_1 - m_1) + p(M_3 - m_3) \tag{3.35}$$

$$(\lambda - (1-p))(M_3 - m_3) = p(M_2 - m_2). \tag{3.36}$$

Note that $\lambda = p$ iff $p = \frac{1}{2}$ and $\lambda = 1 - p$ iff $p = \frac{1}{2}$. Both imply $M_2 = m_2$ as we see from (3.34) and (3.36). Therefore we consider $p \neq \frac{1}{2}$. The case of $p = \frac{1}{2}$ and $M_2 = m_2$ can be found in Step 11.

6. **We give expressions for $\overline{q}, \overline{p}, m_2, M_2$.**
To do this we consider the three expressions for $\overline{q}$ from (3.21) to (3.23). We will exploit consistency of (3.21) and (3.22) and use (3.34). Then we check consistency of (3.35), (3.36) and (3.23) in Step 7.
Equations (3.21) and (3.20) give us

$$\begin{aligned} \overline{q} &= \frac{(1-p)(m_2 - m_1)}{M_1 - m_1} \\ \overline{q} &= \frac{(1-p)m_1 + pm_3 - m_2}{M_2 - m_2}. \end{aligned}$$

Substituting (3.34) and $m_3 = 1 - m_1 - m_2$ in (3.22) we find

$$\overline{q} = \frac{1-p}{\lambda - p}\frac{(1-p)m_1 + p - pm_1 - pm_2 - m_2}{M_1 - m_1}.$$

Hence it must hold that

$$m_2 - m_1 = \frac{1}{\lambda - p}[(1-2p)m_1 + p - (1+p)m_2]$$

which means that

$$\left(1+\frac{1+p}{\lambda-p}\right)m_2=\frac{1}{\lambda-p}[(1+\lambda-3p)m_1+p]$$

giving thus

$$m_2=\frac{(1+\lambda-3p)m_1+p}{1+\lambda}. \tag{3.37}$$

Substituting this in (3.21) returns

$$\begin{aligned}\overline{q}&=\frac{1-p}{M_1-m_1}\frac{(1+\lambda-3p)m_1+p-(1+\lambda)m_1}{1+\lambda}\\&=\frac{p(1-p)(1-3m_1)}{(1+\lambda)(M_1-m_1)}.\end{aligned} \tag{3.38}$$

Then also

$$\begin{aligned}\overline{p}&=\overline{q}+\lambda\\&=\frac{p(1-p)(1-3m_1)+(\lambda+3p^2-3p+1)(M_1-m_1)}{(1+\lambda)(M_1-m_1)}\\&=\frac{p(1-p)+(1+\lambda)(\lambda M_1-m_1)}{(1+\lambda)(M_1-m_1)}.\end{aligned} \tag{3.39}$$

Substituting this $\overline{p}$ in (3.18) we can find M_2:

$$\begin{aligned}M_2&=\frac{p(1-p)(1-3m_1)+\lambda(1+\lambda)(M_1-m_1)+(m_1-pM_1)(1+\lambda)}{(1+\lambda)(1-p)}\\&=\frac{p(1-p)(1-3m_1)+(1+\lambda)[(\lambda-p)M_1+(1-\lambda)m_1]}{(1+\lambda)(1-p)}\\&=\frac{p(1-p)(1-3m_1)+(1+\lambda)(\lambda-p)M_1+(1-3p^2+3p-1)m_1}{(1+\lambda)(1-p)}\\&=\frac{p(1-p)+[(1-p)\lambda+3p^2-4p-1]M_1}{(1+\lambda)(1-p)}\\&=\frac{p(1-p)+(1-p)[\lambda+1-3p]M_1}{(1+\lambda)(1-p)}\\&=\frac{[1+\lambda-3p]M_1+p}{1+\lambda}.\end{aligned} \tag{3.40}$$

7. **We will now check consistency of (3.35), (3.36) and (3.23).**
   First we find

$$\begin{aligned}M_2-m_2&=\frac{[1+\lambda-3p](M_1-m_1)}{1+\lambda}\\M_3-m_3&=-(M_1-m_1)-(M_2-m_2)\\&=\frac{-2-2\lambda+3p}{1+\lambda}(M_1-m_1).\end{aligned}$$

For (3.35) we find that

$$\begin{aligned}\lambda(1+\lambda-3p)&=\lambda+\lambda^2-3p\lambda\\&=\lambda+3p^2-3p+1-3p\lambda\end{aligned}$$

and

$$\begin{aligned}(1-p)(1+\lambda)+p(-2-2\lambda+3p)&=1+\lambda-p-p\lambda-2p-2p\lambda+3p^2\\&=\lambda+3p^2-3p+1-3p\lambda.\end{aligned}$$

For (3.35) we find that

$$\begin{aligned}(\lambda-(1-p))(-2-2\lambda+3p) &= -2\lambda-2\lambda^2+3p\lambda+2+2\lambda-3p-2p-2p\lambda+3p^2\\ &= -2\lambda-6p^2+6p-2+3p\lambda+2+2\lambda-3p-2p-2p\lambda+3p^2\\ &= p+p\lambda-3p^2.\end{aligned}$$

and

$$p(1+\lambda-3p) = p+p\lambda-3p^2.$$

For (3.23) we find

$$\begin{aligned}m_3 &= 1-m_1-m_2\\ &= \frac{(1-m_1)(1+\lambda)-(1+\lambda-3p)m_1-p}{1+\lambda}\\ &= \frac{(-2-2\lambda+3p)m_1-p+1+\lambda}{1+\lambda}\end{aligned}$$

and hence

$$\begin{aligned}m_2-m_3 &= 1-m_1-2m_3\\ &= \frac{(1-m_1)(1+\lambda)+(4+4\lambda-6p)m_1+2p-2-2\lambda}{1+\lambda}\\ &= \frac{(3+3\lambda-6p)m_1+2p-1-\lambda}{1+\lambda}\\ &= \frac{(-1-\lambda+2p)(1-3m_1)}{1+\lambda}.\end{aligned}$$

This gives us

$$\begin{aligned}\overline{q} &= (M_3-m_3)^{-1}p(m_2-m_3)\\ &= \frac{p(-1-\lambda+2p)(1-3m_1)}{(-2-2\lambda+3p)(M_1-m_1)}\\ &= \frac{p(1-p)(1-3m_1)}{(1+\lambda)(M_1-m_1)}.\end{aligned}$$

8. **Now that we found intertwining solutions, we must also impose the conditions (3.26) on these solutions.**
   Note that

   $$0 \le \overline{q}, \overline{p} \le 1$$

   is equivalent with

   $$0 \le \overline{q} \le 1-\lambda.$$

   In this step we shall investigate the consequences of these conditions on $\overline{q}$.
   We first verify when $\overline{q} \ge 0$.
   Suppose $M_1 - m_1 \ge 0$. Then we must have

   $$1-3m_1 \ge 0.$$

   For $\overline{q} \le 1-\lambda$, we must have

   $$\frac{p(1-p)(1-3m_1)}{(1+\lambda)(M_1-m_1)} \le \frac{(1-\lambda)(1+\lambda)(M_1-m_1)}{(1+\lambda)(M_1-m_1)} = \frac{(-3p^2+3p)(M_1-m_1)}{(1+\lambda)(M_1-m_1)}.$$

   If $M_1 - m_1 \ge 0$, then it follows that $M_1 \ge \frac{1}{3}$.
   So the final condition becomes either

   $$0 \le m_1 \le \frac{1}{3} \le M_1 \le 1 \tag{3.41}$$

   or

   $$0 \le M_1 \le \frac{1}{3} \le m_1 \le 1. \tag{3.42}$$

9. **Now we must impose that $M_2, M_3, m_2, m_3 \in [0,1]$.** As $M_3 = 1 - M_1 - M_2$ and $M_1 \in [0,1]$ and similar for m_3, this is equivalent with imposing that

$$0 \leq M_2 \leq 1 - M_1$$

and

$$0 \leq m_2 \leq 1 - m_1.$$

In this step we investigate the first condition $M_2 \geq 0$.
This gives us

$$(1 + \lambda - 3p)M_1 \geq -p.$$

Note that $1 + \lambda - 3p > 0$ for $p < \frac{1}{2}$ and $1 + \lambda - 3p < 0$ for $p > \frac{1}{2}$, giving us that we must impose

$$M_1 \geq \frac{-p}{1 + \lambda - 3p}$$

for $p < \frac{1}{2}$ and

$$M_1 \leq \frac{-p}{1 + \lambda - 3p}$$

for $p > \frac{1}{2}$.

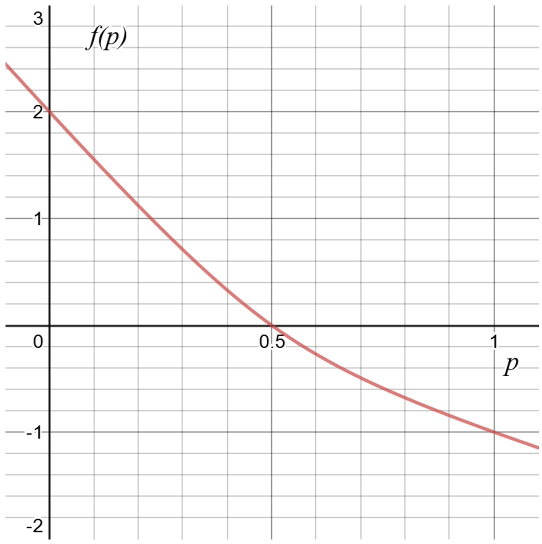


Figure 3.1: We see that the function $f(p) = 1 + \lambda - 3p > 0$ for $p < \frac{1}{2}$ and $f(p) = 1 + \lambda - 3p < 0$ for $p > \frac{1}{2}$.

Note that for $M_1 \in [0,1]$ this is always the case for all $p \in (0,1)$ as we can see in the figure below.

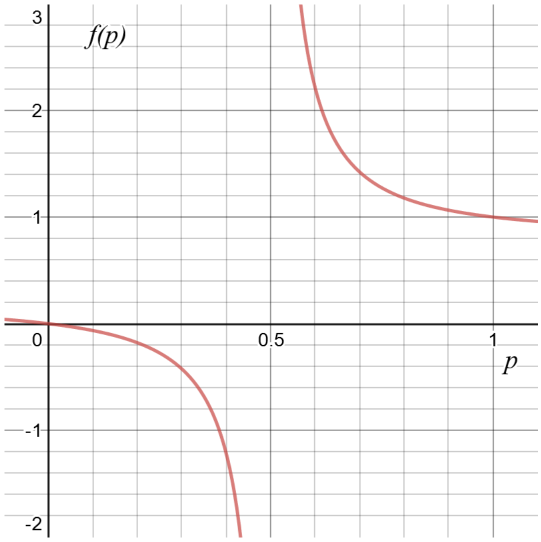


Figure 3.2: The function $f(p) = \frac{-p}{1+\lambda-3p}$. Clearly for all $M_1 \in [0,1]$ we have $M_1 > f(p)$ for $p < \frac{1}{2}$ and $M_1 < f(p)$ for $p > \frac{1}{2}$.

10. **The second condition** $M_2 \leq 1 - M_1$ is equivalent with

$$\frac{(1+\lambda-3p)M_1 + p}{1+\lambda} \leq \frac{(1-M_1)(1+\lambda)}{1+\lambda}$$

which is to say

$$(2+2\lambda-3p)M_1 \leq -p+1+\lambda.$$

We can see in the figure below that $2+2\lambda-3p \geq 0$ for all $p \in (0,1)$.

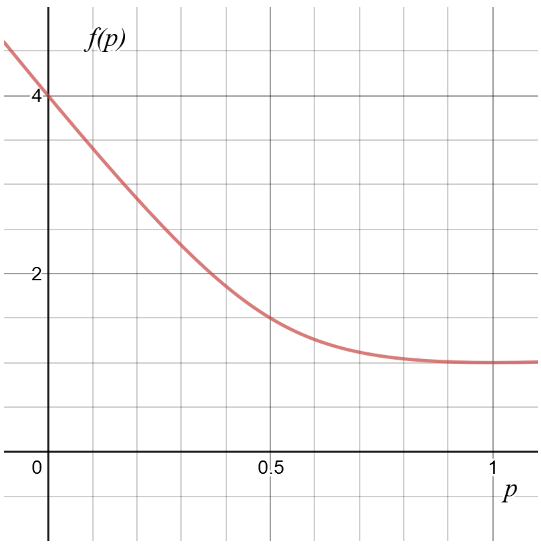


Figure 3.3: We see that the function $f(p) = 2+2\lambda-3p \geq 0$ for all $p \in (0,1)$.

Thus our condition becomes

$$M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}.$$

In the figure below we see $\frac{-p+1+\lambda}{2+2\lambda-3p}$ for $p \in (0,1)$.

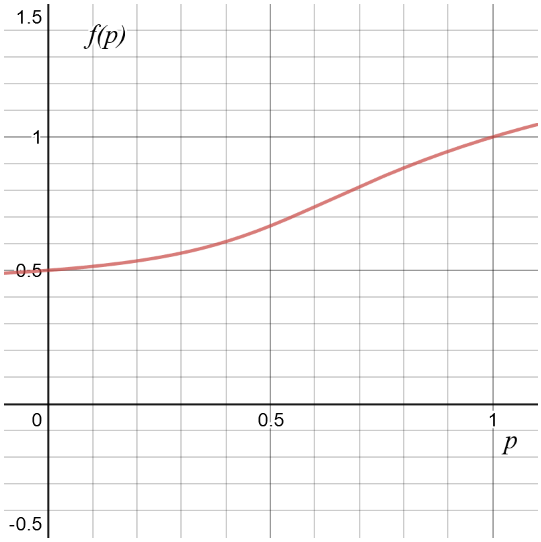


Figure 3.4: We see that the function $f(p) = \frac{-p+1+\lambda}{2+2\lambda-3p}$ gives us a non-trivial bound $M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}$.

Note that due to the similarity in (3.37) and (3.40), we have the same condition for m_1.

Thus our total conditions are either

$$0 \leq m_1 \leq \frac{1}{3} \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p} \tag{3.43}$$

or

$$0 \leq M_1 \leq \frac{1}{3} \leq m_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}. \tag{3.44}$$

11. **Now we investigate the case $p = \frac{1}{2}$.**
Equations (3.18) and (3.22) quickly give us that $M_1 + M_3 = m_1 + m_3$. Imposing $1 = m_1 + m_2 + m_3$, we find that $M_1 + M_2 + M_3 = m_1 + m_3 - M_3 + m_2 + M_3 = 1$.
However, let us consider again equations (3.18) and (3.20). Equation (3.18) becomes

$$\frac{1}{2}M_1 + \frac{1}{2}M_2 - m_1 = \overline{p}(M_1 - m_1)$$

and (3.20) becomes

$$\frac{1}{2}M_2 + \frac{1}{2}M_3 - m_3 = -\overline{p}(M_1 - m_1).$$

So we obtain

$$\frac{1}{2}M_1 + \frac{1}{2}M_2 - m_1 = -\frac{1}{2}M_2 - \frac{1}{2}M_3 + m_3.$$

Enforcing (3.24) and (3.25) again we find

$$\frac{1}{2} = -\frac{1}{2}M_2 + 1 - m_2,$$

or equivalently

$$-\frac{1}{2} = -\frac{3}{2}M_2.$$

Thus we must have that $M_2 = m_2 = \frac{1}{3}$.
Now equations (3.24) and (3.25) give us that

$$M_1 + M_3 = m_1 + m_3 = \frac{2}{3}.$$

Note that the expression for $\overline{q}$ is not affected by the assumptions $p = \frac{1}{2}$ and $M_2 = m_2$, thus the conditions that arise due to (3.27) do not change.

12. **We now consider $\alpha = -\lambda$. Repeating the steps 5 to 7, we find**

$$\begin{aligned}
\overline{q} &= \frac{p(1-p)(1-3m_1)}{(1-\lambda)(M_1-m_1)} \\
M_2 &= \frac{(1-\lambda-3p)M_1+p}{1-\lambda} \\
m_2 &= \frac{(1-\lambda-3p)m_1+p}{1-\lambda}.
\end{aligned}$$

So it only rests to impose the conditions $\overline{q}, \overline{p}, M_i, m_i \in [0,1]$.
In this step we will prove when $\overline{q}, \overline{p} \in [0,1]$.
Note that this is equivalent to

$$\lambda \leq \overline{q} \leq 1.$$

The condition $\overline{q} \geq \lambda$ is equivalent to

$$\overline{q} = \frac{p(1-p)(1-3m_1)}{(1-\lambda)(M_1-m_1)} \geq \frac{\lambda(1-\lambda)(M_1-m_1)}{(1-\lambda)(M_1-m_1)}.$$

As $p(1-p) = \frac{1}{3}(3p-3p^2) = \frac{1}{3}(1-\lambda^2)$, we thus see that this condition translates to (**We assume now $M_1 \geq m_1$ and keep this assumption until the end of the proof, noting that the reverse case can be found doing a similar analysis.**)

$$(1+\lambda)\left(\frac{1}{3} - m_1\right) \geq \lambda(M_1 - m_1),$$

which is the same as

$$M_1 \leq \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}. \tag{3.45}$$

The condition $\overline{q} \leq 1$ we find in a similar way to be the same as

$$M_1 \geq \frac{1}{3}(1+\lambda) - \lambda m_1. \tag{3.46}$$

Note then that these two functions of m_1 intersect when

$$\begin{aligned}
\frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda} &= \frac{1}{3}(1+\lambda) - \lambda m_1 \\
m_1\left(\lambda - \frac{1}{\lambda}\right) &= \frac{1}{3}\left(\lambda - \frac{1}{\lambda}\right) \\
m_1 &= \frac{1}{3}
\end{aligned}$$

for all $p \in [0,1]$.

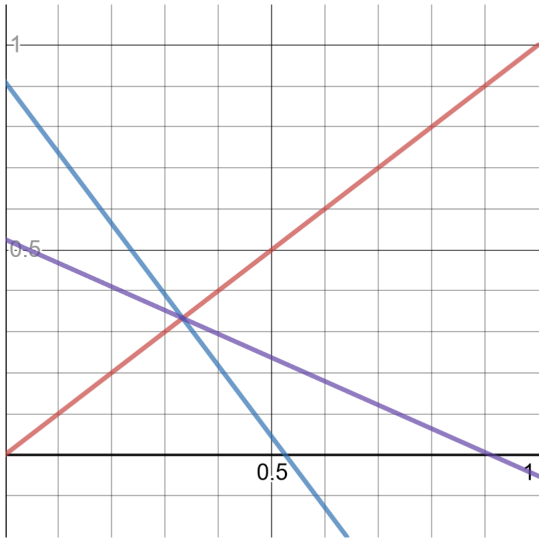


Figure 3.5: On the y-axis we have M_1 as a function of m_1 on the x-axis where $p = \frac{2}{3}$. The red graph is $M_1 = m_1$, the blue graph is $\frac{1}{3}\left(1 + \frac{1}{\lambda}\right) - \frac{m_1}{\lambda}$ and the purple graph is $\frac{1}{3}(1+\lambda) - \lambda m_1$.

From the figure above we see that $\overline{q}, \overline{p} \in [0,1]$ is equivalent to

$$0 \leq m_1 \leq \frac{1}{3} \tag{3.47}$$

$$\frac{1}{3}(1+\lambda) - \lambda m_1 \leq M_1 \leq \frac{1}{3}\left(1 + \frac{1}{\lambda}\right) - \frac{m_1}{\lambda}. \tag{3.48}$$

13. **Now we investigate the condition**

$$0 \leq M_2 \leq 1 - M_1.$$

For $M_2 \geq 0$, we must have $M_1 \leq \frac{-p}{1-\lambda-3p}$.

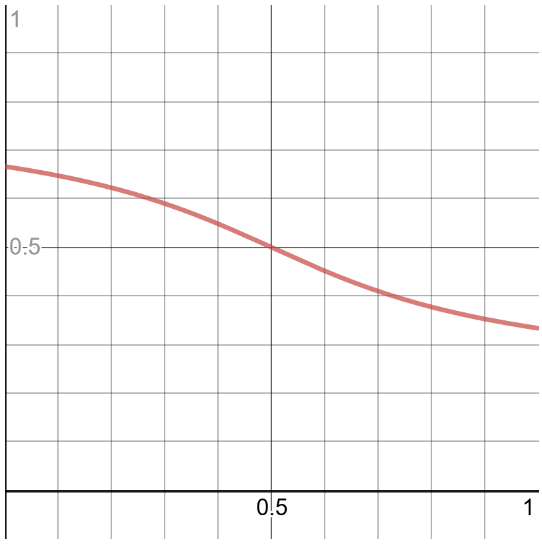


Figure 3.6: The function $\frac{-p}{1-\lambda-3p}$ for $p \in [0,1]$.

Note that $\frac{-p}{1-\lambda-3p} \geq \frac{1}{3}$ for all $p \in [0,1]$.

The condition $M_2 \leq 1 - M_1$ can be rewritten as

$$(1 - \lambda - 3p)M_1 + p \leq 1 - \lambda - (1 - \lambda)M_1,$$

which is equivalent to

$$M_1 \geq \frac{1 - \lambda - p}{2 - 2\lambda - 3p}.$$

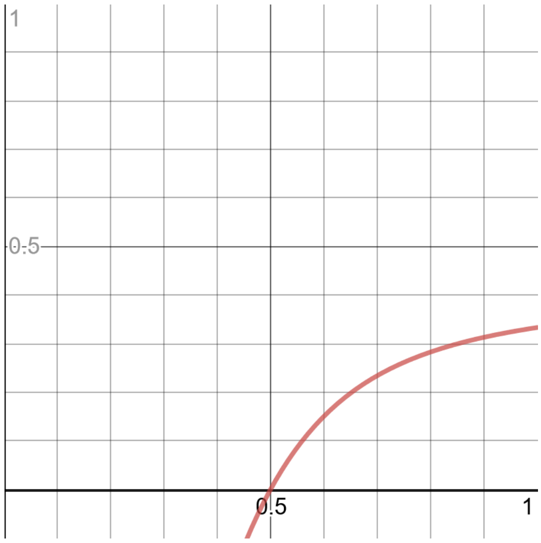


Figure 3.7: The function $\frac{1-\lambda-p}{2-2\lambda-3p}$ for $p \in [0, 1]$.

Note that for $p < \frac{1}{2}$, $\frac{1-\lambda-p}{2-2\lambda-3p}$ is negative and for $p = 1$ it equals $\frac{1}{3}$.

Further we note that we have the same conditions for m_1, due to the expression for m_2.

14. **Now we must combine all these conditions.**
For m_1 we can easily see that we must have

$$0 \leq m_1 \leq \frac{1}{3}$$

for $p \leq \frac{1}{2}$ and

$$\frac{1 - \lambda - p}{2 - 2\lambda - 3p} \leq m_1 \leq \frac{1}{3}$$

for $p > \frac{1}{2}$.

For M_1 we first notice that as $\frac{1-\lambda-p}{2-2\lambda-3p} \leq \frac{1}{3}$ for all p, the only lower bound for M_1 is

$$\frac{1}{3}(1 + \lambda) - \lambda m_1 \leq M_1.$$

For the upper bound we study the figure below.

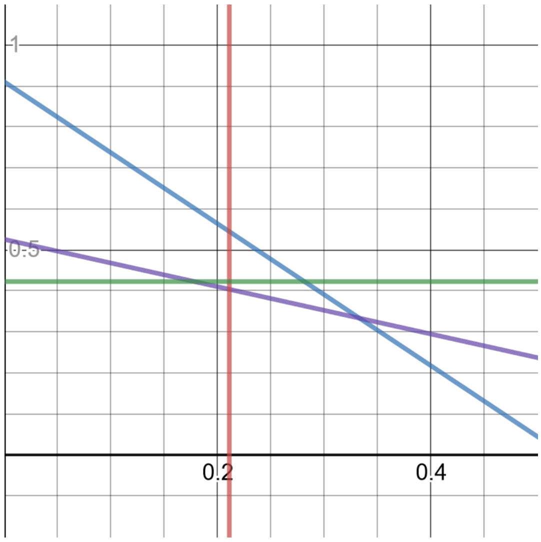


Figure 3.8: On the y-axis we have M_1 as a function of m_1 on the x-axis where $p = \frac{2}{3}$. The blue graph is $\frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}$ and the purple graph is $\frac{1}{3}(1+\lambda) - \lambda m_1$. The green graph is the value $M_1 = \frac{-p}{1-\lambda-3p}$ and the red graph the value $m_1 = \frac{1-\lambda-p}{2-2\lambda-3p}$.

Thus M_1 is bounded by

$$M_1 \le \min\left\{\frac{-p}{1-\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\}.$$

□

3.3 Testing the spectral solutions on the 3-node toy model

In Lemma 3.2 we show that by applying Theorem 2.2 and **perturbing the invariant measure, we can retrieve the non-trivial solutions of Proposition 3.1**.
In Lemma 3.3 we show that the **spectral solutions that are found by constructing a $\overline{P}$ as in Proposition 2.1 can also be non-trivial, but with less freedom than the ones of Proposition 3.1**.

3.3.1 Solutions from the spectral knowledge (Theorem 2.2)

Let us investigate Theorem 2.2 for our model where $n = 3$ and $m = 2$.
In Theorem 2.2 we started from the assumptions that we had non-trivial solutions $(\Lambda, \overline{P})$. Now that we know what they are for our model, we can find the perturbation matrix $C(\overline{x}, j)$ explicitly.
We find the following result.

Lemma 3.2 (Non-trivial perturbed solutions to intertwining)**.** Let $p \ge \frac{1}{2}$.
Consider the eigenvector $\mu_1 = \begin{pmatrix} \frac{1-2p-\lambda}{p} \\ \frac{-1+p+\lambda}{p} \\ 1 \end{pmatrix}$ of P with eigenvalue λ and the eigenvector $\mu_2 = \begin{pmatrix} \frac{1-2p+\lambda}{p} \\ \frac{-1+p-\lambda}{p} \\ 1 \end{pmatrix}$ with eigenvalue $-\lambda$.
Let

$$C_1 = \begin{pmatrix} 1 & a \\ 1 & b \end{pmatrix}$$

be an invertible matrix where

- either $-\frac{1}{3} \le a \le 0 \le b \le \frac{-p}{3(1-2p-\lambda)}$ or $-\frac{1}{3} \le b \le 0 \le a \le \frac{-p}{3(1-2p-\lambda)}$;
- $a \ne b$

and let

$$C_2 = \begin{pmatrix} 1 & a' \\ 1 & b' \end{pmatrix}$$

be an invertible matrix where

- either

$$-\frac{1}{3} \leq b' \leq 0$$
$$-\lambda b' \leq a' \leq \min\left\{\frac{-b'}{\lambda}, -\frac{1}{3}\frac{p}{-1+p-\lambda}\right\}$$

  or

$$0 \leq b' \leq -\frac{1}{3}\frac{p}{-1+p-\lambda}$$
$$-\frac{b'}{\lambda} \leq a' \leq -\lambda b;$$

If $(\Lambda, \overline{P})$ satifies the intertwining equation where $\Lambda = \begin{pmatrix} \nu_{\overline{1}} \\ \nu_{\overline{2}} \end{pmatrix}$ is of rank 2, then

$$\nu_{\overline{1}} = \pi + A(\overline{1}, 1)\mu_1 \tag{3.49}$$
$$\nu_{\overline{2}} = \pi + A(\overline{2}, 1)\mu_1 \tag{3.50}$$

and

$$\overline{P}A(\cdot, 1) = \lambda A(\cdot, 1), \tag{3.51}$$

for either $A = C_1$ or $A = C_2$.

Proof. For the set J defined in Theorem 2.2 we could either have $J = \{0, 1\}$ or $J = \{0, 2\}$.

1. Consider $J = \{0, 1\}$. For the $\nu_{\overline{x}}$ in equation (2.5) we then find

$$\nu_{\overline{x}} = \pi + C_1(\overline{x}, 2)\mu_2.$$

   If we write $a := C(\overline{1}, 2)$ and $b := C(\overline{2}, 2)$, we see

$$\nu_{\overline{1}} = \left(\frac{1}{3} + a\frac{1-2p-\lambda}{p}, \quad \frac{1}{3} + a\frac{-1+p+\lambda}{p}, \quad \frac{1}{3} + a\right),$$

$$\nu_{\overline{2}} = \left(\frac{1}{3} + b\frac{1-2p-\lambda}{p}, \quad \frac{1}{3} + b\frac{-1+p+\lambda}{p}, \quad \frac{1}{3} + b\right).$$

   We impose the constraints that all terms of $\nu_{\overline{x}}$ and all terms of $\overline{P}$ must be in $[0, 1]$, to find constraints on a, b.

   - First of all, we notice

$$-\frac{1}{3} \leq a \leq \frac{2}{3}.$$

   - We now impose

$$0 \leq \frac{1}{3} + a\frac{1-2p-\lambda}{p} \leq 1.$$

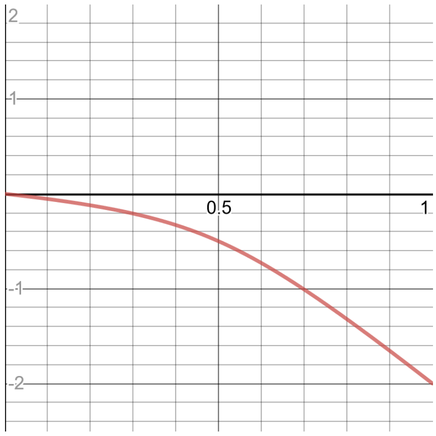


Figure 3.9: The function of $1 - 2p - \lambda$ as a function of $p \in [0, 1]$.

In the figure above we see that $\frac{1-2p-\lambda}{p}$ is negative for all $0 \leq p \leq 1$, so we conclude

$$\frac{2}{3}\frac{p}{1-2p-\lambda} \leq a \leq -\frac{1}{3}\frac{p}{1-2p-\lambda}.$$

- We now impose

$$0 \leq \frac{1}{3} + a\frac{-1+p+\lambda}{p} \leq 1.$$

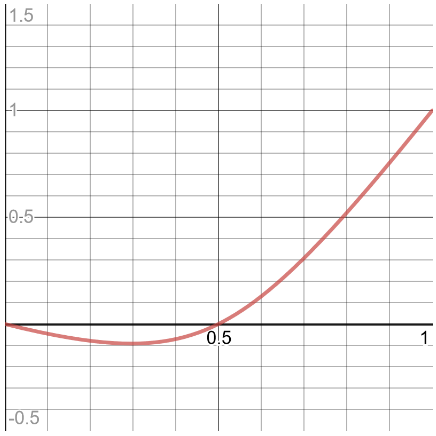


Figure 3.10: The function of $-1 + p + \lambda$ as a function of $p \in [0, 1]$.

In the figure above we see that $\frac{-1+p+\lambda}{p}$ is negative for all $p \leq \frac{1}{2}$ and positive for $p \geq \frac{1}{2}$, so we conclude

$$\frac{2}{3}\frac{p}{-1+p+\lambda} < a < -\frac{1}{3}\frac{p}{-1+p+\lambda}$$

for $p < \frac{1}{2}$ and

$$-\frac{1}{3}\frac{p}{-1+p+\lambda} < a < \frac{2}{3}\frac{p}{-1+p+\lambda}$$

for $p > \frac{1}{2}$.

- First we look at these lower bounds.

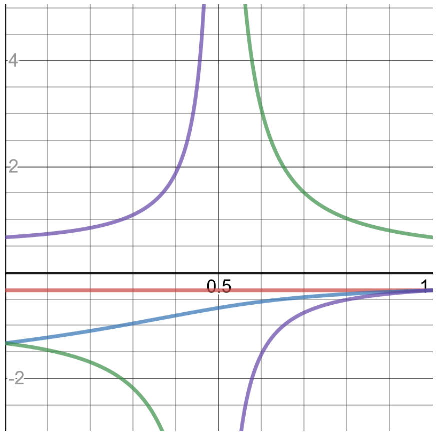


Figure 3.11: The red graph is the constant $-\frac{1}{3}$, the blue graph is $\frac{2}{3}\frac{p}{1-2p-\lambda}$, the green graph is $\frac{2}{3}\frac{p}{-1+p+\lambda}$ (which in this case we only consider for $p < \frac{1}{2}$) and the purple graph is $-\frac{1}{3}\frac{p}{-1+p+\lambda}$ (which we only consider for $p > \frac{1}{2}$).

From the figure above we conclude that we must impose

$$a \geq -\frac{1}{3}.$$

Now we investigate the upper bounds.

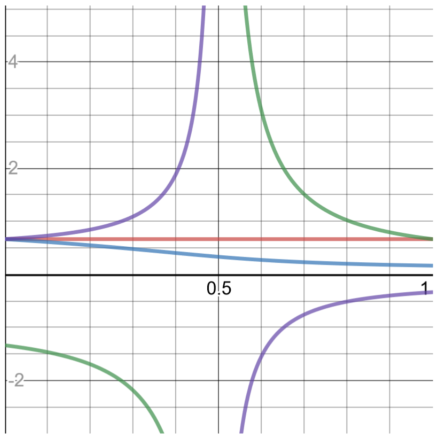


Figure 3.12: The red graph is the constant $\frac{2}{3}$, the blue graph is $-\frac{1}{3}\frac{p}{1-2p-\lambda}$, the green graph is $\frac{2}{3}\frac{p}{-1+p+\lambda}$ (which in this case we only consider for $p > \frac{1}{2}$) and the purple graph is $-\frac{1}{3}\frac{p}{-1+p+\lambda}$ (which we only consider for $p < \frac{1}{2}$).

From the figure above we conclude that we must impose

$$a \leq -\frac{1}{3}\frac{p}{1-2p-\lambda}.$$

Since the discussion for b is similar, we conclude

$$-\frac{1}{3} \leq a, b \leq -\frac{1}{3}\frac{p}{1-2p-\lambda}.$$

- Further we see that

$$\overline{P}C = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix} \begin{pmatrix} 1 & a \\ 1 & b \end{pmatrix} = \begin{pmatrix} 1 & \overline{p}(a-b)+b \\ 1 & \overline{q}(a-b)+b \end{pmatrix}.$$

This must equal

$$\begin{pmatrix} 1 & \lambda a \\ 1 & \lambda b \end{pmatrix},$$

so we find

$$\overline{p} = \frac{\lambda a - b}{a - b}$$
$$\overline{q} = \frac{(\lambda - 1)b}{a - b}.$$

We shall assume now that $a > b$.

- Note that imposing $0 \leq \overline{p} \leq 1$ gives us

$$a \geq \frac{b}{\lambda} \quad \wedge \quad a \geq 0.$$

And imposing $0 \leq \overline{p} \leq 1$ gives us

$$b \leq 0 \quad \wedge \quad b \leq \frac{a}{\lambda}.$$

- **We conclude that our constraints are $a \neq b$ and**

$$-\frac{1}{3} \leq b \leq 0 \leq a \leq \frac{-p}{3(1 - 2p - \lambda)}$$

or

$$-\frac{1}{3} \leq a \leq 0 \leq b \leq \frac{-p}{3(1 - 2p - \lambda)}.$$

2. Consider $J = \{0, 2\}$. For the $\nu_{\overline{x}}$ in equation (2.5) we then find

$$\nu_{\overline{x}} = \pi + C_2(\overline{x}, 2)\mu_2.$$

If we write $a' := C_2(\overline{1}, 2)$ and $b' := C_2(\overline{2}, 2)$, we see

$$\nu_{\overline{1}} = \left(\frac{1}{3} + a' \frac{1 - 2p + \lambda}{p}, \quad \frac{1}{3} + a' \frac{-1 + p - \lambda}{p}, \quad \frac{1}{3} + a' \right),$$
$$\nu_{\overline{2}} = \left(\frac{1}{3} + b' \frac{1 - 2p + \lambda}{p}, \quad \frac{1}{3} + b' \frac{-1 + p - \lambda}{p}, \quad \frac{1}{3} + b' \right).$$

Further

$$\overline{p} = \frac{-\lambda a' - b'}{a' - b'}$$
$$\overline{q} = \frac{(-\lambda - 1)b'}{a' - b'}.$$

- **We assume $a' > b'$, the other case can be investigated in a similar analysis.** The conditions that $\overline{p}, \overline{q} \in [0, 1]$ translate to

$$b' \leq 0$$

and

$$-\lambda b' \leq a' \leq \frac{-b'}{\lambda}.$$

- Now we look at the bounds on $\nu_{\overline{x}}$. It is immediately clear that

$$-\frac{1}{3} \leq a', b' \leq \frac{2}{3}.$$

- Now we must impose
$$0 \leq \frac{1}{3} + a' \frac{1-2p+\lambda}{p} \leq 1$$
and
$$0 \leq \frac{1}{3} + b' \frac{1-2p+\lambda}{p} \leq 1.$$

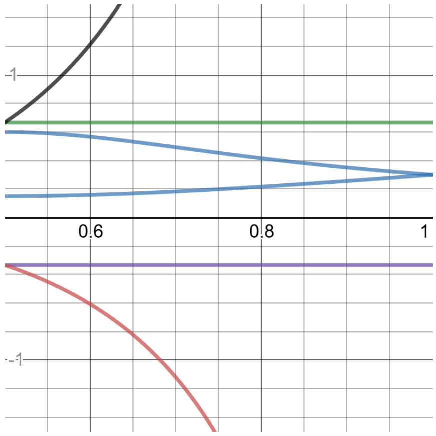


Figure 3.13: We present here functions of p. The red graph is $-\frac{1}{3}\frac{p}{1-2p+\lambda}$, the black graph is $\frac{12}{3}\frac{p}{1-2p+\lambda}$, the upper blue graph is $\frac{-b'}{\lambda}$ and the lower blue graph is $-b'\lambda$, where we took $b' = -\frac{3}{10}$. The constants are $-\frac{1}{3}$ and $\frac{2}{3}$.

From the figure above we see that our conditions on a', b' remain for the moment
$$-\frac{1}{3} \leq b' \leq 0$$
and
$$-\lambda b' \leq a' \leq \frac{-b'}{\lambda}.$$

- Now we must impose
$$0 \leq \frac{1}{3} + a' \frac{-1+p-\lambda}{p} \leq 1$$
and
$$0 \leq \frac{1}{3} + b' \frac{-1+p-\lambda}{p} \leq 1.$$

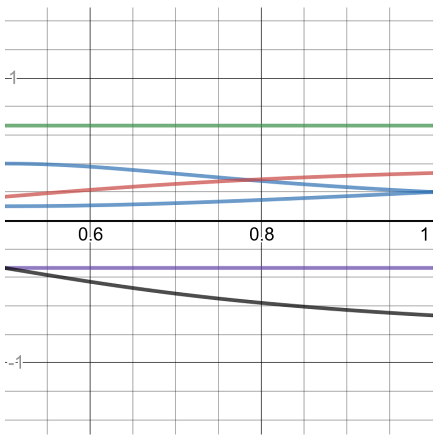


Figure 3.14: We present here functions of p. The red graph is $-\frac{1}{3}\frac{p}{-1+p-\lambda}$, the black graph is $\frac{12}{3}\frac{p}{-1+p-\lambda}$, the upper blue graph is $\frac{-b'}{\lambda}$ and the lower blue graph is $-b'\lambda$, where we took $b' = -\frac{2}{10}$. The constants are $-\frac{1}{3}$ and $\frac{2}{3}$.

In the figure above we see that the conditions on b' do not change, but those on a' do. Now a' has the upperbound

$$a' \leq \min\{\frac{-b}{\lambda}, \frac{-1}{3}\frac{p}{-1+p-\lambda}\}.$$

□

Remark. Note that the solutions for the eigenvalue $+\lambda$ are indeed equivalent to those in Proposition 3.1 where $\overline{p} - \overline{q} = +\lambda$.
If we take $a = \frac{-p}{3(1-2p-\lambda)}$, we see that $M_1 = 0$. If we take $a = 0$, we see that $M_1 = \frac{1}{3}$. If we take $a = -\frac{1}{3}$, we see that

$$\begin{aligned}\frac{1}{3} - \frac{1}{3} \cdot \frac{1-2p-\lambda}{p} &= \frac{1}{3} \cdot \frac{p-1+2p+\lambda}{p} \\ &= \frac{1}{3p} \cdot \frac{(3p-1+\lambda)(2+2\lambda-3p)}{2+2\lambda-3p} \\ &= \frac{1}{3p} \cdot \frac{6p+6p\lambda-9p^2-2-2\lambda+3p+2\lambda+2\lambda^2-3p\lambda}{2+2\lambda-3p} \\ &= \frac{1}{3p} \cdot \frac{-3p^2+3p+3p\lambda}{2+2\lambda-3p} \\ &= \frac{-p+1+\lambda}{2+2\lambda-3p}.\end{aligned}$$

Furthermore if we substitute

$$M_1 = \frac{p/3 + a(1-2p-\lambda)}{p},$$

in the expression (3.18) of M_2, we must have consistency. This gives

$$M_2 = \frac{\frac{1+\lambda-3p}{p}(p/3 + a(1-2p-\lambda)) + p}{1+\lambda}.$$

This must equal the expression of the perturbed M_2 giving us the relation

$$\frac{\frac{1+\lambda-3p}{p}(p/3 + a(1-2p-\lambda)) + p}{1+\lambda} = \frac{p/3 + a(-1+p+\lambda)}{p}$$

that must be verified.
We compute:

$$\begin{aligned}(1+\lambda-3p)(p/3 + a(1-2p-\lambda)) + p^2 =& (1+\lambda)p/3 - p^2 + a(1-2p-\lambda+\lambda-2p\lambda \\ &- \lambda^2 - 3p + 6p^2 + 3p\lambda) + p^2 \\ =& (1+\lambda)p/3 + a(3p^2 - 3p + p\lambda)\end{aligned}$$

and

$$\begin{aligned}(1+\lambda)(p/3 + a(-1+p+\lambda)) =& (1+\lambda)p/3 + a(-1+p+\lambda-\lambda+p\lambda+\lambda^2) \\ =& (1+\lambda)p/3 + a(3p^2 - 2p + p\lambda).\end{aligned}$$

Remark. We can also see that the solutions for the eigenvalue $-\lambda$ are equivalent to those in Proposition 3.1 where $\overline{p} - \overline{q} = -\lambda$.
Consistency with the expression for M_2, m_2 can be seen as it is similar to what we did in the remark above.

We shall focus instead on the bounds. If $b' = 0$, then $m_1 = \frac{1}{3}$. If $b' = -\frac{1}{3}$, then

$$\begin{aligned} m_1 =& -\frac{1}{3}\frac{-p+1-2p+\lambda}{p} \\ =& -\frac{1}{3}\frac{1-3p+\lambda}{p}\frac{2-2\lambda-3p}{2-2\lambda-3p} \\ =& -\frac{1}{3p}\frac{2-2\lambda-3p-6p+6p\lambda+9p^2+2\lambda-2\lambda^2-3p\lambda}{2-2\lambda-3p} \\ =& -\frac{1}{3p}\frac{3p\lambda+9p^2-9p+2-6p^2+6p-2}{2-2\lambda-3p} \\ =& \frac{1-\lambda-p}{2-2\lambda-3p}. \end{aligned}$$

Further if $m_1 = \frac{1}{3} + b'\frac{1-2p+\lambda}{p}$, then $b' = \frac{p}{1-2p+\lambda}\left(m_1 - \frac{1}{3}\right)$.
Then we see that $a' = -\lambda b'$ implies that $M_1 = \frac{1}{3} - \lambda\frac{p}{1-2p+\lambda}\left(m_1 - \frac{1}{3}\right)\frac{1-2p+\lambda}{p} = -\lambda m_1 + \frac{1}{3}(1+\lambda)$.
If $a' = \frac{-b'}{\lambda}$, then $M_1 = \frac{1}{3} - \frac{1}{\lambda}\frac{p}{1-2p+\lambda}\left(m_1 - \frac{1}{3}\right)\frac{1-2p+\lambda}{p} = \frac{-m_1}{\lambda} + \frac{1}{3}\left(1 + \frac{1}{\lambda}\right)$.
If $a' = -\frac{1}{3}\frac{p}{-1+p-\lambda}$, then

$$\begin{aligned} M_1 =& \frac{1}{3}\left[1 - \frac{p}{-1+p-\lambda}\frac{1-2p+\lambda}{p}\right] \\ =& \frac{1}{3}\frac{-1+p-\lambda-1+2p-\lambda}{-1+p-\lambda} \\ =& \frac{-2+3p-2\lambda}{-3+3p-3\lambda}. \end{aligned}$$

To verify that this equals $\frac{-p}{1-\lambda-3p}$, we must check whether

$$-p(-3+3p-3\lambda) = (1-\lambda-3p)(-2+3p-2\lambda).$$

We compute

$$\begin{aligned} (1-\lambda-3p)(-2+3p-2\lambda) =& -2+3p-2\lambda+2\lambda-3p\lambda+2\lambda^2+6p-9p^2+6p\lambda \\ =& 3p\lambda-9p^2+9p-2+6p^2-6p+2 \\ =& 3p-3p^2+3p\lambda. \end{aligned}$$

Thus indeed these are the solutions of Proposition 3.1 for $\overline{p} - \overline{q} = -\lambda$.

Considering these two remarks we may conclude that we have rederived the solutions of Section 3.2.

3.3.2 Testing the universal solutions of Proposition 2.1

Using Proposition 2.1 we can construct an explicit $\overline{P}$. We shall use this $\overline{P}$ to find a Λ that satisfies the intertwining relation $\Lambda P = \overline{P}\Lambda$. This $\overline{P}$ is constructed from the eigenvalue $\theta_0 = 1 > \lambda \geq 0$ in the following way

$$\overline{P} = \begin{pmatrix} \frac{1+\lambda}{2} & \frac{1-\lambda}{2} \\ \frac{1-\lambda}{2} & \frac{1+\lambda}{2} \end{pmatrix}. \tag{3.52}$$

We shall see that **there will be non-trivial solutions that can be found from Proposition 2.1 only for $p = \frac{1}{2}$, but with more constraints than the solutions found in Proposition 3.1.**

Lemma 3.3 (Non-trivial universal solutions to intertwining)**.** The only non-trivial Λ that satisfies intertwining for the explicit $\overline{P}$ given above can be found only when $p = \frac{1}{2}$ and is of the form

$$\Lambda = \begin{pmatrix} M_1 & \frac{1}{3} & \frac{2}{3} - M_1 \\ \frac{2}{3} - M_1 & \frac{1}{3} & M_1 \end{pmatrix},$$

where $0 \leq M_1 \leq \frac{2}{3}$.

Proof. The constraints that we find on solutions of the intertwining are as follows for the given $\overline{p}$ and $\overline{q}$:

$$pM_1 + (1-p)M_2 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(M_1 - m_1) + \frac{1}{2}(M_1 + m_1) \tag{3.53}$$
$$(1-p)M_1 + pM_3 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(M_2 - m_2) + \frac{1}{2}(M_2 + m_2) \tag{3.54}$$
$$pM_2 + (1-p)M_3 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(M_3 - m_3) + \frac{1}{2}(M_3 + m_3) \tag{3.55}$$
$$pm_1 + (1-p)m_2 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(m_1 - M_1) + \frac{1}{2}(M_1 + m_1) \tag{3.56}$$
$$(1-p)m_1 + pm_3 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(m_2 - M_2) + \frac{1}{2}(M_2 + m_2) \tag{3.57}$$
$$pm_2 + (1-p)m_3 = \frac{1}{2}\sqrt{3p^2 - 3p + 1}(m_3 - M_3) + \frac{1}{2}(M_3 + m_3) \tag{3.58}$$
$$M_1 + M_2 + M_3 = 1 \tag{3.59}$$
$$m_1 + m_2 + m_3 = 1 \tag{3.60}$$
$$M_1, M_2, M_3, m_1, m_2, m_3 \in [0, 1]. \tag{3.61}$$

To ease our notation we shall write $\zeta := \frac{1}{2}\sqrt{3p^2 - 3p + 1}$.
We find the solutions to these equations in the following steps.

1. **We first show that** $M_1 + m_1 = M_2 + m_2 = M_3 + m_3 = \frac{2}{3}$**.**
   If we use equations (3.53) and (3.56) to eliminate the term with ζ, we find

   $$pM_1 + (1-p)M_2 - \frac{1}{2}(M_1 + m_1) = \frac{1}{2}(M_1 + m_1) - pm_1 - (1-p)m_2.$$

   This can be rewritten as $(1-p)(M_1 + m_1) = (1-p)(M_2 + m_2)$, giving the condition

   $$M_1 + m_1 = M_2 + m_2.$$

   Doing the same with equations (3.55) and (3.58), we find

   $$M_2 + m_2 = M_3 + m_3.$$

   From (3.59) and (3.60) we find

   $$\begin{aligned} 1 =& M_1 + M_2 + M_3 \\ =& M_1 + M_1 + m_1 - m_2 + M_1 + m_1 - m_3 \\ =& 3(M_1 + m_1) - 1, \end{aligned}$$

   giving

   $$M_1 + m_1 = M_2 + m_2 = M_3 + m_3 = \frac{2}{3} \tag{3.62}$$

   as a new relation that must be satisfied.
2. **Now we show that we must separate the cases** $p \neq \frac{1}{2}$ **and** $p = \frac{1}{2}$**.**
   From equation (3.53) we find M_2 as a function of M_1:

   $$M_2 = (1-p)^{-1}\left[(2\zeta - p)M_1 - \frac{2}{3}\zeta + \frac{1}{3}\right] \tag{3.63}$$

and from (3.55) we find M_2 as a function of M_3:

$$\begin{aligned} M_2 =& p^{-1}\left[\zeta(2M_3-\frac{2}{3})+\frac{1}{3}-(1-p)M_3\right] \\ =& p^{-1}\left[(2\zeta-(1-p))M_3-\frac{2}{3}\zeta+\frac{1}{3}\right]. \end{aligned} \tag{3.64}$$

Note that we assume here

$$2\zeta \neq p.$$

We will consider the case $2\zeta = p$ in Step 6.

3. **For** $p \neq \frac{1}{2}$, we will use the two expressions (3.63) and (3.64) for M_2 in order to **express** M_3 **in terms of** M_1:

$$(1-p)(2\zeta-(1-p))M_3+(1-p)(-\frac{2}{3}\zeta+\frac{1}{3}) = p(2\zeta-p)M_1+p(-\frac{2}{3}\zeta+\frac{1}{3}).$$

Then we find

$$M_3 = \frac{p(2\zeta-p)M_1+(2p-1)(-\frac{2}{3}\zeta+\frac{1}{3})}{(1-p)(2\zeta-(1-p))} = \frac{p(2\zeta-p)M_1+(2p-1)(-\frac{2}{3}\zeta+\frac{1}{3})}{Z}, \tag{3.65}$$

where

$$Z := (1-p)(2\zeta-(1-p)).$$

4. We now check using (3.63) and (3.65), **when we retrieve (3.54).**
We compute

$$\begin{aligned} & -pM_3+\zeta(M_2-m_2)+\frac{1}{2}(M_2+m_2) \\ &= -pM_3+\zeta(2M_2-\frac{2}{3})+\frac{1}{3} \\ &= Z^{-1}\Big[-p^2(2\zeta-p)M_1-p(2p-1)(-\frac{2}{3}\zeta+\frac{1}{3})+(1-p)(2\zeta-(1-p))\zeta(2M_2-\frac{2}{3}) \\ &\quad +\frac{1}{3}(1-p)(2\zeta-(1-p))\Big] \\ &= Z^{-1}\Big[-p^2(2\zeta-p)M_1-p(2p-1)(-\frac{2}{3}\zeta+\frac{1}{3})+2\zeta(2\zeta-(1-p))((2\zeta-p)M_1-\frac{2}{3}\zeta+\frac{1}{3}) \\ &\quad -\frac{2}{3}\zeta(1-p)(2\zeta-(1-p))+\frac{1}{3}(1-p)(2\zeta-(1-p))\Big] \\ &= Z^{-1}\Big[(2\zeta-p)M_1(4\zeta^2-2(1-p)\zeta-p^2)-(2p^2-p)(-\frac{2}{3}\zeta+\frac{1}{3}) \\ &\quad +(4\zeta^2-2(1-p)\zeta)(-\frac{2}{3}\zeta+\frac{1}{3})-(\frac{2}{3}\zeta-p\frac{2}{3}\zeta)(2\zeta-(1-p))+(\frac{1}{3}-\frac{1}{3}p)(2\zeta-(1-p))\Big] \\ &= Z^{-1}\Big[(2\zeta-p)M_1(4\zeta^2-2(1-p)\zeta-p^2)+(-\frac{2}{3}\zeta+\frac{1}{3})(4\zeta^2-2(1-p)\zeta-2p^2+p)) \\ &\quad +(1-p)(2\zeta-(1-p))\frac{1}{3}(1-2\zeta)\Big]. \end{aligned}$$

For (3.54) to hold we must have that this equals $(1-p)M_1$. That is the case when

$$\begin{aligned} \Big[(2\zeta-p)(4\zeta^2-2(1-p)\zeta-p^2)-(1-p)Z\Big]M_1 =& -(-\frac{2}{3}\zeta+\frac{1}{3})(4\zeta^2-2(1-p)\zeta-2p^2+p)) \\ & -(1-p)(2\zeta-(1-p))\frac{1}{3}(1-2\zeta). \end{aligned}$$

5. **We now show that consistency with (3.54) is only possible for the invariant measure.**
We find from the condition in Step 4 that

$$\begin{aligned}
&(2\zeta - p)(4\zeta^2 - 2(1-p)\zeta - p^2) - (1-p)(1-p)(2\zeta - (1-p))\\
&= 8\zeta^3 - 4(1-p)\zeta^2 - 2p^2\zeta - 4p\zeta^2 + 2p(1-p)\zeta + p^3 - 2\zeta + 1 - p + 4p\zeta - 2p + 2p^2\\
&- 2p^2\zeta + p^2 - p^3\\
&= \zeta^2(8\zeta - 4) + \zeta(-2p^2 + 2p(1-p) - 2 - 2p^2 + 4p) + 1 - p + p^2 - 2p + 2p^2\\
&= \zeta^2(8\zeta - 4) + \zeta(-6p^2 + 6p - 2) + 1 - 3p + 3p^2.
\end{aligned}$$

On the other side we find

$$\begin{aligned}
&(-2\zeta + 1)(4\zeta^2 - 2(1-p)\zeta - 2p^2 + p)) + (1-p)(2\zeta - (1-p))(1 - 2\zeta)\\
&= -8\zeta^3 + 4(1-p)\zeta^2 + 4p^2\zeta - 2p\zeta + 4\zeta^2 - 2(1-p)\zeta - 2p^2 + p + (2\zeta(1-p) - (1-p)^2)(1-2\zeta)\\
&= \zeta^2(-8\zeta + 4) + \zeta(4p^2 - 2p + 2 + 2p^2 - 4p) - 2p^2 + p - 1 - p^2 + 2p\\
&= \zeta^2(-8\zeta + 4) + \zeta(6p^2 - 6p + 2) - 3p^2 + 3p - 1.
\end{aligned}$$

Thus we find that (3.54) is satisfied if and only if $M_1 = \frac{1}{3}$. In that case the only solution to the constraints is $M_1 = M_2 = M_3 = m_1 = m_2 = m_3 = \frac{1}{3}$.

6. **We assume now that $p = \frac{1}{2}$.**
From (3.62) and (3.63) it follows then that

$$M_2 = m_2 = \frac{1}{3}.$$

As $M_1 + M_2 + M_3 = 1$, it follows that

$$M_1 + M_3 = \frac{2}{3}.$$

Substitution of $p = \frac{1}{2}$ and the above identities show that equations (3.53) to (3.61) are all satisfied for arbitrary $0 \leq M_1 \leq \frac{2}{3}$.

□

3.4 Evolution of local equilibria for the one-step simple random walk on the 3-node model

Now that we know which non-trivial solutions exist to the intertwining equations in this model, we may adapt Theorem 2.1 to this specific case. This is done in the following proposition.

Proposition 3.2 (Evolution of local equilibria of 1-step distribution on the 3-node model)**.** Let

$$P = \begin{pmatrix} p & 1-p & 0 \\ 1-p & 0 & p \\ 0 & p & 1-p \end{pmatrix}$$

be the transition matrix associated to the process X of a one-step simple random walk on the 3-node model with parameter $p \in (0,1)$ and $\lambda := \sqrt{3p^2 - 3p + 1}$.
Let

$$\Lambda = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix},$$

be such that

- 1. If

$$0 \leq m_1 \leq \frac{1}{3} \leq M_1 \leq \frac{-p + 1 + \lambda}{2 + 2\lambda - 3p} \tag{3.66}$$

or

$$0 \leq M_1 \leq \frac{1}{3} \leq m_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}, \tag{3.67}$$

then $\overline{p} = \overline{q} + \lambda$.

2. If

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}\right\} \leq m_1 \leq \frac{1}{3} \tag{3.68}$$

$$\frac{1}{3}(1+\lambda) - \lambda m_1 \leq M_1 \leq \min\left\{\frac{-p}{1-\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\}, \tag{3.69}$$

or

$$\frac{1}{3} \leq m_1 \leq \frac{-p}{1-\lambda-p} \tag{3.70}$$

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\} \leq M_1 \leq \frac{1}{3}(1+\lambda) - \lambda m_1, \tag{3.71}$$

then $\overline{p} = \overline{q} - \lambda$.

•

$$M_2 = \frac{(1+\alpha-3p)M_1 + p}{1+\alpha} \tag{3.72}$$

and

$$m_2 = \frac{(1+\alpha-3p)m_1 + p}{1+\alpha}; \tag{3.73}$$

•

$$M_3 = 1 - M_1 - M_2 \tag{3.74}$$

and

$$m_3 = 1 - m_1 - m_2; \tag{3.75}$$

and finally given $M_1 \neq m_1$,

•

$$\overline{q} = \frac{p(1-p)(1-3m_1)}{(1+\alpha)(M_1 - m_1)}, \tag{3.76}$$

where $\alpha = \overline{p} - \overline{q}$.

Then there exist two stopping times

$$T_{\overline{1}} \sim \text{Geom}\,(1-\overline{p}) \tag{3.77}$$

$$T_{\overline{2}} \sim \text{Geom}\,(\overline{q}) \tag{3.78}$$

for the process X such that

1. $\nu_{\overline{x}}$ is stationary until time $T_{\overline{x}}$, which means that

$$P_{\nu_{\overline{x}}}\left(X(t) = \cdot | T_{\overline{x}} > t\right) = \nu_{\overline{x}}(\cdot); \tag{3.79}$$

2. for all $\overline{x} \neq \overline{y}$ we have

$$P_{\nu_{\overline{x}}}(X(T_{\overline{x}}) = \cdot) = \nu_{\overline{y}}(\cdot); \tag{3.80}$$

3. $X(T_{\overline{x}})$ and $T_{\overline{x}}$ are independent for all $\overline{x}$.

The non-trivial solutions that we found in Proposition 3.1 allow us to explicitly give the parameters in Theorem 2.1.

We found

$$1 - \overline{P}(\overline{1},\overline{1}) = 1 - \overline{p} = 1 - \frac{p(1-p) + (1+\lambda)(\lambda M_1 - m_1)}{(1+\lambda)(M_1 - m_1)} = \frac{-p(1-p)(1-3M_1)}{(1+\lambda)(M_1 - m_1)};$$
$$1 - \overline{P}(\overline{2},\overline{2}) = \overline{q} = \frac{p(1-p)(1-3m_1)}{(1+\lambda)(M_1 - m_1)}.$$

Since the coarse-grained state space $\overline{\mathcal{X}}$ has only two states $\overline{1}, \overline{2}$, the need for a random variable $\overline{Y}_{\overline{x}}$ disappears. After a random time $T_{\overline{x}}$, the process X that starts from the part of the graph associated to $\overline{x}$ will transition almost surely to the only other part of the graph $\overline{y} \neq \overline{x}$.

3.5 Analysis of the solutions

In consideration of our discussion of metastability, we would like to find $\nu_{\overline{1}}$ and $\nu_{\overline{2}}$ that place most of their mass on the region of the state space that is associated to $\overline{1}$ and $\overline{2}$ respectively.
Proposition 3.1 gives us an easy way of doing that by taking the maximal M_1 and minimal m_1 (or vice versa), i.e. if we take

$$M_1 = \frac{-p+1+\lambda}{2+2\lambda-3p}, \quad m_1 = 0.$$

However it turns out that if we fix M_1 and vary $0 \leq m_1 \leq \frac{1}{3}$, interesting observations can be made.

- In Section 3.5.1 we will see that **the average geometric time $T_{\nu_{\overline{1}}}$ of escaping the course grained state $\overline{1}$ decreases linearly as m_1 increases from 0 to $\frac{1}{3}$.**
- Further we discuss the **role of intertwining in the context of mixing times** in Section 3.5.2.

In this discussion we shall only consider the case $\overline{p} - \overline{q} = +\lambda$.

3.5.1 Behaviour of the geometric escape times

We shall now examine the geometric time of Proposition 3.2.
First let us fix the notation. As we associate the states of the coarse-grained set $\overline{x} \in \overline{\mathcal{X}}$ to the measures $\nu_{\overline{x}}$, we shall write

$$T_{\nu_{\overline{1}}}$$

for the geometric time of Proposition 3.2 that tells us when the process escapes the "metastable" region of the state space that is indicated by $\overline{1}$ (corresponding to the original process starting from the measure $\nu_{\overline{1}}$). Analogously for

$$T_{\nu_{\overline{2}}}.$$

Lemma 3.4 (Dependence of geometric escape time on parameter m_1)**.** For fixed $p \in (0,1)$ consider the two measures

$$\nu_{\overline{1}} = \left(M_1, \frac{(1+\lambda-3p)M_1 + p}{1+\lambda}, \frac{1+\lambda-p+(-2-2\lambda+3p)M_1}{1+\lambda} \right)$$

and

$$\nu_{\overline{2}} = \left(m_1, \frac{(1+\lambda-3p)m_1 + p}{1+\lambda}, \frac{1+\lambda-p+(-2-2\lambda+3p)m_1}{1+\lambda} \right),$$

where $\frac{1}{3} \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}$ is fixed and $0 \leq m_1 \leq \frac{1}{3}$.
If

$$m_1 = 0,$$

then the average time of escaping $\overline{1}$, i.e. $\mathbb{E}[T_{\nu_{\overline{1}}}]$, is maximal and the average time of escaping $\overline{2}$, i.e. $\mathbb{E}[T_{\nu_{\overline{2}}}]$ is minimal.

Proof. Recall that

$$T_{\nu_{\overline{1}}} \sim \text{Geom}(1-\overline{p})$$

and

$$T_{\nu_{\overline{2}}} \sim \text{Geom}(\overline{q}).$$

- We first see that

$$\begin{aligned}\mathbb{E}[T_{\nu_{\overline{1}}}] &= (1-\overline{p})^{-1} \\ &= \frac{(1+\lambda)(M_1-m_1)}{-p(1-p)(1-3M_1)} \\ &= \frac{(1+\lambda)M_1}{-p(1-p)(1-3M_1)} - \frac{(1+\lambda)m_1}{-p(1-p)(1-3M_1)}\end{aligned}$$

  Note that $\frac{(1+\lambda)m_1}{-p(1-p)(1-3M_1)}$ is non-negative for all $0 \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}$ and $p \in (0,1)$. This means we can conclude that $\mathbb{E}[T_{\nu_1}(m_1)]$ decreases linearly as a function of m_1. Thus it will be maximal for $m_1 = 0$.

- For the other escape time we observe that

$$\begin{aligned}\mathbb{E}[T_{\nu_{\overline{2}}}] &= \overline{q}^{-1} \\ &= \frac{(1+\lambda)(M_1-m_1)}{p(1-p)(1-3m_1)}\end{aligned}$$

  As we are constrained to $m_1 \leq \frac{1}{3} \leq M_1$, we see that $\mathbb{E}[T_{\nu_{\overline{2}}}] \to \infty$ as $m_1 \to \frac{1}{3}$ (as we expect since in that case $\nu_{\overline{2}}$ is the invariant measure).

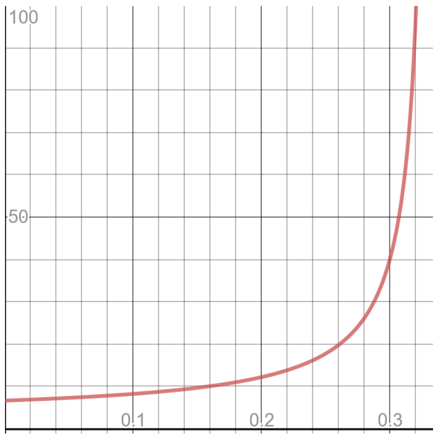


Figure 3.15: The escape time $T_{\nu_{\overline{2}}}$ as a function of m_1, where we have chosen $M_1 = \frac{3}{4}$ and $p = \frac{3}{4}$.

  In the figure above we see that $\mathbb{E}[T_{\nu_{\overline{2}}}]$ is minimal for $m_1 = 0$.

□

Intuitively this result makes sense, as it will take the process a longer time to go from a measure $\nu_{\overline{1}}$ that places most of its mass on the region of the state space associated to $\overline{1}$ to the other extreme measure that lays no mass on state $\overline{1}$. On the other hand escaping from $\overline{2}$ will take less time when $m_1 = 0$, since increasing m_1 will cause $\nu_{\overline{2}}$ to approach the invariant measure.

Example 3.1. We can see this already in the balanced set-up, where $p = \frac{1}{2}$. Let us take $\nu_{\overline{1}}$ to place as much mass as possible on the first vertex, i.e.

$$\nu_{\overline{1}} = \left(\frac{2}{3}, \frac{1}{3}, 0\right).$$

- If we take $m_1 = 0$, i.e.

$$\nu_{\overline{2}} = \left(0, \frac{1}{3}, \frac{2}{3}\right),$$

we see that

$$\mathbb{E}[T_{\nu_{\overline{1}}}] = 4,$$
$$\mathbb{E}[T_{\nu_{\overline{2}}}] = 4.$$

- If we take $m_1 = \frac{1}{3}$, i.e.

$$\nu_{\overline{2}} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right),$$

we see that

$$\mathbb{E}[T_{\nu_{\overline{1}}}] = 2,$$
$$\mathbb{E}[T_{\nu_{\overline{2}}}] = \infty.$$

Thus once the global minimum is reached, the walk will stay there infinitely long.

3.5.2 Bounds on the mixing times in the context of coarse graining

It is a known fact that finite state irreducible aperiodic Markov chains have a unique invariant distribution and that no matter what the initial distribution, the time-t distribution of the chain will converge to the invariant distribution. This is also the case for our 3-node toy model.

In order to be able to speak about convergence of the time-t distribution, we need a notion of distance between the distributions for which we will use the *total variation distance*.

Definition 3.1 (Total variation distance)**.** Let μ, ν be two probability measures on a state space $\mathcal{X}$. The *total variation distance* between μ and ν is given by

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|. \tag{3.81}$$

Thus we know that in our model for any initial distribution μ on $\mathcal{X}$ it holds

$$d_{TV}(\mu P^t, \pi) \to 0$$

as $t \to \infty$.
Now it is interesting to see how quickly the time-t distribution converges to π. That is where the concept of mixing time comes in.

Definition 3.2 (Mixing time)**.** For $\epsilon > 0$ and initial distribution μ, the *mixing time* of the chain *starting from* μ is given by

$$t^{\mu}_{\text{mix}}(\epsilon) = \min\{t \geq 0 : d_{TV}(\mu P^t, \pi) < \epsilon\}. \tag{3.82}$$

The interesting subject that we are able to discuss now is the behaviour of mixing times in the context of intertwining. Since our metastable states are measures $\nu_{\overline{x}}$ for $\overline{x} \in \overline{\mathcal{X}}$, we can look at the mixing time of the process X when starting from such a measure $\nu_{\overline{x}}$, which we will denote by

$$t^{\nu_{\overline{x}}}_{\text{mix}}(\epsilon)$$

for $\epsilon > 0$. Then we can compare this to the mixing time of the coarse-grained process when starting from the measure $\delta_{\overline{x}}$, which we will denote by

$$\overline{t}^{\delta_{\overline{x}}}_{\text{mix}}(\epsilon)$$

for $\epsilon > 0$.

Indeed Proposition 2.1 tells us something about these mixing times. Since we know that thermalization of the coarse-grained process implies that the original process has already thermalized, we can conclude that

$$t_{\text{mix}}^{\nu_{\bar{x}}}(\epsilon) \leq \bar{t}_{\text{mix}}^{\delta_{\bar{x}}}(\epsilon).$$

The question is **how good is this bound?** We will explore this question for the 3-node model by using the explicit intertwining solutions of Proposition 3.1. In particular we see for which choice of parameters M_1, m_1 this bound is optimal.

Lemma 3.5 (Optimal bound on mixing time of the original chain)**.** For fixed $p \in [1/2, 1)$ consider the two measures

$$\nu_{\bar{1}} = \left(M_1, \frac{(1+\lambda-3p)M_1 + p}{1+\lambda}, \frac{1+\lambda-p+(-2-2\lambda+3p)M_1}{1+\lambda} \right)$$

and

$$\nu_{\bar{2}} = \left(m_1, \frac{(1+\lambda-3p)m_1 + p}{1+\lambda}, \frac{1+\lambda-p+(-2-2\lambda+3p)m_1}{1+\lambda} \right),$$

where $\frac{1}{3} \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}$ and $0 \leq m_1 \leq \frac{1}{3}$.
Then

$$\bar{t}_{\text{mix}}^{\delta_{\bar{x}}}(\epsilon) - t_{\text{mix}}^{\nu_{\bar{x}}}(\epsilon) = \frac{\log(M_1 - m_1)}{\log \lambda} \geq 0.$$

Remark. Since we have that

$$\bar{t}_{\text{mix}}^{\delta_{\bar{x}}}(\epsilon) - t_{\text{mix}}^{\nu_{\bar{x}}}(\epsilon) = \frac{\log(M_1 - m_1)}{\log \lambda}.$$

we see that this difference vanishes when $M_1 - m_1 = 1$. Due to the constraints on M_1, m_1, we conclude that the bound is optimal in the case

$$m_1 = 0, \quad M_1 = \frac{-p+1+\lambda}{2+2\lambda-3p}.$$

Proof. (Lemma 3.4)
The proof is given in the following steps:

1. **First we compute $\nu_{\bar{1}} P^t$.**
   We must first **find the t-step distributions** of the Markov chain.
   For the network we have

$$P_{\text{even}}^t = \frac{1}{3} \begin{pmatrix} 1+2\lambda^t & 1-\lambda^t & 1-\lambda^t \\ 1-\lambda^t & 1+2\lambda^t & 1-\lambda^t \\ 1-\lambda^t & 1-\lambda^t & 1+2\lambda^t \end{pmatrix} \tag{3.83}$$

   if t is even and

$$P_{\text{odd}}^t = \frac{1}{3} \begin{pmatrix} 1+(3p-1)\lambda^{t-1} & 1+(-3p+2)\lambda^{t-1} & 1-\lambda^{t-1} \\ 1+(-3p+2)\lambda^{t-1} & 1-\lambda^{t-1} & 1+(3p-1)\lambda^{t-1} \\ 1-\lambda^{t-1} & 1+(3p-1)\lambda^{t-1} & 1+(-3p+2)\lambda^{t-1} \end{pmatrix} \tag{3.84}$$

   if t is odd (to see this derivation we refer the reader to Section 4.1).

   - **We now compute $\nu_{\bar{1}} P_{\text{even}}^t$.**

We find

$$\begin{aligned}
\nu_{\bar{1}} P^t_{\text{even}}(1) &= \frac{1}{3}(M_1 + 2M_1\lambda^t + M_2 - M_2\lambda^t + M_3 - M_3\lambda^t) \\
&= \frac{1}{3}(1 + 3M_1\lambda^t - \lambda^t) \\
&= \frac{1}{3}(1 + (3M_1 - 1)\lambda^t) \\
\nu_{\bar{1}} P^t_{\text{even}}(2) &= \frac{1}{3}(1 + (3M_2 - 1)\lambda^t) \\
\nu_{\bar{1}} P^t_{\text{even}}(3) &= \frac{1}{3}(1 + (3M_3 - 1)\lambda^t),
\end{aligned}$$

where

$$\begin{aligned}
3M_2 - 1 &= \frac{(3 + 3\lambda - 9p)M_1 + 3p - 1 - \lambda}{1 + \lambda} \\
&= \frac{1 + \lambda - 3p}{1 + \lambda}(3M_1 - 1), \\
3M_3 - 1 &= \frac{-(6 + 6\lambda - 9p)M_1 + 2 + 2\lambda - 3p}{1 + \lambda} \\
&= -\frac{2 + 2\lambda - 3p}{1 + \lambda}(3M_1 - 1).
\end{aligned}$$

- **We now compute $\nu_{\bar{1}} P^t_{\text{odd}}$.**
  We find

$$\begin{aligned}
\nu_{\bar{1}} P^t_{\text{odd}}(1) &= \frac{1}{3}(1 + (3p-1)\lambda^{t-1}M_1 + (-3p+2)\lambda^{t-1}M_2 - \lambda^{t-1}M_3) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3}(3pM_1 + (3-3p)M_2 - 1) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}(3p(1+\lambda)M_1 + (3-3p)(1+\lambda-3p)M_1 + (3-3p)p - (1+\lambda)) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}((3p + 3p\lambda + 3 + 3\lambda - 9p - 3p - 3p\lambda + 9p^2)M_1 - 1 - \lambda + 3p - 3p^2) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}(\lambda^2 + \lambda)(3M_1 - 1) \\
&= \frac{1}{3}(1 + (3M_1 - 1)\lambda^t) \\
\nu_{\bar{1}} P^t_{\text{odd}}(2) &= \frac{1}{3} + \frac{\lambda^{t-1}}{3}(3pM_3 + (3-3p)M_1 - 1) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}(3p(1+\lambda) - 3p^2 + (-6p - 6p\lambda + 9p^2 + 3 - 3p + 3\lambda - 3p\lambda)M_1 - (1+\lambda)) \\
&= \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}(-\lambda^2 + 3p\lambda - \lambda + (9p^2 - 9p + 3 - 9p\lambda + 3\lambda)M_1) \\
&= \frac{1}{3} + \frac{\lambda^{t}}{3(1+\lambda)}(-1 - \lambda + 3p + 3(\lambda - 3p + 1)M_1) \\
&= \frac{1}{3}(1 + \frac{1 + \lambda - 3p}{1 + \lambda}(3M_1 - 1)\lambda^t)
\end{aligned}$$

$$
\begin{aligned}
\nu_{\bar{1}}P^t_{\text{odd}}(3) =& \frac{1}{3} + \frac{\lambda^{t-1}}{3}(3pM_2 + (3-3p)M_1 - 1) \\
=& \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}(3p(1+\lambda-3p)M_1 + 3p^2 + (3-3p)(1+\lambda-p) \\
& - (3-3p)(2+2\lambda-3p)M_1 - (1+\lambda)) \\
=& \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}[3M_1(p+p\lambda-3p^2-2-2\lambda+3p+2p+2p\lambda-3p^2) \\
& + 3p^2+3-3p+3\lambda-3p\lambda-3p+3p^2-1-\lambda] \\
=& \frac{1}{3} + \frac{\lambda^{t-1}}{3(1+\lambda)}[3M_1(-2\lambda^2+3p\lambda-2\lambda)+2\lambda^2+2\lambda-3p\lambda] \\
=& \frac{1}{3}(1 - \frac{2+2\lambda-3p}{1+\lambda}(3M_1-1)\lambda^t).
\end{aligned}
$$

2. **Then we calculate $d_{TV}(\nu_{\bar{1}}P^t, \pi)$.**
As we know that $\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, we find that it equals

$$
d_{TV}(\nu_{\bar{1}}P^t, \pi) = \frac{1}{6}\left[1 + \left|\frac{1+\lambda-3p}{1+\lambda}\right| + \left|\frac{2+2\lambda-3p}{1+\lambda}\right|\right]\lambda^t(3M_1-1).
$$

In the figure below we see that $1+\lambda-3p$ is negative for $p \geq \frac{1}{2}$.

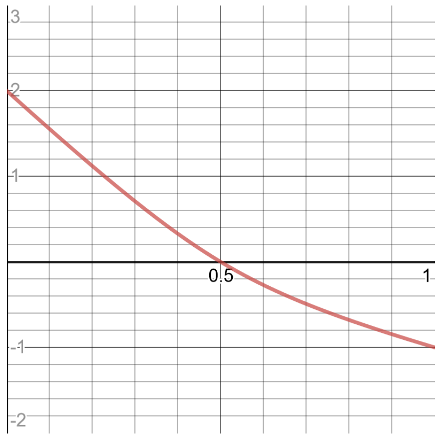


Figure 3.16: The function $1+\lambda-3p$ as a function of $p \in [0,1]$.

And below we see that $2+2\lambda-3p$ is positive for all p.

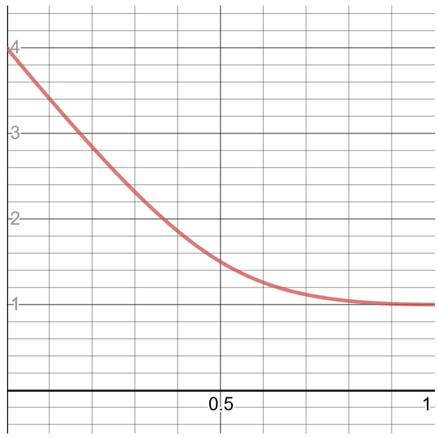


Figure 3.17: The function $2+2\lambda-3p$ as a function of $p \in [0,1]$.

Thus we have

$$d_{TV}(\nu_{\overline{1}}P^t, \pi) = \frac{1}{6}\left[\frac{1+\lambda-1-\lambda+3p+2+2\lambda-3p}{1+\lambda}\right]\lambda^t(3M_1-1)$$
$$= \frac{1}{3}\lambda^t(3M_1-1). \tag{3.85}$$

3. **We compute $\delta_{\overline{1}}\overline{P}^t$.**
First we must find $\overline{P}^t$. This equals

$$\overline{P}^t = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix}^t$$
$$= \begin{pmatrix} 1 & \frac{\overline{p}-1}{\overline{q}} \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \lambda^t \end{pmatrix}\begin{pmatrix} \frac{\overline{q}}{1-\lambda} & \frac{1-\overline{p}}{1-\lambda} \\ \frac{-\overline{q}}{1-\lambda} & \frac{\overline{q}}{1-\lambda} \end{pmatrix}$$
$$= \frac{1}{1-\lambda}\begin{pmatrix} \overline{q}+\lambda^t(1-\overline{p}) & (1-\overline{p})(1-\lambda^t) \\ \overline{q}(1-\lambda^t) & 1-\overline{p}+\lambda^t\overline{q} \end{pmatrix}.$$

Then we find from our expressions of $\overline{p}$ and $\overline{q}$ that

$$\delta_{\overline{1}}\overline{P}^t = \frac{p(1-p)}{(1-\lambda^2)(M_1-m_1)}\left((1-3m_1)+\lambda^t(3M_1-1), (1-\lambda^t)(3M_1-1)\right)$$
$$= \frac{1}{3(M_1-m_1)}\left((1-3m_1)+\lambda^t(3M_1-1), 3M_1-1-\lambda^t(3M_1-1)\right). \tag{3.86}$$

4. **Now we must find the invariant measure $\overline{\pi}$ of the coarse-grained process.**
We want to solve the system of equations

$$\overline{\pi}_1 = \overline{p}\overline{\pi}_1 + \overline{q}\overline{\pi}_2$$
$$\overline{\pi}_2 = (1-\overline{p})\overline{\pi}_1 + (1-\overline{q})\overline{\pi}_2$$
$$\overline{\pi}_1 + \overline{\pi}_2 = 1.$$

This is equivalent to solving

$$(1-\overline{p}+\overline{q})\overline{\pi}_1 = \overline{q}.$$

Solving this returns

$$\overline{\pi}_1 = \frac{1-3m_1}{3(M_1-m_1)}, \quad \overline{\pi}_2 = \frac{-1+3M_1}{3(M_1-m_1)}. \tag{3.87}$$

5. **Then we find that**

$$d_{TV}(\delta_{\overline{1}}\overline{P}^t, \overline{\pi}) = \frac{1}{3}\lambda^t\frac{3M_1-1}{M_1-m_1}. \tag{3.88}$$

6. **Finally we give the respective mixing times.**
For $\epsilon > 0$ we find from Steps 2 and 5 that

$$t_{\text{mix}}^{\nu_{\overline{x}}}(\epsilon) = \frac{\log\frac{3\epsilon}{3M_1-1}}{\log\lambda} \tag{3.89}$$

and that

$$\overline{t}_{\text{mix}}^{\delta_{\overline{x}}}(\epsilon) = \frac{\log\left[\frac{3\epsilon}{3M_1-1}(M_1-m_1)\right]}{\log\lambda}$$
$$= \frac{\log\frac{3\epsilon}{3M_1-1}}{\log\lambda} + \frac{\log(M_1-m_1)}{\log\lambda}$$
$$= t_{\text{mix}}^{\nu_{\overline{x}}}(\epsilon) + \frac{\log(M_1-m_1)}{\log\lambda}. \tag{3.90}$$

□

Chapter 4

T-step random walk on the 3-node toy model

We saw in the previous chapter that intertwining on the 3-node model has non-trivial solutions if we consider the one-step distribution. We expect however for the space of solutions to increase if we examine the T-step distribution on the same model for $T \in \mathbb{Z}_{>1}$. We expect this as P^T looks at the random walk on a certain time scale T. Since now we can fix T from the start and choose it to be large, we will see more solutions and in particular that the dynamics of more local equilibria can be captured. This is what we shall show in practice in this section. In fact the space of solutions depends on the parity of T.

- In Section 4.1 we first compute the T-step distribution.
- In Section 4.2 we show that for T odd, we only find the same non-trivial solutions of Λ as for the one-step distribution but with a different $\overline{P}$ depending on T.
- If T is even, the space of solutions increases. We show in Section 4.3 that there are non-trivial solutions for all $p \in (0, 1)$ and we have more freedom to choose the measure $\nu_{\overline{x}}$ than in the case of T odd. Furthermore the solutions of the one-step are a subset of this space of solutions. (Sections 4.2 and 4.3 find solutions to our intertwining problem through direct computation and thus are analogous to Section 3.2.)
- We reformulate Theorem 2.1 in Section 4.4 to account for the increased number of solutions when T is even.
- Finally it is demonstrated that we can also find non-trivial intertwining solutions using the spectral solutions of Lemma 2.2, but again with extra conditions imposed on them.

Note that we take T here to be finite. If that were not the case, we know that each row of P^T becomes the invariant measure since our model is an irreducible and aperiodic Markov chain on a finite state space.

4.1 Finding the T-step distribution

Before finding solutions of intertwining in this case, we must find the expressions for P^T, where $T \in \mathbb{N}$.

Proposition 4.1 (Closed form expression of P^T)**.** For the transition matrix P of (3.1) and $T \in \mathbb{N}$, we find that P^T equals:

$$P^T = \frac{1}{3}\begin{pmatrix} 1+2\lambda^T & 1-\lambda^T & 1-\lambda^T \\ 1-\lambda^T & 1+2\lambda^T & 1-\lambda^T \\ 1-\lambda^T & 1-\lambda^T & 1+2\lambda^T \end{pmatrix}$$

if T is even and

$$P^T = \frac{1}{3}\begin{pmatrix} 1+(3p-1)\lambda^{T-1} & 1+(-3p+2)\lambda^{T-1} & 1-\lambda^{T-1} \\ 1+(-3p+2)\lambda^{T-1} & 1-\lambda^{T-1} & 1+(3p-1)\lambda^{T-1} \\ 1-\lambda^{T-1} & 1+(3p-1)\lambda^{T-1} & 1+(-3p+2)\lambda^{T-1} \end{pmatrix}$$

if T is odd.

Proof. We can write

$$P = SJS^{-1},$$

with the J the diagonal matrix of P:

$$J = \begin{pmatrix} 1 & & \\ & -\lambda & \\ & & \lambda \end{pmatrix},$$

where

$$\lambda := \sqrt{3p^2 - 3p + 1}, \tag{4.1}$$

and the other matrices are given by

$$S = \begin{pmatrix} 1 & \frac{-2p+\lambda+1}{p} & \frac{-2p-\lambda+1}{p} \\ 1 & \frac{p-\lambda-1}{p} & \frac{p+\lambda-1}{p} \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$S^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6}\left(\frac{1}{\lambda}-1\right) & \frac{-(3p+\lambda-1)}{6\lambda} & \frac{3p+2\lambda-2}{6\lambda} \\ \frac{1}{6}\left(-\frac{1}{\lambda}-1\right) & \frac{3p-\lambda-1}{6\lambda} & \frac{-(3p-2\lambda-2)}{6\lambda} \end{pmatrix}.$$

To find $P^T = SJ^TS^{-1}$, we first compute J^TS^{-1}:

$$J^TS^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6}(-\lambda)^T\left(\frac{1}{\lambda}-1\right) & -(-\lambda)^T\frac{3p+\lambda-1}{6\lambda} & (-\lambda)^T\frac{3p+2\lambda-2}{6\lambda} \\ \frac{1}{6}\lambda^T\left(-\frac{1}{\lambda}-1\right) & \lambda^T\frac{3p-\lambda-1}{6\lambda} & -\lambda^T\frac{-(3p-2\lambda-2)}{6\lambda} \end{pmatrix}.$$

We compute now the nine values in P^T :

$$\begin{aligned}
P_{1,1}^T =& \frac{1}{3} + \frac{\lambda^T}{6p}\left((-1)^T(\frac{1}{\lambda}-1)(-2p+\lambda+1) + (\frac{-1}{\lambda}-1)(-2p-\lambda+1)\right) \\
=& \frac{1}{3} + \frac{\lambda^T}{3}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6p}(1-2p-\lambda^2)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} + \frac{\lambda^T}{3}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}(-3p+1)\left(-1+(-1)^T\right) \qquad (4.2) \\
P_{1,2}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^{T+1}(3p+\lambda-1)(-2p+\lambda+1) + (3p-\lambda-1)(-2p-\lambda+1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6p}(\lambda^2-6p^2+5p-1)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6}(-3p+2)\left(-1+(-1)^T\right) \qquad (4.3)
\end{aligned}$$

$$\begin{aligned}
P_{1,3}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^T(3p+2\lambda-2)(-2p+\lambda+1) - (3p-2\lambda-2)(-2p-\lambda+1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6p}(2\lambda^2-6p^2+7p-2)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}\left(-1+(-1)^T\right) \qquad (4.4) \\
P_{2,1}^T =& \frac{1}{3} + \frac{\lambda^T}{6p}\left((-1)^T(\frac{1}{\lambda}-1)(p-\lambda-1) + (\frac{-1}{\lambda}-1)(p+\lambda-1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6p}(1-p-\lambda^2)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6}(-3p+2)\left(-1+(-1)^T\right) \qquad (4.5) \\
P_{2,2}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^{T+1}(3p+\lambda-1)(p-\lambda-1) + (3p-\lambda-1)(p+\lambda-1)\right) \\
=& \frac{1}{3} + \frac{\lambda^T}{3}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6p}(-\lambda^2+3p^2-4p+1)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} + \frac{\lambda^T}{3}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}\left(-1+(-1)^T\right) \qquad (4.6) \\
P_{2,3}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^T(3p+2\lambda-2)(p-\lambda-1) - (3p-2\lambda-2)(p+\lambda-1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6p}(-2\lambda^2+3p^2-5p+2)\left(-1+(-1)^T\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}(-3p+1)\left(-1+(-1)^T\right) \qquad (4.7) \\
P_{3,1}^T =& \frac{1}{3} + \frac{\lambda^T}{6p}\left((-1)^T(\frac{1}{\lambda}-1) + (\frac{-1}{\lambda}-1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}\left(-1+(-1)^T\right) \qquad (4.8) \\
P_{3,2}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^{T+1}(3p+\lambda-1) + (3p-\lambda-1)\right) \\
=& \frac{1}{3} - \frac{\lambda^T}{6}\left(1+(-1)^T\right) - \frac{\lambda^{T-1}}{6}(3p-1)\left(-1+(-1)^T\right) \qquad (4.9) \\
P_{3,3}^T =& \frac{1}{3} + \frac{\lambda^{T-1}}{6p}\left((-1)^T(3p+2\lambda-2) - (3p-2\lambda-2)\right) \\
=& \frac{1}{3} + \frac{\lambda^T}{3}\left(1+(-1)^T\right) + \frac{\lambda^{T-1}}{6}(3p-2)\left(-1+(-1)^T\right). \qquad (4.10)
\end{aligned}$$

□

Note for our sanity check that if we sum up the rows, we indeed get back 1 thus having a stochastic matrix.

Now that we have the explicit transition probabilities for the T-step random walk, we can write down the **constraints of the intertwining**:

$$\begin{aligned}
\overline{p}(M_1 - m_1) + m_1 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(2M_1 - M_2 - M_3) && (4.11)\\
&+ \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[(-3p+1)M_1 + (3p-2)M_2 + M_3] \\
\overline{p}(M_2 - m_2) + m_2 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(-M_1 + 2M_2 - M_3) && (4.12)\\
&- \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[(-3p+2)M_1 - M_2 + (3p-1)M_3] \\
\overline{p}(M_3 - m_3) + m_3 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(-M_1 - M_2 + 2M_3) && (4.13)\\
&+ \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[M_1 + (-3p+1)M_2 + (3p-2)M_3] \\
\overline{q}(M_1 - m_1) + m_1 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(2m_1 - m_2 - m_3) && (4.14)\\
&+ \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[(-3p+1)m_1 + (3p-2)m_2 + m_3] \\
\overline{q}(M_2 - m_2) + m_2 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(-m_1 + 2m_2 - m_3) && (4.15)\\
&- \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[(-3p+2)m_1 - m_2 + (3p-1)m_3] \\
\overline{q}(M_3 - m_3) + m_3 =& \frac{1}{3} + \frac{\lambda^T}{6}\left(1 + (-1)^T\right)(-m_1 - m_2 + 2m_3) && (4.16)\\
&+ \frac{\lambda^{T-1}}{6}\left(-1 + (-1)^T\right)[m_1 + (-3p+1)m_2 + (3p-2)m_3] \\
1 =& M_1 + M_2 + M_3 && (4.17)\\
1 =& m_1 + m_2 + m_3 && (4.18)\\
[0,1] \ni& M_1, M_2, M_3, m_1, m_2, m_3, \overline{p}, \overline{q}. && (4.19)
\end{aligned}$$

Note that (4.11) to (4.16) reduce to two different sets of equations depending on the parity of T.

4.2 Characterization of intertwining solutions for odd T-step distribution

Let us investigate what happens **if T is odd**. It turns out that the **the solutions found here are consist of the same Λ as those of the one-step distribution, but now with different $\overline{P}$.**

Lemma 4.1 (Characterization of intertwining solutions for odd T-step distribution)**.** If T is odd, the solutions $(\Lambda, \overline{P}_T)$ of the intertwining $\Lambda P^T = \overline{P}_T \Lambda$ of the form

$$\Lambda = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

and

$$\overline{P}_T = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix}$$

are characterized by the difference $\overline{p} - \overline{q}$, which can only be $\pm\lambda^T$. The solutions are identified by

- 1. If $\overline{p} - \overline{q} = +\lambda^T$, then either

  $$0 \le m_1 \le \frac{1}{3} \le M_1 \le \frac{-p+1+\lambda}{2+2\lambda-3p} \tag{4.20}$$

  or

  $$0 \le M_1 \le \frac{1}{3} \le m_1 \le \frac{-p+1+\lambda}{2+2\lambda-3p}; \tag{4.21}$$

  2. If $\overline{p} - \overline{q} = -\lambda^T$, then either

  $$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}\right\} \le m_1 \le \frac{1}{3} \tag{4.22}$$

  $$\frac{1}{3}(1+\lambda) - \lambda m_1 \le M_1 \le \min\left\{\frac{-p}{1-\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\}; \tag{4.23}$$

  or

  $$\frac{1}{3} \le m_1 \le \frac{-p}{1-\lambda-p} \tag{4.24}$$

  $$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\} \le M_1 \le \frac{1}{3}(1+\lambda) - \lambda m_1; \tag{4.25}$$

- $$M_2 = \frac{(1+\alpha-3p)M_1 + p}{1+\alpha} \tag{4.26}$$

  and

  $$m_2 = \frac{(1+\alpha-3p)m_1 + p}{1+\alpha}; \tag{4.27}$$

- $$M_3 = 1 - M_1 - M_2 \tag{4.28}$$

  and

  $$m_3 = 1 - m_1 - m_2; \tag{4.29}$$

and finally given $M_1 \ne m_1$,

- $$\overline{q} = (M_1 - m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1}) + \frac{\lambda^{T-1}p(1-p)}{1+\alpha} + (\alpha^T - 1)m_1\right]; \tag{4.30}$$

- $$\overline{p} = \overline{q} + \alpha^T; \tag{4.31}$$

where $\alpha = +\lambda$ if $\overline{p} - \overline{q} = +\lambda^T$ and $\alpha = -\lambda$ if $\overline{p} - \overline{q} = -\lambda^T$.

Proof. If T is odd, the intertwining constraints (4.11) to (4.16) reduce to:

$$\begin{aligned}
\overline{p}(M_1 - m_1) + m_1 &= -\frac{\lambda^{T-1}}{3}\left[(-3p+1)M_1 + (3p-2)M_2 + M_3\right] \\
\overline{p}(M_2 - m_2) + m_2 &= \frac{\lambda^{T-1}}{3}\left[(-3p+2)M_1 - M_2 + (3p-1)M_3\right] \\
\overline{p}(M_3 - m_3) + m_3 &= -\frac{\lambda^{T-1}}{3}\left[M_1 + (-3p+1)M_2 + (3p-2)M_3\right] \\
\overline{q}(M_1 - m_1) + m_1 &= -\frac{\lambda^{T-1}}{3}\left[(-3p+1)m_1 + (3p-2)m_2 + m_3\right] \\
\overline{q}(M_2 - m_2) + m_2 &= \frac{\lambda^{T-1}}{3}\left[(-3p+2)m_1 - m_2 + (3p-1)m_3\right] \\
\overline{q}(M_3 - m_3) + m_3 &= -\frac{\lambda^{T-1}}{3}\left[m_1 + (-3p+1)m_2 + (3p-2)m_3\right].
\end{aligned}$$

If we use the identities $M_1 + M_2 + M_3 = 1$ and $m_1 + m_2 + m_3 = 1$, we can rewrite these equations to:

$$\overline{p}(M_1 - m_1) + m_1 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[pM_1 + (1-p)M_2\right] \tag{4.32}$$
$$\overline{p}(M_2 - m_2) + m_2 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[(1-p)M_1 + pM_3\right] \tag{4.33}$$
$$\overline{p}(M_3 - m_3) + m_3 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[pM_2 + (1-p)M_3\right] \tag{4.34}$$
$$\overline{q}(M_1 - m_1) + m_1 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[pm_1 + (1-p)m_2\right] \tag{4.35}$$
$$\overline{q}(M_2 - m_2) + m_2 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[(1-p)m_1 + pm_3\right] \tag{4.36}$$
$$\overline{q}(M_3 - m_3) + m_3 = \frac{1}{3}(1 - \lambda^{T-1}) + \lambda^{T-1}\left[pm_2 + (1-p)m_3\right]. \tag{4.37}$$

We find the solutions to these constraints in the following steps.

1. **First we derive a new set of constraints.**
   Combining equation (4.32) with (4.35) to eliminate the term $m_1 - \frac{1}{3}(1 - \lambda^{T-1})$, we find

   $$(\lambda^{T-1}p - \alpha)(M_1 - m_1) + \lambda^{T-1}(1-p)(M_2 - m_2) = 0$$

   where we define

   $$\alpha := \overline{p} - \overline{q}. \tag{4.38}$$

   Similarly from the pairs (4.33)-(4.36) and (4.34)-(4.37) we find respectively

   $$\begin{aligned}
   -\alpha(M_2 - m_2) + \lambda^{T-1}(1-p)(M_1 - m_1) + \lambda^{T-1}p(M_3 - m_3) &= 0 \\
   (\lambda^{T-1}(1-p) - \alpha)(M_3 - m_3) + \lambda^{T-1}p(M_2 - m_2) &= 0.
   \end{aligned}$$

   We rewrite these three equations as

   $$(\alpha - \beta)x - \gamma y = 0 \tag{4.39}$$
   $$\alpha y - \gamma x - \beta z = 0 \tag{4.40}$$
   $$(\alpha - \gamma)z - \beta y = 0, \tag{4.41}$$

where

$$\begin{aligned}
\beta &:= \lambda^{T-1}p \\
\gamma &:= \lambda^{T-1}(1-p) \\
x &:= M_1 - m_1 \\
y &:= M_2 - m_2 \\
z &:= M_3 - m_3.
\end{aligned}$$

2. **Now we consider $\alpha = 0$.**
If $\alpha = 0$, the equations (4.39) to (4.41) become

$$\begin{aligned}
-\beta x - \gamma y = 0 \\
-\gamma x - \beta z = 0 \\
-\gamma z - \beta y = 0.
\end{aligned}$$

Substituting the first two equations in the third, we obtain

$$\left(\frac{\gamma^2}{\beta} + \frac{\beta^2}{\gamma}\right) x = 0.$$

Since $\gamma^2 + \beta^2 = \lambda^{2T-2}((1-p)^2 + p^2) \neq 0$, it follows that $x = 0$ and hence $y = z = 0$ as well. Plugging this into (4.35) to (4.37) we obtain the equations for the invariant measure:

$$\begin{aligned}
m_1 =& \frac{1}{3}(1-\lambda^{T-1}) + \lambda^{T-1}\left[pm_1 + (1-p)m_2\right] \\
m_2 =& \frac{1}{3}(1-\lambda^{T-1}) + \lambda^{T-1}\left[(1-p)m_1 + pm_3\right] \\
m_3 =& \frac{1}{3}(1-\lambda^{T-1}) + \lambda^{T-1}\left[pm_2 + (1-p)m_3\right].
\end{aligned}$$

3. **If now either $\alpha = \beta$ or $\alpha = \gamma$**, then $y = 0$. The other equations return $x = z = 0$, hence we again find the invariant measure.

4. **Now suppose α is neither 0, β nor γ; then we shall show that we must have $\alpha = \lambda^{T-1}$ or $\alpha = \pm\lambda^T$.**
Equations (4.39) and (4.41) tell us that

$$\begin{aligned}
x &= \frac{\gamma}{\alpha - \beta} y \\
z &= \frac{\beta}{\alpha - \gamma} y.
\end{aligned}$$

Plugging this into (4.40) we find

$$\left[\alpha - \frac{\gamma^2}{\alpha - \beta} - \frac{\beta^2}{\alpha - \gamma}\right] y = 0. \tag{4.42}$$

We ignore $y = 0$ (as then $x = z = 0$ and we retrieve the invariant measure) and find

$$\begin{aligned}
0 =& \alpha(\alpha-\beta)(\alpha-\gamma) - \gamma^2(\alpha-\gamma) - \beta^2(\alpha-\beta) \\
=& \alpha^3 - \gamma\alpha^2 - \beta\alpha^2 + \beta\gamma\alpha - \gamma^2\alpha + \gamma^3 - \beta^2\alpha + \beta^3 \\
=& \alpha^3 + \alpha^2(-\gamma-\beta) + \alpha(\beta\gamma - \gamma^2 - \beta^2) + \gamma^3 + \beta^3 \\
=& \alpha^3 - \lambda^{T-1}\alpha^2 + \lambda^{2T-2}\alpha(p - p^2 - 1 - p^2 + 2p - p^2) + \lambda^{3T-3}(1 + 3p^2 - 3p) \\
=& \alpha^3 - \lambda^{T-1}\alpha^2 - \lambda^{2T}\alpha + \lambda^{3T-1} \\
=& (\lambda^{T-1} - \alpha)(\lambda^{2T} - \alpha^2).
\end{aligned}$$

This has solutions

$$\alpha = \lambda^{T-1} \quad \vee \quad \alpha = \pm\lambda^T.$$

5. Let us first consider **the case $\alpha = \lambda^{T-1}$.**
We rewrite (4.39) to (4.41) as

$$\begin{aligned}(1-p)(M_2-m_2) &= (1-p)(M_1-m_1)\\ M_2-m_2 &= (1-p)(M_1-m_1)+p(M_3-m_3)\\ p(M_2-m_2) &= p(M_3-m_3).\end{aligned}$$

Thus $M_1 - m_1 = M_2 - m_2 = M_3 - m_3$ and

$$\begin{aligned}1 =& M_1 + M_2 + M_3\\ =& M_1 + M_1 - m_1 + m_2 + M_1 - m_1 + m_3\\ =& 3(M_1 - m_1) + 1\end{aligned}$$

returns us again $M_1 - m_1 = 0$ and the invariant measure.

6. **We show now that we retrieve the Λ of the one-step walk.**
Let us now **consider**

$$\alpha = \pm\lambda^T. \tag{4.43}$$

We rewrite (4.39) to (4.41) as

$$(1-p)(M_2-m_2) = (\pm\lambda - p)(M_1-m_1) \tag{4.44}$$
$$\pm\lambda(M_2-m_2) = (1-p)(M_1-m_1)+p(M_3-m_3) \tag{4.45}$$
$$p(M_2-m_2) = (\pm\lambda-(1-p))(M_3-m_3). \tag{4.46}$$

But these are exactly the same equations as for the one-step simple random walk. Thus we find the same Λ here.

7. **The following step is to find $\overline{q}$.**
We use for this equation (4.35), and find in the case $\alpha = +\lambda^T$:

$$\begin{aligned}\overline{q} &= (M_1-m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1})+\lambda^{T-1}\left(pm_1+\frac{1-p}{1+\lambda}[(1+\lambda-3p)m_1+p]\right)-m_1\right]\\ &= (M_1-m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1})+\lambda^{T-1}(\frac{p(1-p)}{1+\lambda}+\frac{m_1}{1+\lambda}[1+\lambda-3p-p-p\lambda\right.\\ &\quad\left.+3p^2+p+p\lambda])-m_1\right]\\ &= (M_1-m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1})+\lambda^{T-1}(\frac{p(1-p)+m_1(\lambda+3p^2-3p+1)}{1+\lambda})-m_1\right]\\ &= (M_1-m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1})+\frac{\lambda^{T-1}p(1-p)}{1+\lambda}+(\lambda^T-1)m_1\right].\end{aligned} \tag{4.47}$$

For the case $\alpha = -\lambda^T$, we find similarly

$$\overline{q} = (M_1-m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1})+\frac{\lambda^{T-1}p(1-p)}{1-\lambda}+(-\lambda^T-1)m_1\right]. \tag{4.48}$$

8. **It rests now to discover that the constraints on M_1, m_1 that are induced by the constraints $\overline{q}, \overline{p}, M_i, m_i \in [0,1]$ are the same as in the one-step case.**
Let us first assume that $M_1 \geq m_1$. Then the constraint that $\overline{q} \geq 0$ gives us that

$$m_1 \leq \frac{\frac{1}{3}(1-\lambda^{T-1})+\frac{\lambda^{T-1}p(1-p)}{1+\lambda}}{1-\lambda^T}.$$

Note that as $p(1-p) = \frac{1}{3}(1-\lambda^2)$, we find

$$\begin{aligned}\frac{\frac{1}{3}(1-\lambda^{T-1})+\frac{\lambda^{T-1}p(1-p)}{1+\lambda}}{1-\lambda^T} &= \frac{1}{3}\left(\frac{1-\lambda^{T-1}+\lambda^{T-1}(1-\lambda)}{1-\lambda^T}\right)\\ &= \frac{1}{3}.\end{aligned}$$

Further the constraint $\overline{q} \leq 1 - \lambda^T$ gives us

$$\frac{1}{3}(1-\lambda^{T-1}) + \frac{\lambda^{T-1}p(1-p)}{1+\lambda} + (\lambda^T - 1)m_1 M_1 - m_1 \leq (1-\lambda^T)\frac{M_1 - m_1}{M_1 - m_1}.$$

This gives us

$$M_1 \geq \frac{\frac{1}{3}(1-\lambda^{T-1}) + \frac{\lambda^{T-1}p(1-p)}{1+\lambda}}{1-\lambda^T} = \frac{1}{3}.$$

Note finally that, as the M_i, m_i depend on M_1, m_1 in the same way as in the one-step random walk, we have the same constraints. Thus it must hold that

$$0 \leq m_1 \leq \frac{1}{3} \leq M_1 \leq \frac{-p+1+\lambda}{2+2\lambda-3p}.$$

□

Remark. It was clear from the start that the non-trivial intertwining solutions for the one-step distribution would also be solutions for the T-step distribution as

$$\begin{aligned}\Lambda P^T &= (\Lambda P)P^{T-1} \\ &= (\overline{P}\Lambda)P^{T-1} \\ &= \ldots \\ &= \overline{P}^T \Lambda.\end{aligned}$$

Thus the same Λ can be used in combination with $\overline{P}^T$, which is not the same as $\overline{P}$.

4.3 Characterization of intertwining solutions for even T-step distribution

In this section we will show that **for T even, the solution space increases**. The exact solutions are given in Lemma 4.2.

Lemma 4.2 (Characterization of intertwining solutions for even T-step distribution)**.** If T is even, then there are non-trivial solutions $(\Lambda, \overline{P}_T)$ of the intertwining $\Lambda P^T = \overline{P}_T \Lambda$ of the form

$$\Lambda = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

and

$$\overline{P}_T = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix}$$

iff

- it is satisfied that
$$1 \geq M_1 \geq \frac{1}{3} \geq m_1 \geq 0 \tag{4.49}$$
or
$$1 \geq m_1 \geq \frac{1}{3} \geq M_1 \geq 0; \tag{4.50}$$
- for
$$0 \leq M_2 \leq 1 - M_1 \tag{4.51}$$
$$0 \leq m_2 \leq 1 - m_1 \tag{4.52}$$
it holds that
$$\frac{1}{3}(M_1 - m_1) + \left(m_1 - \frac{1}{3}\right)M_2 + \left(\frac{1}{3} - M_1\right)m_2 = 0. \tag{4.53}$$

- and finally

$$M_3 = 1 - M_1 - M_2 \tag{4.54}$$
$$m_3 = 1 - m_1 - m_2. \tag{4.55}$$

If it holds that $M_1 \neq m_1$, then the transition matrix $\overline{P}$ on the coarse-grained network is given by

$$\overline{P} = \begin{pmatrix} \overline{p} & 1-\overline{p} \\ \overline{q} & 1-\overline{q} \end{pmatrix}, \tag{4.56}$$

where

$$\overline{p} = \frac{\frac{1}{3} - m_1 + \lambda^T \left(M_1 - \frac{1}{3} \right)}{M_1 - m_1}, \tag{4.57}$$
$$\overline{q} = (1 - \lambda^T) \frac{\left(\frac{1}{3} - m_1 \right)}{M_1 - m_1}. \tag{4.58}$$

Proof. If T is even, then we find:

$$\overline{p}(M_1 - m_1) + m_1 = \frac{1}{3} + \frac{\lambda^T}{3}(2M_1 - M_2 - M_3) \tag{4.59}$$
$$\overline{p}(M_2 - m_2) + m_2 = \frac{1}{3} + \frac{\lambda^T}{3}(-M_1 + 2M_2 - M_3) \tag{4.60}$$
$$\overline{p}(M_3 - m_3) + m_3 = \frac{1}{3} + \frac{\lambda^T}{3}(-M_1 - M_2 + 2M_3) \tag{4.61}$$
$$\overline{q}(M_1 - m_1) + m_1 = \frac{1}{3} + \frac{\lambda^T}{3}(2m_1 - m_2 - m_3) \tag{4.62}$$
$$\overline{q}(M_2 - m_2) + m_2 = \frac{1}{3} + \frac{\lambda^T}{3}(-m_1 + 2m_2 - m_3) \tag{4.63}$$
$$\overline{q}(M_3 - m_3) + m_3 = \frac{1}{3} + \frac{\lambda^T}{3}(-m_1 - m_2 + 2m_3). \tag{4.64}$$

We shall find the solution to these constraints in the following steps.

1. **First we derive a new set of equations.**
   We use the equations (4.59) and (4.62) to eliminate the term $m_1 - \frac{1}{3}$ to obtain

$$\left(\overline{p} - \overline{q} - \frac{2\lambda^T}{3} \right)(m_1 - M_1) - \frac{\lambda^T}{3}(M_2 - m_2) - \frac{\lambda^T}{3}(M_3 - m_3) = 0. \tag{4.65}$$

   Doing similarly for the pairs (4.60)/(4.63), (4.61)/(4.64) we obtain

$$\left(\overline{p} - \overline{q} - \frac{2\lambda^T}{3} \right)(m_2 - M_2) - \frac{\lambda^T}{3}(M_1 - m_1) - \frac{\lambda^T}{3}(M_3 - m_3) = 0, \tag{4.66}$$
$$\left(\overline{p} - \overline{q} - \frac{2\lambda^T}{3} \right)(m_3 - M_3) - \frac{\lambda^T}{3}(M_1 - m_1) - \frac{\lambda^T}{3}(M_2 - m_2) = 0. \tag{4.67}$$

   To ease notation, we shall write

$$\tilde{\alpha} := \overline{p} - \overline{q} - \frac{2\lambda^T}{3}$$
$$\beta := \frac{\lambda^T}{3}$$
$$x := M_1 - m_1$$
$$y := M_2 - m_2$$
$$z := M_3 - m_3.$$

Our system of equations (4.65)-(4.67) can be rewritten to be

$$-\tilde{\alpha}x - \beta y - \beta z = 0 \tag{4.68}$$
$$-\tilde{\alpha}y - \beta x - \beta z = 0 \tag{4.69}$$
$$-\tilde{\alpha}z - \beta x - \beta y = 0. \tag{4.70}$$

2. **If we suppose now that $\tilde{\alpha} = 0$, i.e. $\alpha = \frac{2\lambda^T}{3}$, we only find the trivial solutions.**
We see from (4.68) to (4.70) that

$$M_1 - m_1 = M_2 - m_2 = M_3 - m_3 = 0.$$

Plugging in $1 = m_1 + m_2 + m_3$ into (4.62) to (4.64), we find

$$M_1 = m_1 = M_2 = m_2 = M_3 = m_3 = \frac{1}{3}.$$

3. **If we suppose that $\tilde{\alpha} \neq 0$, then we shall find that either $\tilde{\alpha} = \beta$ or $x = y$.**
From (4.68) we find

$$-\beta z = \tilde{\alpha}x - \beta y,$$

plugging this into (4.69) we find

$$(\tilde{\alpha} - \beta)x = (\tilde{\alpha} - \beta)y. \tag{4.71}$$

This is satisfied if $\tilde{\alpha} = \beta$ or if $x = y$.

4. **If $\tilde{\alpha} = \beta$, $M_1 \neq m_1$ and $M_2 \neq m_2$, then**

$$\frac{1}{3}(M_1 - m_1) + \left(m_1 - \frac{1}{3}\right) M_2 + \left(\frac{1}{3} - M_1\right) m_2 = 0.$$

Suppose that $\tilde{\alpha} = \beta$, i.e.

$$\alpha = \lambda^T. \tag{4.72}$$

Substituting $\overline{p} = \overline{q} + \lambda^T$ in (4.59) we obtain

$$\overline{q}(M_1 - m_1) + \lambda^T(M_1 - m_1)m_1 = \frac{1}{3} + \frac{\lambda^T}{3}(3M_1 - 1). \tag{4.73}$$

Rewriting this and doing the same for (4.60) and (4.61) we find

$$\overline{q}(M_1 - m_1) = \left(1 - \lambda^T\right)\left(\frac{1}{3} - m_1\right) \tag{4.74}$$
$$\overline{q}(M_2 - m_2) = \left(1 - \lambda^T\right)\left(\frac{1}{3} - m_2\right) \tag{4.75}$$
$$\overline{q}(M_3 - m_3) = \left(1 - \lambda^T\right)\left(\frac{1}{3} - m_3\right). \tag{4.76}$$

Note that (4.62), (4.63) and (4.64) return the same equations. Further using $M_3 = 1 - M_1 - M_2$ and $m_3 = 1 - m_1 - m_2$ we find that (4.76) is equivalent to (4.74) and (4.75).

- Now suppose wlog that $M_1 \neq m_1$. Then we find

$$\overline{q} = \left(1 - \lambda^T\right)\frac{\left(\frac{1}{3} - m_1\right)}{M_1 - m_1}. \tag{4.77}$$

If also $M_2 \neq m_2$, then we must have

$$\overline{q} = \left(1 - \lambda^T\right)\frac{\left(\frac{1}{3} - m_2\right)}{M_2 - m_2} \tag{4.78}$$

which must equal the quantity in (4.77). Thus we must have $(M_1 - m_1)\left(\frac{1}{3} - m_2\right) = (M_2 - m_2)\left(\frac{1}{3} - m_1\right)$ which gives the condition

$$\frac{1}{3}(M_1 - m_1) + \left(m_1 - \frac{1}{3}\right) M_2 + \left(\frac{1}{3} - M_1\right) m_2 = 0. \tag{4.79}$$

- Now **imposing the condition $\overline{q}, \overline{p} \in [0,1]$** is equivalent to

$$0 \leq \overline{q} \leq 1 - \lambda^T,$$

as we use the pact that $\overline{p} = \overline{q} + \lambda^T$.
To obtain $\overline{q} \geq 0$, we consider (4.77) and see that either

$$m_1 \leq \frac{1}{3} \quad \wedge \quad m_1 \leq M_1$$

or

$$m_1 \geq \frac{1}{3} \quad \wedge \quad m_1 \geq M_1.$$

To obtain $\overline{q} \leq 1 - \lambda^T$, we must have $\frac{\frac{1}{3} - m_1}{M_1 - m_1} \leq 1$ which is the case when

$$m_1 \leq \frac{1}{3} \quad \wedge \quad m_1 \leq M_1$$

or

$$m_1 \geq \frac{1}{3} \quad \wedge \quad m_1 \geq M_1.$$

Combining these two restrictions, we find that M_1 and m_1 must satisfy either

$$M_1 \leq \frac{1}{3} \leq m_1 \tag{4.80}$$

or

$$m_1 \leq \frac{1}{3} \leq M_1. \tag{4.81}$$

5. **If $x = y$, we find a subset of the solution space of Step 4, when $\alpha = \lambda^T$.**
We see that (4.68) and (4.69) both give

$$-(\tilde{\alpha} + \beta)x - \beta z = 0.$$

- If $\tilde{\alpha} = \beta$, then $z = 0$ and (4.70) gives $x = y = 0$, which altogether return the invariant measure.
- Otherwise

$$x = y = \frac{-\beta}{\tilde{\alpha} + \beta} z$$

which using (4.70) returns

$$\left(\frac{2\beta^2}{\tilde{\alpha} + \beta} - \tilde{\alpha} \right) z = 0. \tag{4.82}$$

Again $z = 0$ implies $x = y = 0$, which return the invariant measure.
So we consider when $\frac{2\beta^2}{\tilde{\alpha}+\beta} - \tilde{\alpha} = 0$.
Then we must have that

$$2\beta^2 - \tilde{\alpha}^2 - \tilde{\alpha}\beta = 0$$

which is the same as

$$-\left(\alpha - \frac{2\lambda^T}{3}\right)^2 - \frac{\lambda^T}{3}\left(\alpha - \frac{2\lambda^T}{3}\right) + 2\left(\frac{\lambda^T}{3}\right)^2 = 0.$$

This has solutions

$$\alpha = 0 \quad \vee \quad \alpha = \lambda^T.$$

The case when $\alpha = \lambda^T$ was solved earlier, so we will now consider $\alpha = 0$.
Then combining (4.59) and (4.62) while using $\overline{p} = \overline{q}$ and $1 = M_1 + M_2 + M_3$ and $1 = m_1 + m_2 + m_3$, we find $M_1 = m_1$. Similarly $M_2 = m_2$, $M_3 = m_3$. Earlier we saw that $M_1 - m_1 = M_2 - m_2 = M_3 - m_3$ implies the invariant measure.

□

Remark. Again we want the one-step solutions to be a subset of the space of these solutions. Therefore we verify whether the expression (3.18) for M_2 satisfies the condition (4.48).
When $\overline{p} - \overline{q} = +\lambda$, we have for $M_2 = \frac{(1+\lambda-3p)M_1+p}{1+\lambda}$ and $m_2 = \frac{(1+\lambda-3p)m_1+p}{1+\lambda}$ that

$$\begin{aligned}
&\frac{1}{3}(M_1 - m_1) + (m_1 - \frac{1}{3})\frac{(1+\lambda-3p)M_1 + p}{1+\lambda} + (\frac{1}{3} - M_1)\frac{(1+\lambda-3p)m_1 + p}{1+\lambda} \\
&= \frac{\frac{1}{3}(1+\lambda)(M_1 - m_1) + (m_1 - \frac{1}{3})(1+\lambda-3p)M_1 + (\frac{1}{3} - M_1)(1+\lambda-3p)m_1 + p(m_1 - M_1)}{1+\lambda} \\
&= \frac{\frac{1}{3}(1+\lambda-3p)(M_1 - m_1) + (1+\lambda-3p)[m_1M_1 - \frac{1}{3}M_1 + \frac{1}{3}m_1 - m_1M_1]}{1+\lambda} \\
&= 0.
\end{aligned}$$

The same holds true when $\overline{p} - \overline{q} = -\lambda$. Thus indeed the one-step solutions are also solutions for the even T-step walk.

4.4 Evolution of local equilibria for the T-step simple random walk on the 3-node model

In this section we again state the theorem that follows from combining Theorem 2.1 with the non-trivial solutions of intertwining. First we do this for T odd, as there will not be many changes in comparison with $T = 1$, and then for T even.

Proposition 4.2 (Evolution of local equilibria of odd T-step distribution on the 3-node model)**.** Let X_T be the process of a T-step simple random walk on the 3-node model associated with one-step transition matrix

$$P = \begin{pmatrix} p & 1-p & 0 \\ 1-p & 0 & p \\ 0 & p & 1-p \end{pmatrix}$$

where T is odd.
Let

$$\Lambda = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix},$$

where

- 1. If

$$0 \le m_1 \le \frac{1}{3} \le M_1 \le \frac{-p+1+\lambda}{2+2\lambda-3p} \tag{4.83}$$

or

$$0 \le M_1 \le \frac{1}{3} \le m_1 \le \frac{-p+1+\lambda}{2+2\lambda-3p}, \tag{4.84}$$

then $\overline{p} - \overline{q} = +\lambda^T$.

2. If

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}\right\} \le m_1 \le \frac{1}{3} \tag{4.85}$$

$$\frac{1}{3}(1+\lambda) - \lambda m_1 \le M_1 \le \min\left\{\frac{-p}{1-\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\} \tag{4.86}$$

or

$$\frac{1}{3} \le m_1 \le \frac{-p}{1-\lambda-p} \tag{4.87}$$

$$\max\left\{0, \frac{1-\lambda-p}{2-2\lambda-3p}, \frac{1}{3}\left(1+\frac{1}{\lambda}\right) - \frac{m_1}{\lambda}\right\} \le M_1 \le \frac{1}{3}(1+\lambda) - \lambda m_1, \tag{4.88}$$

then $\overline{p} - \overline{q} = -\lambda^T$.

-
$$M_2 = \frac{(1+\alpha-3p)M_1 + p}{1+\alpha} \tag{4.89}$$

  and
$$m_2 = \frac{(1+\alpha-3p)m_1 + p}{1+\alpha}; \tag{4.90}$$

-
$$M_3 = 1 - M_1 - M_2 \tag{4.91}$$

  and
$$m_3 = 1 - m_1 - m_2; \tag{4.92}$$

and finally given $M_1 \neq m_1$,

-
$$\overline{q} = (M_1 - m_1)^{-1}\left[\frac{1}{3}(1-\lambda^{T-1}) + \frac{\lambda^{T-1}p(1-p)}{1+\alpha} + (\alpha^T - 1)m_1\right]; \tag{4.93}$$

-
$$\overline{p} = \overline{q} + \alpha^T; \tag{4.94}$$

where $\alpha = +\lambda$ if $\overline{p} - \overline{q} = +\lambda^T$ and $\alpha = -\lambda$ if $\overline{p} - \overline{q} = -\lambda^T$.
Then there exist two stopping times

$$T_{\overline{1}}^{\text{odd}} \sim \text{Geom}\,(1-\overline{p}) \tag{4.95}$$
$$T_{\overline{2}}^{\text{odd}} \sim \text{Geom}\,(\overline{q}) \tag{4.96}$$

for the process X_T such that

1. $\nu_{\overline{x}}$ is stationary until time $T_{\overline{x}}^{\text{odd}}$, which means that
$$P_{\nu_{\overline{x}}}\left(X_T(t) = \cdot | T_{\overline{x}}^{\text{odd}} > t\right) = \nu_{\overline{x}}(\cdot); \tag{4.97}$$
2. for all $\overline{x} \neq \overline{y}$ we have
$$P_{\nu_{\overline{x}}}(X_T(T_{\overline{x}}^{\text{odd}}) = \cdot) = \nu_{\overline{y}}(\cdot); \tag{4.98}$$
3. $X_T(T_{\overline{x}}^{\text{odd}})$ and $T_{\overline{x}}^{\text{odd}}$ are independent for all $\overline{x}$.

So we see indeed that the only change compared with the one-step distribution is in the parameters by which we define the geometric random times $T_{\overline{1}}^{\text{odd}}$ and $T_{\overline{2}}^{\text{odd}}$.

Proposition 4.3 (Evolution of local equilibria of even T-step distribution on the 3-node model)**.** Let X_T be the process of a T-step simple random walk on the 3-node model associated with one-step transition matrix

$$P = \begin{pmatrix} p & 1-p & 0 \\ 1-p & 0 & p \\ 0 & p & 1-p \end{pmatrix}$$

where T is even.
Let

$$\Lambda = \begin{pmatrix} M_1 & M_2 & M_3 \\ m_1 & m_2 & m_3 \end{pmatrix},$$

such that

- it is satisfied that
$$1 \geq M_1 \geq \frac{1}{3} \geq m_1 \geq 0 \tag{4.99}$$

or

$$1 \geq m_1 \geq \frac{1}{3} \geq M_1 \geq 0; \tag{4.100}$$

- for

$$0 \leq M_2 \leq 1 - M_1 \tag{4.101}$$
$$0 \leq m_2 \leq 1 - m_1 \tag{4.102}$$

it holds that

$$\frac{1}{3}(M_1 - m_1) + \left(m_1 - \frac{1}{3}\right) M_2 + \left(\frac{1}{3} - M_1\right) m_2 = 0. \tag{4.103}$$

- and finally

$$M_3 = 1 - M_1 - M_2 \tag{4.104}$$
$$m_3 = 1 - m_1 - m_2. \tag{4.105}$$

Then there exist two stopping times

$$T_{\overline{1}}^{\text{even}} \sim \text{Geom}\left((1-\lambda^T)\frac{\left(M_1 - \frac{1}{3}\right)}{M_1 - m_1}\right) \tag{4.106}$$

$$T_{\overline{2}}^{\text{even}} \sim \text{Geom}\left((1-\lambda^T)\frac{\left(\frac{1}{3} - m_1\right)}{M_1 - m_1}\right) \tag{4.107}$$

for the process X_T such that

1. $\nu_{\overline{x}}$ is stationary until time $T_{\overline{x}}^{\text{even}}$, which means that

$$P_{\nu_{\overline{x}}}\left(X_T(t) = \cdot | T_{\overline{x}}^{\text{even}} > t\right) = \nu_{\overline{x}}(\cdot); \tag{4.108}$$

2. for all $\overline{x} \neq \overline{y}$ we have

$$P_{\nu_{\overline{x}}}(X_T(T_{\overline{x}}^{\text{even}}) = \cdot) = \nu_{\overline{y}}(\cdot); \tag{4.109}$$

3. $X_T(T_{\overline{x}}^{\text{even}})$ and $T_{\overline{x}}^{\text{even}}$ are independent for all $\overline{x}$.

Here we see that the parameters by which we define the random times are the same as for when T is odd but for the generality of λ, however the set of measures $\nu_{\overline{x}}$ has increased and we consider all $p \in (0,1)$ allowing us to have many more solutions in our possession.

Example 4.1. In particular we may intertwine with a Dirac measure, where

$$M_1 = 1, \quad M_2 = M_3 = 0.$$

If we take $m_1 = 1$, we see that (4.103) forces

$$m_2 = m_3 = \frac{1}{2}.$$

Equations (4.106) and (4.107) give us that

$$\mathbb{E}\left[T_{\overline{1}}^{\text{even}}\right] = \left(\frac{2}{3} - \frac{2}{3}\lambda^T\right)^{-1}$$
$$\mathbb{E}\left[T_{\overline{2}}^{\text{even}}\right] = \left(\frac{1}{3} - \frac{1}{3}\lambda^T\right)^{-1}.$$

4.5 Testing the universal spectral solutions for the T-step distribution

We shall now demonstrate that the **proposed spectral solutions of Proposition 2.1 now again return the non-trivial solutions with extra conditions**, whereby we refer to the computations in Sections 4.2 and 4.3.

Lemma 4.3. The proposed spectral solutions of Proposition 2.1 in the case of both odd and even T are those of Lemma 4.1 with the extra condition

$$M_1 + m_1 = \frac{2}{3}. \tag{4.110}$$

Proof. Note that the eigenvalues of P^T are just the eigenvalues of P to the power T.

- When T is odd we find the spectrum to be

$$\{1, \lambda^T, -\lambda^T\}.$$

The proposed $\overline{P}$ of Equation (2.8) becomes

$$\overline{P} = \frac{1}{2}\begin{pmatrix} 1+\lambda^T & 1-\lambda^T \\ 1-\lambda^T & 1+\lambda^T \end{pmatrix}. \tag{4.111}$$

Defining $\overline{p}, \overline{q}$ as usual, we see that for $\alpha = \overline{p} - \overline{q}$ we obtain

$$\alpha = \lambda^T.$$

Solutions of intertwining for general $\overline{P}$ with $\alpha = \lambda^T$ have been found in the proof of Lemma 4.1, starting from Equation (4.39). Now we will investigate which solutions correspond to

$$\overline{q} = \frac{1-\lambda^T}{2}.$$

We see that

$$\frac{\left[\frac{1}{3}(1-\lambda^{T-1}) + \frac{\lambda^{T-1}p(1-p)}{1+\lambda} + (\lambda^T - 1)m_1\right]}{M_1 - m_1} = \frac{1-\lambda^T}{2}$$

which we rewrite to be

$$\frac{2}{3}(1-\lambda^{T-1}) + \frac{2\lambda^{T-1}p(1-p)}{1+\lambda} + 2(\lambda^T - 1)m_1 = (1-\lambda^T)(M_1 - m_1)$$

returning

$$\begin{aligned} M_1 + m_1 &= \frac{\frac{2}{3}(1-\lambda^{T-1}) + \frac{2\lambda^{T-1}p(1-p)}{1+\lambda}}{1-\lambda^T} \\ &= \frac{\frac{2}{3}(1-\lambda^{T-1}) - \frac{2\lambda^{T-1}(\lambda^2-1)}{3(1+\lambda)}}{1-\lambda^T} \\ &= \frac{2}{3} \cdot \frac{1-\lambda^{T-1} - \lambda^{T-1}(\lambda - 1)}{1-\lambda^T} \\ &= \frac{2}{3}. \end{aligned}$$

- For even T, we only have eigenvalues

$$\theta_0 = 1, \quad \theta_1 = \lambda^T.$$

The proposed $\overline{P}$ of Equation (2.8) becomes

$$\overline{P} = \frac{1}{2}\begin{pmatrix} 1+\lambda^T & 1-\lambda^T \\ 1-\lambda^T & 1+\lambda^T \end{pmatrix} \tag{4.112}$$

and we see that

$$\alpha = \lambda^T.$$

Again the solutions of intertwining for general $\overline{P}$ with $\alpha = \lambda^T$ have been found in the proof of Lemma 4.2, starting from Equation (4.67). These include the non-trivial solutions given in the statement of Lemma 4.2. Now we will investigate which solutions correspond to

$$\overline{q} = \frac{1-\lambda^T}{2}.$$

We see that

$$(1-\lambda^T)\frac{\frac{1}{3}-m_1}{M_1-m_1} = \frac{1-\lambda^T}{2}$$

which we rewrite to be

$$\frac{2}{3} - 2m_1 = M_1 - m_1.$$

This gives us the condition

$$M_1 + m_1 = \frac{2}{3}.$$

□

Chapter 5

Conclusion

In this thesis the following original work has been done:

- We illustrated how to use the coarse-grained framework introduced in [1] to describe the evolution of local equilibria in all details for an explicit toy example.
- We found meaningful solution to the intertwining equations in this model, i.e. measures that are really localized in different regions of the state space. This shows us how promising the machinery is.
- We explored how "good" the universal spectral solutions of Proposition 2.1 were (for the one-step random walk in Section 3.3 and for the T-step random walk in Section 4.5) and observed that they are not optimal.
- We explored what sort of bounds on the mixing times of the original process can be obtained by looking at the mixing times of the coarse-grained process.
  In particular we found that the bound on the mixing times of the original process is optimal if the two measures that are used to intertwine have as much difference as possible in placing their mass on the first state of the state space.
- The cardinality of the space of intertwining solutions increases as we increase the number of steps T of the random walk that is performed on the network.
  In particular for even $T \geq 2$, we can intertwine with Dirac measures, i.e. measures that place all of their mass on one state. Thus evolution of "very localized states" can be described by intertwining the T-step process.

We may conclude due to Proposition 3.2 and the analysis performed in Section 3.5 that the framework in [1] indeed leads to results that one would intuitively expect. Our further suggestion would be to continue exploring the framework on larger and more challenging models.

Acknowledgements

First and most of all I would like to thank my supervisors, prof. dr. Frank den Hollander and dr. Luca Avena. A lot of gratitude goes out to Luca who introduced me to this leading edge research done by him and his colleagues, and for helping me develop my skills of doing research independently. I enjoyed our discussions and meetings a lot, be they in Leiden, the library or in the train.
I would like to thank prof. Den Hollander as well for all his support and consultations. Whenever it was needed, I could go to him and he was ready to help.

On a personal level I would like to thank my fellow students. Discussing and comparing our research has always been a joy and I look forward to continue doing this in the future.

And last but most certainly not least, I would like to thank my parents. They have supported me all my life and in particular during my studies.

Bibliography

[1] L. Avena, F. Castell, A. Gaudillière, C. Mélot, *Approximate and exact solutions of intertwining equations through random spanning forests*, (2017), arXiv:1702.05992v1.

[2] A. Bovier and F. den Hollander, *Metastability - A Potential-Theoretic Approach*, First Edition (2015), Grundlehren der mathematischen Wissenschaften 351, Springer International Publishing.

[3] L. Avena, F. Castell, A. Gaudillière, C. Mélot, *Random Forests and Network Analysis*, (2018), J. Stat. Phys. 173, 985-1027

[4] J. van 't Hoff, *Études de Dynamiques Chimiques*, (1884), F. Muller, Amsterdam

[5] S. Arrhenius, *On the reaction rate of the inversion of non-refined sugar upon souring*, (1889), Z. Phys. Chem. 4, 226-248

[6] H.A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, (1940), Physica 7, 284-304

[7] S. Glasstone, K. Laidler, H. Eyring, *The Theory of Rate Processes*, (1941), McGraw-Hill, New York

[8] M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, (1984), Grundlehren der mathematischen Wissenschaften 260, Springer, New York

[9] O. Penrose, J. Lebowitz, *Rigorous treatment of metastable states in the van der Waals-Maxwell theory*, (1971), J. Stat. Phys. 3, 211-236

[10] E. Davies, *Dynamical stability of metastable states*, (1982), J. Funct. Anal. 46, 373-386

[11] E. Davies, *Spectral properties of metastable Markov semigroups*, (1983), J. Funct. Anal. 52, 315-329

[12] E. Davies, *Metastability and the Ising model*, (1982), J. Stat. Phys. 27, 657-675

[13] E. Davies, *Metastable states of symmetric Markov semigroups. I*, (1982), Proc. Lond. Math. Soc. 45, 133-150

[14] E. Davies, *Metastable states of symmetric Markov semigroups II*, (1982), J. Lond. Math. Soc. 26, 541-556

[15] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, *Metastability in stochastic dynamics of disordered mean-field models*, (2001), Probab. Theory Relat. Fields 119, 99-161

[16] L. Rogers, J. Pitman, *Markov functions*, (1981), Ann. Probab. 9, no. 4, 573-582

[17] P. Diaconis, J. Fill, *Strong stationary times via a new form of duality*, (1990), Ann. Probab. 18, no. 4, 1483-1522

[18] C. Landim, *Metastable Markov chains*, (2018), arXiv:1807.04144v1.

[19] E. Weinan, E. Vanden-Eijnden, *Towards a theory of transition paths*, (2006), J. Stat. Phys. 123, 503-523

[20] F. Manzo, E. Olivieri, *Dynamical Blume-Capel model: competing metastable states at inïňĄnite volume*, (2001), J. Stat. Phys. 104, 1029-1090 (2001).

[21] M. Cameron, E. Vanden-Eijnden, *Flows in Complex Networks: Theory, Algorithms, and Application to Lennard-Jones Cluster Rearrangement*, (2014), J. Stat. Phys. 156, 427-454

[22] J. Lu, E. Vanden-Eijnden, *Exact dynamical coarse-graining without timescale separation*, (2014), J. Chem. Phys. 141, 044109

[23] L. Avena, A. Gaudillière, *Two applications of random spanning forests*, (2018), J. Theor. Probab. 31, 1975-2004

[24] L. Avena, F. Castell, A. Gaudillière, C. Mélot, *Intertwining wavelets or Multiresolution analysis on graphs through random forests*, (2017), arXiv:1707.04616.