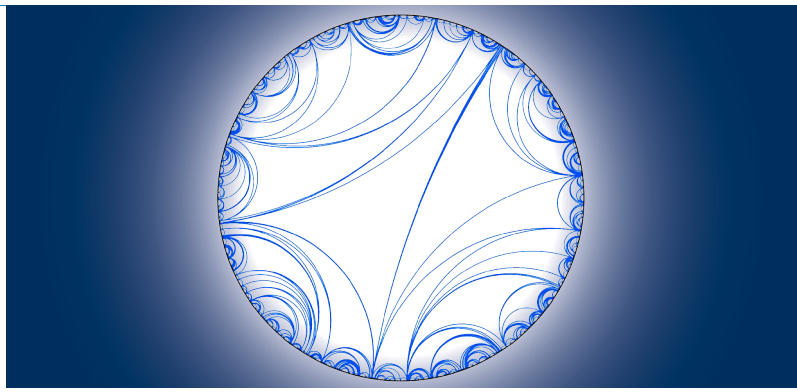
Universität Potsdam



Peter Keller

Mathematical Modeling of Molecular Motors

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 1

Preprints des Instituts für Mathematik der Universität Potsdam

Preprints des Instituts für Mathematik der Universität Potsdam
2 (2013) 1

Peter Keller

Mathematical Modeling of Molecular Motors

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.de abrufbar.

Universitätsverlag Potsdam 2013
http://verlag.ub.uni-potsdam.de/

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Preprints des Instituts für Mathematik der Universität Potsdam** wird herausgegeben vom Institut für Mathematik der Universität Potsdam.

ISSN (online) 2193-6943

Kontakt:
Institut für Mathematik
Am Neuen Palais 10
14469 Potsdam
Tel.: +49 (0)331 977 1028
WWW: http://www.math.uni-potsdam.de

Titelabbildungen:
1. Karla Fritze | Institutsgebäude auf dem Campus Neues Palais
2. Nicolas Curien, Wendelin Werner | Random hyperbolic triangulation
Published at: http://arxiv.org/abs/1105.5089

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2013/6304/
URN urn:nbn:de:kobv:517-opus-63045
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-63045

Institut für Mathematik der Universität Potsdam
Lehrstuhl für Wahrscheinlichkeitstheorie

Max Planck Institut für Kolloide und Grenzflächen
Stochastic Processes in Complex and Biological Systems

Mathematical Modeling of Molecular Motors

Dissertation zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
In der Wissenschaftsdisziplin Wahrscheinlichkeitstheorie.

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Dipl. Math. Peter Keller

Datum der Verteidigung: 22.10.2012

Prüfungskommission/Members of the Jury

Prof. Dr. Gilles Blanchard	*Vorsitz/Chairman*
apl. Prof. Dr. Christine Böckmann	
Prof. Dr. Wilhelm Huisinga	
apl. Prof. Dr. Hannelore Liero	*stellv. Vorsitz/Deputy chairman*
Prof. Dr. Nikolaos Limnios	*Gutachter/Reviewer*
Prof. Dr. Reinhard Lipowsky	
Dr. Pierre-Yves Louis	
Prof. Dr. Sylvie Rœlly (Betreuer/Advisor)	*Gutachter/Reviewer*
Prof. Dr. Philippe Robert	*Gutachter/Reviewer*
Dr. Angelo Valleriani	

Abstract

Amongst the many complex processes taking place in living cells, transport of cargoes across the cytosceleton is fundamental to cell viability and activity. To move cargoes between the different cell parts, cells employ *Molecular Motors*. The motors are responsible fo a huge variety of tasks, ranging from cell division to DNA replication and chemical transport. A subclass of such motors, whose models is the subject-matter of this thesis, operate by transporting cargoes along the so called cellular *micro-tubules*, namely rope-like structures that connect, for instance, the cell-nucleus and outer membrane.

Empirical evidence indicates that the movement exhibited by these motors is stochastic and occurs with constant "step size". This makes Markov Chains with countable state space excellent models for describing the movement of such molecular motors.

The first part of this thesis deals with the extension of reversibility by introducing the concept of *permuted balance*. This new notion allows the reduction of the state space of a Markov Chain by aggregation of states. The result is again a Markov Chain and moreover reversible, i.e. the distribution of the process is invariant under time reversal.

Permuted balance is, besides geometric properties of the transition graph, sufficient for a phenomenon called *Time Duality*, which we introduce in the second part. This property accounts for the equality of certain passage times between states. This equality was observed for several motors empirically and serves therefore as an important tool in understanding the models of molecular motors, using different modeling strategies.

The approach used here is based on the well known phase type distributions, which can be characterized as absorption times of Markov Chains, and an original application of the Doob-h-transform. We use this transform here to characterize processes conditioned for absorption in a designated state, selected out of a set of absorbing states.

In the third part we present a new class of Markov Chains to model molecular motors, which we will refer to as Quasi-Random Walks. This new model is an extension of Quasi-Birth-and-Death Processes by eliminating the boundary condition in zero. We account for the possibility that the motor may detach from the tubule by introducing an additional absorbing state.

We adapt established matrix-geometric tools, normally associated with Quasi-Birth-Death processes, to investigate properties of Quasi-Random Walks processes. Amongst these, we include the total life-time, the total number of specific transitions as well as the maximal distance from zero and the distribution of the last position before detachment.

A following part is devoted to the application of our theoretical results about Quasi-Random Walk for an existing model of the molecular motor *kinesin*.

Zusammenfassung

Neben vielen sehr komplexen Prozessen, ist der Transport von Cargos durch das Zytoskelett fundamental für Aktivität und Lebensfähigkeit einer Zelle. Um Cargos zwischen den verschiedenen Zellteilen zu bewegen, bedienen sich Zellen sogenannter *molekularer Motoren*. Diese Motoren sind verantwortlich für eine große Anzahl unterschiedlichster Aufgaben in der Zelle, angefangen von Zellteilung, Kopieren von DNS bis hin zum gerichteten Transport. Eine Unterklasse dieser Motoren, deren Modelle Gegenstand dieser Arbeit ist, laufen auf Mikrotubuli genannten seilartigen Strukturen, die z. B. zwischen Zellkern und der äußeren Zellmembran gespannt sind.

Da diese Art der Fortbewegung stochastischer Natur ist und experimentell gezeigt wurde, dass die Schrittweiten konstant sind, eignen sich Markovketten in stetiger Zeit mit abzählbarem Zustandsraum hervorragend als Modellierungswerkzeug.

Der erste Teil der Arbeit beschäftigt sich eingehend mit der Einführung eines erweiterten Reversibilitätsbegriffes: des *permutierten Gleichgewichts*. Diese Erweiterung erlaubt es unter anderem den Zustandsraum einer Markovkette durch Zusammenfassen von Zuständen zu verkleinern. Der resultierende Prozess ist wieder eine Markovkette und zudem reversibel, d.h. die Verteilung dieser Kette ist invariant unter Zeitumkehr.

Das permutierte Gleichgewicht ist, neben einigen geometrischen Eigenschaften des Übergangsgraphen, hinreichende Bedingung für ein im zweiten Teil eingeführtes Phänomen namens *Zeitdualität*. Diese Eigenschaft bezeichnet die Verteilungsgleichheit bestimmter Reisezeiten zwischen Zuständen. Diese Verteilungsgleichheit wurde für einige Motormodelle beobachtet. Damit stellt die Zeitdualität ein wichtiges Hilfsmittel zum Verständnis der Funktionsweise von Modellen molekularer Motoren unter Verwendung verschiedener Modellierungsstrategien dar. Die entscheidenden Hilfsmittel sind dabei zum einen die wohlbekannten Phasentypverteilungen, die als Absorptionszeiten von Markovketten dargestellt werden können; zum anderen eine originelle Anwendung der Doob-h-Transformation, um Prozesse zu beschreiben, die darauf bedingt werden in einem vorgegebenen Zustand absorbiert zu werden.

Der dritte Teil befasst sich mit der Modellierung von molekularen Motoren durch eine neu eingeführte Klasse von Markovketten, die hier Quasi-Irrfahrt genannt wird. Diese neue Klasse ist als Erweiterung der aus der Warteschlangentheorie bekannten Quasi-Geburts-und-Todes Prozesse zu verstehen; im Gegensatz zu den Quasi-Geburts-und-Todes Prozessen fehlt der Quasi-Irrfahrt die Randbedingung in Null. Ein zusätzlicher absorbierender Zustand spiegelt die Möglichkeit der Ablösung des Motors vom Mikrotubulus wider.

Die bekannten matrixgeometrischen Methoden der Quasi-Geburts-und-Todes Prozesse werden hier adaptiert, um interessante Eigenschaften dieser Klasse von Prozessen zu analysieren. Darunter fallen etwa die Gesamtlebenszeit, die Gesamtanzahl bestimmter Übergänge, sowie die maximale Distanz vom Startpunkt und die Verteilung der zuletzt besuchten Position vor Absorption.

Der anschliessende Teil befasst sich mit der Anwendung der theoretischen Ergebnisse an ein physikalisches Modell des molekularen Motors *Kinesin*.

Acknowledgment

There are a number of people I owe a lot.

First of all I thank the Max Planck Society and Prof. Lipowsky in particular for the financial support and fruitful discussions.

I further thank Prof. Nikolaos Limnios for his visit here in Potsdam in 2011. He has a great understanding of Semi-Markov Chains and I had the privilege that he shared his wisdom with me. I also thank him for the possibility to meet him in Compiègne for further discussions and for the big amount of work and time he spent as a referee for my thesis.

Prof. Robert deserves also a great thank you, as he also spent a big amount of time for my thesis.

I also want to thank Prof. Dr. Sylvie Rœlly. In the past years she was there for me whenever I had stupid questions, problems and thoughts to share. Her advising was invaluable to me and I might say I am very lucky that I had the chance to work with her. I know how much she invests in her PhD-students and more than one time I wondered how she can manage it all. Chapeau, Sylvie!

I also want to thank Dr. Angelo Valleriani. He was of great help in the past years and faced me with many interesting topics and problems that found its counterpart in this thesis. Without him this thesis would have never been possible.

A huge part of this thesis was written in a little café in Charlottenburg. I want to thank the owners Susan and Mathias Reed-Neumann for the chance letting me transform their coffee into theorems.

Furthermore I want to give a big "thank you" to Dr. Paul Zaslansky who is not only a good friend but also the man who suggested to ask for a PhD-project at the MPI. Who knows what I would have done without his advice. I also want to thank his son Ben for sharing his bilingual knowledge that helped greatly to improve my ludicrous english.

Big thanks go also to Mark Weiland. In the many years we lived together he became invariant under math and I had my own trouble explaining to him what I am actually doing all the time.

Special thanks go to Dr. Michael Högele, self-proclaimed Bavarian in Prussia, for the constant up-to-date information about gold pricing and Latin-American politics and many tips for good literature. Sometimes we even talked about mathematics.

Being in the mad situation that I had two offices, one at the MPI, one at the university, I also want to thank my other office colleague Marko Wehle for many coffee talks and a great work atmosphere.

Another person who deserves a big thank you is my father, he always was there for me and tried his very best to help and support me in any respect.

I would also like to thank my mother for her believe in me and her tireless support since so many years. I wished her illness would have let her the chance to be still here. May she rest in peace.

Contents

Introduction

Amongst the many complex processes taking place in living cells, transport of cargos across the cytosceleton is fundamental to cell viability and activity. To move cargos between the different cell parts, cells employ *Molecular Motors*. The motors are responsible of a huge variety of tasks, ranging from cell division to DNA replication and chemical transport. A subclass of such motors operate by transporting cargos along the so called cellular *microtubules*, namely rope-like structures that connect, for instance, the cell-nucleus and outer membrane.

The diversity of these motors is striking; for the human 14 families of kinesins are identified, where the kinesins are themselves only one of three types of "linear processive motors" (motors that walk on "ropes"), see for instance [Sch03]. The complexity is increased by the fact that in every species different varieties of these motors appear.

Fortunately the concept of how these motors work is quite similar for every motor, see fig. 1. In principle they all use chemical bound energy to perform directed motion.

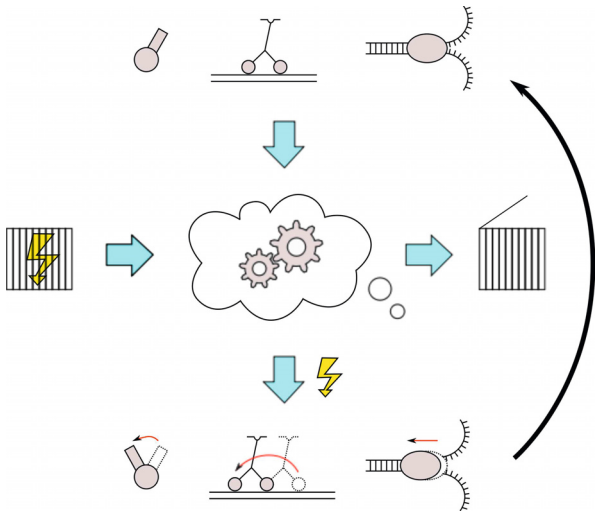


Figure 1: A cartoon of how molecular motors work from a simplified perspective. Depicted are different types of molecular motors, rotation motors, walking motors and a motor that opens piece by piece DNA strands (from left to right).

The aim of this thesis is to find a proper mathematical modeling tool kit for such processes. We find that an extension of the Quasi-Birth-and-Death process fulfills the constraints and is numerically treatable. Alongside we develope interesting new notions, like an extension of

reversibility and a property called *Time Duality*, which arise in some specific models of molecular motors.

The thesis is organized as follows:

In **Chapter 1** we first introduce basic facts about Markov Chains in continuous (MCc) and discrete time (MCd) with at most countable state space.

Then we extend the important but very restrictive notion of reversibility. A reversible measure $\pi = (\pi_i)_{i\in E}$, for a MCd with values in E, is e.g. characterized by the detailed balance equations

$$\pi_i p_{ij} = \pi_j p_{ji},\ i,j \in E$$

where p_{ij} is the transition probability between states i and j. We extend this by

$$\pi_i p_{ij} = \pi_j p_{\sigma(j)\sigma(i)},\ i,j \in E \tag{1}$$

where σ is a permutation of the states. We call (1) *permuted balance equations* and say that π is in permuted balance for the MCd with transitions $(p_{ij})_{i,j\in E}$ if (1) holds. Like for the detailed balance, the permuted balance allows to characterize the stationary distribution π uniquely; we show this in Proposition 1.3.7 together with some other properties implied by permuted balance.

A MCd with a stationary measure in permuted balance has the interesting property that it induces a decomposition of the state space into disjoint subsets on which a new reversible MCd can be defined. The concept of aggregation is known as lumping and was introduced in [KS76]. As permuted balance can hold for different σ at the same time, different aggregations are possible, see Example 1.3.

Another characterization of reversibility is a Criterion introduced by Kolmogoroff in [Kol36] for discrete time and extended later by other authors for continuous time, see e.g. [Kel79], chapter 2. This Criterion is known in the literature as *Kolmogoroff Criterion*. In Theorem 1.3.12 we give the analogue "extended Kolmogoroff Criterion" for permuted balance.

The rest of chapter 1 is devoted to local time and a short introduction of pattern matching. Local time is a notion that has strong connections to the so called "Schur complement". Indeed the Schur complement of a transition matrix or infinitesimal generator for a given partition into blocks can be interpreted as the transition matrix/infinitesimal generator of the process observed only in a subset of states. We adapt here a version for absorbing MCd's. Then again the concept of pattern matching is used in many applications ranging from biology, speech recognition and applications in operation research, see e.g. [Nue08, Fu96] and references therein and [FL03]. We combine both concepts in chapter 4 to gain information about the number of certain transitions before absorption.

In **Chapter 2** we introduce the well known *phase type distributions*, which are defined as absorption times of Markov Chains. This type of "Matrix Exponential Distributions" was introduced rigorously for the first time by Neuts in [Neu94] and applied in the framework of Queueing Theory. Early works in this direction date back to Erlang and Jensen, see [Erl09, Jen54]. The second half of chapter 2 is devoted to the extension to several absorbing states. We show here that a MCc conditioned to absorption in one absorbing state out of several can be described by a linear transform of the transition matrix/infinitesimal generator. This concept is known in older and applied theory as Duality, see [KSK66], and also as Doob-h-transform. This works in our framework as follows. If there is a vector h such that for a sub stochastic matrix P

$$hP \leq h$$

holds, h is called *excessive*. It is then used to transform the matrix P by the linear transformation

$$\tilde{P} = \mathrm{diag}(h)^{-1} P \mathrm{diag}(h),$$

where diag(h) is the diagonal matrix containing the entries of h as diagonal entries, the other entries are zero. Then $\tilde{P}$ is called (in the context of Markov Chains) the h-dual or h-transform. However, the concept of Duality goes a little further; taking $h = \pi$, the reversible measure (if existent) of a stochastic matrix, the π-dual describes the time reversal of the original process. We show here the existence of a special h, such that the associated h-dual describes the original process *conditioned* to absorption in a designated absorbing state, see Lemma 2.2.2. The concept is developed in this thesis for continuous time, but it was already introduced for discrete time as an example in [KS76], see chapter 3, section 3. A reinterpretation of conditioning to absorption in a designated state allows us to describe the process conditioned to leave the transients via specific states or sets thereof. We use this type of condition in chapter 4 to gain insight about the distribution of the last visited state(s) before absorption.

Chapter 3 introduces a new relation called *Time Duality* between states in a MCc with irreducible infinitesimal generator. In a simpler setting and without rigorous proof this appeared already in [LW07, VLL08]. The concept concerns the equality in distribution of the "passage time without return" from i to j and from j to i; see Definition 3.1.1. Thereafter we characterize this property in terms of Laplace transforms of phase type distributions and gain several sufficient conditions for Time Duality, partially based on the structure of the transition graph. One of these sufficient conditions is permuted balance, see section 3.1.2, under some constraints on the neighbourhood of the chosen states.

We found in an early version, that

$$S = D^{-1} S^{\top} D$$

is sufficient, where S denotes the matrix of transitions excluding i and j and D is some invertible matrix. By an algebraic argument there is indeed not much choice for D. It turns out that D must be the product of a diagonal matrix and a permutation matrix. This led directly to the concept of permuted balance, as we note in chapter 1, (1.6). We also answer the interesting question whether the Time Duality relation is an equivalence relation for Birth-and-Death Processes with reflecting boundaries and for reversible Markov Chains on trees in section 4.3.2.

Chapter 4 is devoted to a newly introduced class of Markov Chains, the killed Quasi Random-Walk (kQRW). We are aware of the fact that the name "Quasi Random Walks" appears for other objects for example in Monte Carlo Simulations, see e.g. [dHOS00]. A kQRW is thereby defined as an extension of the well known Quasi-Birth-and-Death process (QBD), introduced by Neuts and described for instance in [Neu94]. The main difference is that we remove the boundary at zero and add an absorbing state. Because of the possible absorption, we added the qualitative "killed". In analogy to a QBD, which is defined on $\mathbb{N} \times M$, the kQRW is defined on $\mathbb{Z} \times M \cup \Delta$ where Δ is an absorbing state and M a finite set. This process is characterized by three matrices A_0, A_1, A_2 (of dimension $|M| \times |M|$) and a $|M|$-dimensional vector Γ, see also fig. 2 (the vector Γ is not denoted there).

Because of the lack of boundary the process can be lumped (aggregation of states) to a finite Markov Chain, see Lemma 4.1.5; the procedure leaves the distribution of the absorption time invariant, see Lemma 4.1.7. The possibility to switch between the different perspectives, i.e. between lumped/unlumped, allows to give theoretical results in terms of the matrices A_i and Γ for a number of questions.

Like in the case of QBD, we partition the state space $\mathbb{Z} \times M \cup \{\Delta\}$ into disjoint subsets $l(z)$, where the set $l(z) := \{z\} \times M$ is called a *level*. The crucial tool is then the introduction of a process that describes only the transition probabilities between neighbouring levels. We call this new process *step-process*. To define it properly we first calculate the probabilities of

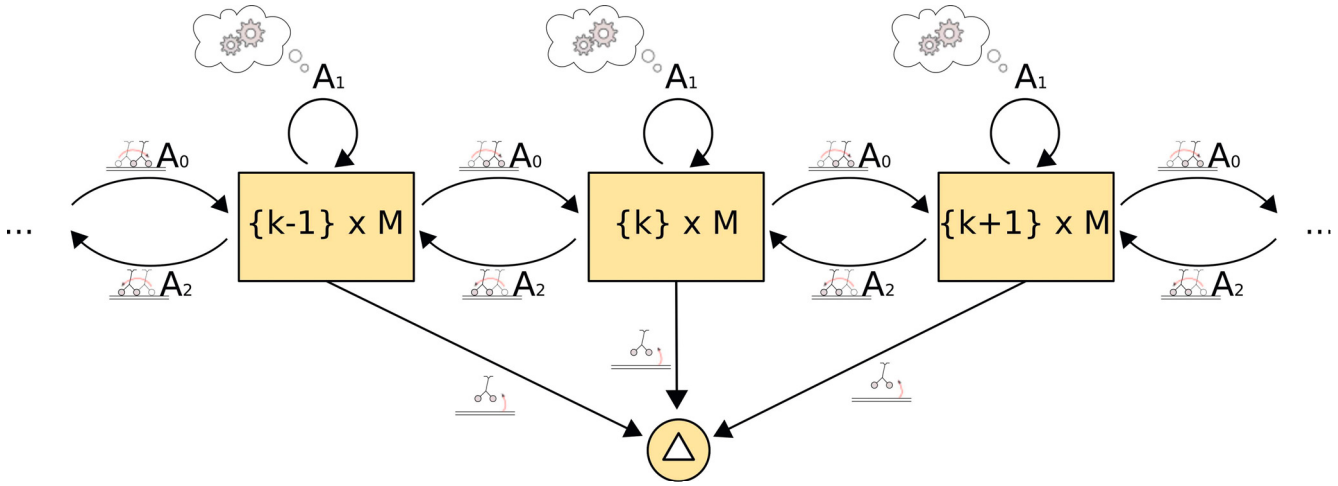


Figure 2: The killed kQRW and its interpretation for walking motors. Compare with fig. 1

moving between levels in Proposition 4.3.2 and define with these probabilities the step-process as a discrete kQRW in Definition 4.3.4. With the help of the step-process we describe in the first section a quantity named runlength, which is the total number of times the original process switches to another level before absorption (times a constant that encodes the displacement, i.e. the "step-length"); it is also the total number of transitions of the step-process before absorption and thus the runlength can be described as a discrete phase type distribution, see Proposition 4.3.5. Combining the notion of step-process, local time and pattern matching we are able to describe the total number of level changes to the left (or to the right) before absorption, see Proposition 4.3.6. In the following section we describe the maximal distance from the start in level zero and the last level visited before absorption, see Propositions 4.4.1 and 4.4.3. We also comment on a possible approximation scheme.

Although the kQRW could also be seen as a *Markov Modulated Process*, see [PTP09] for details, our approach has several advantages. In particular our results are numerically treatable, as matrices can be treated with Mathematica or MatLab, easily. Moreover the kQRW fits very well to the type of approach used for models of Molecular Motors in Biophysics, see e.g. [LL08, VLL08] and references therein.

In **Chapter 5** we apply the results of chapter 4 to an existing model of the molecular motor *kinesin*. We do not interpret the resulting numbers and distributions, but rather want to show that our theoretical results can be applied in a satisfactory manner.

In **Chapter 6** we address questions arising by extending the kQRW to the Semi-Markov regime and propose roughly possible extensions to interacting motors, i.e. coupled kQRW's.

In the appendix we give some results concerning matrix theory and the calculations necessary to characterize phase type distributions as in [Neu94]. For the case of a usual Random Walk with killing we give explicit distributions for the last level distribution and maximal distances from zero. We use here an approximation scheme which has not been used for the derivation of such results before (up to our knowledge), see [Gut09] and references. The approximation argument also yields an idea how to treat the general inverses appearing, in some cases for kQRW.

Chapter 1

Markov Chains in discrete and continuous time with countable state space

In this chapter we first introduce the well known definitions and basic results concerning Markov Chains with countable state space. In the second part we extend the concept of reversibility. The usual notion of reversibility is characterized by the so called *detailed balance equations*. The idea behind this is to compare the probability of a path with the probability of the reversed path, this is expressed in the Kolmogoroff Criterion, see [Kel79]. We relax this comparison here allowing the comparison of two paths which are not necessarily just reversed version of each other; we call this *permuted balance*, see also [KlV11]. This extension is useful in Chapter 3 and can be characterized by an extension of the Kolmogoroff Criterion we give and proof here. The most interesting property of permuted balance is surely that it induces a partition of the state space which can be used to reduce the state space via an aggregation. The resulting process is a *reversible* Markov Chain. Then in the third part we repeat the notion of local time and note that it has connections to the so called *Schur complement* known in many branches of mathematics and physics, see notes in Appendix A. By a simple argument we extend the results given in [LR99] to chains with absorbing states. Indeed local time depends only on the invertibility of certain sub matrices but not on recurrence or transience. The final section is devoted to a short introduction to pattern matching as we need it for chapter 4.

1.1 On notation

Capital letters are usually reserved for matrices or sets, Id is the identity matrix. An exception to this rule are the letters X, Y and Z which will always be associated to random variables. The letter P is reserved for transition matrices and Q will always denote infinitesimal generators (definition below). S denotes both sub stochastic matrices and non conservative infinitesimal generators which will be clear from context which setting is meant. Matrices are given like this $M = (m_{ij})_{i,j\in I}$, where m_{ij} is the (i,j)-th entry of M according to an (totally ordered) index set I. Most often the index set is finite. Matrix transposition is denoted with $M^\top$ and inversion with M^{-1}.

The complement of a set A is denoted with A^c.

A unit vector is denoted with e_i if its only non-zero component is 1 at the i-th component.

Another useful vector is

$$\mathbf{1} := (1, 1, \ldots, 1).$$

For convenience we do not indicate the dimension of $\mathbf{1}$ explicitly but it should always be thought of to have the right dimension such that the proposed multiplication is well defined. We denote vectors or matrices which contain only zeros with $\mathbf{0}$ with the same convention as for $\mathbf{1}$.

For two random variables X, Y equality in distribution is denoted either by $\mathcal{L}(X) = \mathcal{L}(Y)$ or by $X \stackrel{d}{=} Y$.

1.2 Basic facts

In the following we introduce the necessary facts about Markov Chains in discrete and continuous time as needed in this thesis. We omit proofs as there is a huge literature covering this topic, see e.g. the newer literature [Gra08, Nor97, Bré99, LPW09, Beh00] and [KSK66, KS76] for older (but still very useful) references.

Let E be a countable or finite set. We call E the *state space* and the elements of E *states*. We call probability distributions on E in the following just distributions and write them as vectors or sequences. The letters μ, ν and π are reserved for such distributions.

Further let $X := (X_i)_{i \in I}$ be a E-valued stochastic process with I an (totally ordered) index set with minimal element 0. The existence of a suitable probability space $(\Omega, \mathcal{A}, \mathbb{P})$ for X is proven for instance in [Bau02], chapter VIII.

The distribution of X_0 is called *initial distribution* and the E-valued function $X_\cdot(\omega)$ for $\omega \in \Omega$ is called *trajectory*.

Definition 1.2.1 *If we fix a filtration $(\mathcal{F}_i)_{i \in I}$ such that each $\mathcal{F}_i$ denotes the σ-algebra generated by $(X_j)_{0 \leq j \leq i}$ (the canonical filtration), we call $\tau : \Omega \to I$ a* stopping time *if and only if*

$$\{\tau \leq i\} \in \mathcal{F}_i$$

for every $i \in I$.

In the sense of the definition a stopping time is a I-valued random variable with the property that at each $i \in I$ it can be decided whether T is smaller than i or not, as the canonical filtration "carries all the information about the whole history of the process before and including i". All stopping times defined in this thesis will be denoted with τ eventually plus an index.

The stopping time we use in this thesis has several names: *first passage time*, *absorption time*, *first exit time* or *first hitting time*. Although they all appear in different contexts their definition is of the same form:

$$\tau := \inf \{i \geq 0 : X_i \in A\}$$

for $A \subsetneq E$. This stopping time denotes the first time the process *hits* a state of the set A resp. the first time the process *exits* the set A^c.

For this thesis $I = \mathbb{N}$ or $I = \mathbb{R}^+$.

1.2.1 Markov Chains in discrete time

We give here the basic facts about Markov Chains in discrete time, i.e. $I = \mathbb{N}$. We follow the explanations in [Beh00].

Definition 1.2.2 *Let $(X_n)_{n\geq 0}$ be a stochastic process on E with initial distribution μ fixed. If for arbitrary $i_0, i_1, \ldots, i_n \in E$ the* Markov Property*:*

$$\mathbb{P}(X_n = j_n | X_{n-1} = j_{n-1}, X_{n-2} = j_{n-2}, \ldots, X_1 = i_1, X_0 = i_0) \\ = \mathbb{P}(X_n = j_n | X_{n-1} = j_{n-1})$$

holds whenever $\mathbb{P}(X_{n-1} = j_{n-1}, X_{n-2} = j_{n-2}, \ldots, X_1 = i_1, X_0 = i_0) > 0$, $(X_n)_{n\geq 0}$ is called discrete Markov Chain *(MCd). $(X_n)_{n\geq 0}$ is called homogeneous if and only if additionally*

$$\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}) = \mathbb{P}(X_1 = i_n | X_0 = i_{n-1})$$

holds.

Throughout the thesis it is assumed that homogeneity is always fulfilled, we therefore omit the word "homogeneous" whenever it is possible without confusion.

A homogeneous MCd has the useful property that the dependencies between the X_i's do not depend on time. Therefore a MCd can be characterized by a simple matrix.

Definition 1.2.3 *Let $(X_n)_{n\geq 0}$ be a MCd on E. Then the $|E| \times |E|$ matrix $P = (p_{ij})_{i,j\in E}$ given by*

$$p_{ij} := \mathbb{P}(X_1 = j | X_0 = i)$$

is called transition matrix*.*

Together with an initial distribution μ a MCd is readily described by the pair (μ, P). The matrix P is a stochastic matrix.

Definition 1.2.4 *Let $M = (m_{ij})_{i,j}$ be a square matrix. M is called* stochastic *if it has non-negative entries and each row sums up to one, i.e. if*

$$\sum_j m_{ij} = 1$$

or in matrix form

$$M\mathbf{1}^\top = \mathbf{1}^\top$$

holds. M is sub stochastic *if at least one row sum is strictly less than* 1*.*

Remark **1.2.5** The matrix formulation of a stochastic matrix M shows already that 1 is always an eigenvalue to the eigenvector $\mathbf{1}^\top$, thus $Id - M$ is not invertible (because 0 is an eigenvalue of $Id - M$). The inverse of $Id - M$ is an important quantity and defined only when M is sub stochastic, see e.g. [Gan86], chapter 5. In this case the inverse can be characterized by the von Neumann series, i.e.

$$(Id - M)^{-1} = \sum_{n\geq 0} M^n.$$

□
□

Knowing (μ, P) the distribution of X_n for arbitrary n can be reconstructed, i.e.

$$\mathbb{P}(X_n = .) = \mu P^n.$$

We introduce the notion of communication:

Definition 1.2.6 *Let $(X_n)_{n\geq 0}$ be a MCd with transition matrix $P = (p_{ij})_{i,j\in E}$. Let $i, j \in E$ and $i \neq j$. Then i communicates with j if there exist $i_1, i_2, \ldots, i_n \in E$ for some $n \in \mathbb{N}$ such that*

$$p_{ii_1}p_{i_1i_2}\cdots p_{i_{n-1}i_n}p_{i_nj} > 0$$

holds. We write $i \rightsquigarrow j$ if i communicates with j and $i \leftrightsquigarrow j$ if additionally j communicates with i. We impose the convention that every state communicates with itself.

Communication does not need to be symmetric, i.e. $i \rightsquigarrow j \not\Rightarrow j \rightsquigarrow i$.

We can now define:

Definition 1.2.7 *Let $(X_n)_{n\geq 0}$ be a Markov Chain with transition matrix P. If for all choices of pairs of states $i, j \in E$ $i \leftrightsquigarrow j$ the transition matrix P is called* irreducible.

It is easy to show that $\leftrightsquigarrow$ is an equivalence relation on E and thus decomposes the state space into so called *communication classes*, i.e. into sets

$$[i] := \{j \in E : i \leftrightsquigarrow j\}.$$

Definition 1.2.8 *Let $(X_n)_{n\geq 0}$ be a MCd on E. A communication class $[i]$ is* closed *if $p_{ij} = 0$ whenever $i \in [i], j \notin [i]$.*

We note that the closed communication classes do not necessarily exploit the state space. We collect the states not in a closed communication class in the set T.

We now can define the important notions of recurrence and transience:

Definition 1.2.9 *Let $(X_n)_{n\geq 0}$ be a MCd on E with transition matrix $P = (p_{ij})_{i,j\in E}$. Then a state i is called*

- transient*, if $i \in T$.*
- absorbing*, if $|[i]| = 1$ and $p_{ii} = 1$.*
- recurrent*, if $[i]$ is a closed communication class.*

Without a proof we shall give the interpretation of the previous definitions. Transient states are states that are visited only finitely many times during the whole evolution of the process, recurrent and absorbing states are visited infinitely many times with probability one. This definition of recurrence in this form is only valid for finite state space, but in this thesis we will treat only absorbing and transient states. In the general case one distinguishes further null-recurrence and positive-recurrence, see e.g. [Bré99], chapter 3.

We shall give a final result on the limiting behavior of MCd with irreducible transition matrix.

Proposition 1.2.10 *Let $(X_n)_{n\geq 0}$ be a MCd on E with irreducible transition matrix P. Then*

$$\lim_{n\to\infty} \mathbb{P}(X_n = .) = \pi$$

where π is the unique *solution of the linear system*

$$\pi P = \pi,\ \pi \mathbf{1}^\top = 1.$$

The limit distribution π is also called *stationary distribution*. This distribution has the nice property that if $\mathbb{P}(X_0 = .) = \pi$ then

$$\mathbb{P}(X_n = .) = \pi P^n = (\pi P)P^{n-1} = \pi P^{n-1} = \ldots = \pi.$$

1.2.2 Continuous time

We now set $I = \mathbb{R}^+$.

Definition 1.2.11 *A stochastic process $(X_t)_{t\geq 0}$ on E with initial distribution μ, such that for arbitrary $t_0 < t_1 < \ldots < t_n$ and $i_0, i_1, \ldots, i_n$*

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_1} = i_1, X_{t_0} = i_0) = \mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1})$$

holds is called Markov Chain in continuous time *(MCc). If additionally*

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}) = \mathbb{P}(X_{t_n - t_{n-1}} = i_n | X_0 = i_{n-1})$$

holds, the Markov Chain is called homogeneous.

As in the discrete case we avoid writing homogeneous wherever possible and make the assumption that homogeneity holds for all processes appearing in this thesis.

In continuous time the definition of a Markov Chain is more tedious. We thus give only the most important results. We follow [Bré99] Chapter 8.

Let $(X_t)_{t\geq 0}$ be a MCc on E and define $P(t) := (p_{ij}(t))_{i,j\in E}$ with

$$p_{ij}(t) := \mathbb{P}(X_t = j | X_0 = i)$$

the so called *transition semi group*. Such a semi group has the following properties:

- for each $t \geq 0$ $P(t)$ is stochastic.
- $P(0) = Id$.
- $P(t+s) = P(t)P(s)$.

Proposition 1.2.12 *If $(X_t)_{t\geq 0}$ is a MCc on E with transition semi group $P(t)$ as defined above, then there exists a matrix $Q = (q_{ij})_{i,j\in E}$ such that*

$$-q_{ii} := \lim_{\epsilon \to 0} \frac{1 - p_{ii}(\epsilon)}{\epsilon} \in [0, \infty], i \in E$$

and

$$q_{ij} := \lim_{\epsilon \to 0} \frac{p_{ij}(\epsilon)}{\epsilon} \in [0, \infty], i, j \in E, i \neq j.$$

Furthermore $Q\mathbf{1}^\top = \mathbf{0}$.

Definition 1.2.13 *The square matrix $Q = (q_{ij})_{i,j}$ of Proposition 1.2.12 is called* infinitesimal generator.

Remark 1.2.14 In later sections we are interested in inverses of *sub infinitesimal generators*, that is a matrix with non-positive entries at the diagonal and non-negative off-diagonal entries, but where at least one row does not sum up to zero.

A infinitesimal generator is not invertible, as $Q\mathbf{1}^\top = \mathbf{0}$ and thus 0 is an eigenvalue. For sub infinitesimal generators 0 is no eigenvalue as these matrices have the nice property that all eigenvalues have strict negative real part.

□

The name infinitesimal generator is due to the fact that Q describes the probabilities to jump from one state into another within an interval of infinitesimal size, see the preceding proposition. Furthermore the unique minimal solution of the following set of differential equations (the Kolmogoroff forward and backward equations)

$$\frac{d}{dt}P(t) = QP(t),\ \frac{d}{dt}P(t) = P(t)Q$$

is

$$P(t) = \exp(Qt),$$

where $\exp(Qt) := \sum_{n\geq 0} \frac{t^n}{n!}Q^n$. Therefore Q "generates" all necessary information about the distribution of the process. In particular for a fixed initial distribution μ the distribution of X_t is surely given by $\mathbb{P}(X_t = .) = \mu P(t)$ and we can describe this distribution also in terms of the infinitesimal generator Q by

$$\mathbb{P}(X_t = .) = \mu \exp(Qt).$$

The very important notion of *embedded Markov Chain* helps to translate some notions from discrete time to continuous time.

Definition 1.2.15 *Let $(X_t)_{t\geq 0}$ be a MCc on E with fixed initial distribution μ. Then the matrix $P = (p_{ij})_{i,j\in E}$ defined by*

$$p_{ij} := \begin{cases} 0 & i = j \\ \frac{q_{ij}}{-q_{ii}} & i \neq j \end{cases}$$

is stochastic and the pair (μ, P) defines a MCd on E. This MCd is called embedded Markov Chain *(associated to $(X_t)_{t\geq 0}$).*

Then we gain

Proposition 1.2.16 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator Q and initial distribution μ. Then the infinitesimal generator is irreducible if the transition matrix of the embedded Markov Chain is. Moreover a state i is recurrent (transient, absorbing) if the associated state of the embedded Markov Chain is recurrent (transient, absorbing).*

We can also characterize the long time behavior of a MCc.

Proposition 1.2.17 *Let $(X_t)_{t\geq 0}$ be a MCc on E with irreducible infinitesimal generator and initial distribution μ. Then*

$$\lim_{t\to\infty} \mathbb{P}(X_t = .) = \pi,$$

where π is the unique solution of the set of linear equations

$$\pi Q = \mathbf{0}^\top,\ \pi \mathbf{1}^\top = 1.$$

Again π is called *stationary distribution*. It shows the same "stationary behavior" as in the discrete time case, namely if $\mathbb{P}(X_0 = .) = \pi$ then

$$\mathbb{P}(X_t = .) = \pi \exp(Qt) = \pi(Id + \sum_{n\geq 1} \frac{t^n}{n!}Q^n) = \pi + \sum_{n\geq 1} \frac{t^n}{n!} \underbrace{(\pi Q)}_{=\mathbf{0}} Q^{n-1} = \pi.$$

Remark **1.2.18** We have here implicitly imposed regularity conditions, that is $-q_{ii} > \infty$ and summable rows. This is will be the case throughout the thesis.

□
□

How the stationary distribution π for irreducible infinitesimal generator and transition matrices can be calculated by other means is topic of the next section.

1.3 Reversibility, permuted balance and extended Kolmogoroff Criterion

For an arbitrary MCd $(X_n)_{n\geq 0}$ on the finite state space $E = \{1, 2, \ldots, m\}$ with transition matrix $P = (p_{ij})_{i,j\in E}$, a probability measure π on E is *stationary* if and only if

$$\pi P = \pi$$

and *reversible* if and only if for all $i, j \in E$

$$\pi_i p_{ij} = \pi_j p_{ji}. \tag{1.1}$$

The condition (1.1) is called throughout the literature *balance condition*, *detailed balance* or π *is in detailed balance with* P, see e.g. [Rob10] Theorem 4.3. Let for any vector $\nu = (\nu_1, \nu_2, \ldots, \nu_n)$:

$$diag(\nu) := \begin{pmatrix} \nu_1 & 0 & \ldots & 0 & 0 \\ 0 & \nu_2 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & \nu_{n-1} & 0 \\ 0 & 0 & \ldots & 0 & \nu_n \end{pmatrix}$$

the diagonal matrix with the components of ν as diagonal entries; all other entries vanish. For π fixed, define

$$P^R = (p^R_{ij})_{i,j\in E} := \mathrm{diag}(\pi) P \mathrm{diag}(\pi)^{-1},$$

i.e.

$$p^R_{ij} = \frac{\pi_i}{\pi_j} p_{ij}.$$

It is a rescaling of rows and columns of the stochastic matrix P with π, in other words a linear transform.

Then (1.1) is equivalent to the matrix identity:

$$P^\top = P^R \tag{1.2}$$

This shows that the transposed transition matrix is a similarity transform of the original transition matrix in case of reversibility. Then the matrix P^R is also a stochastic matrix with invariant distribution π, but with all transitions $i \to j$ reversed to $j \to i$. The proof is simple:

$$e_i P^R \mathbf{1}^\top = \sum_{j\in E} p^R_{ij} = \sum_{j\in E} \frac{\pi_j}{\pi_i} p_{ji} \overset{(1.1)}{=} \sum_{j\in E} p_{ij} = 1$$

for every $i \in E$ and

$$\pi P^R e_j^\top = \sum_{i \in E} \pi_i p_{ij}^R \overset{(1.1)}{=} \sum_{i \in E} \pi_i \frac{\pi_j}{\pi_i} p_{ji} = \pi_j.$$

The MCd $(X_n^R)_{n \geq 0}$ associated to the transition matrix P^R is called the *time reversal* of $(X_n)_{n \geq 0}$, see e.g. [Nor97], section 1.9. The following relation between the processes $(X_n)_{n \geq 0}$ and $(X_n^R)_{n \geq 0}$ holds:

$$(X_n)_{n \geq 0} \overset{d}{=} (X_n^R)_{n \geq 0}$$

if $\mathcal{L}(X_0) = \pi$.

With Kolmogoroff's extension theorem the process can be extended to negative indices. Then the name *time reversal* becomes clear as the following property holds:

$$(X_n)_{n \geq 0} \overset{d}{=} (X_{-n})_{n \geq 0}.$$

Thus reversibility also states informally that future and past can be exchanged without changing the distribution.

By the preceding argumentation reversibility turns out to be a path-wise property. Indeed it means that in equilibrium (that is $\mathcal{L}(X_0) = \pi$) there is no "$\mathbb{P}$-measurable" difference whether the dynamics defined by the chain takes the path $(j_1, j_2, \ldots, j_k)$ from j_1 to j_k or the reversed one from j_k to j_1 for any $j_1, \ldots, j_k \in E$.

In [Kol36] this property of exchangability is characterized via a condition on loops. A non-trivial loop is here a sequence of at least three distinct states. A modernized form of this statement is proved in [Kel79], section 1.5, Theorem 1.7. (here in the notation of this thesis):

Theorem 1.3.1 *Let $(X_n)_{n \geq 0}$ be a MCd with irreducible transition matrix $P = (p_{ij})_{i,j \in E}$ on E. Then the unique stationary measure π is reversible with respect to P if and only if*

$$p_{j_1 j_2} p_{j_2 j_3} \cdots p_{j_{k-1} j_k} p_{j_k j_1} = p_{j_1 j_k} p_{j_k j_{k-1}} \cdots p_{j_3 j_2} p_{j_2 j_1} \tag{1.3}$$

for any finite sequence of states $j_1, j_2, \ldots, j_k \in E$.

Proof See [Kel79], section 1.5, Theorem 1.7. ■

The Kolmogoroff criterion (1.3) allows to check for reversibility without knowing π explicitly. But calculating all loops for a given transition graph can be a very tedious, because there might be many loops. Thus the criterion is better for determining that a given dynamics is *not* reversible.

For two states j_1, j_2 (1.3) is always fulfilled. That is why we call loops of length two *trivial loops*.

Remark 1.3.2 An obvious consequence of Theorem 1.3.1 is that any MCd without non-trivial loops is *automatically* reversible. Two such important classes of reversible MCd's are Birth-and-Death-Processes and MCd's on trees. □

Reversibility is a very strong condition on a MCd and thus seldom fulfilled. Therefore we introduce a new weaker form of reversibility. First we fix some notations. Let $\mathcal{S}_E$ be the

symmetric group defined on $E = \{1, 2, \ldots, n\}$. An element of $\mathcal{S}_E$ is denoted with σ and called *permutation*. We interpret σ as a bijective function $\sigma : E \to E$ and write $\sigma(i)$ for the image of $i \in E$ under σ. If σ is the identity, we write $\sigma = id$. Furthermore let $P_\sigma := (p_{ij}^\sigma)_{i,j\in E}$ a $|E| \times |E|$ matrix such that

$$p_{ij}^\sigma = \begin{cases} 1 & \text{if } j = \sigma(i) \\ 0 & else \end{cases}.$$

As permutations are bijective P_σ is a bi-stochastic matrix for arbitrary $\sigma \in \mathcal{S}_E$ and $P_{id} = Id$ the identity matrix. We say $i \in E$ is a *fixed point* of σ if $\sigma(i) = i$. The matrix P_σ is then irreducible if and only if σ contains no fixed points. By definition of σ, the associated matrix P_σ is also invertible (because σ is invertible).

The idea of a weaker form of reversibility is the following. Theorem 1.3.1 states that the path weight of any loop $\nu = (j_1, j_2, \ldots, j_n)$ is equal to the path weight of the reversed loop

$$\nu^R := (j_n, j_{n-1}, \ldots, j_1).$$

If we define

$$\nu^\sigma := (\sigma(j_1), \sigma(j_2), \ldots, \sigma(j_k))$$

we can demand that the path weight of ν is equal to the path weight of $(\nu^\sigma)^R$.

Definition 1.3.3 *Let $(X_n)_{n\geq 0}$ be an MCd on E with irreducible transition matrix $P = (p_{ij})_{i,j\in E}$ and π be a probability distribution. If there exists a permutation $\sigma \in \mathcal{S}_E$ such that for all $i, j \in E$*

$$\pi_i p_{ij} = \pi_{\sigma(j)} p_{\sigma(j)\sigma(i)} \tag{1.4}$$

holds, π is in permuted balance *with P.*

For the neutral element $\sigma = id$ (1.1) and (1.4) coincide.

The following example shows that definition 1.3.3 is non-empty. We will give more sophisticated examples in chapter 3:

Example 1.3.4 Let $(X_n)_{n\geq 0}$ be a MCd on $E := \{1, 2, 3, 4\}$ with transition matrix:

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

and initial distribution μ. The transition graph is given in figure 1.1.

By Theorem 1.3.1 it is immediate that this MCd can not have a reversible stationary distribution, as the only existing non-trivial loop has no reversed counterpart with positive weight.

Let now $\sigma = (24)$, thus σ leaves 1 and 3 invariant, while it maps 2 to 4 and vice versa. Then (1.4) together with $\pi \mathbf{1}^\top = 1$ has the unique solution:

$$\pi = \frac{1}{4}(1, 1, 1, 1),$$

and is also the invariant measure of P (this is not implied by the definition of permuted balance).

Note that the choice of σ is not unique, $\sigma = (13)$ also fulfills (1.4), but $\sigma = (12)$ does not. □

□

We proceed with a list of properties of permuted balance, but before:

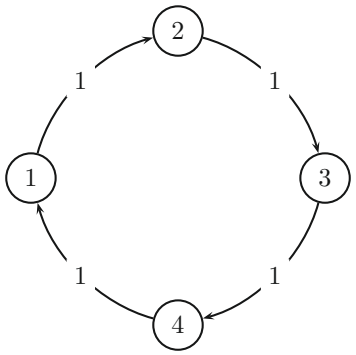


Figure 1.1: Transition graph of example 1.3.

Definition 1.3.5 *For a given permutation $\sigma \in \mathcal{S}_E$, the set*

$$\langle i \rangle_\sigma := \bigcup_{k \geq 0} \{\sigma^k(i)\}$$

is the cycle generated by $i \in E$ under σ.

The cycles under σ generate a canonical decomposition of E.

Lemma 1.3.6 *Let σ be a permutation over E. Then there exists a minimal list of states*

$$i_0, i_1, \ldots, i_k, \ldots \in E$$

with respect to σ such that

$$E = \bigcup_{k \geq 0} \langle i_k \rangle_\sigma$$

is a union of disjoint sets.

Proof The fact that σ is a bijection ensures the existence of such a (minimal) decomposition.

■

The following proposition is a collection of algebraic properties implied by permuted balance.

Proposition 1.3.7 *Let P be an irreducible transition matrix and π be a distribution in permuted balance with P for some permutation σ. Assume that $\pi_i = \pi_{\sigma(i)}$ for all $i \in E$. Then the following properties hold*

(i) π is the (unique) stationary distribution of P.

(ii) π is constant on cycles of σ, i.e. for all $i \in E$ and $k \in \mathbb{N}$

$$\pi_i = \pi_{\sigma^k(i)}.$$

(iii) for $i, j \in E$

$$p_{ij} = \begin{cases} p_{\sigma^k(i)\sigma^k(j)} & \text{if } k \text{ even} \\ \frac{\pi_j}{\pi_i} p_{\sigma^k(j)\sigma^k(i)} & \text{if } k \text{ odd} \end{cases}$$

in particular if $i, j \in \langle i \rangle_\sigma$

$$p_{ij} = \begin{cases} p_{\sigma^k(i)\sigma^k(j)} & \text{if } k \text{ even} \\ p_{\sigma^k(j)\sigma^k(i)} & \text{if } k \text{ odd} \end{cases}$$

(iv) if $\langle i \rangle_\sigma = \{i\}$ and $\langle j \rangle_\sigma$ is such that $\#\langle j \rangle_\sigma = k > 0$ and k odd, then

$$p_{ij} = p_{i\sigma^m(j)} = \frac{1 - \sum_{k \notin \langle j \rangle_\sigma} p_{ik}}{\#\langle j \rangle_\sigma}$$

for all $m \geq 0$.

(v) For $k \in E$ such that $\langle i \rangle_\sigma \cap \langle k \rangle_\sigma = \emptyset$ and $i, j \in \langle i \rangle_\sigma$:

$$\sum_{m \in \langle k \rangle_\sigma} p_{im} = \sum_{m \in \langle k \rangle_\sigma} p_{jm}$$

(vi) Let $E_\sigma = \{\langle k \rangle_\sigma, k \in E\}$ and

$$p_{\langle i \rangle_\sigma \langle j \rangle_\sigma} := \sum_{m \in \langle j \rangle_\sigma} p_{im}$$

Then the matrix $\tilde{P} = (p_{\langle i \rangle_\sigma \langle j \rangle_\sigma})_{\langle i \rangle_\sigma \langle j \rangle_\sigma \in E_\sigma}$ is stochastic with stationary measure $\tilde{\pi}$ given by

$$\tilde{\pi}_{\langle i \rangle_\sigma} = \#\langle i \rangle_\sigma \pi_i.$$

Moreover $\tilde{\pi}$ is reversible with respect to $\tilde{P}$.

Proof

(i) Direct computation shows:

$$\pi P e_j^\top = \sum_{i \in E} \pi_i p_{ij} \overset{(1.4)}{=} \sum_{i \in E} \pi_{\sigma(j)} p_{\sigma(j)\sigma(i)} = \pi_j \sum_{i \in E} p_{\sigma(j)\sigma(i)} = \pi_j$$

(ii) Since $\sigma^k(j) \in E$ for any $k \geq 0$:

$$\pi_i = \pi_{\sigma(i)} = \pi_{\sigma^2(i)} = \dots$$

by assumption.

(iii) By iteration of (1.4) the first formula for arbitrary $i, j \in E$ follows. Since $i, j \in \langle i \rangle_\sigma$ (and thus $i \in \langle j \rangle_\sigma$) there exists a k_0 such that $\sigma^{k_0}(i) = j$ and therefore by (i) $\pi_i = \pi_j$.

(iv) Since k is odd, the set $\langle j \rangle_\sigma$ is equal to $\langle \sigma^2(j) \rangle_\sigma$. Then (iii) proves the statement.

(v) Since i and j are in the same cycle, i.e. $j \in \langle i \rangle_\sigma$ there exists a minimal k_0 such that $\sigma^{k_0}(j) = i$. Since $2k_0$ is even it follows by (iii)

$$p_{jm} = p_{i\sigma^{2k_0}(m)}.$$

Summing on both side over the elements of $\langle k \rangle_\sigma$ gives

$$\sum_{m \in \langle k \rangle_\sigma} p_{im} = \sum_{m \in \langle k \rangle_\sigma} p_{j\sigma^{2k_0}(m)} = \sum_{m \in \langle k \rangle_\sigma} p_{jm}$$

since by definition of $\langle k \rangle_\sigma$ the images $\sigma^{2k_0}(m) \in \langle k \rangle_\sigma$ for all $m \in \langle k \rangle_\sigma$.

(vi) By direct computation we obtain for the probability to move from set $\langle i \rangle_\sigma$ to $\langle k \rangle_\sigma$ (with rescaling):

$$\frac{1}{\#\langle i \rangle_\sigma} \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} p_{nm} \overset{(iv)}{=} \frac{1}{\#\langle i \rangle_\sigma} \#\langle i \rangle_\sigma \sum_{m \in \langle j \rangle_\sigma} p_{im} = p_{\langle i \rangle_\sigma \langle j \rangle_\sigma}$$

The stochasticity is directly inherited from P and the definition of the $p_{\langle i \rangle_\sigma \langle j \rangle_\sigma}$. We check the balance conditions by direct computation:

$$\begin{aligned}
\tilde{\pi}_{\langle i \rangle_\sigma} p_{\langle i \rangle_\sigma \langle j \rangle_\sigma} &= \#\langle i \rangle_\sigma \pi_i \sum_{m \in \langle j \rangle_\sigma} p_{im} = \pi_i \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} p_{nm} \\
&\overset{(i)}{=} \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} \pi_n p_{nm} = \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} \pi_{\sigma(m)} p_{\sigma(m)\sigma(n)} \\
&\overset{(i)}{=} \pi_j \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} p_{\sigma(m)\sigma(n)} \overset{(\star)}{=} \pi_j \sum_{n \in \langle i \rangle_\sigma} \sum_{m \in \langle j \rangle_\sigma} p_{mn} \\
&\overset{(v)}{=} \#\langle j \rangle_\sigma \pi_j \sum_{n \in \langle i \rangle_\sigma} p_{jn} = \tilde{\pi}_{\langle j \rangle_\sigma} p_{\langle j \rangle_\sigma \langle i \rangle_\sigma}
\end{aligned}$$

where $(\star)$ is just a change of summation order. Therefore $\tilde{\pi}$ is reversible (according to (1.1)) for $\tilde{P}$.

■

Proposition 1.3.7 shows that the permuted balance as defined in (1.4) allows to compute the stationary distribution if and only if $\pi_i = \pi_{\sigma(i)}$, (see (i) in the proposition). Conversely, for any path $\nu = (i_1, \ldots, i_n)$ from i_1 to i_n we have by permuted balance:

$$\begin{aligned}
\pi_{i_1} \prod_{k=1}^{n-1} p_{i_k i_{k+1}} &= \pi_{i_1} \prod_{k=1}^{n-1} \frac{\pi_{\sigma(i_{k+1})}}{\pi_{i_k}} p_{\sigma(i_{k+1})\sigma(i_k)} \\
&= \pi_{\sigma(i_n)} \prod_{k=1}^{n-1} \frac{\pi_{\sigma(i_k)}}{\pi_{i_k}} \cdot \prod_{k=1}^{n-1} p_{\sigma(i_{k+1})\sigma(i_k)}
\end{aligned}$$

Conversely if we know the stationary distribution, then the model can be only in permuted balance with respect to permutation σ if it exchanges states with equal stationary distribution. Therefore if all components of the stationary distribution are pairwise distinct, permuted balance is impossible.

Since the aim of the definition of permuted balance was that any path can be replaced by the permuted reversed one, the "extra cost" of doing so $\prod_{k=1}^{n-1} \frac{\pi_{\sigma(i_k)}}{\pi_{i_k}}$ has to be equal to one. As paths of length two should also have this property $\pi_i = \pi_{\sigma(i)}$ is implied.

In the following permuted balance has thus to be understood under the implicit assumption $\pi_i = \pi_{\sigma(i)}$ for each $i \in E$, to ensure that a measure in permuted balance is also stationary. Thus a measure π is in permuted balance, if and only if

$$\forall i, j \in E : \pi_i p_{ij} = \pi_j p_{\sigma(j)\sigma(i)} \tag{1.5}$$

holds. In this case condition (1.4) can be rewritten in matrix form as:

$$P_\sigma^{-1}\mathrm{diag}(\pi)P = P^\top P_\sigma^{-1}\mathrm{diag}(\pi). \tag{1.6}$$

Therefore the transpose of P is (like in the reversible case) a similarity transform of P:

$$P^\top = (P_\sigma^{-1}\mathrm{diag}(\pi))P(P_\sigma^{-1}\mathrm{diag}(\pi))^{-1}. \tag{1.7}$$

A reversible MCd can be associated to the matrix $\tilde{P}$ and the invariant vector $\tilde{\pi}$ given in (vi) of Proposition 1.3.7 for some initial condition. But this does not explain under which conditions the image of a MCd in permuted balance by aggregation of states is still a MCd; this *does* depend on the choice of initial condition of the original chain. The following results are the condensate of Definition 6.3.1 and Theorem 6.3.2, p. 124 and Theorem 6.4.1, p. 133 of [KS76].

Definition 1.3.8 *Let $(X_n)_{n\geq 0}$ be a MCd on E with initial distribution μ and let $E_1 \cup \ldots \cup E_k$ be a partition of E. Then the process $(Y_n)_{n\geq 0}$ with values in $\{E_1, \ldots, E_k\}$ defined by*

$$\begin{aligned}&\mathbb{P}_{\mu'}(Y_1 = E_i | Y_0 = E_j) = \mathbb{P}_\mu(X_1 \in E_i | X_0 \in E_j)\\ &\mathbb{P}_{\mu'}(Y_n = E_i | Y_0 = E_j, Y_1 = E_{j_1}, \ldots, Y_{n-1} = E_{j_{n-1}})\\ &= \mathbb{P}_\mu(X_n \in E_i | X_0 \in E_j, X_1 \in E_{j_1}, \ldots, X_{n-1} \in E_{j_{n-1}})\end{aligned}$$

with initial distribution $\mu' = (\mu'_{E_1}, \ldots, \mu'_{E_k})$ with $\mu'_{E_i} := \mathbb{P}(X_0 \in E_i)$ is called the lumped process.

The following proposition determines under which conditions the lumped process is again a MCd, depending on the original chain.

Proposition 1.3.9 *Let $(X_n)_{n\geq 0}$ be MCd on E with initial distribution $\mu = (\mu_1, \mu_2, \ldots)$ and transition matrix $P = (p_{ij})_{i,j\in E}$. Let $E = E_1 \cup \ldots \cup E_n \cup \ldots$ be a partition of the E into pairwise disjoint sets. Then the lumped process as defined in Definition 1.3.8 is a MCd, if and only if for every E_i and E_j and all $l, m \in E_i$*

(i) $\sum_{n\in E_j} p_{ln} = \sum_{n\in E_j} p_{mn}$

(ii) $\sum_{n\in E_j} \mu_l p_{ln} = \sum_{n\in E_j} \mu_m p_{mn}$

hold.

We now want to connect lumpability and permuted balance. It turns out that under special conditions on the initial distribution a MCd with a transition matrix in permuted balance with some π and a permutation $\sigma \in \mathcal{S}_E$ can be reduced to a reversible MCd on the aggregated (smaller) state space given by the canonical decomposition of E under σ.

Theorem 1.3.10 *Let $(X_n)_{n\geq 0}$ be a MCd on E with transition matrix $P = (p_{ij})_{i,j\in E}$ in permuted balance with π for some $\sigma \in \mathcal{S}_E$ and $E = E_1 \cup \ldots \cup E_k$ the canonical decomposition of E with respect to σ and μ the initial condition. Then the lumped process (as defined in Definition 1.3.8) is a MCd with reversible stationary distribution μ' if and only if for every E_i and E_j and for all $l, m \in E_i$*

$$\sum_{n\in E_j} \mu_l p_{ln} = \sum_{n\in E_j} \mu_m p_{mn}. \tag{1.8}$$

This holds in particular for $\mu = \pi$.

Proof By Proposition 1.3.7, (v), (vi), the first condition of Proposition 1.3.9 is under the assumptions of the theorem always fulfilled and the stationary measure of the lumped chain is reversible. Thus (1.8) is necessary and sufficient for the lumped chain to be a MCd. If $\mu = \pi$ then the conditions of Proposition 1.3.9 are both fulfilled, as the invariant distribution of $(X_n)_{n\geq 0}$ is uniform on cycles, by Proposition 1.3.7, (ii).

■

Example 1.3.11

Let $(X_n)_{n\geq 0}$ be a MCd on $E = \{1, 2, 3, 4, 5\}$ with transition matrix

$$P = \frac{1}{6}\begin{pmatrix} 0 & 2 & 0 & 2 & 2 \\ 3 & 0 & 3 & 0 & 0 \\ 0 & 2 & 0 & 2 & 2 \\ 3 & 0 & 3 & 0 & 0 \\ 3 & 0 & 3 & 0 & 0 \end{pmatrix}$$

and transition graph as given in fig. 1.2.

By irreducibility of P the stationary distribution is unique and given as

$$\pi = \frac{1}{12}(3, 2, 3, 2, 2).$$

Note that this model is reversible and thus in permuted balance with respect to $\sigma = id$. In general this does not imply that it is also in permuted balance with respect to other permutations. We show here that there may be several different permutations σ leading to very different lumped reversible chains. We omit the details of the calculations and give the results directly. The resulting lumped transition graphs are given in fig. 1.3.

(i) $\underline{\sigma = (245)}$ We have

$$E_\sigma = \{\langle 1\rangle_\sigma, \langle 3\rangle_\sigma, \langle 2\rangle_\sigma\} = \{\{1\}, \{3\}, \{2, 4, 5\}\}$$
$$\pi = \frac{1}{4}(1, 2, 1)$$

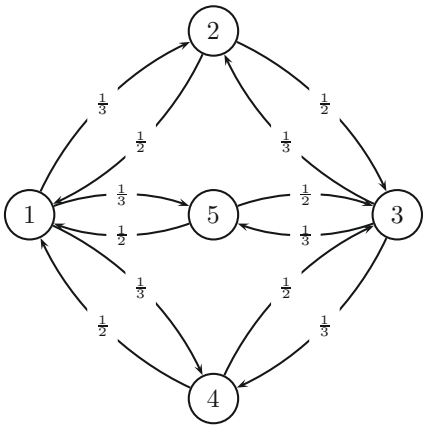


Figure 1.2: Transition graph of Example 1.3.

(ii) $\underline{\sigma = (13)}$ We have

$$E_\sigma = \{\langle 1\rangle_\sigma, \langle 2\rangle_\sigma, \langle 4\rangle_\sigma, \langle 5\rangle_\sigma\} = \{\{1,3\},\{2\},\{4\},\{5\}\}$$
$$\pi = \frac{1}{6}(3,1,1,1)$$

(iii) $\underline{\sigma = (13)(245)}$ We have

$$E_\sigma = \{\langle 1\rangle_\sigma, \langle 2\rangle_\sigma\} = \{\{1,3\},\{2,4,5\}\}$$
$$\pi = \frac{1}{2}(1,1)$$

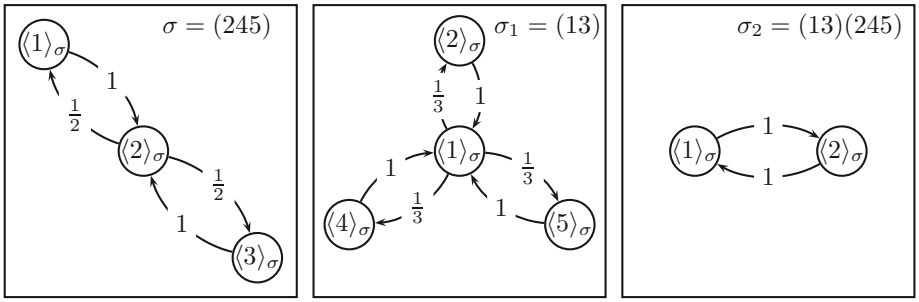


Figure 1.3: Lumped chains for Example 1.3.

□
□

To characterize the permuted balance using only the transition probabilities we extend the Kolmogoroff criterion given in Theorem 1.3.1 and call it *extended Kolmogoroff Criterion for permuted balance*.

Theorem 1.3.12 *Let $(X_n)_{n\geq 0}$ be a MCd on E with irreducible transition matrix $P = (p_{ij})_{i,j\in E}$. The unique stationary distribution π is in permuted balance with respect to a given permutation σ over E if and only if*

$$p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_k i_1} = p_{\sigma(i_1)\sigma(i_k)} \cdots p_{\sigma(i_3)\sigma(i_2)} p_{\sigma(i_2)\sigma(i_1)} \tag{1.9}$$

for each finite sequence of states $i_1, i_2, \ldots, i_k \in E$.

Proof The proof is analogue to the one given in [Kel79], section 1.5, Theorem 1.7 for the usual Kolmogoroff Criterion.

Let π be in permuted balance. It is also the stationary distribution by Proposition 1.3.7, (i). Then (1.5) implies for a sequence $i_1, \ldots, i_n$:

$$\begin{aligned}
\pi_{i_1} p_{i_1 i_2} &= \pi_{i_2} p_{\sigma(i_2)\sigma(i_1)} \\
\pi_{i_2} p_{i_1 i_3} &= \pi_{i_3} p_{\sigma(i_3)\sigma(i_2)} \\
&\vdots \\
\pi_{i_{k-1}} p_{i_{k-1} i_k} &= \pi_{i_k} p_{\sigma(i_k)\sigma(i_{k-1})} \\
\pi_{i_k} p_{i_k i_1} &= \pi_{i_1} p_{\sigma(i_1)\sigma(i_k)}
\end{aligned}$$

Multiplying all left hand sides and all right hand sides together and canceling the $\pi_i' s$ out gives (1.9).

Conversely, suppose (1.9) holds and fix $i_0 \in E$. Then for arbitrary $i \in E$ there exists by irreducibility a sequence of states $i_1, i_2, \ldots, i_k$ with

$$p_{\sigma(i),\sigma(i_k)} p_{\sigma(i_k)\sigma(i_{k-1})} \cdots p_{\sigma(i_1)\sigma(i_0)} > 0$$

from $\sigma(i)$ to $\sigma(i_0)$. If there is another path $j_1, \ldots, j_m$ such that

$$p_{\sigma(i),\sigma(j_k)} p_{\sigma(j_k)\sigma(j_{k-1})} \cdots p_{\sigma(j_1)\sigma(i_0)} > 0,$$

identity (1.9) implies

$$\frac{p_{i_0 i_1} \cdots p_{i_{n-1} i_n} p_{i_n i}}{p_{\sigma(i),\sigma(i_n)} p_{\sigma(i_n)\sigma(i_{n-1})} \cdots p_{\sigma(i_1)\sigma(i_0)}} = \frac{p_{i_0 j_1} \cdots p_{j_{m-1} j_n} p_{j_m i}}{p_{\sigma(i),\sigma(j_m)} p_{\sigma(j_m)\sigma(j_{m-1})} \cdots p_{\sigma(j_1)\sigma(i_0)}}.$$

Thus, for fixed $B > 0$,

$$\pi_i := B \frac{p_{i_0 i_1} \cdots p_{i_{n-1} i_n} p_{i_n i}}{p_{\sigma(i),\sigma(i_n)} p_{\sigma(i_n)\sigma(i_{n-1})} \cdots p_{\sigma(i_1)\sigma(i_0)}}$$

is defined for every $i \in E$ and independent of the chosen path.

We now need to show the permuted balance equation (1.5). Therefore for a state $k \in E$ with $p_{\sigma(k)\sigma(i)} > 0$ we can write

$$\pi_k = B \frac{p_{i_0 i_1} \cdots p_{i_{n-1} i_n} p_{i_n i} p_{ik}}{p_{\sigma(k)\sigma(i)} p_{\sigma(i),\sigma(i_n)} p_{\sigma(i_n)\sigma(i_{n-1})} \cdots p_{\sigma(i_1)\sigma(i_0)}} = \pi_i \frac{p_{ik}}{p_{\sigma(k)\sigma(i)}}$$

and thus the permuted balance equations are fulfilled. π as defined must be also stationary by Proposition 1.3.7 and B is uniquely determined by $\pi \mathbf{1}^\top = 1$.

■

Remark **1.3.13** The proof of Theorem 1.3.12 shows that the stationary resp. the permuted balanced distribution can be directly reconstructed from (1.9).

□

Remark **1.3.14** The results of Lemma 1.3.6, Propositions 1.3.9 and 1.3.7, Theorems 1.3.10 and 1.3.12 still hold if E is countable infinite. We interpret the symmetric group on E as the collection of all bijective functions on E (which is also a group with the usual composition). Note also that positive recurrence is needed for the uniqueness and positivity of the invariant distribution π in the countable case.

□

With a method called *uniformization* it is simple to transport the preceding results on MCd's to the time continuous case under the mild assumption that the infinitesimal generator has summable rows in the case of countable state space, see e.g. [Ros96].

Definition 1.3.15 *Let $Q = (q_{ij})_{i,j\in E}$ be an infinitesimal generator. Then for*

$$\gamma = \max_{i\in E}\{-q_{ii}\}$$

the stochastic matrix $P = (p_{ij})_{i,j\in E}$ defined by

$$p_{ij} = \begin{cases} \frac{1}{\gamma}q_{ij} & i \neq j \\ 1 - \sum_{j,j\neq i}\frac{1}{\gamma}q_{ij} & i = j \end{cases}$$

is called the uniformization *of Q.*

Proposition 1.3.16 *Let $(X_t)_{t\geq 0}$ be a MCc on E with irreducible infinitesimal generator. Then P, the uniformization of Q, has the same stationary distribution as Q.*

Proof Let π be the stationary distribution of Q, i.e.

$$\pi Q = 0,\ \pi \mathbf{1}^\top = 1.$$

We need to show $\pi P = \pi$. To do so, we rephrase the definition of P in matrix notation:

$$P = \frac{1}{\gamma}Q + Id.$$

Now the statement is obvious.

■

Remark **1.3.17** Thanks to the uniformization method Proposition 1.3.7 and Theorem 1.3.12 turn into their continuous time counterpart just by replacing p_{ij} by the corresponding q_{ij} and taking care of the fact that

$$\pi Q e_j^\top = 0$$

rather than π_j.

□

For further reference we define permuted balance in the continuous case:

Definition 1.3.18 *Let $(X_t)_{t\geq 0}$ be a MCc on E with irreducible infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. If there exist a probability measure π and a permutation σ on E such that for all $i, j \in E$*

$$\pi_i q_{ij} = \pi_j q_{\sigma(j)\sigma(i)} \tag{1.10}$$

then π is in permuted balance *with Q.*

1.4 Local time

For some applications it turns out to be useful to observe a given MCd or MCc only when it is in a finite subset D of its state space. The idea is to stop the global clock when the process leaves D and to let it run again when the process re-enters D. Naturally for any state j that can be left to D^c one would add the "expected excursion time" in D^c to the sojourn time in j. How this is done unveils the following theorem given in [LR99], section 5.3. A detailed proof is given in [KSK66], chapter 6, Proposition 6.4 and Lemmas 6.6 and 6.7 as pointed out in the proof in [LR99]. The continuous time part is stated in [LR99], section 5.5, Theorem 5.5.3 together with a proof.

For a finite subset $D \subsetneq E$ a transition matrix P over E resp. an infinitesimal generator Q can be partitioned in the following way according to D:

$$P = \begin{pmatrix} P_D & P_{DD^c} \\ P_{D^cD} & P_{D^c} \end{pmatrix} \tag{1.11}$$

resp.

$$Q = \begin{pmatrix} Q_D & Q_{DD^c} \\ Q_{D^cD} & Q_{D^c} \end{pmatrix}. \tag{1.12}$$

Theorem 1.4.1 *For a homogeneous, irreducible and positive recurrent MCd with transition matrix P on E and $D \subsetneq E$ with partition (1.11), the process $(X_n^D)_{n\geq 0}$ restricted to D is again homogeneous, irreducible and positive recurrent with transition matrix*

$$P_{|D} := P_D + P_{DD^c}(Id - P_{D^c})^{-1}P_{D^cD}.$$

For a homogeneous, irreducible and positive recurrent MCc with infinitesimal generator Q on E with partition (1.12) for the same set D, the process $(X_t^D)_{t\geq 0}$ restricted to D is again homogeneous, irreducible and positive recurrent with infinitesimal generator

$$Q_{|D} := Q_D - Q_{DD^c}Q_{D^c}^{-1}Q_{D^cD}.$$

The matrices $(Id - P)^{-1}$ in discrete case and $-Q^{-1}$ for the continuous time case are identified as matrices of expected passage times, see e.g. [DS65] and [DS67]. The (i, j)-th entry of these matrices is the expectation of the first passage time from i to j resp. the expectation of the first return time to i if $i = j$. Due to this interpretation it is clear that the transition matrix/infinitesimal generator of the restricted process are indeed just "corrected" by the expectation of the time of possible excursion to D^c departing from $i \in D$ to D^c and returning via state $j \in D^c$.

Remark **1.4.2** The notion of local time is in fact the stochastic interpretation of the so called *Schur complement* associated to the decomposition into a block matrix of an infinitesimal

generator as in Theorem 1.4.1. As long as D^c is not the empty set the Schur complement always exists, is invertible and an infinitesimal generator. For transition matrices resp. stochastic matrices $P_{|D}$ from Theorem 1.4.1 is the Schur complement associated to the matrix $Id - P$ with partition as in (1.11).

In both cases (continuous and discrete time) the Schur complement exists and is non-singular even in the case that absorbing states are contained in D, as the matrices associated to D^c are still invertible as long as all states in D^c are transient. In fact a review of the proofs in [KSK66] and [LR99] shows that the absence of absorbing states in D^c is the only constraint for using local time, i.e. recurrence of the chain is *not* necessary.

□

1.5 Pattern matching

For a given trajectory of a Markov Chain it is interesting to know the distribution of the occurrence of a certain transition or a given sequence of transitions. Such a sequence is often called *pattern*. The method which we shortly introduce is standard in many fields of applications, such as speech recognition or sequence analysis of DNA (base pair sequences), see e.g. [Nue08] and references therein.

The basic idea to analyze patterns in trajectories would be to define a Markov Chain with a "memory" long enough to find patterns of a certain length. For a MCd $(X_n)_{n\geq 0}$ with transition matrix $P = (p_{ij})_{i,j\in E}$ a (homogeneous) memory of length $m > 1$ would have the form:

$$\begin{aligned}
&\mathbb{P}(X_{n+m} = i_{n+m} | X_{n+m-1} = i_{n+m-1}, \dots, X_0 = i_0) \\
&\quad = \mathbb{P}(X_{n+m} = i_{n+m} | X_{n+m-1} = i_{n+m-1}, \dots, X_n = i_n) \\
&\quad = \mathbb{P}(X_m = i_{n+m} | X_{m-1} = i_{n+m-1}, \dots, X_0 = i_n) \\
&\quad = p_{i_n i_{n+1}} \cdot \dots \cdot p_{i_{n+m-1} i_{n+m}}.
\end{aligned}$$

The memory complicates the analysis as such a process is *not* a Markov Chain anymore. This problem can be circumvented by defining a new Markov Chain on the product state space

$$E' = E \times E \times \dots \times E =: E^{\times m}$$

with an appropriate transition matrix. This procedure turns every sequence of transitions of length m of the original Chain into a *state*.

Naturally a pattern of length m can first appear after m steps of the original Markov Chain, thus we fix:

Definition 1.5.1 *Let $(X_n)_{n\geq 0}$ be a MCd on E with transition matrix $P = (p_{ij})_{i,j\in E}$ and initial distribution μ. For a fixed pattern length $m \geq 1$ define $P' = (p'_{ij})_{i,j\in E^{\times m}}$ and $\mu' = (\mu'_i)_{i\in E^{\times m}}$ component-wise by*

$$p'_{(i_1,i_2,\dots,i_m)(j_1,j_2,\dots,j_m)} := \begin{cases} p_{i_2 i_3} \cdot \dots \cdot p_{i_m j_m} & \text{if } j_1 = i_2, \dots, j_{m-1} = i_m \\ 0 & \text{else} \end{cases}$$

and

$$\mu'_{i_1,i_2,\dots,i_m} := \mu_{i_1} p_{i_1 i_2} \dots p_{i_{m-1} i_m}.$$

We call the MCd associated to the pair (μ', P') the m-pattern chain.

By definition of the m-pattern chain the transition matrix P' is extremely sparse if m is big (i.e. the transition matrix contains many zeros), but for the applications of this thesis we choose $m = 2$.

Remark **1.5.2** In the case when the chain has an absorbing state, say Δ, we can restrict the transformation to the set of transient states, as it is in general not very enlightening to count the number of occurrences of transitions like $(j, \Delta), j \in E \setminus \{\Delta\}$, because they occur at most only once.

Therefore in the case of a MCd with only one absorbing state the m-pattern chain can be derived for the sub-matrix S associated to the transient states first; for the moment call the associated matrix S'. After that a m-pattern chain with absorbing state can be constructed through the definition

$$P' := \begin{pmatrix} 1 & 0 \\ (Id - S')\mathbf{1}^\top & S' \end{pmatrix}.$$

In a sense this is an aggregation of all states in $E^{\times m}$ containing a Δ. That this is possible is also a consequence of the lumping notion defined in chapter 2.

□

In chapter 4 we combine the pattern matching method with local time and the results on absorbing Markov Chains from chapter 2 to get insight in the distribution of the number of certain transitions before absorption, see Proposition 4.3.6.

Chapter 2

Absorption in Markov Chains

In this chapter we introduce the well known *phase type distributions*, which are defined as absorption times of Markov Chains. This type of "Matrix Exponential Distributions" was introduced rigorously for the first time by Neuts in [Neu94] and applied in the framework of Queueing Theory. We give only the basic characteristics, the detailed computations are in appendix B. The second half of chapter 2 is devoted to the situation of several absorbing states. We show here that a MCc conditioned to absorption in one absorbing state out of several can be described by a linear transform of the transition matrix/infinitesimal generator. This concept is known in older and applied theory as Duality, see [KSK66], and also as Doob-h-transform. We show here that a Markov Chain *conditioned* to absorption is characterized as a linear transformation of the transition matrix/infinitesimal generator. The concept is developed in this thesis for continuous time, but it was already introduced for discrete time as an example in [KS76], see chapter 3, section 3. A reinterpretation of conditioning to absorption in a designated state allows us to describe the process conditioned to leave the transients via specific states or sets thereof. We use this type of condition in chapter 4 to gain insight about the distribution of the last visited state(s) before absorption.

2.1 Phase type distributions

The notion *phase type Distribution* or shorter *PH-type distribution* was first introduced in [Neu94] as the time to absorption in a finite state Markov Chain (both in discrete and continuous time).

2.1.1 PH-type distribution as absorption time distribution

Consider a Markov Chain $(X_i)_{i\in I}$ with finite state space

$$E := \{\Delta\} \cup M$$

with $M := \{1, 2, \ldots, m\}$ for some index set I, such that Δ is an absorbing state and can be reached from every state in M. We extend the usual order "$<$" on $\mathbb{N}$ by defining $\Delta < k$ for all $k \in \mathbb{N}$. This order is used for the order of entries in the transition matrices and infinitesimal generators, except otherwise noted.

The *stopping time* (with respect to the canonical filtration)

$$\tau_\Delta := \inf\{i \in I : X_i = \Delta\}$$

is called *first exit time* of M or *absorption time* (in Δ).

The distribution of τ_Δ admits, depending on $(X_i)_{i\in I}$, a so called *representation*. We need some notation to define a representation.

For the discrete time case (i.e. $I = \mathbb{N}_0$) consider a MCd on E with transition matrix P and initial distribution μ. There is a natural decomposition of P into a block matrix:

$$P := \begin{pmatrix} 1 & \mathbf{0} \\ R_\Delta^\top & S \end{pmatrix}, \tag{2.1}$$

where S is a $m \times m$ matrix containing the transition probabilities between the states in M and R_Δ is a vector of length m with

$$R_\Delta = (p_{1\Delta}, p_{2\Delta}, \ldots, p_{m\Delta}).$$

For the initial distribution we write

$$\mu = (\mu_\Delta, \mu_M),$$

where $\mu_\Delta = \mathbb{P}(X_0 = \Delta)$ and

$$\mu_M := (\mu_1, \ldots, \mu_m)$$

with $\mu_i = \mathbb{P}(X_0 = i)$. We note that, depending on I, the matrix S is a sub stochastic matrix if $I = \mathbb{N}_0$, i.e. there is at least one row with sum less than one. If $I = \mathbb{R}^+$, the matrix S is a non conservative infinitesimal generator, i.e. there is at least one row with sum less then zero. We note that both matrices are invertible in the sense that S^{-1} exists if $I = \mathbb{R}^+$ and $(Id - S)^{-1}$ exists if $I = \mathbb{N}_0$, see appendix A for a proof.

Definition 2.1.1 *Let $(X_n)_{n\geq 0}$ be a MCd on E with transition matrix P and decomposition (2.1). Then*

$$\tau_\Delta := \inf\{n \geq 0 : X_n = \Delta\}$$

is of PH-type *or has* PH-type distribution *with* representation*: $PH(\mu_M, S)$. We also write*

$$\tau_\Delta \sim PH(\mu_M, S).$$

For $I = \mathbb{R}^+$ we decompose the infinitesimal generator Q analogue to the discrete case as:

$$Q := \begin{pmatrix} 0 & \mathbf{0} \\ R_\Delta^\top & S \end{pmatrix}. \tag{2.2}$$

Definition 2.1.2 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator Q and decomposition (2.2). Then*

$$\tau_\Delta := \inf\{t \geq 0 : X_t = \Delta\}$$

is of PH-type *or has* (continuous) PH-type distribution *with* representation*: $PH(\mu_M, S)$. We also write*

$$\tau_\Delta \sim PH(\mu_M, S).$$

For further computations we note the following trivial yet useful properties:

Corollary 2.1.3 *(i) Let $(X_n)_{n\geq 0}$ be a MCd on E with transition matrix P and decomposition (2.1), then*

$$(Id - S)^{-1} R_\Delta^\top = \mathbf{1}^\top.$$

(ii) Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator Q and decomposition (2.2), then

$$-S^{-1} R_\Delta^\top = \mathbf{1}^\top.$$

Proof The properties $P\mathbf{1}^\top = \mathbf{1}^\top$ and $Q\mathbf{1}^\top = \mathbf{0}^\top$ together with Lemma A.1.1 imply the statements.

■

2.1.2 Basic Properties

The basic properties of PH-type distributions are given in Tables 2.1 (discrete) and 2.2 (continuous) for further reference. Very brief calculations of the characteristics of PH-type distributions are given in [Neu94] and [LR99], in both references chapter 2. We included detailed calculations in appendix B. We use the notation introduced before and consider

$$\tau_\Delta \sim PH(\mu_M, S)$$

according to Definitions 2.1.1 and 2.1.2.

To avoid confusion we note again the basic definitions we used for the computations of the characteristics of PH-type distributions.

Definition 2.1.4 (basic characteristics) *Let X be some random variable with (cumulative) distribution function $F_X(t)$, with $X \geq 0$ almost sure.*

(i) If X takes its values in $\mathbb{N}_0$ the generating function of X is defined as the power series

$$G_X(z) := \sum_{n\geq 0} \mathbb{P}(X = n) z^n.$$

(ii) The moment generating function is defined as

$$M_X(s) := \int_0^\infty \exp(st) dF_X(t).$$

(iii) The characteristic function is defined as

$$\phi_X(s) := \int_0^\infty \exp(ist) dF_X(t)$$

where i is the imaginary unit.

We have to note that Neuts used another definition of moment generating functions, which results into a sign difference, see chapter 2 of [Neu94].

Quantity	Symbol	Expression
Support	$\text{supp}(\tau_\Delta)$	$\mathbb{N}_0$
Distribution	$\mathbb{P}(\tau_\Delta = k)$	$\begin{cases}\mu_\Delta & k=0\\ \mu_M S^{k-1}R^\top & k\geq 1\end{cases}$
Distribution function	$F_{\tau_\Delta}(t)$	$F_{\tau_\Delta}(t) = (1-\mu_M S^{\lfloor t\rfloor}\mathbf{1}^\top)\mathbb{1}_{\mathbb{R}^+}(t)$
n-th Factorial Moment	$\mathbb{E}\left(\frac{\tau_\Delta!}{(\tau_\Delta-n)!}\right)$	$n!\mu_M S^{n-1}(Id-S)^{-n}\mathbf{1}^\top$
Expectation	$\mathbb{E}(\tau_\Delta)$	$\mu_M(Id-S)^{-1}\mathbf{1}^\top$
Variance	$\text{Var}(\tau_\Delta)$	$\mu_M(Id+S)(Id-S)^{-2}\mathbf{1}^\top$ $-(\mu_M(Id-S)^{-1}\mathbf{1}^\top)^2$
Generating function	$G_{\tau_\Delta}(z)$	$\mu_\Delta + z\mu_M\left(Id-zS\right)^{-1}R^\top$

Table 2.1: Properties of $\tau_\Delta \sim PH(\mu_M, S)$ in the discrete time setting.

Quantity	Symbol	Expression
Support	$\text{supp}(\tau_\Delta)$	$\mathbb{R}^+$
Distribution	$\mathbb{P}(\tau_\Delta = .)$	$\mu_M \exp(St)R^\top dt + \mu_\Delta\delta_0$
Distribution function	$F_{\tau_\Delta}(t)$	$1-\mu_M\exp(St)\mathbf{1}^\top$
n-th Moments	$\mathbb{E}(\tau_\Delta^n)$	$\mu_M(-1)^n n!S^{-n}\mathbf{1}^\top$
Expectation	$\mathbb{E}(\tau_\Delta)$	$-\mu_M S^{-1}\mathbf{1}^\top$
Variance	$Var(\tau_\Delta)$	$2\mu_M S^{-2}\mathbf{1}^\top - (\mu_M S^{-1}\mathbf{1}^\top)^2$
moment generating function	$M_{\tau_\Delta}(s)$	$-\mu_M\left(sId+S\right)^{-1}R^\top + \mu_\Delta$
characteristic function	$\varphi_{\tau_\Delta}(s)$	$-\mu_M\left(isId+S\right)^{-1}R^\top + \mu_\Delta$

Table 2.2: Properties of $\tau_\Delta \sim PH(\mu_M, S)$ in the continuous time setting.

2.1.3 Typical examples of PH-type distributions

In this section we briefly discuss examples of common distribution seen as PH-type distributions. The most prominent and simple ones are the geometric distribution and the exponential distribution. Starting from these distributions other distributions can be constructed by defining suitable absorbing Markov Chains.

2.1.4 Examples for discrete distributions

2.1.4.1 The geometric distribution

A Bernoulli trial is a random experiment with two outcomes, "success" and "failure". The geometric distribution describes independently repeated Bernoulli trials until the first "success". Such a sequence of Bernoulli trials is often interpreted as repeated coin tossing, where the event "success" is (for example) "the coin showed heads after throwing".

This repeated experiment can be modeled by a MCd with a single absorbing state in the following way.

Suppose $E = \{\Delta\} \cup \{1\}$ and initial distribution $\mu = (0, 1)$ and the transition matrix

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

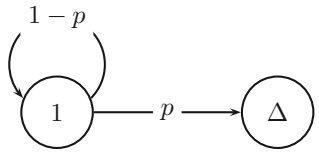


Figure 2.1: Transition graph for the geometric distribution as phase type distribution.

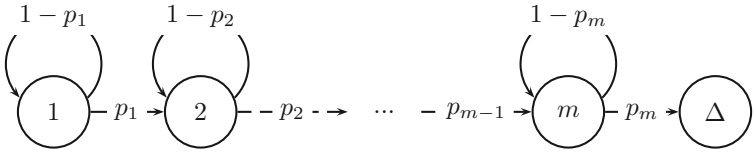


Figure 2.2: The sum of m geometric distributions modeled by a MCd.

We interpret here the state 1 as "experiment failed" and Δ as "experiment succeeded". The transition graph is given in fig. 2.1.

The n-th power of the transition matrix is

$$P^n = \begin{pmatrix} 1 & 0 \\ (1-p)^n & p\sum_{k=0}^{n-1}(1-p)^k \end{pmatrix}$$

and therefore by induction on n

$$\mathbb{P}(\tau_\Delta = n) = p(1-p)^{n-1}.$$

This shows that τ_Δ has geometric distribution to the parameter p, i.e.

$$\mathcal{L}(\tau_\Delta) = Geo(p).$$

2.1.4.2 Generalized geometric distributions

The geometric distribution is interpreted as the (random) number of repetitions until an experiment succeeds for the first time. A simple extension is to start another experiment and wait for its first success. This can be repeated m times, each time with possibly different coins (see the coin interpretation in the discussion of the geometric distribution). This distribution can be interpreted as a PH-type distribution in the following sense.

Let $E = \{\Delta\} \cup \{1, 2, \ldots, m\}$ and $(X_n)_{n\geq 0}$ be a MCd with transition matrix P. According to

the decomposition (2.1) we have

$$S = \begin{pmatrix} 1-p_1 & p_1 & 0 & \dots & 0 & 0 \\ 0 & 1-p_2 & p_2 & \dots & 0 & 0 \\ 0 & 0 & 1-p_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-p_{m-1} & p_{m-1} \\ 0 & 0 & 0 & \dots & 0 & 1-p_m \end{pmatrix}$$

and $R_\Delta = p_m e_m$.

It is an immediate consequence of the form of the transition graph given in fig. 2.1.4.2, that the absorption time τ_Δ is just the sum of m random variables, each geometric distributed with parameter p_i, if $\mu_M = \delta_1$. It is also immediate that exchanging two parameters p_j and p_i does not change the distribution of τ_Δ as convolution is commutative. That also shows that for a given distribution there might be different Markov Chains with the very same absorption time distribution.

2.1.5 Examples for continuous distributions

2.1.5.1 Exponential distribution

Consider a continuous time Markov chain on $E = \{\Delta\} \cup \{1\}$ with initial distribution $\mu = (0, 1)$ and infinitesimal generator

$$Q = \begin{pmatrix} 0 & 0 \\ w & -w \end{pmatrix}.$$

See also fig. 2.3 for the transition graph. This distribution is clearly the continuous time analogue of the geometric distribution, there is no (structural) difference regarding the transition graph.

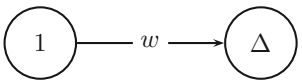


Figure 2.3: transition graph for the exponential distribution to the parameter w as phase type distribution.

Here $S = -w$ and $R_\Delta = w$. If $\mu_M = 1$, i.e. $\mu_\Delta = 0$, the moment generating function (see table 2.2) is

$$M_{\tau_\Delta}(s) = -\mu_M(sId + S)^{-1}R = (w - s)^{-1}w = \left(1 - \frac{s}{w}\right)^{-1}$$

and thus $\mathcal{L}(\tau_\Delta) = \mathcal{E}(w)$, where $\mathcal{E}(w)$ is the exponential distribution to the parameter w.

2.1.5.2 Erlang distribution

The Erlang distribution is just the sum of exponential distributions with the same parameter and models the time spend until the m-th exponentially distributed arrival in for instance a

queue. Sometimes the Erlang distribution $Erl(m, \lambda)$ is also denoted as gamma distribution $\Gamma(m, \lambda)$.

The Erlang distribution is modeled as a MCc on $E = \{\Delta\} \cup \{1, 2, \ldots, m\}$ and by (2.2)

$$S := \begin{pmatrix} -w & w & \ldots & 0 & 0 \\ 0 & -w & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -w & w \\ 0 & 0 & \ldots & 0 & -w \end{pmatrix}$$

with $R_\Delta = we_m$ and initial distribution $\mu = \delta_1$. The graphical representation is given in fig. 2.4.

Figure 2.4: Transition graph for the representation of an Erlang distribution via phase type distributions.

The density of τ_Δ with respect to Lebesgue measure is (after some algebraic manipulations):

$$f_{\tau_\Delta}(t) = \frac{w^m t^{m-1}}{(m-1)!} \exp(-wt).$$

The density $f_{\tau_\Delta}(t)$ is according to the picture the m-times convolution of an exponential distribution with parameter w.

2.1.5.3 Hypo exponential or generalized Erlang distribution

The hypo exponential distribution is a generalization of the Erlang distribution, instead of summing exponential distributions with the same parameter, different parameters for each summand can be allowed. For the model see fig. 2.5 and compare also with the generalized geometric distribution and the Erlang distribution. The Hypo exponential distribution is in this sense the direct continuous time analogue of the generalized geometric distribution.

According to the decomposition (2.2):

$$S = \begin{pmatrix} -w_1 & w_1 & 0 & \ldots & 0 & 0 \\ 0 & -w_2 & w_2 & \ldots & 0 & 0 \\ 0 & 0 & -w_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -w_{m-1} & w_{m-1} \\ 0 & 0 & 0 & \ldots & 0 & -w_m \end{pmatrix}.$$

and R_Δ is given by $w_m e_m$ and $\mu = \delta_1$.

Figure 2.5: Transition graph for the representation of a generalized Erlang distribution via phase type distributions.

As it is possible to exchange the rates between transitions without changing the absorption time, the transition rates between states in M can be arbitrarily interchanged. This shows (again) that a given representation of τ_Δ is only one out of several possible choices to model an absorption time.

2.1.5.4 Cox distribution

Let E be $E := \{\Delta\} \cup \{1, \ldots, m\}$ with Δ absorbing and initial distribution $\mu = \delta_1$. Furthermore

$$S = \begin{pmatrix} -w_1 & p_1 w_1 2 & 0 & \ldots & 0 & 0 \\ 0 & -w_2 & p_2 w_2 & \ldots & 0 & 0 \\ 0 & 0 & -w_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -w_{m-1} & p_{m-1} w_{m-1} \\ 0 & 0 & 0 & \ldots & 0 & -w_m \end{pmatrix},$$

with $p_i \in (0, 1]$. Then

$$R_\Delta^\top = ((1-p_1)w_1, (1-p_2)w_2, \ldots, (1-p_{m-1})w_{m-1}, w_m)^\top$$

according to the decomposition (2.2). The corresponding transition graph is displayed in fig. 2.6.

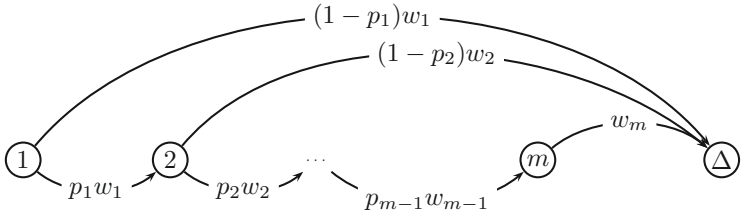


Figure 2.6: The transition graph for the phase type representation of the Cox distribution.

The idea behind the definition of a Cox distribution comes from queueing theory. Imagine that a certain type of service takes m steps to completion and each step takes an exponential

time. This is the framework of generalized Erlang distribution. The Cox distribution extends this idea in the sense that shortcuts are allowed, such that at each step i the service can be immediately finished. This picture directly leads to an expression for the density $f_{\tau_\Delta}(t)$ of the Cox distribution (remember that the density with respect to Lebesgue measure exists as there is no initial weight on Δ, see table 2.2):

$$f_{\tau_\Delta}(t) = \sum_{k=1}^{m}(1-p_k)f_{X_k}(t)\prod_{l=1}^{k-1}p_l f_{X_l}(t)$$

with the convention that $p_m = 0$ and $f_{X_i}(t)$ is the density of an exponential distribution with parameter w_i.

2.2 Extensions to several absorbing states

In the preceding section we introduced PH-type distributions as the distribution of the time until absorption in the single absorbing state Δ. A natural generalization is to allow several absorbing states, say $\Delta_1, \ldots, \Delta_k$. This does not change the time until absorption, i.e. the time until absorption is the same if the set of absorbing states would be replaced by a single absorbing state or vice versa. But another question makes this generalization more interesting. What is the distribution of the time until absorption in a specific absorbing state or a subset of absorbing states? Clearly at any fixed time t the condition to absorption in a specific state/set is a condition on the whole future of the process after t and does, as we will see, change the whole dynamics while the process remains to be markovian. This condition can be expressed by a so called *Doob-h-transform* and turns out to be a simple "rescaling" of the transition probabilities/rates. This rescaling can be interpreted as defining a new chain on the subset of all trajectories of the original Markov Chain ending in the designated absorbing state/s. First we precise the heuristic approach. Then we unveil the distribution of the time until absorption in the set of all absorbing states as a mixture of PH-type distributions, each describing the absorption in a different Δ_i. The mixing parameter for each distribution is exactly the probability to be absorbed in Δ_i before reaching any other absorbing state.

En passant we give a method how to condition the process on leaving the set of transient states via given states and thus to specific absorbing states. It is therefore a condition connected to the last time a state is visited.

We cover here only the time continuous case as it is needed for the following chapters. Discrete time analogs can be easily derived by following the lines of the proofs in continuous time and the uniformization method introduced in Definition 1.3.15, chapter 1.

2.2.1 Absorption in a given subset of absorbing states

Let

$$E := \Delta \cup M$$

for $\Delta := \{\Delta_1, \ldots, \Delta_k\}$ and $M := \{1, 2, \ldots, m\}$. We extend the order relation "¡" such that $\Delta_1 < \Delta_2 < \ldots < \Delta_k < 1 < \ldots$ and use this order as order for the sequence of states for the infinitesimal generator, i.e. the first row relates to the smallest element of E, the second to the second largest and so on.

Assume that M is a single communication class and all states in Δ are absorbing. If $k = 1$ we identify the *set* Δ with the *state* Δ; this is the case treated in the preceding section with a single absorbing state. Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$ and initial distribution μ. In analogy to (2.2), we decompose Q into a block matrix:

$$Q := \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ R_{\Delta_1}^\top & \dots & R_{\Delta_k}^\top & S \end{pmatrix} \tag{2.3}$$

where S is a $m \times m$ matrix with $S = (q_{ij})_{i,j \in M}$ and

$$R_{\Delta_i} = (q_{1\Delta_i}, \dots, q_{m\Delta_i})$$

the vector encoding the transitions to Δ_i. Furthermore we decompose the initial distribution $\mu = (\mu_\Delta, \mu_M)$, where

$$\mu_\Delta = (\mu_{\Delta_1}, \dots, \mu_{\Delta_k}),\ \mu_M = (\mu_1, \dots, \mu_m).$$

Let

$$\tau_{\Delta_i} := \inf\{t \geq 0 : X_t = \Delta_i\}$$

be the absorption time in Δ_i and for $A \subsetneq \Delta$ let

$$\tau_A := \inf\{t \geq 0 : X_t \in A\}$$

the time until absorption in A or the hitting time of A.

With $\tau_\Delta := \inf\{t \geq 0 : X_t \in \Delta\}$ it is immediate that

$$\tau_\Delta = \inf\{\tau_{\Delta_i}, \Delta_i \in \Delta\}.$$

Remark **2.2.1** It is a standard result that τ_Δ is almost surely finite in our setting. But it has to be pointed out, that the τ_{Δ_i} have a non vanishing probability to be infinite.

□

We continue with the determination of the infinitesimal generator $Q_{|A}$ of the process conditioned to absorption in a fixed set $A \subsetneq \Delta$ (the case $A = \Delta$ does not need to be considered as absorption in Δ is already almost sure). For the derivation of the infinitesimal generator of the conditioned process we need to calculate

$$Q_{|A} := \lim_{t \to 0} \frac{d}{dt} \mathbb{P}_i(X_t = j | \tau_A < \infty). \tag{2.4}$$

We adopt the calculations from [KS76], chapter 3, section 3, for continuous time:

Lemma 2.2.2 *Let $(X_t)_{t \geq 0}$ be a MCc on E with infinitesimal generator Q and initial distribution $\mu_M = e_i$, resp. $\mu = \delta_i$. Then for fixed $\Delta_l \in \Delta$:*

$$\mathbb{P}_i(X_t = j | \tau_{\Delta_l} < \infty) = \frac{h_j^l}{h_i^l} \mathbb{P}_i(X_t = j) \tag{2.5}$$

with $h^l = (h_j^l)_{j \in M} := (\mathbb{P}_j(\tau_{\Delta_l} < \infty))_{j \in M} = (-S)^{-1} R_{\Delta_l}^\top$.

Proof We use the definition of the conditional probability and the Markov property:

$$\begin{aligned}
&\mathbb{P}\left(X_t = j \,|X_0 = i, \tau_{\Delta_l} < \infty\right) \\
&= \mathbb{P}\left(X_t = j \,|X_0 = i, \exists s \geq 0 : X_{t+s} = \Delta_l\right) \\
&= \frac{\mathbb{P}(\exists s \geq 0 : X_{t+s} = \Delta_l | X_t = j)\mathbb{P}(X_t = j | X_0 = i)}{\mathbb{P}(\exists s \geq 0 : X_{t+s} = \Delta_l | X_0 = i)} \\
&= \frac{\mathbb{P}(\exists s \geq 0 : X_s = \Delta_l | X_0 = j)}{\mathbb{P}(\exists s' \geq 0 : X_{s'} = \Delta_l | X_0 = i)} \mathbb{P}(X_t = j | X_0 = i) \\
&= \frac{\mathbb{P}_j(\tau_{\Delta_l} < \infty)}{\mathbb{P}_i(\tau_{\Delta_l} < \infty)} \mathbb{P}(X_t = j | X_0 = i) \\
&= \frac{h_j^l}{h_i^l} \mathbb{P}(X_t = j | X_0 = i)
\end{aligned}$$

We further compute

$$h_i^l = e_i \int_0^\infty \exp(St) dt R_{\Delta_l}^\top = e_i S^{-1}[\mathbf{0} - Id] R_{\Delta_l}^\top \tag{2.6}$$

■

We note also that the h^l vectors form a stochastic matrix in the following sense:

Corollary 2.2.3 *In the setting of Lemma 2.2.2 the $m \times k$ matrix*

$$H := \begin{pmatrix} h_1^1 & \dots & h_1^k \\ \vdots & \ddots & \vdots \\ h_m^1 & \dots & h_m^k \end{pmatrix} = (-S)^{-1} \begin{pmatrix} R_{\Delta_1}^\top & \dots & R_{\Delta_k}^\top \end{pmatrix}$$

is stochastic and strict positive.

Proof Simple computation and Lemma 2.2.2 yield:

$$H\mathbf{1}^\top = \sum_{\Delta_i \in \Delta} (-S)^{-1} R_{\Delta_i}^\top = \mathbf{1}^\top$$

by generator property and invertibility of S, i.e.

$$S\mathbf{1}^\top + \sum_{\Delta_i \in \Delta} R_{\Delta_i}^\top = \mathbf{0}$$

holds. The strict positivity of H follows by the assumption that M is a single communication class and all states in Δ can be reached.

■

The condition to absorption in a set $A \subsetneq \Delta$ is a simple corollary of Lemma 2.2.2.

Corollary 2.2.4 *Let $(X_t)_{t \geq 0}$ be a MCc on E with infinitesimal generator Q with decomposition as in (2.3) and initial distribution $\mu_M = e_i$. Then for fixed $A \subsetneq \Delta$:*

$$\mathbb{P}_i(X_t = j | \tau_A < \infty) = \frac{h_j^A}{h_i^A} \mathbb{P}_i(X_t = j) \tag{2.7}$$

with $h^A = (h_j^A)_{j \in M} := (\mathbb{P}_j(\tau_A < \infty))_{j \in M} = \sum_{\Delta_l \in A} h^l$. Furthermore $h^\Delta = \mathbf{1}^\top$.

Proof We first note that the event $\{\tau_A < \infty\}$ can be decomposed into disjoint events:

$$\{\tau_A < \infty\} = \bigsqcup_{\Delta_i \in A} \{\tau_{\Delta_i} < \infty\},$$

as $\{\tau_{\Delta_i} < \infty\} = \{\tau_{\Delta_i} < \infty\} \cup \{\tau_{\Delta_k} = \infty : \Delta_k \in \Delta, k \neq i\}$. Analogue to the proof of Lemma 2.2.2

$$\begin{aligned}\mathbb{P}_i(X_t = j | \tau_A < \infty) &= \frac{\mathbb{P}_j(\tau_A < \infty)}{\mathbb{P}_i(\tau_A < \infty)} \mathbb{P}_i(X_t = j) \\ &= \frac{\sum_{\Delta_i \in A} \mathbb{P}_j(\tau_{\Delta_i} < \infty)}{\sum_{\Delta_i \in A} \mathbb{P}_i(\tau_{\Delta_i} < \infty)} \mathbb{P}_i(X_t = j).\end{aligned}$$

The rest follows by (2.6). If $A = \Delta$ then obviously the condition has no influence and absorption in Δ is almost sure, thus $h^A = h^\Delta = \mathbf{1}^\top$. This can also be seen from the definition of h^A, because

$$S\mathbf{1}^\top + \sum_{\Delta_j \in \Delta} R_{\Delta_j}^\top = 0$$

by generator property of Q.

■

Lemma 2.2.2 and Corollary 2.2.4 give the transition function of the conditioned processes. In order to calculate $Q_{|\Delta_i}$ or $Q_{|A}$ the infinitesimal generators of the process conditioned to absorption in Δ_i resp. in $A \subsetneq \Delta$ we gain the following result.

Proposition 2.2.5 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator Q and decomposition (2.3). Furthermore let $A \subsetneq \Delta$ be a set of absorbing states. Then the infinitesimal generator of $(X_t)_{t\geq 0}$ conditioned to absorption in A is*

$$Q_{|A} =: \begin{pmatrix} 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ R_{\Delta_1|A}^\top & \dots & R_{\Delta_k|A}^\top & S_{|A} \end{pmatrix}, \tag{2.8}$$

where

$$\begin{aligned} R_{\Delta_i|A}^\top &:= \mathbb{1}_{\Delta_i \in A} \mathit{diag}(h^A)^{-1} R_{\Delta_i}^\top \\ S_{|A} &:= \mathit{diag}(h^A)^{-1} S \mathit{diag}(h^A) \end{aligned}$$

and h^A is defined as in Corollary 2.2.4.

Proof Equation (2.4) shows by Lemma 2.2.4

$$\lim_{t\to 0} \frac{d}{dt} \mathbb{P}_i(X_t = j | \tau_A < \infty) = \frac{h_j^A}{h_i^A} \lim_{t\to 0} \frac{d}{dt} \mathbb{P}_i(X_t = j) = \frac{h_j^A}{h_i^A} q_{ij}$$

by definition of the infinitesimal generator of the original unconditioned process. Equation (2.8) is then the matrix version of this calculation.

■

Remark **2.2.6** The assumption that M forms a single communication class is necessary for the invertibility of $\mathrm{diag}(h^i)$, see the section about the role of pseudo inverses.

□

Example **2.2.7** We give an example how h^A transforms Q and how this relates to the transition graph. Consider a $\mathrm{MCc}(X_t)_{t\geq 0}$ on the state space $E = \{\Delta_1, \Delta_2\} \cup \{1, 2\}$ defined by the transition graph given in fig. 2.7 and with infinitesimal generator

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & -2 \end{pmatrix}.$$

This process is a simple version of a Birth-and-Death Process with absorbing boundaries.

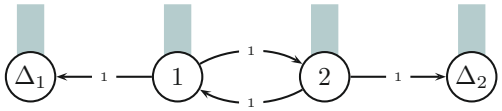


Figure 2.7: An example of a MCc with several absorbing states. The bars indicate the (pointwise) values of $(h_i^{\Delta})^{-1}$ (not to scale).

According to the decomposition (2.2)

$$R_{\Delta_1}^{\top} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad R_{\Delta_2}^{\top} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad S = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}.$$

We fix $A = \{\Delta_1\}$ as absorbing set. According to Lemma 2.2.2

$$\begin{aligned} h^A = h^1 &= (-S)^{-1} R_{\Delta_1}^{\top} \\ &= \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix} \end{aligned}$$

and thus

$$Q_{|A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{3}{2} & 0 & -2 & \frac{1}{2} \\ 0 & 0 & 2 & -2 \end{pmatrix}.$$

We used equal transition rates for all states to show that the condition to absorption in Δ_1 introduces a "drift" which forces the dynamic away from Δ_2. Indeed one could interpret

$$((h_{\Delta_1}^A)^{-1}, (h_1^A)^{-1}, (h_2^A)^{-1}, (h_{\Delta_2}^A)^{-1}) = (1, \frac{3}{2}, 3, \infty)$$

as a potential on E forcing the dynamic to reach the minimum of the potential, which is attained in Δ_1.

We also observe that $\mathrm{diag}(h^A)$ is a projection from $E = \{1, 2\} \cup \{\Delta_1, \Delta_2\}$ onto $\{1, 2\} \cup \{\Delta_1\}$, thus the "connection" to Δ_2 is literally cut off; the transition graph of the conditioned process in fig. 2.8, amplifies this interpretation.

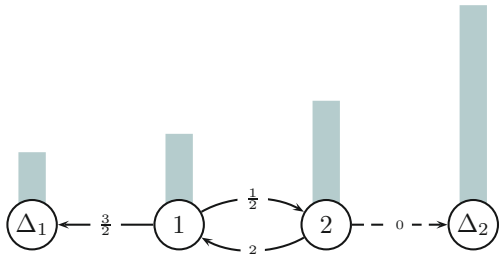


Figure 2.8: The model from fig. 2.7 after the conditioning. Note that Δ_2 is now disconnected from the system. The bars indicate the (point-wise) values of $(h_i^{\Delta_1})^{-1}$ (not to scale).

□
□

The results of Lemma 2.2.2 and Corollary 2.2.4 fall into a class of transformations called *h-transform* or depending on the author *Doob-h-transform*. Doob introduced an early version in []. In case of countable state space this is treated in [KSK66] chapter 8 and [SS67] and has relations to potential theory for Markov Chains.

2.2.2 Absorption law as mixture of PH-type distributions

In this chapter we show that indeed the time to absorption in a set of several absorbing states is a mixture of absorption times in single absorbing states. Each of these absorption times describes the time until absorption in one specific absorbing state.

The distribution of the time to absorption in the set Δ can be described as PH-type distribution with representation $PH(\mu_M, S)$, when the set Δ is identified with the *single* state Δ. A direct consequence of this observation is the following lemma.

Proposition 2.2.8 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator Q and decomposition (2.3). Then*

$$\mathbb{P}_\mu(\tau_\Delta \in dt) = \mu_\Delta \delta_0 + \sum_{\Delta_l \in \Delta} \mathbb{P}_{\mu'}(\tau_{\Delta_l} \in dt)$$

for $h^l = (-S)^{-1} R_{\Delta_l}^\top$ and $\mu' := \mu_M \, diag(h^l)$.

Proof As M is a single communication class every absorbing state can be reached from every state in M with positive probability. Thus by table 2.2, Proposition 2.2.5 and Corollary 2.2.3

$$\begin{aligned}
\mathbb{P}_\mu(\tau_\Delta \in dt) &= \mu_\Delta \delta_0 + \mu_M \exp(St) R_\Delta^\top dt \\
&= \mu_\Delta \delta_0 + \sum_{\Delta_l \in \Delta} \mu_M \mathrm{diag}(h^l) \mathrm{diag}(h^l)^{-1} \exp(St) \mathrm{diag}(h^l) \mathrm{diag}(h^l)^{-1} R_{\Delta_l}^\top dt \\
&= \mu_\Delta \delta_0 + \sum_{\Delta_l \in \Delta} \mu_M \mathrm{diag}(h^l) \exp(S_{|\Delta_l}) R_{|\Delta_l}^\top dt.
\end{aligned}$$

As $\mu_M \mathrm{diag}(h^l)$ is still a (sub) stochastic vector on M the result follows.

■

Proposition 2.2.8 shows that the absorption in the set Δ of absorbing states can be indeed described as a mixture of the times to absorption in each Δ_l. In particular if $\mu_M = \delta_i$ then the lemma states

$$\mathbb{P}_i(\tau_\Delta \in dt) = \sum_{\Delta_l \in \Delta} e_i(-S)^{-1} R_{\Delta_l}^\top \mathbb{P}_i(\tau_{\Delta_l} \in dt) = \sum_{\Delta_l \in \Delta} h_i^l \mathbb{P}_i(\tau_{\Delta_l} \in dt),$$

see (2.6), and thus is a convex combination of the densities of the absorption times in the absorbing states Δ_l. The property $\sum_{\Delta_l \in \Delta} h_i^l = 1$ is ensured by Corollary 2.2.3, as $(h_i^l)_{\Delta_l \in \Delta}$ is itself a probability distribution on Δ for each $i \in M$.

The main idea of Proposition 2.2.8 was to use the property

$$R_\Delta = \sum_{\Delta_l \in \Delta} R_{\Delta_l}$$

to show that the absorption in Δ is a mixture. This decomposition yields more information. We need definitions to understand how the mixture property relates to conditions on leaving M via fixed states.

Definition 2.2.9 *Let $(X_t)_{t\geq 0}$ be MCc on E with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Then the set of states*

$$\partial M := \{i \in M | \exists l : q_{i\Delta_l} > 0\}$$

is called boundary of M.

Definition 2.2.10 *Let $(X_t)_{t\geq 0}$ be MCc on E with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Then the set*

$$\rhd_l := \{i \in \partial M : q_{i\Delta_l} > 0\}$$

is called the gate to Δ_l *or* Δ_l-gate. *Accordingly for any $A \subset \Delta$:*

$$\rhd_A := \bigcup_{\Delta_l \in A} \rhd_l$$

is the gate to A. If there is only one absorbing state we write just $\rhd$.

It is clear that the Δ_l-gate is just the set of states from which absorption in $\Delta_l \in \Delta$ takes place. Thus the process is absorbed in Δ_l only via the Δ_l-gate. In other words, h_i^l is the probability to start in i and leave the system via $\rhd_l$.

If we assume that Δ is a singleton, we can replace Δ by $|\partial M|$ absorbing states, such that for every $i \in \partial M$ there is a unique Δ_i, see fig. 2.9. This procedure does not change the time to absorption. The boundary ∂M is then a disjoint union gates with cardinality one. Therefore the probability to leave M via a specific boundary state $b \in \partial M$ is just

$$\mu_M(-S)^{-1} e_b^\top (e_i R_\Delta^\top),$$

the probability to be absorbed in Δ_b.

If the set of absorbing states is not a singleton, we can just replace this set by a single state Δ and apply the same method as before.

In fact we have already shown:

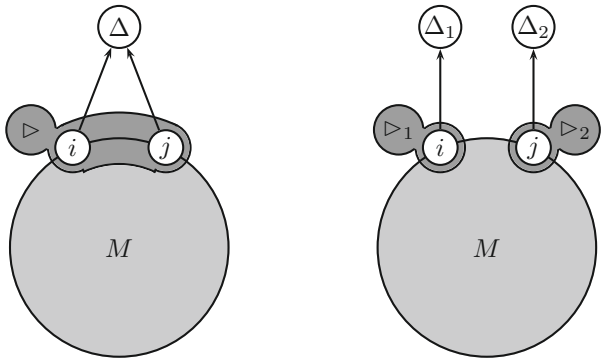


Figure 2.9: Introducing an absorbing state for each state of the gate.

Theorem 2.2.11 *Let $(X_t)_{t\geq 0}$ be a MCc on $E = \Delta \cup M$ where Δ is the only absorbing state and (2.2) decomposition of its infinitesimal generator Q. Let the initial condition be $\mu = (\mu_\Delta, \mu_M)$ and $B \subsetneq \delta M$ an arbitrary set of boundary states of M. Furthermore define R_B by*

$$\forall i \in M : e_i R_B^\top := \begin{cases} e_i R_\Delta^\top & i \in B \\ 0 & else \end{cases}.$$

For $h^B = (-S)^{-1} R_B^\top$ the MCc associated to

$$Q_{|B} := \begin{pmatrix} 0 & 0 \\ diag(h^B)^{-1} R_\Delta^\top & diag(h^B)^{-1} S\, diag(h^B) \end{pmatrix}$$

and $\mu_{|B} := (\mu_\Delta, \mu_M\, diag(h^B))$ is the original chain conditioned on leaving M via B.

2.2.3 Role of Pseudo Inverses and conditional absorbing states

We have made throughout the last section the assumption that M is a single communication class. This condition is convenient as shown in following example. We need the so called Pseudo Inverse, as defined for example in [BIG03], chapter 1. By this notion it is possible to find matrices for singular or even rectangular matrices that behave very much like the inverse for quadratic non-singular matrices. Omitting the details of their definition and their properties, we want to draw attention to problems (and their possible solution) that arise if the single communication class assumptions fails. The following example illustrates that.

Example **2.2.12** Let $(X_t)_{t\geq 0}$ be a MCc on $E = \{\Delta_1, \Delta_2\} \cup \{1, 2, 3\}$ with initial condition μ and infinitesimal generator

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 1 \\ 1 & 0 & 1 & -2 & 0 \\ 0 & 1 & 0 & 0 & -1 \end{pmatrix}.$$

The transition graph is given in figure 2.10.

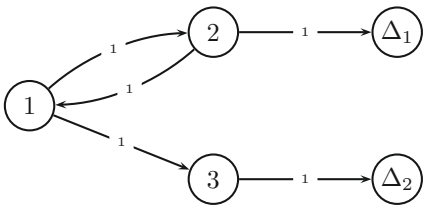


Figure 2.10: Transition graph where $M = \{1, 2, 3\}$ consists of the two communication classes $\{1, 2\}$ and $\{3\}$.

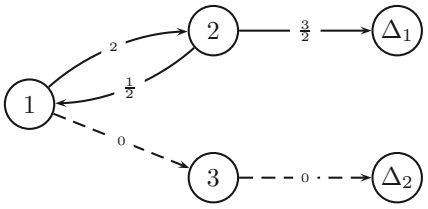


Figure 2.11: Transition after conditioning on absorption in Δ_1

Furthermore

$$R_{\Delta_1} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}^\top, \; R_{\Delta_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}^\top, \; S = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Heuristically the condition on absorption in Δ_1 or equivalently leaving M via state 2 turns state 3 into an absorbing state, as 3 can only be left to Δ_2. Computation of h^1 yields

$$h^1 = (-S)^{-1} R_{\Delta_1}^\top = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3}(1, 2, 0)^\top.$$

Obviously $\mathrm{diag}(h^1)$ can not be inverted in the usual sense, but taking the pseudo inverse fixes that. Indeed the dynamic conditioned to absorption in Δ_1 has according to Proposition 2.2.5 the form

$$Q_{|\Delta_1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 2 & 0 \\ \frac{3}{2} & 0 & \frac{1}{2} & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 2.11 shows the new transition graph. □

□

Chapter 3

Time Duality in Markov Chains

In the following we introduce a new relation called *Time Duality* between states of a MCc with irreducible infinitesimal generator. In a simpler setting and without rigorous proof this appeared already in [LW07, VLL08]. The concept concerns the equality in distribution of the "passage time without return" from i to j and from j to i. We characterize this property in terms of Laplace transforms of phase type distributions and gain several sufficient conditions for Time Duality, partially based on the structure of the transition graph. One of these sufficient conditions is permuted balance, see section 3.1.2, under some constraints on the neighbourhood of the chosen states.

We also gaze the interesting question whether the Time Duality relation is an equivalence relation for Birth-and-Death Processes with reflecting boundaries and for reversible Markov Chains on trees.

3.1 Definitions and preliminaries

Let $(X_t)_{t\geq 0}$ be a MCc on a finite state space E with irreducible infinitesimal generator and initial distribution in $i \in E$.

For an arbitrary state $j \neq i$ the stopping time

$$\tau_{ij} := \inf\{t : X_t = j\}$$

is known as *first passage time* from i to j. We define another random time

$$\sigma_{ij} := \sup\{t < \tau_{ij} : X_t = i\}.$$

It is the time when i has been left the last time before reaching j. Naturally σ_{ij} can not be a stopping time, as σ_{ij} depends on the future development of X_t.

Now we can define the pure passage time in the following way:

Definition 3.1.1 *We call the difference*

$$\tau^*_{ij} := \tau_{ij} - \sigma_{ij}$$

the pure passage time *from i to j, see also fig. 3.1.*

Due to a trick we can describe τ^*_{ij} as a PH-type distribution. For fixed states j in E, turn j into an absorbing state by cutting off all outgoing transitions and replace i by the two states

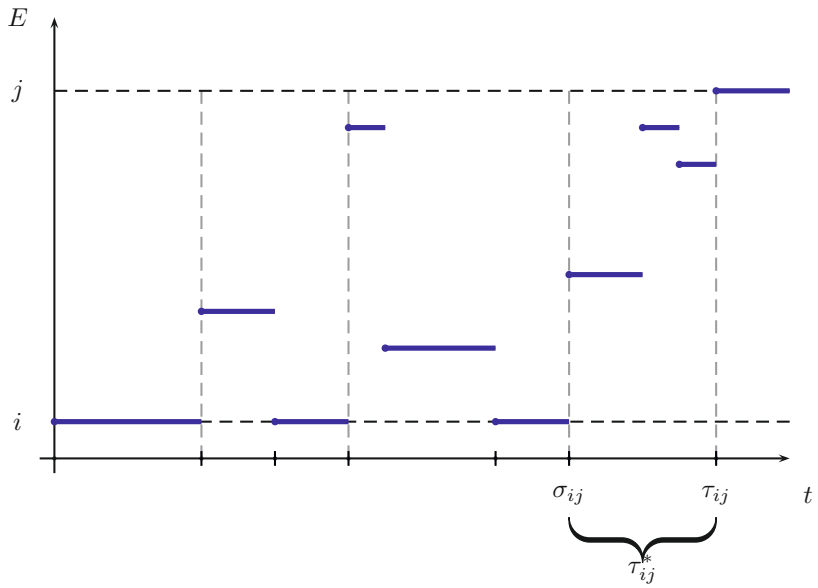


Figure 3.1: Example trajectory of a passage from i to j.

i_{in} and i_{out}, where $q_{ki_{in}} = q_{ki}$ but $q_{i_{in}k} = 0$ for all $k \in E \setminus \{i\}$. Thus i_{in} inherits only the "ingoing" transitions and is therefore absorbing. The other state i_{out} is such that $q_{ik} = q_{i_{out}k}$ but $q_{ki_{out}} = 0$ for all $k \in E \setminus \{i\}$, i.e. i_{out} inherits only the "outgoing" transitions. See also fig. 3.2.

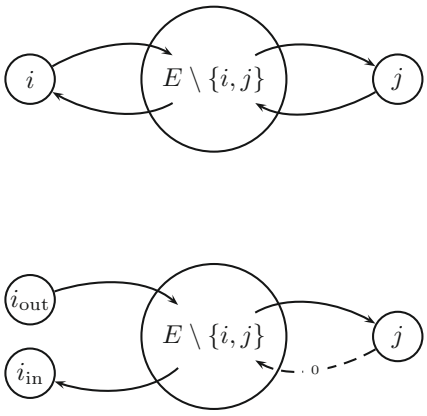


Figure 3.2: Splitting method to characterize the pure passage time τ^*_{ij}.

If we now can compute an initial distribution on the direct neighbours of i, the time to absorption in j under this new distribution has then the same distribution as τ^*_{ij}.

As there might be some states in E which can be reached in a single step from i, but are not connected with j by a path of strict positive probability avoiding i, we need to refine the notion of neighbourhood.

Definition 3.1.2 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Then for a state $i \in E$ the set*

$$\mathcal{N}_i := \{j \in E : q_{ij} > 0\}$$

is called the neighbourhood *of i.*

Note that Definition 3.1.2 is very similar but not equivalent to the definition of *gates* given in Definition 2.2.10.

Definition 3.1.3 *Let $(X_t)_{t\geq 0}$ be a MCc on E with infinitesimal generator $Q = (q_{kl})_{k,l\in E}$. For $i, j \in E$ let the set $\mathcal{N}_{i\to j}$ be the set of states in the neighbourhood of i that are connected with j via at least one path of non-vanishing weight with states in $E \setminus \{i\}$, i.e.*

$$\mathcal{N}_{i\to j} := \{k \in \mathcal{N}_i : \exists n \geq 1 \exists i_1, i_2, \dots i_n \in E \setminus \{i, j\} : q_{ki_1} q_{i_1 i_2} \dots q_{i_n j} > 0\}.$$

We say $\mathcal{N}_{i\to j}$ is the connected component of the neighbourhood of i with respect to j *or, if the context is clear,* the connected neighbourhood of i, *see fig. 3.3. We call $\mathcal{N}_{i\to j}$* simple *if it is a singleton and write n_{ij} instead of $\mathcal{N}_{i\to j}$.*

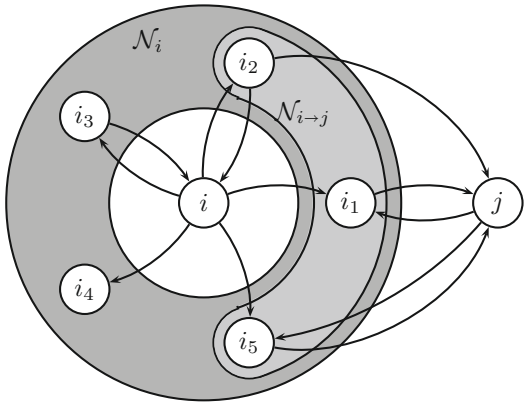


Figure 3.3: Neighbourhood of i (dark gray) and connected neighbourhood of i (light gray) with respect to j.

With these definitions the pure passage time between two states is characterized as PH-type distribution by the following proposition.

Proposition 3.1.4 *Let $(X_t)_{t\geq 0}$ be a recurrent MCc on E with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Fix $k, l \in E$, $k \neq l$. Let the initial condition be $\mu = \delta_k$ and define the sub matrix $S := (q_{ij})_{i,j\in E\setminus\{k,l\}}$ of Q and $R_{\Delta_s} := (q_{is})_{i\in E\setminus\{k,l\}}$.*

Then the first pure passage time τ^*_{kl} from k to l is of PH-type with representation $\sim PH(\nu_r, S_{|l})$ where

$$\nu_k := \frac{1}{\sum_{i\in\mathcal{N}_{k\to l}} q_{ki}} \begin{cases} q_{ki} & \text{if } i \in \mathcal{N}_{k\to l} \\ 0 & \text{else} \end{cases}$$

and

$$S_{|l} := diag(h^l)^{-1} S diag(h^l),$$

with $h^l := (-S)^{-1} R^\top_{\Delta_l}$.

Proof The procedure given in the introduction of this chapter proofs the statement by chapter 2, Proposition 2.2.5 (the set Δ of absorbing states would contain i_{in} and j). ■

We now compare the distributions of the pure passage times τ^*_{ij} and τ^*_{ji}. Intuitively only in very symmetric models $\tau^*_{ij} \stackrel{d}{=} \tau^*_{ji}$ should hold. The following example shows that this intuition is wrong.

Example 3.1.5
Consider $(X_t)_{t\geq 0}$ the MCc on $E = \{1, 2, 3, 4\}$ with infinitesimal generator Q given by

$$Q = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & w & -1-w & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

with $w > 1$.

The associated transition graph, see fig. 3.4, shows that this example is a simple Birth-and-Death Process with reflecting boundaries. Intuitively the passage from 1 to 4 should be slower than the passage from 4 to 1 for $w > 1$. Surprisingly the pure passage times have the same distribution for *all* $w > 0$, as we will now prove.

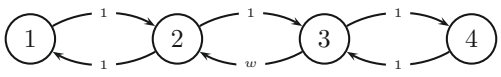


Figure 3.4: Transition graph for Example 3.5.

By Proposition 3.1.4 the pure passage time τ^*_{14} resp. τ^*_{41} has PH-type distribution with representation $PH(\delta_2, S_{|4})$ resp. $PH(\delta_3, S_{|1})$, with

$$S_{|4} = \text{diag}(h^4)^{-1} S \text{diag}(h^4)$$

resp.

$$S_{|1} = \mathrm{diag}(h^1)^{-1} S \mathrm{diag}(h^1),$$

where

$$S = (q_{ij})_{i,j\in\{2,3\}} = \begin{pmatrix} -2 & 1 \\ w & -1-w \end{pmatrix}$$

and

$$h^1 = (-S)^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; h^4 = (-S)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We have further

$$(-S)^{-1} = \frac{1}{2+w} \begin{pmatrix} 1+w & 1 \\ w & 2 \end{pmatrix}$$

and therefore

$$S_{|1} = \begin{pmatrix} -2 & \frac{w}{1+w} \\ 1+w & -1-w \end{pmatrix}, \; S_{|4} = \begin{pmatrix} -2 & 2 \\ \frac{w}{2} & -1-w \end{pmatrix}.$$

Apparently $S_{|1} \neq S_{|4}$. To show

$$\tau_{14}^* \overset{d}{=} \tau_{41}^*$$

we utilize the moment generating function given in table 2.2. We gain

$$\begin{aligned} M_{\tau_{14}^*}(s) &= -(1,0)(sId + S_{|4})^{-1} R_{|4}^\top \\ &= -(1,0) \left(sId + \begin{pmatrix} -2 & 2 \\ \frac{w}{2} & -1-w \end{pmatrix} \right)^{-1} \begin{pmatrix} 0 \\ 1+\frac{w}{2} \end{pmatrix} \\ &= \frac{2+w}{2-3s+s^2+w-sw} \\ &= -(0,1) \left(sId + \begin{pmatrix} -2 & \frac{w}{1+w} \\ 1+w & -1-w \end{pmatrix} \right)^{-1} \begin{pmatrix} 2-\frac{w}{1+w} \\ 0 \end{pmatrix} \\ &= -(0,1)(sId + S_{|1})^{-1} R_{|1}^\top \\ &= M_{\tau_{41}^*}(s) \end{aligned}$$

We note further that

$$\begin{aligned} M_{\tau_{14}^*}(s) = M_{\tau_{41}^*}(s) &= \frac{2+w}{s^2 - s(w+3) + 2 + w} \\ &= \frac{2+w}{(1-s)(2+w-s)} = (1-s)^{-1} \left(1 - \frac{s}{2+w}\right)^{-1} = M_{X+Y}(s) \end{aligned}$$

where X and Y are independent exponentially distributed random variables to the parameters 1 and $2+w$. Indeed, the parameters are the eigenvalues of $-S$ and also of $-S_{|4}, -S_{|1}$. This is a general property of Birth-and-Death Processes, see e.g. [DM09] for a probabilistic proof and [Kei79] for an early proof based on Laplace transforms. □

□

The preceding example clearly shows that $\tau_{ij}^* \overset{d}{=} \tau_{ji}^*$ is difficult to understand intuitively. We thus give explicit sufficient conditions for the equality (in distribution) of the pure passage times. The permuted balance (and thus also reversibility), defined in chapter 1, (1.5) implies Time Duality between two states under certain assumptions.

Definition 3.1.6 *Let $(X_t)_{t\geq 0}$ be a recurrent MCc on the finite state space E. Fix $i, j \in E$. If*

$$\tau_{ij}^* \stackrel{d}{=} \tau_{ji}^*$$

we say i and j are in Time Duality.

Some previous work of Kijima on such a problem has been worked out under certain symmetry properties on the infinitesimal generator, see Propositions 4.1 and 4.2. in [Kij88]. His main assumptions for Time Duality are reversibility and (independently) a symmetry on the infinitesimal generator Q which states:

$$Q = K^{-1}QK,$$

where K is a special permutation matrix sometimes called *anti identity*:

$$K = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \dots & 0 & 0 \end{pmatrix}.$$

This last "invariance under permutation" is indeed covered by our definition of permuted balance.

We show in the following sections that the notion of permuted balance includes and extends the symmetry conditions given by Kijima and give other conditions under which Time Duality between two states hold.

Examples show that permuted balance does not exploit all cases in which Time Duality appears. We therefore give several results on how to connect the geometry of the transition graph with Time Duality in more general situations.

3.2 Sufficient conditions for Time Duality in case of simple neighbourhood

In this first section we study the pure passage time between two states whose connected neighbourhoods are singletons. This not only allows to characterize Time Duality straightforward via Proposition 3.1.4 and the moment generating function of a PH-type distribution, but also delivers the basis for the generalization to connected neighbourhoods which are no more restricted to singletons.

3.2.1 A Condition based on moment generating functions

Let i, j be two states of the state space E and assume that $|\mathcal{N}_{i\to j}| = |\mathcal{N}_{j\to i}| = 1$. In this case we identify the singleton $\mathcal{N}_{i\to j}$ resp. $\mathcal{N}_{j\to i}$ with n_{ij} resp. n_{ji}, according to Definition 3.1.3. This also means that a passage from i to j has to pass through n_{ji} and analogue for the passage from j to i. Thus by Proposition 3.1.4 and table 2.2 the moment generating function $M_{ij}(s)$ resp. $M_{ji}(s)$ of the first pure passage time τ_{ij}^* resp. τ_{ij}^* is given as

$$M_{ij}(u) = -e_{n_{ij}}(uId + S_{|j})^{-1}R_{|\Delta_j}^\top = \frac{e_{n_{ij}}(uId + S)^{-1}e_{n_{ji}}^\top}{e_{n_{ij}}S^{-1}e_{n_{ji}}^\top}$$

resp.

$$M_{ji}(s) = -e_{n_{ji}}(uId + S_{|i})^{-1}R^{\top}_{|\Delta_i} = \frac{e_{n_{ji}}(uId+S)^{-1}e^{\top}_{n_{ij}}}{e_{n_{ji}}S^{-1}e^{\top}_{n_{ij}}}.$$

Obviously Time Duality between i and j holds if the moment generating functions of τ^*_{ij} and τ^*_{ji} are equal. This leads to the following criterion for Time Duality in the case of simple neighbourhoods of i and j.

Theorem 3.2.1 *Let $(X_t)_{t\geq 0}$ be an irreducible MCc on a finite state space E, with infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Suppose that for $i, j \in E$ the connected neighbourhoods are both singletons and denote the connected neighbourhood of i with n_{ij} and the one of j with n_{ji} according to Definition 3.1.3. Then Time Duality between i and j holds if and only if*

$$\forall u \leq 0 : \frac{R_{ij}(u)}{R_{ji}(u)} = \frac{R_{ij}(0)}{R_{ji}(0)} \tag{3.1}$$

holds, where $R_{ij}(u) := e_{n_{ij}}(uId + S)^{-1}e^{\top}_{n_{ji}}$.

Proof The statement is just an equivalent formulation of the fact that two random variables have the same distribution if their moment generating function are equal.

■

Remark **3.2.2** In condition 3.1 there is actually no restriction on u as long as it does not equal an eigenvalue of $-S$.

□

3.2.2 Time Duality via permuted balance

If the infinitesimal generator of a MCc is in permuted balance for some permutation σ with i and j fixed points Time Duality between i and j holds if their neighbourhoods are simple.

Theorem 3.2.3 *Let $(X_t)_{t\geq 0}$ be an irreducible MCc on E with infinitesimal generator Q in permuted balance with respect to some permutation σ. For any two distinct states $i, j \in E$ such that i, j are fixed points of σ and their neighbourhoods are simple, Time Duality holds.*

Proof The invariant distribution π is unique due to irreducibility. Permuted balance is then equivalent to the matrix equation

$$Q^{\top} = (\mathrm{diag}(\pi)P_\sigma)^{-1}Q\mathrm{diag}(\pi)P_\sigma,$$

see chapter 2. This also holds for the matrix $S = (q_{kl})_{k,l\in E\setminus\{i,j\}}$ (with an obvious restriction of π and P_σ to the states encoded in S).

We compute for every $u \leq 0$:

$$\begin{aligned} R_{ij}(u) &= e_{n_{ij}}(uId+S)^{-1}e^{\top}_{n_{ji}} = e_{n_{ji}}(uId+S^{\top})^{-1}e^{\top}_{n_{ij}} \\ &= e_{n_{ji}}(\mathrm{diag}(\pi)P_\sigma)^{-1}(uId+S)^{-1}\mathrm{diag}(\pi)P_\sigma e^{\top}_{n_{ij}} \\ &= \frac{\pi_{n_{ij}}}{\pi_{n_{ji}}}R_{ji}(u), \end{aligned}$$

thus

$$\forall u \leq 0 : \frac{R_{ij}(u)}{R_{ji}(u)} = \frac{\pi_{n_{ij}}}{\pi_{n_{ji}}}.$$

The condition of Theorem 3.2.1 is therefore fulfilled and i and j are in Time Duality.

■

The following example shows that the condition that i and j are fixed points to have Time Duality between them is necessary for the setting as in Theorem 3.2.3.

Example **3.2.4** Let $(X_t)_{t\geq 0}$ be a MCc on $E = \{1, 2, 3, 4, 5, 6, 7\}$, defined by the transition graph given in fig. 3.5 and with infinitesimal generator

$$Q := \begin{pmatrix} -4 & 2 & 1 & 1 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & -3 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -3 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & -4 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & -4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -2 \end{pmatrix}.$$

We note that the model is not reversible by the Kolmogoroff criterion, see Theorem 1.3.1, as

$$2 = q_{12}q_{25}q_{54}q_{41} \neq q_{14}q_{45}q_{52}q_{21} = 1.$$

But permuted balance holds, e.g. for $\sigma = (34)$.

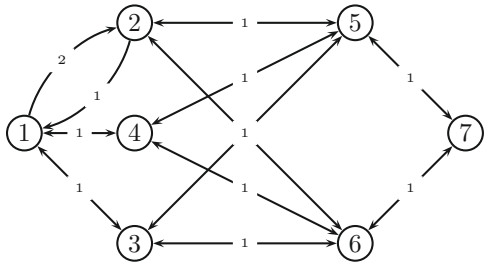


Figure 3.5: All transition rates are 1 except otherwise noted.

Without giving the exact calculations, we find that between 1 and 7 occurs Time Duality. Time Duality between 3 and 7 does not hold, as

$$\frac{\nu_3 \text{diag}(h^7)^{-1}(S + uId)^{-1}R_{\Delta_7}^\top}{\nu_7 \text{diag}(h^3)^{-1}(S + uId)^{-1}R_{\Delta_3}^\top} = \frac{11}{27}\frac{27 - 14u + 2u^2}{11 - 7u + u^2}$$

which is not constant. □

□

3.2.3 Further results based on the structure of the transition graph

In this section we give further results on Time Duality between two states i and j as consequence of a particular geometry of the transition graph. To simplify the argumentation we shall assume that with $E\setminus\{i,j\}$ is actually the communication class(es) of n_{ij} and n_{ji}. Note that the number of communication classes in $E\setminus\{i,j\}$ can be greater than one depending on i and j, while there is by assumption only one communication class in E.

To shorten notation we also fix for any square matrix K the following convention:

$$K_u := K + uId.$$

The basic idea of all following proofs in this section is to use the generalized Frobenius formula given in Appendix A. It allows to compute certain parts of the inverse of a matrix, if the original matrix is in block form. The formula is especially useful if one of the blocks is 1×1, because in this case the appearing Schur complements are real numbers. If there is no 1×1 block then the complements are again matrices and non-commutativity matrices comes into play which makes some calculations more difficult.

Theorem 3.2.5 *Let $(X_t)_{t\geq 0}$ be a MCc on a E with irreducible infinitesimal generator $Q=(q_{ij})_{i,j\in E}$. Let $i,j\in E$ be fixed and with simple connected neighbourhood. Let $\hat{S}=(q_{kl})_{k,l\in E\setminus\{n_{ij},i,j\}}$ and for arbitrary states r,s:*

$$\hat{R}_{rs}(u) := e_r\hat{S}_u^{-1}e_s^\top.$$

If the neighbourhood of n_{ij} is of cardinality two, i.e. $\mathcal{N}_{n_{ij}}=\{i,x\}$, see also fig. 3.6, Criterion (3.1) for Time Duality between i and j reduces to:

$$\forall u\leq 0: \frac{\hat{R}_{xj}(u)}{\hat{R}_{jx}(u)} = \frac{\hat{R}_{xj}(0)}{\hat{R}_{jx}(0)}.$$

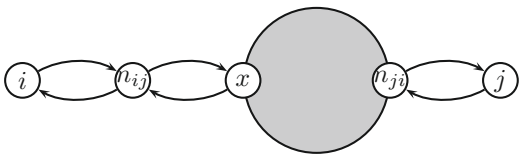


Figure 3.6: In this situation i and j are in Time Duality when n_{ij} and j are.

Proof Let us assume that the neighbours of n_{ij} are i and x. Then we can decompose the matrix $S=(q_{kl})_{k,l\in E\setminus i,j}$ further:

$$S=\begin{pmatrix} q_{n_{ij}n_{ij}} & q_{n_{ij}x}e_x \\ q_{xn_{ij}}e_x^\top & \hat{S}\end{pmatrix},$$

where $\hat{S}$ is a further restriction of the transient states without state n_{ij}. By the generalized

Frobenius formula for block inverses (see appendix) we have:

$$S_u^{-1} = \begin{pmatrix} q_{n_{ij}n_{ij}} + u & q_{n_{ij}x}e_x \\ q_{xn_{ij}}e_x^\top & \hat{S}_u \end{pmatrix}^{-1}$$
$$= K^{-1}\begin{pmatrix} 1 & -q_{n_{ij}x}e_x\hat{S}_u^{-1} \\ -q_{xn_{ij}}\hat{S}_u^{-1}e_x^\top & K\hat{S}_u^{-1} + q_{n_{ij}x}q_{xn_{ij}}\hat{S}_u^{-1}e_x^\top e_x\hat{S}_u^{-1} \end{pmatrix}.$$

where $K = q_{n_{ij}n_{ij}} + u - q_{n_{ij}x}q_{xn_{ij}}e_x\hat{S}_u^{-1}e_x^\top \in \mathbb{R}^+$.

Therefore for any u not an eigenvalue of $-S$:

$$\frac{R_{ij}(u)}{R_{ji}(u)} = \frac{K^{-1}(1, -q_{n_{ij}x}e_x\hat{S}_u^{-1})e_{n_{ji}}^\top}{K^{-1}e_{n_{ji}}\begin{pmatrix} 1 \\ -q_{xn_{ij}}\hat{S}_u^{-1}e_x^\top \end{pmatrix}} = \frac{q_{n_{ij}x}e_x\hat{S}_u^{-1}e_{n_{ji}}^\top}{q_{xn_{ij}}e_{n_{ji}}\hat{S}_u^{-1}e_x^\top} = \frac{q_{n_{ij}x}\hat{R}_{xj}(u)}{q_{xn_{ij}}\hat{R}_{jx}(u)}.$$

■

Theorem 3.2.5 can obviously be used iteratively to reduce the detection of Time Duality to the "essential" set of states, that is the set of states without the states belonging to Birth-and-Death like parts attached to the connected (simple) boundaries, see fig. 3.7.

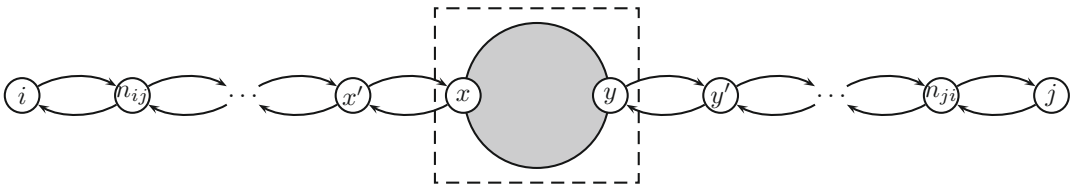


Figure 3.7: Time Duality between i and j can be reduced to Time Duality between x' and y'.

The preceding Theorem has a simple consequence, which will be interpreted later.

Corollary 3.2.6 *In a Birth-and-Death Process each two states are in Time Duality.*

Proof The proof is simply done by iteration of Theorem 3.2.5.

■

In some sense Theorem 3.2.5 is also a transitivity result. In the situation of fig. 3.7 i is in Time Duality with x and j with y by Corollary 3.2.6; if x and y are also in Time Duality then i and j are, too.

A similar situation is depicted in fig. 3.8. As the next theorem shows the forced passage through x, y, z suffices to check Time Duality between i and j just by checking Time Duality between i and y and y and j.

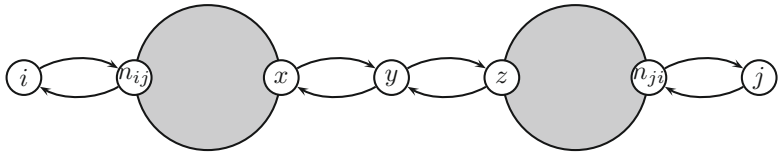


Figure 3.8: The structure of the graph forces any passage from i to j and vice versa through y.

Theorem 3.2.7 *Let $(X_t)_{t\geq 0}$ be a MCc on E with irreducible infinitesimal generator $Q = (q_{kl})_{k,l\in E}$. Fix two states $i, j \in E$ and assume that their connected neighbourhood is simple. Let $S = (q_{kl})_{k,l\in E\setminus\{i,j\}}$. If there exist three states $x, y, z \in E \setminus \{i, j\}$ such that S can be decomposed into*

$$S = \begin{pmatrix} q_{yy} & q_{yx}e_x & q_{yz}e_z \\ q_{xy}e_x^\top & S_i & 0 \\ q_{zy}e_z^\top & 0 & S_j \end{pmatrix} \tag{3.2}$$

Time Duality between i and j holds if and only if

$$\forall u \leq 0 : \frac{R_{ix}(u)R_{zj}(u)}{R_{jz}(u)R_{xi}(u)} = \frac{R_{ix}(0)R_{zj}(0)}{R_{jz}(0)R_{xi}(0)}$$

holds, where

$$R_{ix}(u) := e_{n_{ij}} S_{i,u}^{-1} e_x^\top, \quad R_{xi}(u) := e_x S_{i,u}^{-1} e_{n_{ij}}^\top$$
$$R_{jz}(u) := e_{n_{ji}} S_{j,u}^{-1} e_z^\top, \quad R_{zj}(u) := e_z S_{j,u}^{-1} e_{n_{ji}}^\top.$$

Proof The proof uses the generalized Frobenius formula for inverses of block matrices. It suffices to compute only one entry of the inverse of (3.2):

$$S_u^{-1} = \begin{pmatrix} * & & * \\ * & K \begin{pmatrix} S_{i,u}^{-1} & 0 \\ 0 & S_{j,u}^{-1} \end{pmatrix} + \begin{pmatrix} q_{xy}q_{yx}S_{i,u}^{-1}e_x^\top e_x S_{j,u}^{-1} & q_{xy}q_{yz}S_{i,u}^{-1}e_x^\top e_z S_{j,u}^{-1} \\ q_{zy}q_{yx}S_{j,u}^{-1}e_z^\top e_x S_{i,u}^{-1} & q_{zy}q_{yz}S_{j,u}^{-1}e_z^\top e_z S_{i,u}^{-1} \end{pmatrix} \end{pmatrix}.$$

The exact expression for the real constant K is not relevant as it cancels out after the application of Theorem 3.2.1.

■

The last results were in the flavor of serial network. The next theorem shows what to do in a "parallel situation", like in figure 3.9.

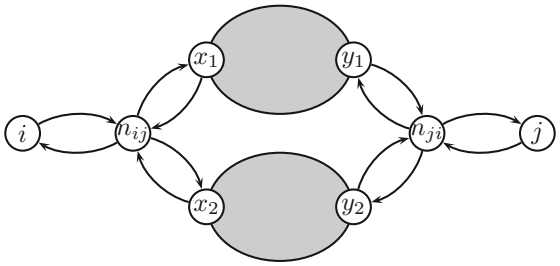


Figure 3.9: Situation of Theorem 3.2.8

Theorem 3.2.8 *Let $(X_t)_{t\geq 0}$ be a MCc on E with irreducible infinitesimal generator $Q = (q_{kl})_{k,l\in E}$. Let i, j, x_1, x_2, y_1, y_2 be fixed states such that the matrix $S = (q_{kl})_{k,l\in E\setminus\{i,j\}}$ has the following form*

$$S = \begin{pmatrix} q_{n_{ij}n_{ij}} & q_{n_{ij}x_1}e_{x_1} & q_{n_{ij}x_2}e_{x_2} & q_{n_{ij}n_{ji}} \\ q_{x_1n_{ij}}e_{x_1}^\top & S_1 & 0 & q_{y_1n_{ji}}e_{y_1}^\top \\ q_{x_2n_{ij}}e_{x_2}^\top & 0 & S_2 & q_{y_2n_{ji}}e_{y_2}^\top \\ q_{n_{ji}n_{ij}} & q_{n_{ji}y_1}e_{x_1} & q_{n_{ji}y_2}e_{y_2} & q_{n_{ji}n_{ji}} \end{pmatrix},$$

see also fig. 3.9 for reference. Let

$$T(u) := \frac{q_{n_{ij}x_1}q_{y_1n_{ji}}R^{(1)}_{x_1y_1}(u) + q_{n_{ij}x_2}q_{y_2n_{ji}}R^{(2)}_{x_2y_2}(u) - q_{n_{ij}n_{ji}}}{q_{n_{ji}y_1}q_{x_1n_{ij}}R^{(1)}_{y_1x_1}(u) + q_{n_{ji}y_2}q_{x_2n_{ij}}R^{(2)}_{y_2x_2}(u) - q_{n_{ji}n_{ij}}}$$

with

$$R^{(1)}_{x_1y_1}(u) := e_{x_1}S^{-1}_{1,u}e^\top_{y_1},\ R^{(2)}_{x_2y_2}(u) := e_{x_2}S^{-1}_{2,u}e^\top_{y_2},$$
$$R^{(1)}_{y_1x_1}(u) := e_{y_1}S^{-1}_{1,u}e^\top_{x_1},\ R^{(2)}_{y_2x_2}(u) := e_{y_2}S^{-1}_{2,u}e^\top_{x_2}.$$

Then Time Duality holds if and only if

$$\forall u \leq 0 : T(u) = T(0)$$

holds.

Proof This time we need to use both variants of the generalized Frobenius formula. We first calculate the quantities of interest in a more general way, i.e. for a, c, g, i real numbers (stars denote quantities we do not need for the following computation and thus we shall not give them explicitly):

$$M := \begin{pmatrix} a & B & c \\ D & E & F \\ g & H & i \end{pmatrix}^{-1} = K_1^{-1}\begin{pmatrix} & 1 & -(B,c)\begin{pmatrix} E & F \\ H & i \end{pmatrix}^{-1} \\ -\begin{pmatrix} E & F \\ H & i \end{pmatrix}^{-1}\begin{pmatrix} D \\ g \end{pmatrix} & & * \end{pmatrix}.$$

Further

$$\begin{pmatrix} E & F \\ H & i \end{pmatrix}^{-1} = K_2^{-1}\begin{pmatrix} * & -E^{-1}F \\ -HE^{-1} & 1 \end{pmatrix}.$$

In both cases K_1 and K_2 are real numbers. Thus the upper right corner of the inverse of M is given as

$$K_1^{-1}K_2^{-1}(BE^{-1}F - c)$$

and the lower left corner:

$$K_1^{-1}K_2^{-1}(HE^{-1}D - g).$$

If we make the following identification

$$a := q_{n_{ij}n_{ij}}, \qquad B := (q_{n_{ij}x_1}e_{x_1}, q_{n_{ij}x_2}e_{x_2}),\ c := q_{n_{ij}n_{ji}},$$
$$D := \begin{pmatrix} q_{x_1n_{ij}}e_{x_1}^\top \\ q_{x_2n_{ij}}e_{x_2}^\top \end{pmatrix}, \qquad E := \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix},\ F := \begin{pmatrix} q_{y_1n_{ji}}e_{y_1}^\top \\ q_{y_2n_{ji}}e_{y_2}^\top \end{pmatrix},$$
$$g := q_{n_{ji}n_{ij}}, \qquad H := (q_{n_{ji}y_1}e_{y_1}, q_{n_{ji}y_2}e_{y_2}),\ c := q_{n_{ji}n_{ji}}$$

we get

$$\frac{e_{n_{ij}}S_u^{-1}e_{n_{ji}}^\top}{e_{n_{ji}}S_u^{-1}e_{n_{ij}}^\top} = \frac{q_{n_{ij}x_1}q_{y_1n_{ji}}e_{x_1}S_{1,u}^{-1}e_{y_1}^\top + q_{n_{ij}x_2}q_{y_2n_{ji}}e_{x_2}S_{2,u}^{-1}e_{y_2}^\top - q_{n_{ij}n_{ji}}}{q_{n_{ji}y_1}q_{x_1n_{ij}}e_{y_1}S_{1,u}^{-1}e_{x_1}^\top + q_{n_{ji}y_2}q_{x_2n_{ij}}e_{y_2}S_{2,u}^{-1}e_{x_2}^\top - q_{n_{ji}n_{ij}}}.$$

The statement then follows by Theorem 3.2.1.

■

The question may arise why such a complicated theorem is useful. Indeed if $q_{n_{ij}n_{ji}} = q_{n_{ji}n_{ij}} = 0$ and with some extra conditions on S_1 and S_2 the behavior of models which arise in the context of Molecular Motors can be explained.

We derive the following corollaries of Theorem 3.2.8 and take over the definitions and assumptions from there.

If we assume that S_2 is the zero matrix in Theorem 3.2.8 the situation is like checking Time Duality between n_{ij} and n_{ji} in a system with simple connected neighbourhood similar to the one above where x_1 is the simple connected neighbourhood of n_{ij} and y_1 the one of n_{ji}. The next corollary shows that indeed the transition rates that determine how to reach x_1 and y_1 from n_{ij} and n_{ji} do not play *any* role for Time Duality.

Corollary 3.2.9 *Suppose $S_2 = 0$ and $q_{n_{ij}n_{ji}} = q_{n_{ji}n_{ij}} = 0$ then Time Duality is equivalent to*

$$\forall u \leq 0 : \frac{R_{x_1y_1}^{(1)}(u)}{R_{y_1x_1}^{(1)}(u)} = \frac{R_{x_1y_1}^{(1)}(0)}{R_{y_1x_1}^{(1)}(0)}.$$

In particular it is independent of the choice of the transition rates $q_{n_{ij}x_1}$ and $q_{n_{ji}y_1}$.

If there is a similarity relation between S_1 and S_2 Time Duality between i and j can be reduced drastically as the following corollary shows.

Corollary 3.2.10 *If*

$$R^{(1)}_{x_1y_1}(u) = R^{(2)}_{x_2y_2}(u)$$

and

$$R^{(1)}_{y_1x_1}(u) = R^{(2)}_{y_2x_2}(u)$$

then Time Duality between i and j is equivalent to

$$\frac{R^{(1)}_{x_1y_1}(u)}{R^{(1)}_{y_1x_1}(u)} = \frac{R^{(1)}_{x_1y_1}(0)}{R^{(1)}_{y_1x_1}(0)}$$

and is independent of the choice of $q_{n_{ij}x_1}$, $q_{n_{ij}x_2}$, $q_{n_{ji}y_1}$ and $q_{n_{ji}y_2}$.

Finally we note that the preceding corollaries are extendable trivially to the case when there are more than two "simple connected" subsystems. The corollaries also show that "parallel circuits" are way more demanding to allow Time Duality.

3.3 Extension to non-simple boundary

In the general case where the connected neighbourhoods are not simple anymore the criterion 3.1 becomes more complicated. Naturally one can define Time Duality between two states i, j by comparing the associated moment generating functions. This yields:

Theorem 3.3.1 *Let $(X_t)_{t\geq 0}$ be MCc on E, with irreducible infinitesimal generator $Q = (q_{ij})_{i,j\in E}$. Then i is in Time Duality with j if and only if*

$$\forall u \leq 0 : \frac{\nu_i \, diag(h^j)^{-1}(uId + S)^{-1}R^\top_{\Delta_j}}{\nu_j \, diag(h^i)^{-1}(uId + S)^{-1}R^\top_{\Delta_i}} = 1,$$

with the definitions of Proposition 3.1.4.

Obviously Theorem 3.3.1 coincides with Theorem 3.2.1 if the connected neighbourhoods of i and j are simple.

Let i, j be fixed states of E. As in the preceding section we identify $E \setminus \{i, j\}$ with the set of states that are in the same communication class(es) as $\mathcal{N}_{i\to j}$ and $\mathcal{N}_{j\to i}$. Then we define for $S = (q_{kl})_{k,l\in E\setminus\{i,j\}}$

$$\tilde{S} := \begin{pmatrix} \bar{q}_{ii} & R_{iS} & 0 \\ R^\top_{Si} & S & R^\top_{Sj} \\ 0 & R_{jS} & \bar{q}_{jj} \end{pmatrix} \tag{3.3}$$

where $R_{iS} = \nu_i \mathrm{diag}(h^j)^{-1}$, $R_{jS} = \nu_j \mathrm{diag}(h^i)^{-1}$, $R_{Si} := R_{\Delta_i}$, $R_{Sj} := R_{\Delta_j}$, $\bar{q}_{ii} > R_{iS}\mathbf{1}^\top$ and $\bar{q}_{jj} > R_{jS}\mathbf{1}^\top$.

With this definition we derive the following description of Time Duality between i and j:

Theorem 3.3.2 *Let $(X_t)_{t\geq 0}$ be a MCc on E, with irreducible infinitesimal generator $Q = (q_{kl})_{k,l\in E}$. Two fixed states i and j of E are in Time Duality if and only if*

$$\forall u \leq 0 : \frac{\tilde{R}_{ij}(u)}{\tilde{R}_{ji}(u)} = \frac{\tilde{R}_{ij}(0)}{\tilde{R}_{ji}(0)}$$

holds, where

$$\tilde{R}_{ij}(u) := e_i \tilde{S}^{-1}_u e^\top_j, \ \tilde{R}_{ji}(u) := e_j \tilde{S}^{-1}_u e^\top_i.$$

Proof By Lemma 3.2.9 the condition of the theorem is just a reformulation of the condition of Theorem 3.3.1.

■

The trick to treat the general case is thus a reformulation of the equality of the moment generating functions in the flavor of simple connected neighbourhood, see also fig. 3.10. Thus all results of the preceding section can also be applied to the general case.

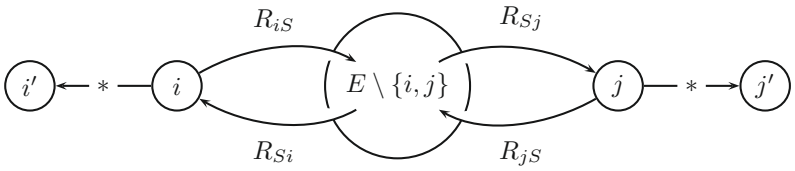


Figure 3.10: Non simple boundary reformulated in the flavour of simple boundary.

D

The following example shows that the reinterpretation done in Theorem 3.3.1 can simplify the verification of Time Duality.

Example **3.3.3** Let $(X_t)_{t\geq 0}$ be a MCc on $E := \{1,2,3,4,5\}$ with infinitesimal generator

$$Q := \begin{pmatrix} -2 & 1 & 1 & 0 & 0 \\ 1 & -2 & 0 & 0 & 1 \\ 1 & 0 & -2 & 0 & 1 \\ 0 & 1 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

The transition graph is given in fig. 3.11.

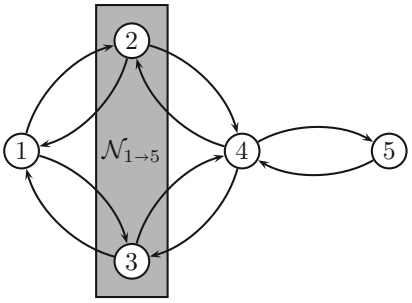


Figure 3.11: Transition graph to Example 3.3.

Now $\tilde{S}$ is given by

$$\tilde{S} = \begin{pmatrix} -5 & 2 & 2 & 0 & 0 \\ 1 & -2 & 0 & 1 & 0 \\ 1 & 0 & -2 & 1 & 0 \\ 0 & 1 & 1 & -3 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}.$$

If we now interpret $\tilde{S}$ as derived from the system displayed in fig. 3.12 (the parameters q^*_{11} and q^*_{55} can obviously chosen greater then $R_{\Delta_1}\mathbf{1}^\top$ and $R_{\Delta_5}\mathbf{1}^\top$), we already know that Time Duality appears between $1'$ and $5'$, as the new system is reversible by Theorem 1.3.1 chapter 2 and Theorem 3.2.3.

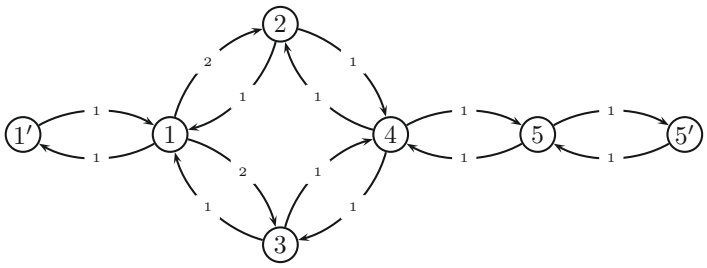


Figure 3.12: Reinterpretation of the original dynamics in terms of simple boundary.

In this sense Time Duality between 1 and 5 is the consequence of reversibility in the modified and extended model and thus also implied by the Time Duality between $1'$ and $5'$ (compare fig. 3.11 with fig. 3.12).

□
□

3.4 Some examples

3.4.1 Time Duality without permuted balance

In this subsection we want to give an example that permuted balance does not exploit all cases where Time Duality appears. Corollaries 3.2.9 and 3.2.10 indicate that indeed there are some situations where this might be the case.

Example **3.4.1** Let $(X_t)_{t\geq 0}$ be the MCc associated to the transition graph in fig. 3.13. We keep the transitions as abstract as possible in the sense that every off-diagonal entry of the infinitesimal generator is positive if there is a transition depicted in the transition graph.

Time Duality between any two direct neighbours is given, as the communication class of the connected neighbourhood forms a Birth-and-Death process. If we want Time Duality between 1 and 3 we need to investigate under which conditions

$$\frac{\nu_1 \mathrm{diag}(h^3)^{-1}(S+uId)^{-1}R^\top_{\Delta_3}}{\nu_3 \mathrm{diag}(h^1)^{-1}(S+uId)^{-1}R^\top_{\Delta_1}} = \frac{q_{12}q_{22}(u+q_{44})+q_{14}q_{44}(u+q_{22})}{q_{32}q_{22}(u+q_{44})+q_{34}q_{44}(u+q_{22})}\frac{q_{33}}{q_{11}}$$

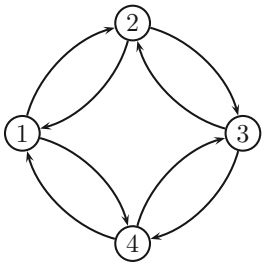


Figure 3.13:

is a constant. As

$$\frac{q_{33}}{q_{11}} = \frac{q_{32} + q_{34}}{q_{12} + q_{14}}$$

the only constraint to fulfill the condition in Theorem 3.3.1 is $q_{22} = q_{44}$.

In this situation it is easier to check the reinterpretation as model with simple boundary, see fig. 3.14, as Theorem 3.2.8 can be applied. The necessary conditions of this Theorem also reduce to just $q_{22} = q_{44}$.

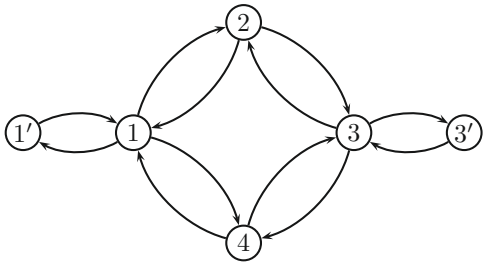


Figure 3.14: The transition rates to the new states $1'$ and $3'$ can be chosen arbitrarily non-zero.

Though the dynamic defined is irreducible, a unique stationary distribution π exists. The exact value is too complicated, but the freedom of choice of transition rates is big enough to choose the rates such that $\pi_1 \neq \pi_2 \neq \pi_3 \neq \pi_4$ and $q_{12}q_{23}q_{34}q_{41} \neq q_{14}q_{43}q_{32}q_{21}$ (this prevents reversibility). For such a choice of rates permuted balance cannot occur, as for every permutation not the identity at least two states must be exchanged. But by chapter 1, Proposition 1.3.7, (ii) , this would imply that π is constant on such a transposition, which was already excluded by the choice of rates.

One choice of these parameters could be:

$$Q = (q_{ij})_{i,j\in\{1,2,3,4\}} = \begin{pmatrix} -3 & 2 & 0 & 1 \\ 2 & -5 & 3 & 0 \\ 0 & 1 & -3 & 2 \\ 4 & 0 & 1 & -5 \end{pmatrix}$$

with associated stationary distribution

$$\pi = \frac{1}{136}(50, 27, 35, 24).$$

This choice is also not reversible as

$$48 = q_{12}q_{23}q_{34}q_{41} \neq q_{14}q_{43}q_{32}q_{21} = 2.$$

□
□

3.4.2 Birth-and-Death Processes with reflecting boundary and processes on trees

Birth-and-Death Processes with reflecting boundaries enjoy a very curious property. Corollary 3.2.6 states that any two arbitrary states i and j are in Time Duality. This is also true for biased random walks on $\mathbb{Z}$, which are in some sense an extension of Birth-and-Death Processes.

If we interpret Time Duality as a relation and write $i \overset{TD}{\leftrightarrow} j$ for "i is in Time Duality with j", then $\overset{TD}{\leftrightarrow}$ defines an equivalence relation on the state space (in the case of Birth-and-Death processes). Reflexivity is trivial and also symmetry by definition. A more difficult problem is transitivity. Theorem 3.2.7 delivers the answer and shows that $\overset{TD}{\leftrightarrow}$ is transitive. The associated Hasse diagram to the Time Duality relation would be thus the complete graph.

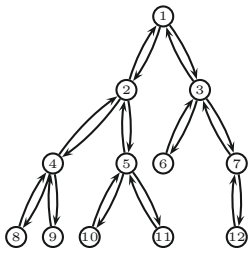


Figure 3.15: An example transition graph of a Markov Chain on a tree, transition rates not shown here.

After all the property that any two states of a Birth-and-Death Process are in Time Duality is a kind of "super-symmetry" that is not shared by all processes. A close relative to Birth-and-Death Processes are processes on trees (under the assumption that the state space is a single communication class). Processes on trees also have the property that any connected neighbourhood with respect to some arbitrary distinct states is simple. For an example of a Markov Chain on a tree see fig. 3.15. The assumption of a single communication class implies that for any transition rate $q_{ij} > 0$ also $q_{ji} > 0$. The tree property ensures that any two states are connected in a "loop free" way such that the whole process is reversible. Interestingly this is already sufficient for Time Duality together with the simple connected neighbourhoods, see section 3.1.2. Thus also Processes on trees are super-symmetric in the sense introduced above.

The question whether Time Duality is also an equivalence relation in the general case or is only a so called *tolerance relation*, i.e. just reflexive and symmetric, remains open.

Chapter 4

Killed Quasi-Random-Walk

In living cells transport mechanisms are crucial for the survival of the cell. For these transport mechanisms proteins called *molecular motors* play a fundamental role. As we detail in chapter 5 their movement on a rope-like structure called "microtubule" is very similar to a Radom Walk on $\mathbb{Z}$. But at each position $z \in \mathbb{Z}$ the further movement is a consequence of the passage through different "internal" states. Additionally the motor can detach itself from the tubule and the process stops. Several perspectives can be taken to model such a process. We decide here for the matrix analytic way by extending the well known Quasi-Birth-and-Death Process (QBDP) introduced in [Neu94]. They are described process a few matrices of finite dimension. Many interesting properties can also be expressed with these matrices which make them an important tool in Queueing Theory, see [Neu94] and [LR99]. The main advantage is surely the accessibility to numerical methods and the possibility to apply results from Matrix Theory. We call our extension *killed Quasi-Random-Walk* (kQRW). "Killed" is a terminology from, for instance, [DZ05b] and [DZ05a], where usual Birth and Death Processes are extended by a "killing rate", possibly unequal for each $i \in \mathbb{N}$ (we do not allow position dependent killing rates).

Lemma 4.1.5 given in section 1 provides a very powerful tool for the analysis of kQRW's. The lemma allows it to reduce the process to a finite state Markov Chain, where certain quantities stay invariant under the aggregation. We will use this reduction tool extensively in the following sections. However, similar methods can not be applied to QBDP's as the absence of a boundary in a kQRW is crucial for our methods.

4.1 Definitions and preliminaries

In the following we adapt the definition of Quasi-Birth-and-Death Process of [LR99], chapter 6, to define a kQRW. Thus let

$$E := \{(z,i) : z \in \mathbb{Z}, i \in M\} \cup \{\Delta\},$$

where Δ is an absorbing state and $M := \{1, 2, \ldots, m\}$ a finite set. For (z,i) we refer to z as level. In the literature, see e.g. [LR99] i is often called *phase*; we say in most cases states rather than phases. The decomposition

$$\mathbb{Z} \times M = \bigsqcup_{z \in \mathbb{Z}} l(z) := \bigsqcup_{z \in \mathbb{Z}} \{(z,1), (z,2), \ldots, (z,m)\}$$

divides the state space into so called *levels* denoted by $l(z)$. A transition (z,i) to Δ is called *killing* and the rate γ_i associated to this transition is a *killing rate*.

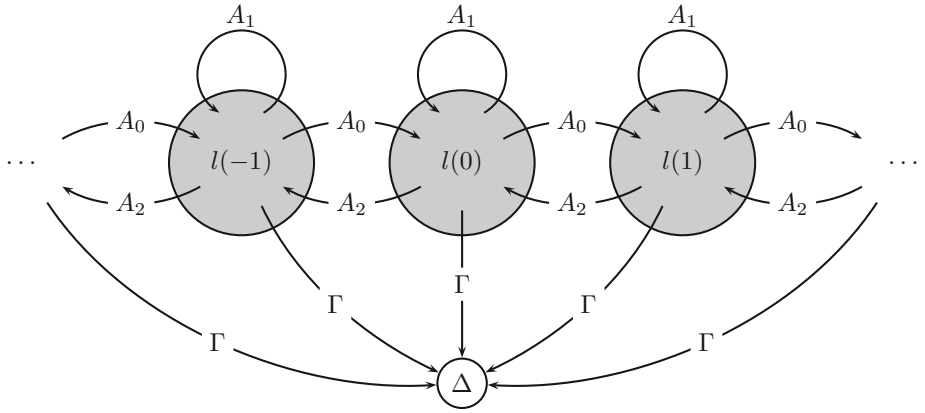


Figure 4.1: Schematic display of the concept of killed Quasi-Random-Walk.

Definition 4.1.1 *Let $(X_t)_{t\geq 0}$ be a MCc on E with initial condition μ concentrated on level 0. Define the infinite-dimensional sub generator*

$$S := \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \\ \dots & A_1 & A_0 & 0 & 0 & \dots \\ \dots & A_2 & A_1 & A_0 & 0 & \dots \\ \dots & 0 & A_2 & A_1 & A_0 & \dots \\ \dots & 0 & 0 & A_2 & A_1 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

with A_0, A_1, A_2 matrices of order $m \times m$, A_0 and A_2 non negative and A_1 a sub generator. The vector $R_\Delta := (\dots, \Gamma, \Gamma, \Gamma, \dots)$ where

$$\Gamma := (\gamma_1, \gamma_2, \dots, \gamma_m)$$

is the vector of killing rates for each level. Further assume that

$$(\Gamma + A_0 + A_1 + A_2) = \mathbf{0}.$$

If the infinitesimal generator Q of $(X_t)_{t\geq 0}$ has the form

$$Q = \begin{pmatrix} 0 & \mathbf{0} \\ R_\Delta^\top & S \end{pmatrix} \tag{4.1}$$

we say $(X_t)_{t\geq 0}$ is a killed Quasi-Random-Walk *(kQRW), see also fig. 4.1.*

The matrices A_i characterize the behavior of the process; while A_1 contains the transition rates between states of a fixed level, the matrix A_2 contains the transition to change from level z to level $z-1$ for each $z \in \mathbb{Z}$, accordingly A_0 contains the transition rates for the change from z to $z+1$. The notation is passed on by Neuts original definition, which is not ideal but standard.

Note that if R_Δ is zero in every component than the *kQRW* is a generalization of the usual Random Walk (without killing). Indeed, this is not a very interesting object: the process is either transient or zero-recurrent. This can be characterized by a simple extension of the drift conditions developed in [LT03].

We also describe the attained level at time t, that is the projection on the first coordinate.

Definition 4.1.2 *Let $(X_t)_{t\geq 0}$ be a kQRW. The process $(L_t)_{t\geq 0}$ is defined by projection on the first coordinate of $(X_t)_{t\geq 0}$, i.e.*

$$L_t := \begin{cases} \Delta, & \text{if } X_t = \Delta \\ \pi_1(X_t), & \text{else} \end{cases},$$

where π_1 is the projection on the first coordinate, i.e. $\pi_1((z,i)) = z$). We call $(L_t)_{t\geq 0}$ the level process *associated to $(X_t)_{t\geq 0}$.*

It is clear that in general the level process is not Markovian, as the original process can jump several times between states of the same level leading to non exponential holding times in each level. In particular the probability to be in level $z \in \mathbb{Z}$ at time t under start in $i \in l(0)$ is given by the sum

$$\mathbb{P}_i(L_t = z) = \sum_{k\in l(z)} e_i \exp(Qt) e_k^\top .$$

In analogy to chapter 2 about absorption times and their distribution we introduce the time to absorption or the time until the random walker is killed in the usual way.

Definition 4.1.3 *Let $(X_t)_{t\geq 0}$ be a kQRW. The stopping time*

$$\tau_\Delta := \inf\{t \geq 0 : X_t = \Delta\}$$

is the time to absorption in Δ. We call this time also killing time.

Remark **4.1.4** In an obvious way we adapt the Definitions 4.1.1, 4.1.2 and 4.1.3 for discrete time where the only difference to continuous time is that A_0, A_1, A_2 and Γ are non negative in each component and satisfy $(A_0 + A_1 + A_2 + \Gamma)\mathbf{1}^\top = \mathbf{1}^\top$. Also the level process defined in Definition 4.1.2 can be adapted in analogue form to discrete time, which will then be denoted with $(L_n)_{n\geq 0}$..

In our context we always want to investigate the behavior of a continuous time kQRW. If we compute properties of a discrete time version we have in mind to investigate properties of the *embedded MCd* associated to the continuous time kQRW.

□

We defined the kQRW in a very homogeneous way in the sense that each level has the same states and killing rates. This sort of "spatial homogeneity" allows to reduce the state space to a finite Markov Chain with an absorbing state. This reduction is very useful in order to compute properties that are not related with the position of the Random Walker.

The result of the following lemma is well known in the physics literature as "closing method", see e.g. the series of papers by T. Hill [Hil88b], [Hil88a] and [Hil88c]. We want to give a proof here because (up to our knowledge) there is no mathematical justification of the method, even not in the mentioned references, although the argumentation is (heuristically) convincing. The closing method also regards the alteration of absorbing states, in the sense that transitions are

added from the absorbing state back to the transients according to the initial distribution. The idea behind this is a renewal process, where renewals happen according to the distribution of the absorption time. We do not use this addition here.

Lemma 4.1.5 *Let $(X_t)_{t\geq 0}$ be a kQRW. Define*

$$[i] := \bigcup_{z\in\mathbb{Z}} (z,i),$$

$M' = \{[i] : i \in M\}$ and $E' := M' \cup \{\Delta\}$. With respect to this decomposition $(X_t)_{t\geq 0}$ is lumpable.

Proof In Proposition 1.3.9 we characterized under which conditions a discrete time Markov Chain is lumpable depending on the initial distribution. The extension to continuous time is straight forward via either the uniformization method or by lifting the discrete time conditions to continuous time explicitly, this is done for instance in [TK06]. We want to give here only the conditions as they are analogue to the ones given in Proposition 1.3.9 as we need them for the proof.

Given the decomposition of the original state space E into $\{\Delta\} \cup \bigcup_{i\in M}[i]$ the infinitesimal generator Q associated to $(X_t)_{t\geq 0}$ is lumpable if and only if

$$\sum_{z\in\mathbb{Z}} q_{(z_1,i)(z,k)} = \sum_{z\in\mathbb{Z}} q_{(z_2,i)(z,k)} \tag{4.2}$$

and

$$\sum_{z\in\mathbb{Z}} q_{(z_1,i)(z,\Delta)} = \sum_{z\in\mathbb{Z}} q_{(z_2,i)(z,\Delta)} \tag{4.3}$$

for arbitrary $(z_1,i),(z_2,i) \in [i]$ and $[k]$. As direct transitions are only allowed between the same level or the nearest neighboring levels the condition is equivalent to

$$q_{(z_1,i)(z_1,k)} + q_{(z_1,i)(z_1-1,k)} + q_{(z_1,i)(z_1+1,k)} = q_{(z_2,i)(z_2,k)} + q_{(z_2,i)(z_2-1,k)} + q_{(z_2,i)(z_2+1,k)},$$

which is obviously fulfilled as the transition rates are defined level independent. Though Δ is a state which is not lumped together with other states the condition of lumpability for the set $\{\Delta\}$ is also automatically fulfilled. Theorem 2.11 in [TK06] shows that (4.2) and (4.3) are necessary and sufficient for lumpability. As the initial distribution is concentrated on level $l(0)$ there is no incompatibility which finishes the proof.

■

As we use the lumped chain from Lemma 4.1.5 in the following we need to define it:

Definition 4.1.6 *We will refer to the lumped Markov Chain of Lemma 4.1.5 as* reduction *of $(X_t)_{t\geq 0}$ the* reduced chain. *If not otherwise denoted we associate $(Y_t)_{t\geq 0}$ to the reduced chain on state space $E' = M' \cup \{\Delta\}$.*

The preceding Lemma 4.1.5 induces a very useful property: the killing time of the kQRW has the same distribution as the time to absorption in the reduced model. This can be easily proven by reformulating the lumping procedure as a matrix operation.

Lemma 4.1.7 *Let $(X_t)_{t\geq 0}$ be a kQRW and $(Y_t)_{t\geq 0}$ its reduced version, given in Lemma 4.1.5. Let furthermore be*

$$\tilde{\tau}_\Delta := \inf\{t \geq 0 : Y_t = \Delta\}$$

then

$$\tau_\Delta \stackrel{d}{=} \tilde{\tau}_\Delta$$

if $\mu_{|l(0)} = \mu'$ holds.

Proof By Lemma 4.1.5 we already know that the infinitesimal generator Q associated to $(X_t)_{t\geq 0}$ is lumpable. By Theorem 2.10 in [TK06] there exist two matrices U, V such that Q is lumpable if and only if

$$VUQV = QV. \tag{4.4}$$

The extension of (4.4) to countable state space is problem free here, because each row of the original process contains only finitely many non-zero entries. The lumpability criterion can by the spatial homogeneity property of a kQRW be characterized on the transient states via the $\infty \times m$-matrix $U := (\ldots, 0, Id, Id, Id, 0, \ldots)$ non-zero only at levels -1, 0 and 1 where it coincides with the $m \times m$ identity matrix and V the $m \times \infty$ matrix zero everywhere except of level 0 where it is also the $m \times m$ identity. This also works for any other "reference levels" different from 0.

The matrices have the useful property that $UV = Id$. The lumped infinitesimal generator (restricted to the transients) is then given by

$$\tilde{S} = USV.$$

With 4.4 we get that

$$\tilde{S}^n = (USV)^n = US^nV,$$

where $\tilde{S}$ is the infinitesimal sub generator associated with the transients of $(Y_t)_{t\geq 0}$. We also have by induction

$$VUS^nV = S^nV.$$

The complete infinitesimal generator of $(Y_t)_{t\geq 0}$ is given as

$$\tilde{Q}\begin{pmatrix} 0 & 0 \\ UR_\Delta^\top & USV \end{pmatrix}$$

and the initial condition transforms to μV which is just the projection onto level $l(0)$. Then

$$\begin{aligned} F_{\tau_\Delta}(t) &= 1 - \mu V \exp(USVt)\mathbf{1}^\top = 1 - \mu V \sum_{n\geq 0} \frac{t^n}{n!}(USV)^n\mathbf{1}^\top \\ &= 1 - \mu \sum_{n\geq 0} \frac{t^n}{n!} \underbrace{VUS^nV}_{=\tilde{S}^nV}\mathbf{1}^\top = 1 - \mu \exp(\tilde{S}t)\underbrace{V\mathbf{1}^\top}_{=\mathbf{1}^\top} \\ &= F_{\tilde{\tau}_\Delta}(t) \end{aligned}$$

The last step of the computation turns the m-dimensional vector $\mathbf{1}^\top$ into the infinite dimensional vector $\mathbf{1}^\top$ by multiplying with V.

■

Corollary 4.1.8 *The statements of Lemma 4.1.5 and Lemma 4.1.7 are analogue for the discrete kQRW.*

4.2 Lifetime

By Lemma 4.1.7 the lifetime of a kQRW with killing is simply the time to absorption in the reduced model introduced in Lemma 4.1.5. The time to absorption is thus of PH-type distribution.

Proposition 4.2.1 *Let $(X_t)_{t\geq 0}$ be a kQRW. Then τ_Δ is PH-type distributed with representation*

$$\tau_\Delta \sim PH(\mu_{|l(0)}, A_0 + A_1 + A_2)$$

where $\mu_{|l(0)}$ is the initial distribution μ of $(X_t)_{t\geq 0}$ restricted to level zero.

Proof Lemma 4.1.5 states that the absorption time of $(X_t)_{t\geq 0}$ coincides with the absorption time of the reduction $(Y_t)_{t\geq 0}$. Thus by chapter 2 the distribution of τ_Δ, the time until absorption in Δ, is a PH-type distribution with representation as given in the statement of this proposition. ■

4.3 The notion of forward and backward steps and their distribution

The kQRW contains a lot of information regarding both the time and the spatial development of the process. Furthermore the associated infinitesimal generator resp. transition matrix is of countable dimension and thus explicit calculations are difficult. The level process on the other hand does not contain *enough* information in some sense. It totally ignores the fact that inside levels several transitions can be done before changing level, thus it can not be Markovian.

We thus want to introduce an intermediate process (in discrete time), which we introduce as *step-process*, that contains only information about the transitions between levels. We use it extensively to characterize certain properties of the generalized Random Walker in the following, which are difficult or impossible to achieve otherwise.

The idea behind the step-process is the observation that each trajectory can be split between excursions to new levels. That means beginning at some state of $l(0)$ the Random Walker reaches after some time another level, say $l(1)$ for the first time. After some transitions in $l(1)$ the Walker approaches a new level and so on and so forth. We derive a process that describes exactly the probabilities of jumping from one level to another in such a way that the transition probabilities describe the passage probabilities to another level from the state the process has entered the level.

In analogy to the boundary notion defined in Definition 2.2.9 in chapter 2 we introduce a sort of boundary set for the levels. We define them according to the states which can be reached from either the left or the right level.

Definition 4.3.1 *Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW on $E = \mathbb{Z}\times M$. For level $l(k)$ define the following sets:*

1. *Arrival from the left:*

$$[\![\mapsto]\!]^{(k)} := \{i \in M : \exists j \in M : p_{(k-1,j),(k,i)} > 0\}$$

2. *Arrival from the right:*

$$[\![\leftarrow]\!]^{(k)} := \{i \in M : \exists j \in M : p_{(k+1,j),(k,i)} > 0\}$$

We call the elements of these sets accordingly arrival *states. The definition is independent of the level by the spatial homogeneity of $(X_n)_{n\geq 0}$, thus we just write $[\![\mapsto]\!]$ ($[\![\leftarrow]\!]$) if the context is clear.*

We note that $[\![\mapsto]\!]$ and $[\![\leftarrow]\!]$ do not need to be disjoint.

We are interested in the probabilities of starting in level 0 and hitting an arrival set of level $l(1)$ or $l(-1)$ before killing and the probabilities of being killed before reaching another arrival set. This is done by a modification of the process. To calculate these "passage probabilities" between arrival states of neighboring levels, we set $[\![\leftarrow]\!]^{(-1)}$ and $[\![\rightarrow]\!]^{(1)}$ absorbing and use the results of chapter 2, section 2.1 to condition to absorption in one of the states of the neighboring arrival sets or Δ; this delivers the requested probabilities.

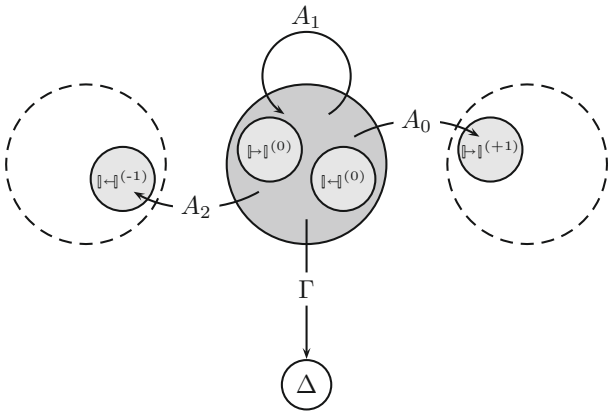


Figure 4.2: Setting the arrival sets of $l(-1)$ and $l(+1)$ absorbing allows to compute the passage probabilities between arrival states of neighboring levels.

Proposition 4.3.2 *Let $(X_t)_{t\geq 0}$ be a kQRW and infinitesimal generator as introduced in Definition 4.1.1. Let i be an arrival state, i.e. $i \in [\![\leftarrow]\!]^{(0)} \cup [\![\mapsto]\!]^{(0)}$, then the probability to reach an arrival state (or Δ) $j \in [\![\leftarrow]\!]^{(-1)} \cup [\![\mapsto]\!]^{(1)} \cup \{\Delta\}$ of a neighboring level is given by*

$$p_{(0,i),(\pm 1,j)} := e_i(-A_1)^{-1}R_j^\top,$$

where R_j is the vector containing the transition rates of leaving the level $l(0)$ and arriving at the state j.

Proof If we define the set $\Delta' := [\![\leftarrow]\!]^{(-1)} \cup [\![\mapsto]\!]^{(1)} \cup \{\Delta\}$ as a set of absorbing states, see also fig. 4.2, we can interpret the problem of calculating the passage probabilities as the computation of the absorption probability into a specific state of the set Δ'. The statement is then a direct consequence of Lemma 2.2.2, chapter 2.

■

Remark 4.3.3 If we consider the continuous time kQRW we can also calculate the time distribution of a passage from an arrival state to another in a neighboring level. These times are given by recognizing that the absorption time of the conditioned process given in Lemma 2.2.2 have phase type distribution. Thus the passage time starting form $(0,i)$ to (k,j), $k \in \{-1,1\}$, $i \in [\![\leftarrow]\!]^{(0)} \cup [\![\mapsto]\!]^{(0)}$, $j \in [\![\leftarrow]\!]^{(-1)} \cup [\![\mapsto]\!]^{(1)} \cup \{\Delta\}$ is given by the distribution function

$$F_{T_{(0,i),(k,j)}}(t) = 1 - \frac{1}{p_{(0,i),(k,j)}} e_i \exp(A_1 t)\mathbf{1}^\top$$

in the notation of Proposition 4.3.2.

□

With the passage probabilities computed in Proposition 4.3.2 we can build a new Quasi-Random-Walk. Though we are only interested in the steps between levels, we restrict the new state space to the arrival states.

Definition 4.3.4 *Let $M_A := ([\![\mapsto]\!]^{(0)} \cup [\![\leftarrow]\!]^{(0)})$ and $E := \mathbb{Z} \times M_A$. Let $m := |M_A|$ and define $A_1 := 0$ (zero matrix of dimension $m \times m$) and the $m \times m$-matrices $A_0 := (a_{ij}^{(0)})_{i,j\in M_A}$, $A_2 := (a_{ij}^{(2)})_{i,j\in M_A}$ with*

$$a_{ij}^{(0)} := p_{(0,i),(1,j)},$$

$$a_{ij}^{(2)} := p_{(0,i),(-1,j)}$$

and

$$\Gamma := (p_{(0,i),\Delta})_{i\in M_A}$$

as given in 4.3.2. If the initial condition is concentrated on the arrival states of level zero, we call the kQRW defined by the matrices A_0, A_1, A_2 and Γ the step-process.

The step-process has now the nice property that every transition encodes a step from one level to another, or killing. Thus it is natural that the number of total transitions of the step-process is just the number of levels the original process visits before being killed, thus $L_{n-1} \neq L_n$ before absorption for every $n \leq \tau_\Delta$. Then the number of transitions until absorption can be described by a discrete PH-distribution.

Proposition 4.3.5 *Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW and let $(Y_n)_{n\geq 0}$ be its associated step-process with initial distribution μ. Then the number of levels visited, i.e. the length of the random vector*

$$(L_0, L_1, \ldots, L_{\tau_\Delta - 1}),$$

before absorption is equal in law to the absorption time of the step-process, i.e.

$$\mathbb{P}(dim((L_0, L_1, \ldots, L_{\tau_\Delta - 1})) = k) = \mu(A_0 + A_2)^{k-1}\Gamma^\top$$

for $k \geq 1$.

Proof The result follows immediately by Lemma 4.1.5, as the absorption time of $(Y_n)_{n\geq 0}$ is identical with the one of its reduced version and is therefore a PH-type distribution with representation $PH(\mu, A_0 + A_2)$.

■

With the total number of changes of levels a notion called *run-length* can be calculated, it is the number of level changes multiplied by a displacement constant. We calculate this quantity and others in chapter 5 for a model from biophysics.

We have introduced the step-process to answer more precise questions, for instance: "How many transitions (steps) $(k, i) \to (k + 1, j)$ have occured before killing?". We answer this question by using the pattern matching introduced in chapter 1, section 3, together with the notion of local time introduced in chapter 1, section 2. The idea is to first define a new Markov chain on $M_A \times M_A$ which allows to identify transitions of the original process as states of the new chain, called 2-pattern chain in chapter 1, section 3. Then the local time allows to further reduce the resulting Chain to a subset of states $T \subsetneq M_A \times M_A$, i.e. transitions of the original process. The number of visits of states in T is then the number of transitions before killing.

For the following result we refer also to Remark 1.4 in chapter 1, section 3, where we explain how to use the notion of local time in the presence of absorbing states.

Proposition 4.3.6 *Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW on $\mathbb{Z}\times M$. Assume that the 2-pattern chain associated to the step-process of $(X_n)_{n\geq 0}$ has transition matrix P with, decomposed as follows*

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \Gamma_T^\top & S_T & S_{TT^c} \\ \Gamma_{T^c}^\top & S_{T^cT} & S_{T^c} \end{pmatrix}$$

for a subset $T \subsetneq M_A \times M_A$, where M_A is the set of arrival states. Let $\mu = (\mu_T, \mu_{T^c})$ be the initial condition of the step process. Let further be

$$\tilde{S} := S_T + S_{TT^c}(Id - S_{T^c})^{-1}S_{T^cT},$$

$$\tilde{\Gamma}^\top := \Gamma_T^\top + S_{TT^c}(Id - S_{T^c})^{-1}\Gamma_{T^c}^\top$$

and

$$\tilde{\mu} = \mu_T + \mu_{T^c}(Id - S_{T^c})^{-1}S_{T^cT}.$$

Then the total number of steps in T before killing has the following distribution:

$$\mathbb{P}(\#(\textit{steps in } T) = k) = \begin{cases} \mu_\Delta & k = 0 \\ \tilde{\mu}\tilde{S}^{k-1}\tilde{\Gamma}^\top & k \geq 1 \end{cases},$$

where Δ is the absorbing state of the 2-pattern chain.

Proof The step-process is a (discrete time) kQRW defined on $\mathbb{Z} \times M_A \cup \Delta$. Then by the discrete time version of Lemma 4.1.5, the step-process can be reduced to a chain with $m + 1$ states, where $m := |M_A|$. The 2-pattern chain associated to the reduced version of the step process is a Markov Chain with $m^2 + 1$ states, as the transitions from i to Δ and from Δ to Δ where aggregated into a new absorbing state, also denoted with Δ, see 1.4. Then the total number of steps in T is the time to absorption in Δ and thus a PH-type distribution with representation $PH(\tilde{\mu}, \tilde{S})$.

■

Note that the same approach can be used to compute the number of certain transitions of the original chain. It is then not necessary to switch to the step-process.

Remark 4.3.7 Throughout the whole chapter we assumed for simplicity that the initial distribution μ is concentrated on the arrival states of level $l(0)$, thus the initial condition of the step-process is just the restriction of μ to the arrival states. If the initial condition of the original chain is concentrated in $i \notin [\![\leftarrow]\!]^{(0)} \cup [\![\rightarrow]\!]^{(0)}$, a new "initial distribution" on the arrival states can be constructed, by computing the passage probabilities from i to the arrival states before leaving $l(0)$. This is done by turning the states in $[\![\leftarrow]\!]^{(0)} \cup [\![\rightarrow]\!]^{(0)}$ absorbing and following the approach in Proposition 4.3.2.

□

We now give an example to illustrate how to derive the step-process and some of the properties mentioned above, as the preceding results are complicated to write down.

Example **4.3.8** Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW with $M := \{1,2,3\}$ and

$$A_2 := \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad A_1 := \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}, \qquad A_0 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix},$$

and let $\Gamma := (0, \frac{1}{3}, 0)$. Let the initial condition be concentrated on 1 in level $l(0)$. The transition graph of this process is given in fig. 4.3.

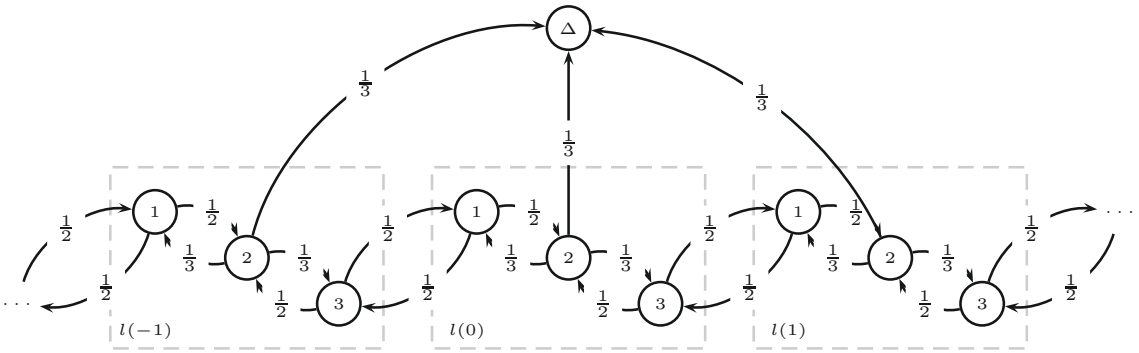


Figure 4.3: Transition graph for a kQRW with three states per level.

Surely $[\![\mapsto]\!] = \{1\}$ and $[\![\leftarrow]\!] = \{3\}$, thus $M_A = \{1,3\} \subsetneq \{1,2,3\} = M$ and furthermore

$$(Id - A_1)^{-1} = \frac{1}{4}\begin{pmatrix} 5 & 3 & 1 \\ 2 & 6 & 2 \\ 1 & 3 & 5 \end{pmatrix}.$$

Accordingly $R_1 = \frac{1}{2}(0,0,1)$ and $R_3 = \frac{1}{2}(1,0,0)$, therefore the hitting probabilities between arrival states of neighboring levels are:

$$\begin{aligned} p_{(k,1),(k+1,1)} &= (1,0,0)(Id - A_1)^{-1}R_1^\top = \frac{1}{8}, \\ p_{(k,3),(k+1,1)} &= (0,0,1)(Id - A_1)^{-1}R_1^\top = \frac{5}{8}, \\ p_{(k,1),(k-1,3)} &= (1,0,0)(Id - A_1)^{-1}R_3^\top = \frac{5}{8}, \\ p_{(k,3),(k-1,3)} &= (0,0,1)(Id - A_1)^{-1}R_3^\top = \frac{1}{8}, \\ p_{(k,1),\Delta} &= (1,0,0)(Id - A_1)^{-1}\Gamma^\top = \frac{1}{4}, \\ p_{(k,3),\Delta} &= (0,0,1)(Id - A_1)^{-1}\Gamma^\top = \frac{1}{4}. \end{aligned}$$

Thus the step-process $(Y_n)_{n\geq 0}$ of $(X_n)_{n\geq 0}$ is defined on $\mathbb{Z} \times \{1,3\} \cup \{\Delta\}$ and characterized by

$$A_0' = \frac{1}{8}\begin{pmatrix} 1 & 0 \\ 5 & 0 \end{pmatrix}, \qquad A_2' = \frac{1}{8}\begin{pmatrix} 0 & 5 \\ 0 & 1 \end{pmatrix}, \qquad \Gamma' = \frac{1}{4}(1,1).$$

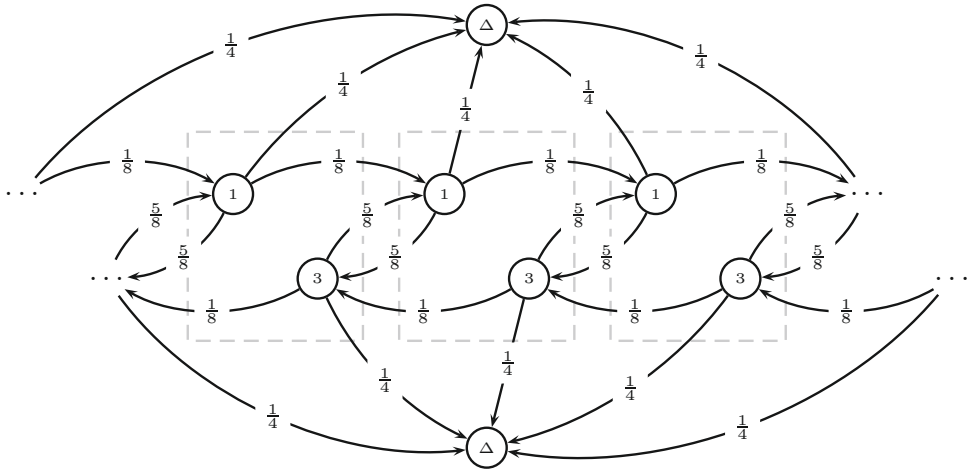


Figure 4.4: For graphical reasons the single state Δ has been split in two.

The transition graph of the step-process is depicted in fig. 4.4.
The reduced Chain associated to the step-process has transition matrix

$$P = \begin{pmatrix} 1 & 0 \\ \Gamma' & A'_0 + A'_2 \end{pmatrix}$$

with

$$A'_0 + A'_2 = \frac{1}{8}\begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}$$

and $R_\Delta = \frac{1}{4}(1,1)$ according to Lemma 4.1.5. It is a very simple three state model with transition graph given in fig. 4.5

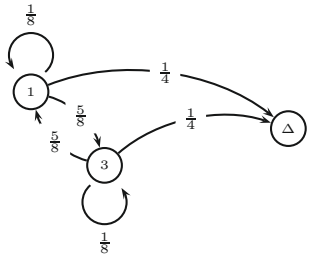


Figure 4.5: Reduced step-process

Now with Proposition 4.3.5 we gain

$$\mathbb{P}(dim(L_0,\ldots,L_{\tau_\Delta-1}) = k) = (1,0)\begin{pmatrix} \frac{1}{8} & \frac{5}{8} \\ \frac{5}{8} & \frac{1}{8} \end{pmatrix}^{k-1}\begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \frac{1}{4}\left(1-\frac{1}{4}\right)^{k-1},$$

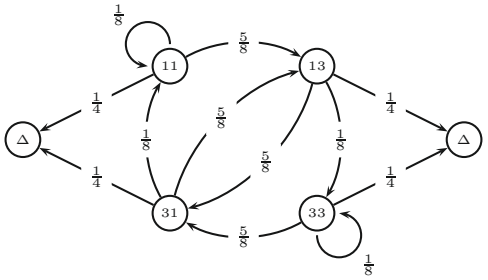


Figure 4.6: The 2-pattern chain. Δ appears twice only for graphical reasons.

which is a geometric distribution to the parameter $\frac{1}{4}$.

The number of changing level to the right in the original process can be calculated as follows. Note that this is the same as the number of transitions to the right in the step process and also the number of transitions from 1 to 1 and from 1 to 3 in the reduced step-process. First we calculate the transition matrix of the 2-pattern chain associated to the *reduced* step-process, see chapter 1 section 3. The result is

$$P' = \frac{1}{8}\begin{pmatrix} 8 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 5 & 0 \\ 2 & 1 & 0 & 5 & 0 \\ 2 & 0 & 5 & 0 & 1 \\ 2 & 0 & 5 & 0 & 1 \end{pmatrix}.$$

We order here like $\Delta, 11, 31, 13, 33$. As the initial condition was concentrate on state $(0,1)$ we have to compute a new initial distribution for the 2-pattern chain. The probability to start in 11 is just the probability of the going from 1 to 1 in a single step, which occurs with probability $\frac{1}{8}$. The other probabilities can be computed in the same way. Note that as the initial condition is concentrated on 1 there is no weight on 33 and 31 initially in the initial distribution for the 2-pattern chain. The initial distribution is therefore $\mu' = \frac{1}{8}(2,1,5,0,0)$. Note that there is a positive probability to start in the absorbing state, as a transition from the starting state 1 of the step-process to the absorbing state Δ is possible. That reflects quantitatively the possibility that there are no forward steps at all before killing.

If we now want to count the number of right transitions, i.e. the number of steps from $l(k)$ to $l(k+1)$, we have to set $M' = \{(11), (21), \Delta\}$ and gain according to Proposition 4.3.6:

$$\begin{aligned} \tilde{S} &= \begin{pmatrix} \frac{1}{8} & 0 \\ \frac{1}{8} & 0 \end{pmatrix} + \begin{pmatrix} \frac{5}{8} & 0 \\ \frac{5}{8} & 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{8} \\ 0 & \frac{7}{8} \end{pmatrix}^{-1} \begin{pmatrix} 0 & \frac{5}{8} \\ 0 & \frac{5}{8} \end{pmatrix} = \begin{pmatrix} \frac{1}{8} & \frac{25}{56} \\ \frac{1}{8} & \frac{25}{56} \end{pmatrix}, \\ \tilde{\Gamma} &= \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} + \begin{pmatrix} \frac{5}{8} & 0 \\ \frac{5}{8} & 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{8} \\ 0 & \frac{7}{8} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{7} \\ \frac{3}{7} \end{pmatrix}, \\ \tilde{\mu} &= \frac{1}{8}(1,5). \end{aligned}$$

Thus

$$\mathbb{P}(\text{number of right steps} = k) = \mathbb{P}(\#\{11, 31\} = k) = \begin{cases} \frac{1}{4} & k = 0 \\ \frac{9}{28}\left(\frac{4}{7}\right)^{k-1} & k \geq 1 \end{cases},$$

which is again very similar to a geometric distribution. It is however a very simple observation,

that the total number of right transitions is dominated by the total number of steps (transitions between levels resp. run-length) before absorption.

□
□

4.4 The last visited level and the minimal and maximal level attained before killing

By the definition of a kQRW the event that absorption takes place has probability one. Theorem 2.2.11 showed how to condition a Markov Chain on leaving the set of transients via a specific set of states. Thus we can easily condition the process to leave via a specific level by identifying the states of a level that have a direct connection to the absorbing state Δ. In chapter 2 we called the collection of those states the boundary, see Definition 2.2.9. Clearly the event that a level $l(z)$ was the last level before absorption regards only the position of the Random Walker. We can thus restrict the investigation to discrete time.

Proposition 4.4.1 *Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW. Define R_k as vector of countable size containing only zeros except for the entries associated with $l(k)$ where it coincides with Γ, the m-dimensional vector of killing rates. Let the initial distribution μ be concentrated on $l(0)$. Then the last visited level $L_{\tau_\Delta - 1}$ has distribution:*

$$\mathbb{P}_\mu(L_{\tau_\Delta-1} = k) = \mu(Id - S)^{-1}R_k^\top$$

holds.

Proof That the last level is $l(k)$ is the condition that the process leaves via the states in $l(k)$. The statement is thus a consequence of Theorem 2.2.11, chapter 2, where we introduced conditioning on leaving via some designated states.

■

Remark **4.4.2** Although the distribution of the last level before killing has an abstract form, i.e. an infinite matrix needs to be inverted, we compute this distribution in Appendix C for the simple case if $|M| = 1$. This is the case of a usual killed Random Walk we there identify $\mathbb{Z} \times M$ with $\mathbb{Z}$, in other words $l(k) = k$.

□

The distribution of the maximum and the minimum before killing is easily calculated, if M contains only one phase, i.e. $E = \mathbb{Z} \times \{\Delta\}$. It is a killed Random Walk on $\mathbb{Z}$. The probability of starting from $l(0)$ and reaching $l(1)$ for the first time before killing is the probability p_+ that $l(1)$ is the new maximum. After that a further first passage to the right would make $l(2)$ be the new maximum. But by the spatial homogeneity of the (here: discrete time) kQRW both probabilities are equal. Therefore it is clear that in this case the maximum has geometric distribution with parameter $1 - p_+$. However computing the exact value of p_+ even in this simple case is not simple, for this reason we have deferred this computation to Appendix C.

The situation changes if $|M| \geq 2$. Then the initial distribution on $l(0)$ comes into play and the argument, that the passage probability from $l(0)$ to $l(1)$ is the same as from $l(1)$ to $l(2)$ does not hold anymore. Fortunately we still can describe this case with similar arguments.

Proposition 4.4.3 *Let $(X_n)_{n\geq 0}$ be a (discrete time) kQRW with initial distribution μ concentrated on $l(0)$. Then the distribution of the maximal level attained before absorption is given as*

$$\mathbb{P}(max\{L_0, L_1, \dots, L_{\tau_\Delta - 1}\} = k) = \begin{cases} (1-p_{k+1})p_k & k \geq 1 \\ 1 - p_1 & k = 0 \end{cases} \tag{4.5}$$

with $p_i = \mu((Id - S_{|k\leq 0})^{-1}A)^i$ and $S_{|k\leq 0}$ the matrix gained from S by replacing in S as defined in Definition 4.1.1 all entries associated to positive levels with zeros and accordingly A is gained from S by deleting all entries except the A_0 at level $l(1)$.

Proof Suppose we change the states of $l(1)$ to absorbing states. Then we can calculate the probability distribution of being absorbed in a state of $l(1)$ before killing by

$$\nu := \mu(Id - S_{|k\leq 0})^{-1}A,$$

where ν is then distribution on $l(1)$. This is a consequence of Corollary 2.2.3, chapter 2, where we characterized absorption in a specific subset of absorbing states. That means also p_1, the probability to reach any of the states in $l(1)$ is given as $\nu\mathbf{1}^\top$.

Starting now in $l(1)$ and taking $l(2)$ absorbing under start with the new distribution ν is by spatial homogeneity the same as starting with ν (shifted to $l(0)$) and taking $l(1)$ absorbing as before. Now the new distribution on $l(1)$ (or $l(2)$) has the form

$$\nu(Id - S_{|k\leq 0})^{-1}A = \mu((Id - S_{|k\leq 0})^{-1}A)^2$$

and thus

$$p_2 = \mu((Id - S_{|k\leq 0})^{-1}A)^2\mathbf{1}^\top.$$

In this interpretation it is clear that by iteration $p_i = \mu((Id - S_{|k\leq 0})^{-1}A)^i$ is a probability distribution on level i. Then the probability that $l(k), k \geq 0$ was the maximum level attained before killing, is the probability to reach at some point level k and then being killed before a first visit in $k+1$, which is exactly $1 - p_{k+1}$ by spatial homogeneity.

■

The minimum level ever attained has naturally the same form as (4.5), but with A_0 and A_2 exchanged.

The time to start in $l(0)$ and attain $l(1)$ resp. $l(-1)$ for the first time is known in the literature as first *ascending* resp. . *descending ladder epoch*. Especially for Random Walks this is very well known and relates to many interesting properties. For a good and exploiting overview see e.g. the book "Stopped Random Walks" by Gut ([Gut09]). We make some additional notes on the relations with ladder epochs in Appendix C, where we give exact expressions for minimum and maximum level distribution before killing.

The methods applied there also give an approximation scheme for the general case. This approximation is done by restricting the process to a finite number of levels symmetrically around $l(0)$, say $l(-n), l(-n+1), \dots, l(n-1), l(n)$, and alternating the transitions at $l(n)$ and $l(-n)$ in such a way that absorption is ensured whenever the Walker tries to leave that restriction. That means instead of reaching $l(n+1)$ or $l(-n-1)$ the Walker is redirected to Δ. The restriction can then be described by a finite state version of the original Chain by removing all non reachable states from the state space. This makes it possible to evaluate the quantities

given in Propositions 4.4.1 and 4.4.3 numerically, but it is not immediately clear how to choose a proper n. Note also that the derived expressions are meant for infinite matrices, thus a rescaling is necessary after the restriction to a finite number of levels.

However, for the killed Random Walk we can take the limit $n \to \infty$ to get the distributions of maximum level visited and last level before killing as weak limit of the restricted, finite state versions.

The following remark indicates that there is "reduction" trick to circumvent the inversion of an infinite matrix for the distribution of the maximal distances from zero and the last level before absorption.

Remark **4.4.4** One could ask whether there is a finite modification (i.e. a finite state model) which describes the first passage probability from $l(0)$ to $l(1)$. That this is not possible is a consequence of automaton theory and regular languages. As it is not in the scope of this thesis to relate automaton theory with absorbing Markov Chains, we just want to give a sketch of a possible proof. For every absorbing Markov Chain with a finite set of states we can interpret the sequence of transition probabilities between transient states as a word over the alphabet $\Sigma := \{p_{ij}, i, j \in E \setminus \{\Delta\}\}$. That each word is finite is assured if we assume that absorption is certain, this holds for instance if the transient set is a single communication class. Then the computer scientists call the set of all words L generated by the Markov Chain the associated *language* L. Markov Chains are in this interpretation finite automatons. The theory now states that L generated by a absorbing Markov Chain is a so called *regular language*; these languages are in a sense languages with words that have no memory about how many letters appear before a specific one (analog to the Markov Property). If we now want to compute the probability of the passage from $l(0)$ to $l(1)$ before killing, say in the very simple case $|M| = 1$, then each trajectory starting at 0 and arriving at 1 after a number of transitions would make left and right transitions, denoted here with $+$ and $-$, in a very specific way. The total number of $+$ is one more than $-$, but while $--+++$ is a correct sequence (going 2 left, 2 right and then going to level 1) the sequence $++-+-$ is not (the first transition leads already to level 1), although the number of left and right transitions fits. Such expressions (up to the last $+$) are known as *correct bracket expression*, for instance "$--++-+$" can be assigned to "(())()". It is known that such expressions can not be generated by finite automatons, as these expressions do *not* form a regular language. That is (after some reasonable effort to proof the "equivalence" of absorbing Markov Chains and regular automatons) the reason that there is *no* finite state Markov Chain allowing to compute the first passage to $l(1)$ starting in $l(0)$, easily in particular not in the general case of $|M| > 1$. Thus there is no (easy) way to circumvent the inversion of an infinite matrix in this setting.

The theory that belongs to this argumentation can be found in [EP02] and references therein and has also some applications in the optimization of pattern matching chains, see e.g. [Nue08]. □

Chapter 5

Modelling of Molecular Motors and applications to Kinesin V

Every living being consists at least of one cell. As Hirokawa and Takemura write in [Sch03], p. 79 (shortened), a cell is quite complex:

> *A cell is not simply a bag filled with cytoplasmic fluid surrounded by the plasma membrane in which membranous organelles ... float and through which newly synthesized proteins diffuse to reach their destination. Instead, cells transport and sort proteins and lipids after their synthesis to their proper destinations at appropriate velocities in membranous organelles and protein complexes using various kinds of motor proteins.*

Those motor proteins are also called *molecular motors*. They fulfill several transport tasks, among others carrying cargoes from one end of the cell to another. We consider one of them, *kinesin*, in the following and apply the tools developed in the preceding sections to a physicists model of this protein.

Kinesin is a relatively large protein, $\sim 100nm$ in size and consists of several parts, see also fig. 5.1.

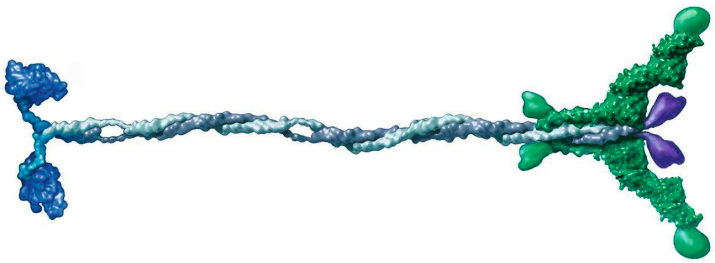


Figure 5.1: The two heads (dark blue) on the left attach to the microtubule and function like foots walking on the tubule. The right clamp-like structure attaches to cargoes. The picture is part of fig. 1 in [Val03].

This molecular motor has two heads, that actually work like foots. They are called "heads" because the chemical reactions that drive the process happens there; it is a usual chemists notion for chemical active parts.

Kinesin walks (in a stochastic fashion) along little roads, called *micro-tubules*, that are spanned between cell core and cell rim. These roads themselves are polymers, made out of several identical parts. These building blocks have a size of 8nm. Each of kinesin's heads can attach to exactly one of this building blocks and walks step by step on the filament in a "hand-over-hand" fashion, see [YTVS04, CC05]. Thus the steps kinesin makes have an equal constant size of 8nm. The clamp like structure on the other end of the motor can bind to a cargo. By stepping along the filament the cargo is pulled in one direction, as stepping forwards is very probable, i.e. backwards step are very rare events under normal biological conditions, unless an additional force pulls them into another direction.

The mechanism how kinesin actually walks along the tubules is known to work as follows. Each of the heads has a so called "binding pocket". This is the place where adenosine triphosphate (ATP) can be hydrolyzed. The ATP is used in many processes and can be thought of as "fuel". If a binding pocket is occupied by an ATP molecule the head can bind very strongly to the tubules; this state is called "T". Now the ATP can be broken into adenosine diphosphate (ADP) and an (anorganic) phosphate ion (that is the hydrolyzing process). After the release of the phosphate the head is in a state called "D", where a head is only weakly bound to the filament. The further release of ADP brings the head into a state called "E", where now the binding pocket is empty and the head is again strongly bound to the filament. If a new ATP arrives the head can return into state T and the cycle can begin anew. As every chemical reaction is reversible, the process can run also in the other direction, depending on the concentrations of ATP, ADP and phosphate. However, under normal conditions reversing the process is not very probable. As each kinesin has two heads this makes a total number of 9 possible combinations of states. The actual stepping can take place if one head is in state D and one in state T (that would be state TD or DT). Then the head in state D has just broken up an ATP into ADP and an phosphate ion, which is then released. The energy that gets freed by the consumption of ATP can be used to rip off the weakly bound head from the filament and push it over the other strongly bound one, where it can bind again to the filament. The motor has then made a single step. This process works in both directions along the filament, but most of the motors have a preferred direction of movement.

In [LL07] the authors identify for kinesin's stepping six essential states (DT, TD, ET, TE, DE, ED) for the walking mechanism, under the condition that the motor never unbinds from the filament completely. They state that the states EE, DD and TT are never reached. We might remark that, as computed in [VLL08], this model shows Time Duality. This can also be proven by Corollary 3.2.10 from chapter 3; we omit the details here.

The six state model of [LL07] can be extended by the possibility of unbinding from the filament when the motor attains state EE, see [LL08]. Including unbinding is a more natural situation and fits nicely into the introduced framework of killed Quasi-Random Walk (kQRW).

We introduce the model of kinesin here as a kQRW $(X_t)_{t\geq 0}$ on

$$E := \mathbb{Z} \times \{1, 2, 3, \ldots, 7\} \cup \{\Delta\}$$

by its transition graph, see fig. 5.2, the details about what states code which motor head configuration are given in [LL08]. The resulting matrices A_i and Γ are given as follows:

$$A_2 := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{52} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \; A_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{25} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_1 := \begin{pmatrix} 0 & w_{12} & 0 & 0 & 0 & w_{16} & w_{17} \\ w_{21} & 0 & w_{23} & 0 & 0 & 0 & w_{27} \\ 0 & w_{32} & 0 & w_{34} & 0 & 0 & 0 \\ 0 & 0 & w_{43} & 0 & w_{45} & 0 & w_{47} \\ 0 & 0 & 0 & w_{54} & 0 & w_{56} & w_{57} \\ w_{61} & 0 & 0 & 0 & w_{65} & 0 & 0 \\ w_{71} & w_{72} & 0 & w_{74} & w_{75} & 0 & 0 \end{pmatrix},$$

and

$$\Gamma = (0, 0, 0, 0, 0, 0, w_{7\Delta}),$$

see fig. 5.2 for the definition of the transition rates w_{ij}.

We assume that the initial condition is concentrated on state 7 in level 0 and denote that by $\mu_0 = e_7$. This is assumed to be the state when the motor arrives at the tubule for the first time.

The introduced transition rates w_{ij} depend on concentrations of ATP, ADP, (anorganic) phosphate P and a force F^*. We choose here for these quantities the following values:

$$[ATP] = 10\mu M, \; [ADP] = 0.5\mu M, \; [P] = 0.5\mu M, \; F^* \in \{0, 10\},$$

(where the brackets denote concentrations). In experiments F^* is a force that pulls the motor against its preferred walking direction, to force the motor to make back steps which are else wise very seldom and thus difficult to observe, see e.g. [SSSB93]. For the following computations we shall consider the undisturbed case of $F^* = 0$ and the case $F^* = 10$.

5.1 Lifetime

According to Proposition 4.2.1, the distribution of the life time τ_Δ is given as a PH-type distribution with representation

$$\tau_\Delta \sim PH(e_7, A_0 + A_1 + A_2)$$

and thus

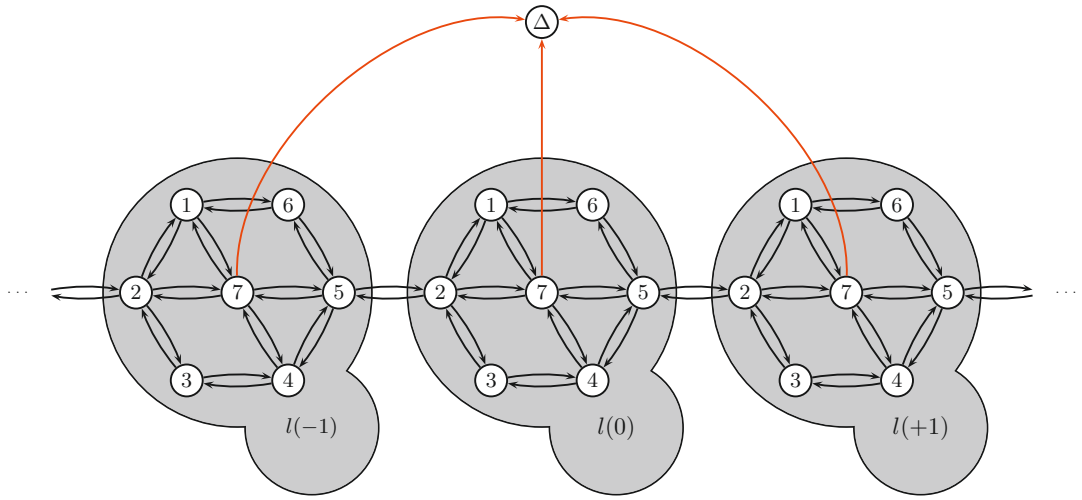
$$F_{\tau_\Delta}(t) = 1 - e_7 \exp((A_0 + A_1 + A_2)t)\mathbf{1}^\top$$

and

$$f_{\tau_\Delta}(t) = e_7 \exp((A_0 + A_1 + A_2)t)\Gamma^\top$$

according to Table 2.2. A numerical evaluation shows that F_{τ_Δ} is not exponential as fig. 5.3 (left side) would suggest. A closer look at the density (same figure right side) unveils a "kink" which is atypical for exponential distributions. The expectation of the killing resp. absorption time is

	$F^* = 0$	$F^* = 10$
$\mathbb{E}(\tau_\Delta)$	$6.0657s$	$4.127s$



$$
\begin{array}{lll}
w_{56} = \frac{200}{1+\exp(0.15F^*)}, & w_{12} = \frac{4[ATP]}{1+\exp(0.25F^*)}, & w_{25} = 2.9^5 \exp(-0.65F^*) \\
w_{65} = 0.06\frac{[ADP]}{1+\exp(0.15F^*)}, & w_{16} = 0.02\frac{[P]}{1+\exp(0.15F^*)}, & w_{21} = \frac{200}{1+\exp(0.25F^*)}, \\
w_{52} = 0.36\exp(0.35F^*), & w_{54} = \frac{3.08\times10^{-10}}{1+\exp(0.25F^*)}, & w_{17} = 2.8\frac{[ADP]}{1+\exp(0.15F^*)} \\
w_{71} = \frac{100}{1+\exp(0.15F^*)}, & w_{75} = \frac{1.2\times10^{-4}[P]}{1+\exp(0.15F^*)}, & w_{7\Delta} = 3\exp(0.1F^*) \\
w_{57} = w_{71}, & w_{61} = w_{56}, & w_{23} = w_{56} \\
w_{34} = w_{61}, & w_{45} = w_{12}, & w_{32} = w_{65} \\
w_{43} = w_{16}, & w_{27} = w_{57}, & w_{72} = w_{75} \\
w_{74} = w_{71}, & w_{47} = w_{17} &
\end{array}
$$

Figure 5.2: A transition from i to j exists if an arrow connects them in the graph. The transition rate is denoted by w_{ij} and given below the graph. The exact values of the various parameters are given in [VLL08, LL08].

5.2 The step-process and derived quantities

The notion of steps is quite important in the analysis of motor models. Surely the state at which the motors has completed a step (there are two possibilities, namely TD or DT) has an influence of the further development. Once the motor has completed a step forwards it is in state TD, but to make another forwards step, it has to return to the state DT of the same level before it can make a step to the right. Otherwise it is not possible to toss the trailing head forwards.

The step-process introduced in chapter 4, see Definition 4.3.4, is designed to take care of this form of "correlation". Omitting the details of computation we give the resulting quantities $A'_0, A'_1, A'_2, \Gamma'$ for $F^* = 0$ and $A''_0, A''_1, A''_2, \Gamma''$ for $F^* = 10$ that describe the step-process on $E = \mathbb{Z} \times \{2, 5\}$ (2 and 5 are the states that are reached first by switching level, we called these states in chapter 4 *arrival states*, see also fig. 5.2, thereby $2 \equiv TD$ and $5 \equiv DT$):

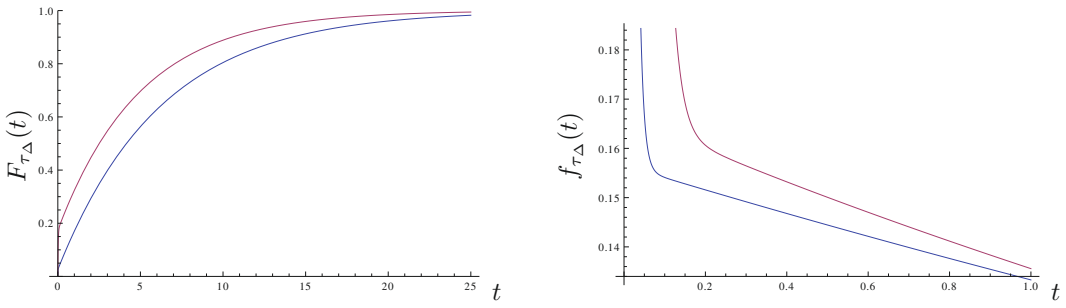


Figure 5.3: Distribution function (left) and density (right) of τ_Δ, for $F^* = 0$ (blue) and $F^* = 10$ (red). Note that the distribution functions are both zero at $t = 0$, but grow very fast near zero. The reason is a kink in the density (see left).

$F^* = 0$	$F^* = 10$
$A'_0 = \begin{pmatrix} 0 & 2.8938 \times 10^{-3} \\ 0 & 1.2381 \times 10^{-6} \end{pmatrix}$	$A''_0 = \begin{pmatrix} 0 & 0.2190 \\ 0 & 0.0193 \end{pmatrix}$
$A'_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$A''_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$
$A'_2 = \begin{pmatrix} 0.9842 & 0 \\ 0.9999 & 0 \end{pmatrix}$	$A'_2 = \begin{pmatrix} 0.6986 & 0 \\ 0.9644 & 0 \end{pmatrix}$
$\Gamma'^{\top} = \begin{pmatrix} 0.0129 \\ 1.1423 \times 10^{-5} \end{pmatrix}$	$\Gamma''^{\top} = \begin{pmatrix} 0.0825 \\ 0.0163 \end{pmatrix}$

Thus the transition matrices P' and P'' of the reduced model are given as:

$$P' := \begin{pmatrix} 1 & 0 \\ \Gamma' & A'_0 + A'_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1.2928 \times 10^{-2} & 0.9842 & 2.8938 \times 10^{-3} \\ 1.1423 \times 10^{-5} & 0.9999 & 1.2381 \times 10^{-6} \end{pmatrix} \tag{5.1}$$

and

$$P'' := \begin{pmatrix} 1 & 0 \\ \Gamma' & A''_0 + A''_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.0825 & 0.6986 & 0.2190 \\ 0.0163 & 0.9644 & 0.0193 \end{pmatrix} \tag{5.2}$$

by the discrete time version of Lemma 4.1.5.

We give all computed quantities in Table 5.1.

As we shortly discussed in Remark 4.3, the initial distribution on state 7 of the original model makes it necessary to compute a new initial distribution μ on the arrival states $2, 5$ which reflect the probabilities to start in state 7 and to reach $2, 5$ and Δ before a new level is reached. We do this by setting 2 and 5 absorbing and calculating the passage probabilities starting from state 7. Note that this implies an initial weight on the absorbing state Δ. This reflects the event that a molecular motor detaches itself from the tubule without making any step.

The transition matrix P' is readily seen to privilege steps to the right with a very high probability (the matrix A'_2 contains non-zero probabilities *very* close to 1). Thus the run-length distribution is identical with the maximal level distribution, the number of forward steps (transitions $(k, 2) \to (k + 1, 2)$ and $(k, 5) \to (k + 1, 2)$) and also with the distribution of the

	$F^* = 0$	$F^* = 10$
expected lifetime $\mathbb{E}(\tau_\Delta)$	$6.0657s$	$4.1247s$
initial distribution μ	$(0.0302, 0.4849, 0.4849)$	$(0.1951, 0.4025, 0.4025)$
expected number of steps	75.7248	11.7535
expected run-length	$605.7984nm$	$94.0279nm$
velocity ν	$100\frac{nm}{s}$	$23\frac{nm}{s}$

Table 5.1: All values are rounded.

last level visited before absorption. Therefore we compute only the run-length for $F^* = 0$. The preceding problem does not exist for $F^* = 10$, thus we discuss this case separately.

The number of steps the motor does before it detaches from the tubule is distributed as the absorption time of $(Y_n)_{n\geq 0}$ generated by (μ, P') resp. (μ, P''), where μ is the new initial distribution, see Table 5.1 for $F^* = 0$ resp. $F^* = 10$. Thus the absorption time τ_Δ has a discrete PH-type distribution with representation $PH(\mu, A_0' + A_2')$ resp. $PH(\mu, A_0'' + A_2'')$. The distribution of the run-length is depicted in fig. 5.4.

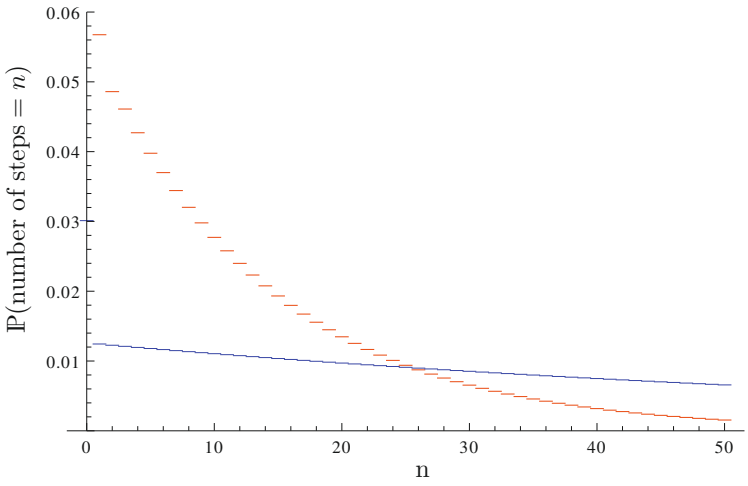


Figure 5.4: Distribution of the (total) number of steps before absorption for $F^* = 0$ (blue) and $F^* = 10$ (red). Note that the value in 0 for the distribution drawn in red is about 0.2 and not displayed.

As the distance a single step bridges is exactly $8nm$ for kinesin the expected value of the run-length is the expected number of steps times this distance, see Table 5.1.

The ratio between expected run-length and expected life time is called "velocity" and denoted here with ν.

Without further note we fix now $F^* = 10$ and continue applying the theoretical results from chapter 4. As mentioned before Remark 4.4 in chapter 4, we have to approximate some expressions as we cannot numerically invert an infinite matrix. The approximation is carried out by taking only a finite number of levels into account, symmetrical around zero, applying the results of chapter 4 and then rescaling to gain a probability distribution. This is an approximation (actually a truncation) which works quite good and fits to the results of simulations of the original process (not shown here).

The distributions of the maximal and the minimal distance from zero to the right and left are depicted in fig. 5.5 and 5.6, see also Proposition 4.4.3 in chapter 4. The transition matrix of the step-process P'' given in (5.2) shows that transitions $(k, 2) \to (k-1, 5)$ have a probability of about 0.2 (contrastingly, in the case of $F^* = 0$ this was almost zero). Such a transition would be a backwards step after a forwards step (a left step after a right step). Still, the probability to make another step to the left is very small (about 0.01). Thus the minimal level ever visited should be very close to zero; this is confirmed by fig. 5.6. For the same reason the distribution of the maximal level ever visited is broader, see fig. 5.5.

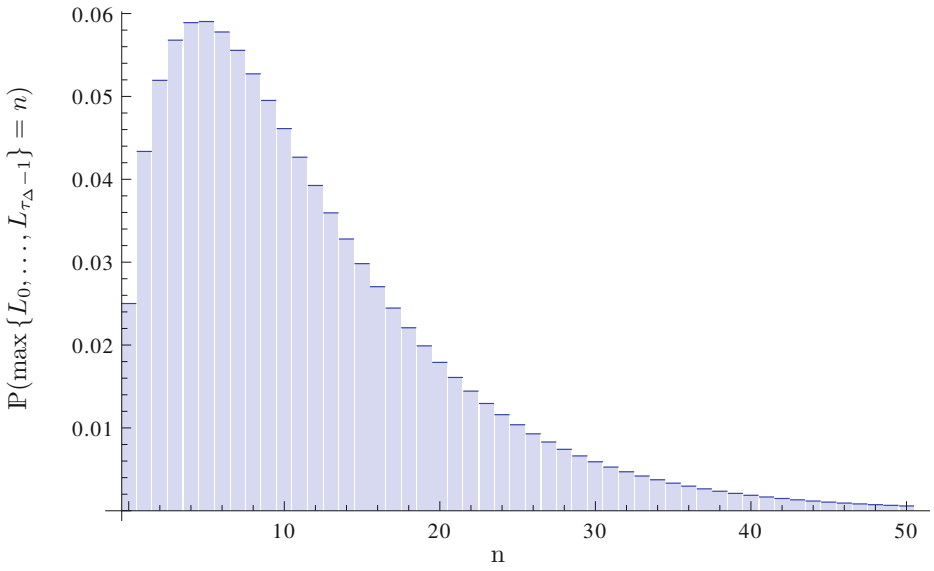


Figure 5.5: Approximated distribution of the maximum level before absorption.

The distribution of the position at which the motor unbinds can be calculated by Proposition 4.4.1 with the same approximation scheme as before. As already mentioned, the probability to make two steps to the left in sequence is very low, hence the level from which detachment takes place should be positive with high probability. Fig. 5.7 confirms that heuristic assumption.

Fig. 5.8 shows the convergence behavior of the approximation.

We have seen, therefore, that some of the tools developed in chapter 4 allow to derive exact distributions. Most of these distributions, especially the run time, run length and the maximum distance have never been computed before, to the best of our knowledge. As these distributions should be experimentally accessible, they may provide a further level of verification of the model assumptions. From the scope of this work, we thus provide useful mathematically rigorous tools to extend the range of predictability of present and future models for molecular motors.

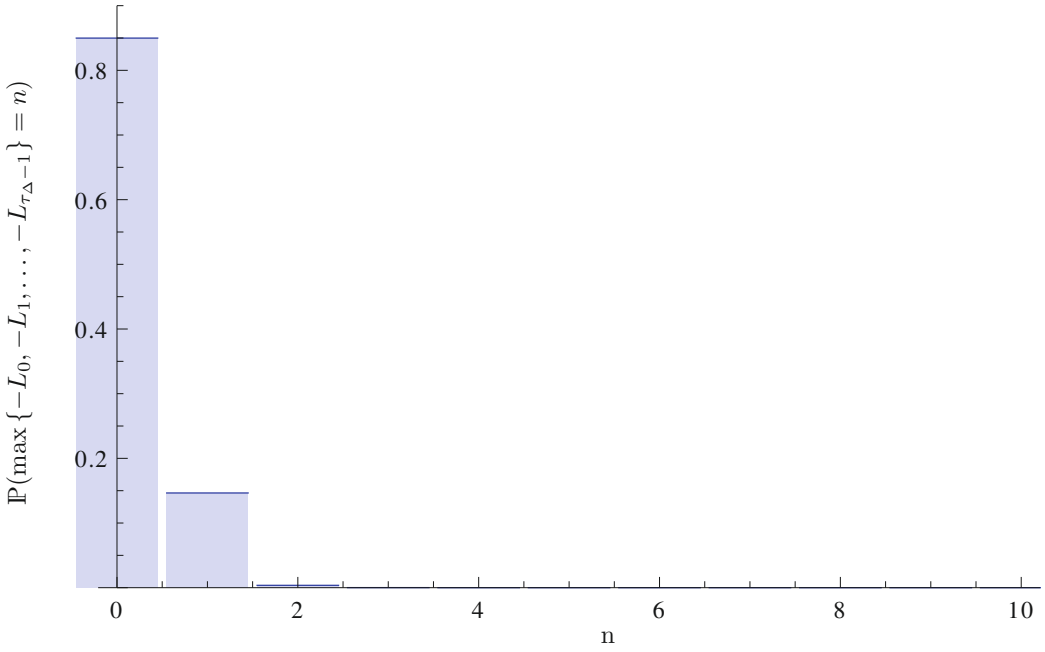


Figure 5.6: Distribution of the minimum level before absorption. Note that here the absolute of the maximal distance from zero to the left is depicted, i.e. the maximum distribution when A_0 and A_2 are exchanged.

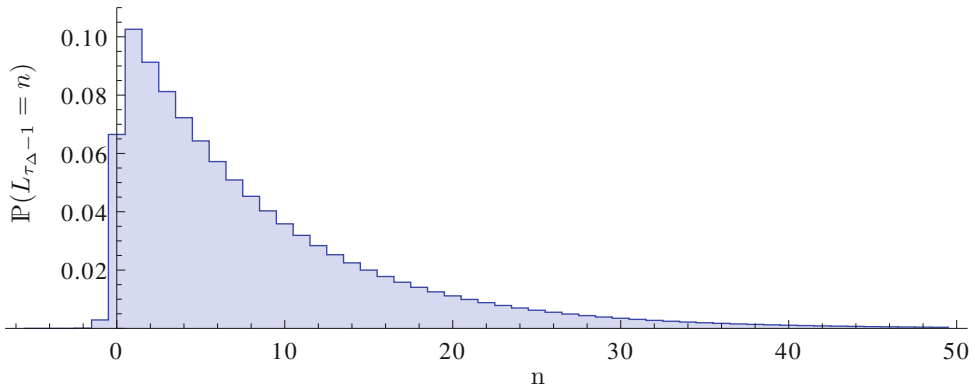


Figure 5.7: Approximated distribution of the last visited level before absorption.

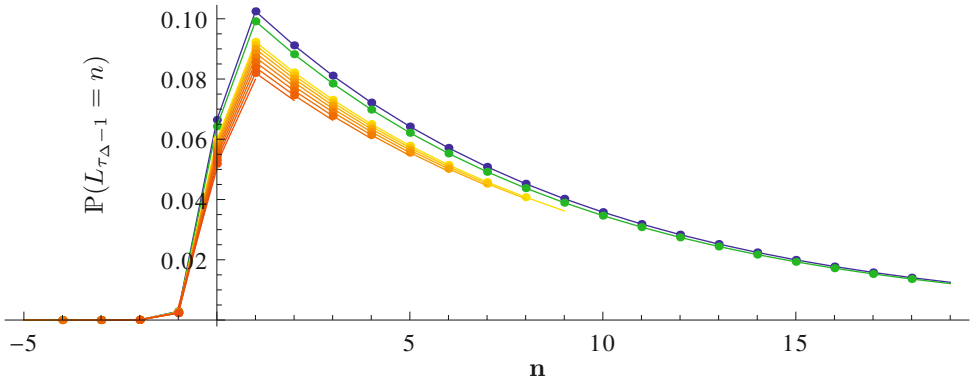


Figure 5.8: Plot of the approximation of the last level visited for $2n + 1$ levels and $n \in \{1, 2, \ldots, 9, 10, 20, 50\}$ (from red to blue). Note that here convergence is quite good and errors appear only at the artificial boundaries introduced for approximation (not shown). For $n > 50$ the resulting curve does not differ significantly form the blue one.

Chapter 6

Perspectives and Outlook

In this section we want to give a short overview what kind of questions did remain open and could be starting point for further investigations.

6.1 Time Duality in Semi-Markov Chains

In chapter 3 we have introduced a relation between two states i and j named Time Duality. This relation describes (roughly spoken) that the passage from i to j without return to i has the same distribution as the passage from j to i without return to j.

The investigation of Time Duality relation was inspired by the papers [LW07, VLL08]; in the first paper Time Duality was discussed for several possible motor models and in the second Time Duality was found in an explicit computation for kinesin. We gave in this thesis a rigorous mathematically treatment of this phenomenon based on the Laplace transforms of the *pure passage times*, see sections 1 and 2 of chapter 3. The main advantage of this treatment is that Time Duality can be reduced to a simple condition involving only elements of inverses of finite matrices. This contrasts with the proof in [LW07], which involves (implicitly) a deep graph theoretic argument about gradient functions which can be found e.g. in [BL92].

None of these methods gives an intuitive explanation about the reason why Time Duality appears. Although we identified several geometrical settings for Time Duality, see chapter 3, section 3.1.3, and also made connections to reversibility and the newly introduced permuted balance, the examples in section 3.3.2 are not very intuitive in the following sense.

We continue the explanations with Example 3.1, see fig. 6.1.

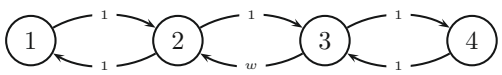


Figure 6.1: Transition graph of Example 3.1, in chapter 3, for arbitrary $w > 0$.

We identified that the pure passage times τ^*_{14} and τ^*_{41} have PH-type distribution with repre-

sentations as follows:

$$\tau_{14}^* \sim PH(\delta_2, S_{|4})$$
$$\tau_{41}^* \sim PH(\delta_3, S_{|1})$$

where

$$S_{|1} = \begin{pmatrix} -2 & \frac{w}{1+w} \\ 1+w & -1-w \end{pmatrix}, \qquad S_{|4} = \begin{pmatrix} -2 & 2 \\ \frac{w}{2} & -1-w \end{pmatrix},$$

and that

$$\tau_{14}^* \stackrel{d}{=} \tau_{41}^* \stackrel{d}{=} X + Y$$

where X, Y two independent exponentially distributed random variables. The parameters are 1 and $2 + w$, which are the eigenvalues of $-S_{|1}$ and $-S_{|4}$. Surprisingly the embedded Markov Chains of these two MCc's conditioned to be absorbed in a designated state are identical (up to the order of columns and rows), while $S_{|1}$ and $S_{|4}$ are not. The equivalence of the embedded Markov Chains of the conditioned processes thus seems to be a necessary condition for Time Duality.

Time Duality is indeed the comparison of absorption times. Unfortunately two absorption times equal in distribution can be generated by very different Markov Chains, especially the number of states can be different. In [O'C99] some problems and open questions are discussed, amongst others the question of the minimal number of states necessary to represent a given PH-type distribution. One way to solve such questions is purely algebraic and proposed in [Ryd96, IAK92]. The conditions given in the papers involve twelve matrix conditions which are difficult to understand in a probabilistic setting.

In chapter 4 we have introduced a special (discrete time) kQRW, the so called *step-process*, see Definition 4.3.4. A continuous version, where transitions between states are not exponential but given by the step times mentioned in Remark 4.3 is very interesting to the experimentalists, as they can observe only this process, for an example see [LW07] and [VLL08]. Such a process is a so called *Semi Markov Chain*, which is defined as follows. We give here only a rough introduction adapted from the introduction to Semi Markov Chains in discrete time as given in [BL08]. Details and applications (also for continuous time) can be found e.g. in [JM06, LO01].

Definition 6.1.1 *Let $(Z_t)_{t\geq 0}$ be a stochastic process on the countable state space E. Let $(J_n)_{n\geq 0}$ be the sequence of visited states, $(S_n)_{n\geq 0}$ the sequence of jump times and $(X_n)_{n\geq 1}$ the sequence of sojourn times in each state J_n; note that $X_n = S_{n+1} - S_n$. Then $(Z_t)_{t\geq 0}$ is a* homogeneous Semi-Markov Chain *if and only if*

$$\begin{aligned}\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | (J_0, \ldots, J_n) = (j_0, \ldots, j_{n-1}, i); (S_0, \ldots, S_n) = (s_0, \ldots, s_n)) \\ = \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n = i) \\ = \mathbb{P}(J_1 = j, S_1 - S_0 \leq t | J_0 = i) = \mathbb{P}(J_1 = j, X_0 \leq t | J_0 = i)\end{aligned}$$

holds whenever the condition has probability greater zero for states $j_0, \ldots, j_{n-1}, i, j \in E$ and $s_0 < s_1 < \ldots < s_n$. The matrix-valued function $Q(t) = (q_{ij}(t))_{i,j\in E}$ defined via

$$q_{ij}(t) := \mathbb{P}(J_1 = j, X_0 \leq t | J_0 = i)$$

is called Semi-Markov Kernel.

Very much like in the Markov Chain regime we have a kind of independence from the past, but the strong Markov Property does not hold anymore. Instead the Markov Property holds only at jump times of the chain; at all other points the process "feels" a dependency on the time how long it already stayed in a state; this is usually called "aging". Thus the name *Semi*-Markov Chain.

Without proof we just remark (see references above) that

$$q_{ij}(t) = p_{ij}F_{ij}(t)$$

where p_{ij} are the transition probabilities of the (embedded) homogeneous MCd $(J_n)_{n\geq 0}$ and $F_{ij}(t)$ are distribution functions.

Like in the Markov case it would be interesting to know how Time Duality would be defined in the Semi-Markov regime. The question seems to be quite difficult. The main tool in the Markov case was to determine how a MCc conditioned to absorption in a specific state can be characterized. We used Lemma 2.2.2, which characterizes the conditioned process by a linear transform of the original infinitesimal generator. The computation in the said lemma clearly shows that the Markov Property plays a fundamental role.

In [LO01] Chapter 5, section 4 there is an interesting result regarding the notion of *reliability*. Adopting the notation they use, we define

$$R_i(t) := \mathbb{P}_i(\{\forall u \in [0,t] : Z_u \in E_0\})$$

the probability to be in a chosen subset $E_0 \subsetneq E$, $|E_0| = r$, up to time t starting in $i \in E_0$. According to Proposition 5.3 in [LO01] this quantity can be calculated as follows:

$$R_i(t) = e_i(Id - Q_{00}(t))^{(-1)} * (Id - H_0(t))\mathbf{1}_r^\top \tag{6.1}$$

where $Q_{00}(t)$ is the sub matrix of the Semi-Markov Kernel associated to the states in E_0, H_0 is the diagonal matrix containing the survival probability in each state of E_0 up to time t on the diagonal and zero elsewhere and $\mathbf{1}_r^\top$ the r-dimensional vector containing only ones. In contrast to Markov Chains there appears a convolution inverse, which is characterized by

$$(Id - Q_{00}(t))^{(-1)} = \sum_{n\geq 0}(Q_{00}(t))^{(n)}$$

where $Q_{00}^{(n)}(t)$ is the n-fold matrix convolution product of $Q_{00}(t)$ with itself. This latter equation is the convolution analogue to the van Neumann series for non-negative matrices with spectral radius strictly smaller than 1.

There is a remarkable similarity to the absorption probabilities given in Proposition 4.3.2 in the form of the expressions. Indeed the proof of this expression characterizes $R_i(t)$ as the probability that the first hitting time of E_0^c is greater than t.

How exactly this can be used for conditioning a Semi Markov Chains is not clear, but it seems feasible that reliability is a good starting point.

6.2 Extension to coupling of killed Quasi Random-Walks in the modeling of Molecular Motors

In living cells cargoes are not transported by just a single molecular motor but rather by a whole collection of possibly even different motor types. This makes both the mathematical and experimental investigation more difficult, see e.g. [KL05, KML06] and references therein.

6.2. EXTENSION TO COUPLING OF KILLED QUASI RANDOM-WALKS IN THE MODELING OF MOLECULAR MOTORS

Investigations of possible models of cooperative motor transport begin usually with the case of two motors bound to one cargo. As the motors are assumed to be irreversibly attached to their cargo this regime implies some additional forces acting on the motors. For instance there is a symmetric force between the leading motor and the one behind, pulling the second motor forwards and the leading one backwards, where the strength of the force depends on the distance of both motors. This force is usually modeled via a spring between the two motors or between each motor and the cargo, see e.g. [BKM$^+$11, BKKL12]. As in chapter 5 for the kinesin the transition rates between the chemical states depend on this force. Furthermore the structure of the microtubule the motors walk on implies that the steps are of equal constant size and thus the forces that the pair of motors can only take a finite number of values. There also exists a maximal force a motor can stand before it is ripped off the tubule.

One could define an extended version of a motor model by replacing the set of phases M (aka the set of chemical states) by $M' := M \times F$, where F is a set of possible values for the force. This only extends the number of phases while the results of chapter 4 are still applicable, i.e. the extension is still a Quasi-Random Walk with killing on $\mathbb{Z} \times M'$.

Coupling is a useful method to model pairs of molecular motors mathematically, see e.g. [Lin92] and for a series of nice examples and algorithmic applications [LPW09, Hae08]. Such a coupling (between two motors) would be defined as follows:

Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be two Quasi Random Walks with killing. Then the vector $(Z_t)_{t\geq 0} = (Z_t^{(1)}, Z_t^{(2)})_{t\geq 0}$ is called *coupling of* $(X_t)_{t\geq 0}$ *and* $(Y_t)_{t\geq 0}$ if and only if

$$Z_t^{(1)} \stackrel{d}{=} (X_t)_{t\geq 0},$$
$$Z_t^{(2)} \stackrel{d}{=} (Y_t)_{t\geq 0}.$$

The definition of coupling allows a lot of freedom for the dependence between $(Z_t^{(1)})_{t\geq 0}$ and $(Z_t^{(2)})_{t\geq 0}$. It is therefore possible to change the value of the force (in the extended motor model on $\mathbb{Z} \times (M \times F)$) whenever one of the two motors moves and thus inducing an increase or decrease of the force between the motors.

Such a coupling can be defined for an arbitrary finite number of killed Quasi Random-Walks. As the general treatment of such coupled processes surely is of interest to physicists it is not clear if such coupled Quasi-Random Walks are mathematically treatable and what kind of (exact) results can be obtained. The hope is of course that the defined coupling can be interpreted as a new kQRW (seemingly with a very complicated state space), to be able to apply the results of chapter 4.

Appendix A

Matrix Theory

In this minimal review chapter we collect some properties and useful identities for matrix manipulations.

A.1 On the invertibility of subgenerators and substochastic matrices

We rephrase here a lemma with its proof from [Neu94], chapter 2. The lemma concerns the invertibility of any sub generator. It is a property of most importance, as the expressions for the PH-type distributions, see Appendix B, rely heavily on this property, as well as the Schur complement method to inverse block matrices in the next section.

Lemma A.1.1 *Let Q be the infinitesimal generator of a MCc and S be the submatrix of Q which contains all transitions between the transient states (if existent). Then S is invertible.*

Proof Suppose that S is not invertible, then there exists a vector $y \neq 0$ such that $yS = 0$ (i.e. 0 is an eigenvalue of S). Then

$$y \exp(S\xi)\mathbf{1}^\top = y\mathbf{1}^\top$$

for each $\xi \in \mathbb{R}$ and thus

$$\lim_{\xi \to \infty} y \exp(S\xi)\mathbf{1}^\top = y\mathbf{1}^\top \neq 0,$$

which is a contradiction to the assumption, as for each $\xi \in \mathbb{R}^+$ the quantity $\exp(S\xi)\mathbf{1}^\top$ is the probability that the process is still in the set of transients for each $\xi \geq 0$.

■

A similar proof for a substochastic matrix S shows that $Id - S$ is invertible, see e.g. [KS76], Theorem 3.2.1 and 1.11.1. It is also important to note that these matrices have a probabilistic interpretation in terms of expectations, see [DS65] and [DS67].

With the Perron-Frobenius calculus it is possible to show that already the property of sub stochasticity resp. being a sub generator is enough to make a (finite) matrix invertible, see [Zha05].

A.2 The Frobenius Formula for block matrices

The Frobenius Formula, as it is called in [Gan86], chapter 2, is well known in the literature under the name *Banachiewicz inversion formula*. It has numerous applications ranging from numerics, algebra to applications in statistics and probability theory.

The inverse of a block Matrix

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

can be derived according to [Gan86], (86)-(89) via the identities

$$M^{-1} = \begin{pmatrix} M_{11}^{-1} + M_{11}^{-1} M_{12} H^{-1} M_{21} M_{11}^{-1} & -M_{11}^{-1} M_{12} H^{-1} \\ -H^{-1} M_{21} M_{11}^{-1} & H^{-1} \end{pmatrix} \tag{A.1}$$

if M_{11} is invertible, resp.

$$M^{-1} = \begin{pmatrix} K^{-1} & -K^{-1} M_{12} M_{22}^{-1} \\ -M_{22}^{-1} M_{21} K^{-1} & M_{22}^{-1} + M_{22}^{-1} M_{21} K^{-1} M_{12} M_{22}^{-1} \end{pmatrix} \tag{A.2}$$

if M_{22} is invertible. The matrices

$$H := M_{22} - M_{21} M_{11}^{-1} M_{12}, \qquad K := M_{11} - M_{12} M_{22}^{-1} M_{21}.$$

are called *Schur complements* throughout the literature. The book [Zha05] gives a good overview over the vast applications of these matrices.

Appendix B

On PH-type distributions

In the following we give the explicit computations of characteristica of PH-type distributions. The computations follow the lines of [Neu94] and [LR99], but in more detail. Note also that we used another definition of Laplace transform, thus our results differ by sign from the results of Neuts.

We used the following definitions for the computations:

Definition B.0.1 *Let X be some random variable with (cumulative) distribution function $F_X(t)$, with $X \geq 0$ almost sure.*

(i) If X takes its values in $\mathbb{N}_0$ the generating function of X is defined as the power series

$$G_X(z) := \sum_{n \geq 0} \mathbb{P}(X = n) z^n.$$

(ii) The moment generating function is defined as

$$M_X(s) := \int_0^\infty \exp(st) dF_X(t).$$

(iii) The characteristic function is defined as

$$\phi_X(s) := \int_0^\infty \exp(ist) dF_X(t)$$

where i is the imaginary unit.

B.1 Discrete time

B.1.1 Distribution and distribution function

Lemma B.1.1 *The distribution of X_n is given by*

$$\mathbb{P}(X_n = .) = \left(\mu_M S^n, 1 - \mu_M S^n \mathbf{1}^\top\right) \in [0,1]^{k+1}$$

Proof The n-th power of the transition matrix P is given by an induction argument by

$$P^n = \begin{pmatrix} S^n & \left(\sum_{k=0}^{n-1} S^k\right) R \\ 0 & 1 \end{pmatrix}.$$

Though P is a stochastic matrix $S\mathbf{1}^\top + R = \mathbf{1}^\top$ (the row sum of P must be one) and therefore

$$R = (Id - S)\mathbf{1}^\top.$$

Then direct calculation unveals

$$\begin{aligned}
\mathbb{P}(X_n = .) = \mu P^n &= \left(\mu_M S^n, \mu_M \left(\sum_{k=0}^{n-1} S^k\right) R + \mu_\Delta\right) \\
&= \left(\mu_M S^n, \mu_M (Id - S^n)(Id - S)^{-1} R + \mu_\Delta\right) \\
&= \left(\mu_M S^n, \mu_M (Id - S^n)\mathbf{1}^\top + \mu_\Delta\right) \\
&= \left(\mu_M S^n, 1 - \mu_M S^n \mathbf{1}^\top\right).
\end{aligned}$$

■

Lemma B.1.2 *Let $T \sim PH(\mu_M, S)$. Then T has distribution*

$$\mathbb{P}(T = n) = \begin{cases} \mu_\Delta & n = 0 \\ \mu_M S^{n-1} R & n \geq 1 \end{cases}.$$

Proof The probability to be already absorbed at time 0 is given by μ_Δ. To be absorbed exactly at time $n > 0$ requires that the Markov Chain was in $E \setminus \{k\}$ until time $n - 1$ before transition at time n to the absorbing. state k. Lemma B.1.1 gives therefore

$$\mathbb{P}(T = n) = \mathbb{P}(X_0 \in E \setminus \{k\}, \ldots, X_{n-1} \in E \setminus \{k\}, X_n = k) = \mu_M S^{n-1} R.$$

■

Lemma B.1.3 *Let $T \sim PH(\mu_M, S)$. Then the distribution function $F_T(t) := \mathbb{P}(T \leq t)$ of T is given by*

$$F_T(t) = (1 - \mu_M S^{\lfloor t \rfloor} \mathbf{1}^\top) \mathbb{1}_{\mathbb{R}^+}(t)$$

for $t \in \mathbb{R}$.

Proof Direct calculation shows

$$\begin{aligned}
F_T(n) = \mathbb{P}(T \leq n) = \sum_{l=0}^{n} \mathbb{P}(T = l) &= \mu_M \left(\sum_{l=0}^{n-1} S^l\right) R + \mu_\Delta \\
&= \mu_M \left(\sum_{l=0}^{n-1} S^l - \sum_{l=1}^{n} S^l\right) \mathbf{1}^\top + \mu_\Delta = \mu_M (Id - S^n)\mathbf{1}^\top + \mu_\Delta \\
&= 1 - \mu_M S^n \mathbf{1}^\top
\end{aligned}$$

Alternative

$$\mathbb{P}(T \leq n) = \mathbb{P}(X_n = k)$$

delivers the same result.

■

Remark **B.1.4** Though S contains transition probabilities between transient states only, the spectral radius of S is strictly less than one and therefore

$$\lim_{n\to\infty} S^n = \underline{\mathbf{0}}$$

which is sufficient to show

$$\lim_{t\to\infty} F_T(t) = 1.$$

For $t = 0$ the distribution function takes the value

$$F_T(0) = 1 - \mu_M \mathbf{1}^\top = 1 - 1 + \mu_\Delta = \mu_\Delta.$$

□

B.1.2 Generating function and factorial moments

Lemma B.1.5 *The generating function of T with representation (μ_M, S) is given by*

$$G_T(z) = \mu_\Delta + z\mu_M \left(1 - zS\right)^{-1} R$$

for $z \in \mathbb{R}$ with $|z| < 1$.

Proof Using von Neumann series in the version for matrices shows

$$\begin{aligned} G_T(z) &= \sum_{k=0}^{\infty} z^k \mathbb{P}(T = k) = \mu_\Delta + z\mu_M \left(\sum_{k=0}^{\infty} (zS)^k\right) R \\ &= \mu_\Delta + z\mu_M \left(Id - zS\right)^{-1} R \end{aligned}$$

■

Lemma B.1.6 *The kth factorial moment of T with representation (μ_M, S) is given by*

$$\mathbb{E}\left(\frac{T!}{(T-k)!}\right) = k!\mu_M S^{k-1}(Id - S)^{-k}\mathbf{1}^\top$$

for $k \geq 1$.

Proof Lemma B.1.5 delivers an expression for the factorial moments of T by derivation with respect to z of the generating function M_T of T:

$$\begin{aligned} \mathbb{E}\left(\frac{T!}{(T-k)!}\right) &:= \mathbb{E}(T(T-1)(T-2)\dots(T-k+1)) \\ &= \frac{d^k}{dz^k} G_T^{(k)}(z)\Big|_{z=1^-} = \mu_M k! S^{k-1}(Id - S)^{-k-1} R \\ &= k!\mu_M S^{k-1}(Id - S)^{-k}\mathbf{1}^\top \end{aligned}$$

■

Corollary B.1.7 *The expectation and variance of T with representation (μ_M, S) are given by*

$$\mathbb{E}(T) = \mu_M(Id - S)^{-1}\mathbf{1}^\top$$

and

$$Var(T) = \mu_M(Id + S)(Id - S)^{-2}\mathbf{1}^\top - (\mu_M(Id - S)^{-1}\mathbf{1}^\top)^2$$

Proof Setting $k = 1$ in lemma B.1.6 delivers the expression for the expectation of T. For the variance observe

$$\begin{aligned}\mathrm{Var}(T) &= \mathbb{E}\left(\frac{T!}{(T-2)!}\right) + \mathbb{E}(T) - \mathbb{E}(T)^2 \\ &= 2\mu_M S(Id - S)^{-2}\mathbf{1}^\top + \mu_M(Id - S)(Id - S)^{-2}\mathbf{1}^\top - \mathbb{E}(T)^2 \\ &= \mu_M(S + Id)(Id - S)^{-2}\mathbf{1}^\top - \mathbb{E}(T)^2\end{aligned}$$

Replacing $\mathbb{E}(T)$ by the expression given in lemma B.1.6 for $k = 1$ finishes the calculations. ■

B.2 Continuous time

B.2.1 Distribution and distribution function

Lemma B.2.1 *The distribution of the process $(X_t)_{t\geq 0}$ at a given time $t \geq 0$ is given by*

$$\mathbb{P}(X_t = .) = (\mu_M \exp(St), 1 - \mu_M \exp(St)\mathbf{1}^\top).$$

Proof Recall

$$\mathbb{P}(X_t = .) = \mu \exp(Qt).$$

Using the definition of the matrix exponential delivers

$$\begin{aligned}\exp(Qt) = \sum_{k\geq 0} Q^k \frac{t^k}{k!} &= \begin{pmatrix} Id & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k\geq 1} \begin{pmatrix} S^k & S^{-1}S^k R \\ 0 & 0 \end{pmatrix} \frac{t^k}{k!} \\ &= \begin{pmatrix} \sum_{k\geq 0} S^k \frac{t^k}{k!} & S^{-1}\left(-Id + \sum_{k\geq 0} S^k \frac{t^k}{k!}\right) R \\ 0 & 1 \end{pmatrix} \qquad \text{using } -S^{-1}R = \mathbf{1}^\top \\ &= \begin{pmatrix} \exp(St) & \mathbf{1}^\top - \exp(St)\mathbf{1}^\top \\ 0 & 1 \end{pmatrix}\end{aligned}$$

Multiplying from the left with the initial distribution yields

$$\mu \exp(Qt) = (\mu_M \exp(St), 1 - \mu_M \exp(St)\mathbf{1}^\top)$$

as claimed. ■

Lemma B.2.2 *The distribution function of $T \sim PH(\mu_M, S)$ is given by*

$$F_T(t) = 1 - \mu_M \exp(St)\mathbf{1}^\top.$$

Proof Write

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leq t) = 1 - \mathbb{P}(T > t) = 1 - \mathbb{P}(X_t \in \{0, 1, \ldots, k-1\}) \\ &= 1 - \mu_M \exp(St)\mathbf{1}^\top \end{aligned}$$

■

Lemma B.2.3 *The distribution of $T \sim PH(\mu_M, S)$ is*

$$\mathbb{P}_T(dt) = \mu_M \exp(St)Rdt + \mu_\Delta \delta_0(dt).$$

Proof Direct calculation shows

$$\begin{aligned} dF_T(t) &= -\mu_M \exp(St)S\mathbf{1}^\top dt + \mu_\Delta \delta_0(t) \\ &= \mu_M \exp(St)Rdt + \mu_\Delta \delta_0(t) \end{aligned}$$

■

Note that the distribution function is not absolutely continuous with respect to Lebesgue-measure as long there is an initial weight on the absorbing state.

B.2.2 Characteristic and moment generating function, moments

The calculations are continued with the characteristic function of T to derive a formula for the moments of T.

Lemma B.2.4 *The characteristic function of a phase type random variable T with representation $PH(\mu_M, S)$ is*

$$\varphi_T(s) = -\mu_M \left(isId + S\right)^{-1} R + \mu_\Delta$$

for $|s| < 1$.

Proof Direct calculation leads to

$$\begin{aligned} \varphi_T(s) &= \int_0^\infty \exp(ist) dF_T(t) = \int_0^\infty \exp(ist)(\mu_M \exp(St)R)dt + \mu_\Delta \\ &= \mu_M \sum_{n \geq 0} \frac{S^n}{n!} (-is)^{-1} \int_0^\infty (-is)\exp(-(-is)t)t^n dt R + \mu_\Delta \\ &= -(is)^{-1}\mu_M \sum_{n \geq 0} \frac{S^n}{(-is)^n} R + \mu_\Delta = -\mu_M \left(isId + S\right)^{-1} R + \mu_\Delta \end{aligned}$$

■

Corollary B.2.5 *The moment generating function of a phase type random variable T with representation $PH(\mu_M, S)$ is*

$$M_T(s) = -\mu_M \left(sId + S\right)^{-1} R + \mu_\Delta$$

and exists always.

Proof The proof is analogue to the proof of the characteristic function, as

$$M_T(s) = \int_0^\infty \exp(st) dF_T(t).$$

■

Corollary B.2.6 *All moments of a phase type random variable T with representation $PH(\mu_M, S)$ exist and are given by*

$$\mathbb{E}(T^n) = \mu_M (-1)^n n! S^{-n} \mathbf{1}^\top.$$

Proof Existence is ensured by the existence of the moment generating function. For $n \geq 1$

$$\begin{aligned}
\mathbb{E}(T^n) &= \frac{d^n}{ds^n} M_T(s)\Big|_{s=0} \\
&= \frac{d^n}{ds^n} \left(-\mu_M \left(sId + S \right)^{-1} R + \mu_\Delta \right)\Big|_{s=0} \\
&= \mu_M (-1)^{n+1} n! S^{-n} \underbrace{S^{-1} R}_{=-\mathbf{1}^\top} = \mu_M (-1)^n n! S^{-n} \mathbf{1}^\top
\end{aligned}$$

as claimed.

■

Lemma B.2.7 *The expectation of a phase type random variable T with representation $PH(\mu_M, S)$ is given by*

$$\mathbb{E}(T) = -\mu_M S^{-1} \mathbf{1}^\top$$

and the variance is

$$Var(T) = 2\mu_M S^{-2} \mathbf{1}^\top - (\mu_M S^{-1} \mathbf{1}^\top)^2.$$

Proof Both quantities are given by the moment generating function and

$$\mathrm{Var}(T) = \mathbb{E}(T^2) - \mathbb{E}(T)^2.$$

■

Appendix C

The simple killed Random Walk

In this short section we want to present applications of the results of chapter 4 to the simple killed Random Walk, that is a killed Quasi-Random-Walk were the number of phases is one, i.e. $|M| = 1$. In this case we identify $\mathbb{Z} \times \{1\}$ with $\mathbb{Z}$. The main tool of this section is an approximation scheme in the sense that we introduce additional killing if the Random Walk leaves the symmetric interval $[-n, n]$. The limit $n \to \infty$ gives then the exact result as limit of the approximation. We are aware of the fact that some of the results already exist in the literature, some of them in [Gut09], but we do not find any references that use such an approximation scheme. This scheme is also applied to the computations in chapter 5.

We restrict the attention to the distribution of the last state before killing and the distribution of the maximum and minimum level ever attained before absorption and thus discrete time. All other notions introduced in chapter 4 are solvable without treating infinite matrices, as for these properties the chain can be reduced to a finite state version, see Lemmas 4.1.5 and 4.1.7. We also just give the results without much commentary and too detailed proofs, some of them have to be understood as sketches.

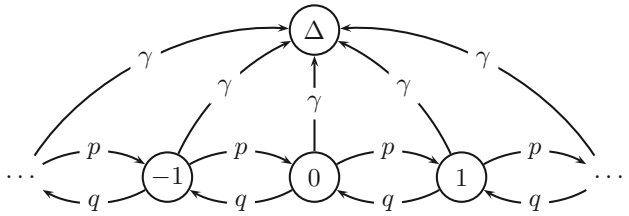


Figure C.1: The simple killed Random Walk.

We consider $(X_n)_{n\geq 0}$ the killed Random Walk as defined on $\mathbb{Z}\cup\{\Delta\}$ with

$$S := \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots & \\ \dots & 0 & p & 0 & \dots & 0 & 0 & \dots \\ \dots & q & 0 & p & \dots & 0 & 0 & \dots \\ \dots & 0 & q & 0 & \dots & 0 & 0 & \dots \\ & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \\ \dots & 0 & 0 & 0 & \dots & 0 & p & \dots \\ \dots & 0 & 0 & 0 & \dots & q & 0 & \dots \\ & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots \end{pmatrix},$$

and (countable infinite) vector

$$R_\Delta^\top = -S\mathbf{1}^\top = \gamma\mathbf{1}^\top$$

where $\gamma + p + q = 1$.

Lemma C.0.8 *Define the $n \times n$ matrix*

$$M_n := \begin{pmatrix} 1 & -p & 0 & \dots & 0 & 0 \\ -q & 1 & -p & \dots & 0 & 0 \\ 0 & -q & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -p \\ 0 & 0 & 0 & \dots & -q & 1 \end{pmatrix}$$

and fix

$$J(n) := \det(M_n).$$

Then

$$\begin{aligned} J(n) &= \frac{1}{2^{n+1}} \sum_{k=1}^{n+1} \binom{n+1}{k} (1-(-1)^{k+1})(1-4pq)^{\frac{k-1}{2}} \\ &= \frac{1}{2^{n+1}\sqrt{1-4pq}} \left(\left(1+\sqrt{1-4pq}\right)^{n+1} - \left(1-\sqrt{1-4pq}\right)^{n+1} \right). \end{aligned}$$

Proof By Laplace's formula for determinants:

$$J(n) = J(n-1) - pqJ(n-2) \tag{C.1}$$

and $J(0) = 1$, $J(1) = 1$. This recursion is solved by the formula in the statement and is a standard result.

■

We now can give by conditioning on the last state, see Theorem 2.2.11, the exact result for the distribution of the last visited level before absorption resp. the position at which the Random Walk is killed.

Proposition C.0.9 *Let $(X_n)_{n\geq 0}$ be (discrete time) kQRW with $|M| = 1$, $A_0 = p$, $A_1 = 0$, $A_2 = q$ and $\Gamma = 1 - p - q$. Then*

$$\mathbb{P}(\textit{last visited level} = k) = \frac{(1-\tilde{p})(1-\tilde{q})}{1-\tilde{p}\tilde{q}} \begin{cases} \tilde{p}^k & k \geq 0 \\ \tilde{q}^{|k|} & k \leq 0 \end{cases}$$

with

$$\tilde{p} := \frac{1-d}{2q}, \quad \tilde{q} := \frac{1-d}{2p}$$

and $d := \sqrt{1-4pq}$.

Proof We fix first a finite version. Let therefore $S_{|n}$ be the $(2n+1) \times (2n+1)$ submatrix of S associated to the states in $\{-n, -n+1, \ldots, -1, 0, 1, \ldots, n\}$. To this matrix we can assign a Markov Chain on $\{\Delta\} \cup \{-n, -n+1, \ldots, n-1, n\}$ in discrete time with transition matrix

$$P_{|n} := \begin{pmatrix} 1 & 0 \\ \Gamma_{|n} & S_{|n} \end{pmatrix}$$

and $\Gamma_{|n}$ the restriction of the original infinite vector Γ to $[-n, n]$ with the addition that $\Gamma_{|n}(n) = p + \gamma$ and $\Gamma_{|n}(-n) = q + \gamma$ ensuring that the Random Walker is absorbed if it tries to leave the interval $[-n, n]$.

As the Random Walk starts almost surely in 0 only one row of the inverse is needed. By the inversion formula relating the inverse of a matrix with its determinant and adjoint matrix we immediately get

$$e_0 M_{2n+1}^{-1} e_k^\top = \begin{cases} p^k \frac{J(2n+1-k)}{J(2n+1)} & 0 \leq k \leq n \\ q^{|k|} \frac{J(2n+1-k)}{J(2n+1)} & -n \leq k \leq 0 \end{cases} \tag{C.2}$$

and thus by 2.2.11 a discrete distribution on $[-n, n]$; the distribution κ_n of the last visited level in $[-n, n]$:

$$\kappa_n(k) = \frac{\gamma e_0 M_{2n+1}^{-1} e_k^\top}{\gamma e_0 M_{2n+1}^{-1} \mathbf{1}^\top} = \begin{cases} \frac{1}{Z} q^{|k|} J(2n+1-k) & -n \leq k \leq 0 \\ \frac{1}{Z} p^k J(2n+1-k) & n \geq k \geq 0 \end{cases}$$

where

$$Z := J(n) + \sum_{k=1}^{n} (p^k + q^k) J(n-k)$$

acts as renormalization constant, to ensure that κ is a probility measure.

It remains to calculate for every fixed k

$$\lim_{n\to\infty} \frac{J(2n+1-k)}{Z} = \lim_{n\to\infty} \frac{J(2n+1-k)}{J(2n+1) + \sum_{m=1}^{n} (p^m + q^m) J(2n+1-m)}.$$

Define $L(n,k) := \frac{J(2n+1-k)}{Z}$, then $\frac{(2pq)^k}{p+q-1} L(n,k)$ reduces to

$$L(n) = \frac{(1+d)^k(1-d)^{n+1} + (1-d)^k(1+d)^{n+1}}{2^{n+1}d(p^{n+1}+q^{n+1}) - (1-d)^{n+1}d - (1+d)^{n+1}d}$$
$$= \frac{(1+d)^k\left(\frac{1-d}{1+d}\right)^{n+1} + (1-d)^k}{\frac{2^{n+1}d(p^{n+1}+q^{n+1})}{(1+d)^{n+1}} - d\left(\frac{1-d}{1+d}\right)^{n+1} - d}.$$

Certainly $\frac{1-d}{1+d} < 1$ and thus

$$\lim_{n\to\infty}\left(\frac{1-d}{1+d}\right)^n = 0$$

and

$$\lim_{n\to\infty}\frac{2^{n+1}(p^{n+1}+q^{n+1})}{(1+d)^{n+1}} = 0$$

because

$$\frac{2p}{1+\sqrt{1-4pq}} = \frac{2p(1-\sqrt{1-4pq})}{4pq} = \frac{1-\sqrt{(1-2pq)^2-4p^2q^2}}{2q}$$
$$< \frac{1-\sqrt{(1-2pq)^2}}{2q} = p < 1$$

and analogue for $\frac{2q}{1+\sqrt{1-4pq}}$.

Thus

$$\lim_{n\to\infty} L(n) = \frac{1-p-q}{(2pq)^k}\frac{(1-d)^k}{d}$$

and the limit distribution is given as

$$\kappa_\infty(k) = \frac{1-p-q}{d}\left(\frac{1-d}{2pq}\right)^k \begin{cases} p^k & k \geq 0 \\ q^{|k|} & k \leq 0 \end{cases}$$
$$= \frac{\gamma}{d}\begin{cases} \left(\frac{1-d}{2q}\right)^k & k \geq 0 \\ \left(\frac{1-d}{2p}\right)^{|k|} & k \leq 0 \end{cases}.$$

A simple calculation ensures that

$$\frac{\gamma}{d} = \frac{(1-\tilde{p})(1-\tilde{q})}{1-\tilde{p}\tilde{q}}.$$

What we (implicitly) calculated here is the weak limit for $n \to \infty$ of the distributions of the last visited levels for the finite approximations. The limit is a distribution on $\mathbb{Z}$ which resembles the last visited level distribution of the original killed Random Walk.

■

The distribution of the last level is in its discrete version not very well known, see e.g. [KI06] for basic properties. Conversely its continuous analogue is the famous First Error Law of Laplace. This distribution is also known as double sided exponential distribution, as described for example in [Fel68].

We want to compute the exact distribution of the minimal and maximal state attained before killing:

Proposition C.0.10 *Let $(X_n)_{n\geq 0}$ be (discrete time) generalized Random Walk with $|M| = 1$, $A_0 = p$, $A_1 = 0$, $A_2 = q$ and $\Gamma = 1 - p - q$, see also fig. C.1. Then*

$$\mathbb{P}(\max\{X_0, X_1, \dots, X_{\tau_\Delta - 1}\} = k) = (1 - p_+)p_+^k$$

and

$$\mathbb{P}(\min\{X_0, X_1, \dots, X_{\tau_\Delta - 1}\} = k) = (1 - p_-)p_-^k$$

where $p_+ := \frac{2p}{1+d}$ and $p_- := \frac{2q}{1+d}$ with $d := \sqrt{1-pq}$.

Proof For the calculation of the exact form of the distribution of the maximal and minimal level attained before killing we can refer back to chapter 4. There we argued that because of the spatial homogeneity the maximum distribution must be geometric with some parameter p_+, where p_+ refers to the probability to start in 0 and reaching 1 before killing.

We set state 1 absorbing and calculate the probability to be absorbed in 1 before absorption in Δ can take place. This is by chapter 2, section 3 given as

$$p_+ = pe_0(Id - S_{|k\leq 0})^{-1}e_0^\top$$

where $S_{|k\leq 0}$ is the restriction of S to the non positive integers. We approximate this infinite matrix like in the proof of Proposition C.0.9 (and omit the details of the exact definition as it is one to one to the preceding one, except that we take $[-n, 0]$ as state space of the sequence of approximating Markov Chains). With the notation of Lemma C.0.8 we have

$$e_0(Id - S_{|k\leq 0})^{-1}e_0^\top = \frac{\det(M_{n-1})}{\det(M_n)} = \frac{J(n-1)}{J(n)} = 2\frac{\frac{1}{1+d} - \frac{(1-d)^n}{(1+d)^{n+1}}}{1 - \frac{(1-d)^{n+1}}{(1+d)^{n+1}}} \overset{n\to\infty}{\longrightarrow} \frac{2}{1+d}.$$

The proof for the minimum distribution is the same except that we need to compute

$$p_- = pe_0(Id - S_{|k\geq 0})^{-1}e_0^\top$$

and thus need to approximate via considering intervals of the form $[0, n]$.

■

As a final remark we want to draw attention to what happens if we compute the distribution of X_T where $(X_n)_{n\geq 0}$ is a usual Random Walk (without killing) on $\mathbb{Z}$ moving forward with probability p and backwards with q, $p+q = 1$, and T is a geometric distributed random variable to the parameter $1 - \gamma$ independent of $(X_n)_{n\geq 0}$. Then we gain

$$\mathbb{P}(X_T = k) = \frac{1-\gamma}{\sqrt{1-4\gamma^2 pq}} \begin{cases} \left(\frac{1-\sqrt{1-4\gamma^2 pq}}{2p\gamma}\right)^k & k \leq 0 \\ \left(\frac{1-\sqrt{1-4\gamma^2 pq}}{2q\gamma}\right)^{|k|} & k \geq 0 \end{cases}$$

which is very similar to the distribution we gained in the previous case for the killed Random Walk. The proof is made by combinatorial means:

$$\mathbb{P}(X_T = 0) = \sum_{n\geq 0} \mathbb{P}(T = 2n)\binom{2n}{n} p^n q^n = \sum_{n\geq 0} \gamma^n (1-\gamma)\binom{2n}{n} p^n q^n$$
$$= (1-\gamma)\sum_{n\geq 0}\binom{2n}{n}(\gamma^2 pq)^n = \frac{1-\gamma}{\sqrt{1-4\gamma^2 pq}}. \tag{C.3}$$

More generally ending up in $k \geq 0$ means that there is an excess of k steps. Thus

$$\mathbb{P}(X_T = k) = \sum_{n\geq 0} \mathbb{P}(T = 2n+k)\binom{2n+k}{n} p^{n+k} q^n$$
$$= (p\gamma)^k (1-\gamma)\sum_{n\geq 0}\binom{2n+k}{n}(\gamma^2 pq)^n$$
$$\overset{(*)}{=} \frac{1-\gamma}{\sqrt{1-4\gamma^2 pq}}\left(\frac{1-\sqrt{1-4\gamma^2 pq}}{2q\gamma}\right)^k$$

The equal sign $(*)$ can be justified in the following way. In [GKP94], page 203, (5.72) states

$$\sum_{n\geq 0}\binom{2n+k}{n} z^k = \frac{B_2(z)^k}{\sqrt{1-4z}}$$

with

$$B_2(z) = \frac{1-\sqrt{1-4z}}{2z}$$

as given in (5.68) in the reference for $|z| < \frac{1}{4}$. Setting $z = \gamma^2 pq$ gives the desired result.

It is noteworthy that indeed

$$G_{\tau_+}(\gamma) := \frac{1-\sqrt{1-4\gamma^2 pq}}{2p\gamma}$$

is the generating function of the so called "first ascending ladder epoch", see e.g. [Gut09] section 2.9, evaluated at γ, the failure probability of the geometric distribution τ defined earlier.

The first ascending ladder epoch τ_+ is the time needed to attain a new maximum of the random walk (not stopped). This stopping time is defined as

$$\tau_+ := \inf\{n \geq 0 | X_n > 0\}.$$

This is clearly also the first time that there is an excess of exactly one step to the right to the total number of steps left and right and thus an extremum. By this combinatorial interpretation we can calculate the generating function of τ_+ in the same manner as the distribution of the position after stopping according to a geometric time, but under the additional constraint that the state at $\tau_+ - 1$ equals 0:

$$G_{\tau_+}(s) = \sum_{n\geq 0}\binom{2n+1}{n}\frac{1}{2n+1} q^n p^{n+1} s^{2n+1} = sp\sum_{n\geq 0}\binom{2n+1}{n}\frac{(s^2 pq)^n}{2n+1}$$
$$\overset{(**)}{=} spB_2(s^2 pq) = sp\frac{1-\sqrt{1-4s^2 pq}}{2s^2 pq} = \frac{1-\sqrt{1-4s^2 pq}}{2qs}.$$

Equation (∗∗) is justified as before by (5.68) in [GKP94]. In the same manner it is easy to see that the generating function of the first descending ladder height

$$\tau_- := \inf\{n \geq 0 | X_n < 0\}$$

has analogue form:

$$G_{\tau_-}(s) = \frac{1 - \sqrt{1 - 4s^2pq}}{2ps}.$$

It is now obvious that the n-th ascending ladder epoch $\tau_+^{(n)}$ has by spatial homogeneity and the strong Markov Property (the process before $\tau_+^{(n)}$ is independent from the process after that time) the generating function:

$$G_{\tau_+^{(n)}}(s) = \left(G_{\tau_+}(s)\right)^n.$$

Indeed $(\tau_+^{(n)})_{n\geq 0}$ is a renewal process. The exact behaviour of this renewal process depends on whether the random walk oscillates or has drift to $\pm\infty$, i.e. the renewal process could be terminating.

The above formulas are valid for $|s^2pq| < \frac{1}{4}$. As $pq \leq \frac{1}{4}$ the parameter s can take any value in $[0, 1]$.

Going back to the interpretation of the skewed discrete Laplace distribution we gained for X_τ, $\tau \sim Geo(1-\gamma)$, we can now write

$$\mathbb{P}(X_\tau = k) = \mathbb{P}(X_\tau = 0) \begin{cases} G_{\tau_-^{(|k|)}}(\gamma) & k \leq 0 \\ G_{\tau_+^{(k)}}(\gamma) & k \geq 0 \end{cases} \tag{C.4}$$

$$= \mathbb{P}(X_\tau = 0) \begin{cases} \left(G_{\tau_-}(\gamma)\right)^{|k|} & k \leq 0 \\ \left(G_{\tau_+}(\gamma)\right)^{k} & k \geq 0 \end{cases} \tag{C.5}$$

As we know that the resulting distribution is a discrete skewed Laplace distribution, we have also:

$$\mathbb{P}(X_\tau = 0) = \frac{(1 - G_{\tau_-}(\gamma))(1 - G_{\tau_+}(\gamma))}{1 - G_{\tau_-}(\gamma)G_{\tau_+}(\gamma)}$$

We must note that $G_{\tau_-}(s)$ and $G_{\tau_+}(s)$ can be defective if the random walk has a drift. Because then the probability to never hit any negative or positive state is non zero, except in the symmetric case $p = q = \frac{1}{2}$. This behaviour is (of course) directly reflected by the behaviour of $G_{\tau_\pm}(1)$ depending on p.

It is this kind of flexibility of perspectives on Random Walks that allows explicit calculations in many cases and thus a huge literature evolved over the past decades. Unfortunately none of the above approaches can be easily adapted to the killed Quasi-Random-Walk as defined in chapter 4.

Nomenclature

$\operatorname{diag}(\mu)$	diagonal matrix with with the entries of μ as diagonal
$\mathbf{0}$	$(0, 0, 0, \ldots, 0)$
^{c}	complement of a set
$\stackrel{d}{=}$	equality in distribution
$^{-1}$	inversion of a matrix
$\sim$	same distribution as, relation
$^\top$	transposition of a matrix or vector
$\mathbf{1}$	$(1, 1, 1, \ldots, 1)$
$\mathbb{1}_A$	indicatorfunction of the event A
$*$	convolution
$\llbracket \mapsto \rrbracket$	states in a level that can be arrived from the left
$\llbracket \leftarrow \rrbracket$	states in a level that can be arrived from the right
A_0	matrix of transitions to the right in a kQRW or QBD
A_1	matrix of transitions within a level in a kQRW or QBD
A_2	matrix of transitions to the left in a kQRW or QBD
ADP	adenosine diphosphate
ATP	adenosine triphosphate
BD	Birth-and-Death Process
$\langle i \rangle_\sigma$	cylce generated by i with respect to σ
$\mathcal{N}_{i \to j}$	connected neighborhood of i and j
$\mathbb{E}(X)$	expection of the r.v. X
$\mathcal{E}(\lambda)$	density of the exponential distribution to the parameter $\lambda > 0$
E	countable set of states

e_i	unit vector, i-th component 1
e.g.	for instance
et al.	et alii
etc.	et cetera
$\varphi_X(t)$	characteristic function of the r.v. X
$F_X(t)$	distribution function of the r.v. X
$f_X(t)$	density of the r.v. X, if existing
$G_X(t)$	generating function of the discrete r.v. X
Γ	killing rates for one level
$\triangleright$	gate
i.e.	in explanation
$J(n)$	determinant of Id-Q for a simple killed Random Walk
kQRW	killed Quasi-Birth-and-Death Process
$(L_n)_{n\geq 0}$	level process in discrete time
$(L_t)_{t\geq 0}$	level process in continuous time
$\mathcal{L}(X)$	law of r.v. X
Λ	the absorbing state of kQRW
MCc	Markov Chain in continuous time
MCd	Markov Chains in discrete time
$M_X(t)$	momentgenerating function of the r.v. X
μM	micromolar, molar concentration, unit
M	finite set of phases for a kQRW
$\mathbb{N}$	natural numbers
$\mathbb{N}^*$	natural numbers without zero
$\mathbb{N}_0$	natural numbers with zero
$\mathcal{N}_i$	neighborhood of i
$\mathbb{P}(X = .)$	distribution of the r.v. X
$\mathbb{P}_X(dt)$	distribution of the r.v. X
$PH(\mu_M, S)$	representation of a phase type distribution
π	the stationary distribution

π_i	i-th component of the stationary distribution
P	transition matrix
p_{ij}	transition probability from i to j, entry of P
P	phosphate ion
Q	infinitesimal generator
$Q(t)$	Semi-Markov kernel
q_{ij}	transition rate from i to j, entry of Q
QBD	Quasi-Birth-and-Death Process
$\mathbb{R}$	real numbers
$\mathbb{R}^+$	positive real numbers
R_Δ	vector with transition rates/probabilities from the transients to Δ
resp.	respectively
r.v.	random variable
$\mathcal{S}_n$	symmetric group of order n
S	sub stochastic matrix of sub generator
supp	support
σ	permutation
σ_{ij}	last exit time from i before reaching j
τ	a stopping time
τ_{ij}	first passage time from i to j
τ^*_{ij}	pure passage time from i to j
$\mathrm{Var}(X)$	variance of the r.v. X
$(X_n)_{n\geq 0}$	Markov Chain in discrete time
$(X_t)_{t\geq 0}$	Markov Chain in continuous time
$(Y_n)_{n\geq 0}$	step-process in discrete time
$(Y_t)_{t\geq 0}$	step-process in continuous time
$\mathbb{Z}$	integers

Eidesstattliche Erklärung

Ich erkläre hiermit, dass die vorliegende Arbeit selbstständig und ausschließlich mit den angebenen Mitteln angefertigt wurde. Die Arbeit wurde bisher an keiner anderen Hochschule eingereicht.

Peter Keller

Bibliography

[Bau02] Heinz Bauer. *Wahrscheinlichkeitstheorie*. De Gruyter-Lehrbuch, 5 edition, 2002.

[Beh00] Ehrhard Behrends. *Introduction to Markov Chains with Special Emphasis to Rapid Mixing*. Advanced Lectures in Mathematics. Vieweg Verlag, 2000.

[BIG03] Adi Ben-Israel and Thomas Greville. *Generalized Inverses, Theory and Applications*. Springer, 2003.

[BKKL12] Florian Berger, Corina Keller, Stefan Klumpp, and Reinhard Lipowsky. Distinct transport regimes for two elastically coupled molecular motors. *Phys. Rev. Lett.*, 108:208101, May 2012.

[BKM+11] Florian Berger, Corina Keller, Melanie J. I. Mueller, Stefan Klumpp, and Reinhard Lipowsky. Co-operative transport by molecular motors. *Biochemical Society Transactions*, 39(5):1211–1215, 2011.

[BL92] Jonathan M. Borwein and Adrian Stephen Lewis. Decomposition of multivariate functions. *Canadian Journal of Mathematics*, 44(3):463–482, 1992.

[BL08] Vlad Stefan Barbu and Nikolaos Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications, Their Use in reliability and DNA Analysis*, volume 191 of *Lecture Notes in Statistics*. Springer, 2008.

[Bré99] Pierre Brémaud. *Markov Chains – Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, 1999.

[CC05] Nick J. Carter and Robert A. Cross. Mechanics of the kinesin step. *Nature*, pages 308–312, 2005.

[dHOS00] F. den Hollander, E. Olivieri, and E. Scoppola. Metastability and nucleation for conservative dynamics. *Journal of Mathematical Physics*, 41(3), 2000.

[DM09] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *Journal of Theoretical Probability*, 22:558–586, 2009.

[DS65] John N. Darroch and Eugene Seneta. On quasi-stationary distributions in absorbing discrete-time finite markov chains. *Journal of Applied Probability*, 2:88–100, 1965.

[DS67] John N. Darroch and Eugene Seneta. On quasi-stationary distributions in absorbing continuous-time markov chains. *Journal of Applied Probability*, 4:192–196, 1967.

[DZ05a] Erik A. Van Doorn and Alexander I. Zeifman. Birth-death processes with killing. *Statistics and Probability Letters*, 72:33–42, 2005.

[DZ05b] Erik A. Van Doorn and Alexander I. Zeifman. Extinction probability in a birth-death process with killing. *Journal of Applied Probability*, 42(1):185–198, 2005.

[EP02] Katrin Erk and Lutz Priese. *Theoretische Informatik*. Springer, 2 edition, 2002.

[Erl09] Agner Krarup Erlang. Sandsynlighedsregning og telefonsamtaler. *Nyt Tidsskrift for Matematik B*, 1909.

[Fel68] William Feller. *An introduction to probability theory and its applications, Vol. 1*, volume 1 of *Wiley series in probability and mathematical statistics*. Wiley, 1968.

[FL03] James C. Fu and W. Y. Wendy Lou. *Distribution Theory of Runs and Patterns and Its Applications*. World Scientific Pub Co Inc, 1 edition, 2003.

[Fu96] James C. Fu. Distribution theory of runs and patterns associated with a sequence of multi-state trials. *Statistica Sinica*, 6:957–974, 1996.

[Gan86] Felix R. Gantmacher. *Matrizentheorie*. Springer-Verlag, 1986.

[GKP94] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company, Reading, MA, second edition, 1994. A foundation for computer science.

[Gra08] Carl Graham. *Chaînes de Markov: Cours, exercices et corrigés détaillés*. Dunod, 2008.

[Gut09] Allan Gut. *Stopped Random Walks: Limit Theorems and Applications*. Springer Series in Operations Research and Financial Engineering, Operations Research. Springer, 2 edition, 2009.

[Hae08] Olle Haeggstroem. *Finite Markov Chains and algorithmic applications*. Cambridge Univ. Press, 2008.

[Hil88a] Terrell L. Hill. Further properties of Random Walks on diagrams (graphs) with and without cycles. *Proc. Natl. Acad. Sci. USA*, 85(9):3271–3275, 1988.

[Hil88b] Terrell L. Hill. Interrelations between Random Walks on diagrams (graphs) with and without cycles. *Proc. Natl. Acad. Sci. USA*, 85(9):2879–2883, 1988.

[Hil88c] Terrell L. Hill. Number of visits to a state in a random walk, before absorption, and related topics. *Proc. Natl. Acad. Sci. USA*, 96:4577–4581, 1988.

[IAK92] Hisashi Ito, Shun-Ichi Amari, and Kingo Kobayashi. Identifiability of hidden markov information sources and their minimum degrees of freedom. *IEEE Transaction on Information Theory*, 38:324–333, 1992.

[Jen54] Arne Jensen. *A Distribution Model Applicable to Economics*. Munks- gaard, 1954.

[JM06] Jacques Janssen and Raimondo Manca. *Applied Semi-Markov Processes*. Springer, 2006.

[Kei79] Julian Keilson. *Markov Chain Models – Rarity and Exponentiality*, volume 28 of *Applied Mathematical Sciences*. Springer-Verlag, 1979.

[Kel79] Frank P. Kelly. *Reversibility and Stochastic Networks*. Wiley, 1979.

[KI06] Tomasz Kozubowski and Seidu Inusah. A skew laplace distribution on integers. *Annals of the Institute of Statistical Mathematics*, 58:555–571, 2006.

[Kij88] Masaaki Kijima. On passage and conditional passage times for markov chains in continuous time. *Journal of Applied Probability*, 25:279–290, 1988.

[KL05] Stefan Klumpp and Reinhard Lipowsky. Cooperative cargo transport by several molecular motors. *Proceedings of the National Academy of Sciences*, 102(48):17284–17289, 2005.

[KlV11] Peter Keller, Sylvie Rœlly, and Angelo Valleriani. On time duality in markov chains with absorbing boundary. 2011. submitted to Journal of Applied Probability.

[KML06] Stefan Klumpp, Melanie Mueller, and Reinhard Lipowsky. cooperative transport by small teams of molecular motors. *Biophysical Reviews and Letters*, 1(4):353–361, 2006.

[Kol36] Andrei Nikolajewitsch Kolmogoroff. Zur Theorie der Markoffschen Ketten. *Mathematische Annalen*, 112:155–160, 1936. 10.1007/BF01565412.

[KS76] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains - reprint*. Undergraduate Texts in Mathematics. Springer-Verlag, 1976.

[KSK66] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp. *Denumerable Markov Chains*. The University Series in Higher Mathematics. D. van Nostrand Company, Inc., 1966.

[Lin92] Torgny Lindvall. *Lectures on the coupling method*. Wiley series in probability and mathematical statistics : Probability and mathematical statistics. Wiley, 1992.

[LL07] Steffen Liepelt and Reinhard Lipowsky. Kinesin's network of chemomechanical motor cycles. *Phys. Rev. Lett.*, pages 258102–1–4, 2007.

[LL08] Reinhard Lipowsky and Steffen Liepelt. Chemomechanical coupling of molecular motors: Thermodynamics, network representations, and balance conditions. *Journal of Statistical Physics*, 130:39–67, 2008. 10.1007/s10955-007-9425-7.

[LO01] Nikolaos Limnios and G. Oprisan. *Semi-Markov Processes and Reliability*. Statistics For Industry And Technology. Birkhäuser Boston, 2001.

[LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times – With a chapter on "Coupling from the past" by James G. Propp and David B. Wilson*. American Mathematical Society, 2009.

[LR99] Guy Latouche and Vaidyanathan Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modelling*. ASA & SIAM, 1999.

[LT03] Guy Latouche and Peter G. Taylor. Drift conditions for matrix-analytic models. *Mathematics of Operations Research*, 28(2):346–360, 2003.

[LW07] Martin Lindén and Mats Wallin. Dwell time symmetry in random walks and molecular motors. *Biophysical Journal*, 92:3804–3816, 2007.

[Neu94] Marcel F. Neuts. *Matrix Geometric Solutions in Stochastic Models*. Dover Publications, Inc., 1994.

[Nor97] James R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[Nue08] Gregory Nuel. Pattern Markov Chains: optimal Markov Chain embedding through deterministic finite automata. *Journal of Applied Probability*, 45(1):226–243, 2008.

[O'C99] Colm A. O'Cinneide. Phase-type distributions: open problems and a few properties. *Communications in Statistics – Stochastic Models*, 15(4):731–757, 1999.

[PTP09] António Pacheco, Loon Ching Tang, and Narahari Umanath Prabhu. *Markov-modulated processes and semiregenerative phenomena*. World Scientific, 2009.

[Rob10] Philippe Robert. *Stochastic Networks and Queues*. Stochastic Modelling and Applied Probability. Springer, 2010.

[Ros96] Sheldon M. Ross. *Stochastic Processes*. Wiley Series in Probability and Mathematical Statistics, 2 edition, 1996.

[Ryd96] Tobias Rydén. On identifiability and order of continuous-time aggregated markov chains, markov-modulated poisson processes, and phase-type distributions. *Journal of Applied Probability*, 33(3):640–653, Sept. 1996.

[Sch03] Manfred Schliwa, editor. *Molecular Motors*. Wiley-VCH, 2003.

[SS67] Kemeny Snell and Laurie Snell. Excessive functions of continuous time markov chains. *Journal of Combinatorial Theory*, (3):256–278, 1967.

[SSSB93] Karel Svoboda, Christoph F. Schmidt, Bruce J. Schnapp, and Steven M. Block. Direct observation of kinesin stepping by optical trapping interferometry. *Nature*, pages 721–727, 1993.

[TK06] Jianjun Paul Tian and D. Kannan. Lumpability and Commutativity of Markov Processes. *Stochastic Analysis and Applications*, 24(3):685–702, 2006.

[Val03] Ronald D. Vale. The molecular motor toolbox for intracellular transport. *The Cell*, 122:467–480, 2003.

[VLL08] Angelo Valleriani, Steffen Liepelt, and Reinhard Lipowsky. Dwell time distributions for kinesin's mechanical steps. *EPL (Europhysics Letters)*, 82(2):28011–p1–28011–p6, 2008.

[YTVS04] Ahmet Yildiz, Michio Tomishige, Ronald D. Vale, and Paul R. Selvin. Kinesin walks hand-over-hand. *Science*, pages 676–678, 2004.

[Zha05] Fuzhen (ed.) Zhang. *The Schur complement and its applications*. New York, NY: Springer, 2005.