

UNIVERSITÄT POTSDAM
LEHRSTUHL FÜR WAHRSCHEINLICHKEITSTHEORIE

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI MATEMATICA

BERLIN MATHEMATICAL SCHOOL

Reciprocal classes of continuous time Markov Chains

DISSERTATION EINGEREICHT VON
GIOVANNI CONFORTI
ZUR ERLANGUNG DES AKADEMISCHEN GRADES
"DOCTOR RERUM NATURALIUM"
(DR. RER. NAT.)
IN DER WISSENSCHAFTSDISZIPLIN STOCHASTIK

Candidate:

Giovanni CONFORTI

Advisors:

Prof. Dr. Paolo DAI PRA

Prof. Dr. Sylvie ROELLY

This work is licensed under a Creative Commons License:
Attribution 4.0 International
To view a copy of this license visit
<http://creativecommons.org/licenses/by/4.0/>

Published online at the
Institutional Repository of the University of Potsdam:
URN [urn:nbn:de:kobv:517-opus4-82255](http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-82255)
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-82255>

Summary

In this thesis we study *reciprocal classes* of Markov chains. Given a continuous time Markov chain on a countable state space, acting as reference dynamics, the associated reciprocal class is the set of all probability measures on path space that can be written as a mixture of its bridges. These processes possess a conditional independence property that generalizes the Markov property, and evolved from an idea of Schrödinger, who wanted to obtain a probabilistic interpretation of quantum mechanics.

Associated to a reciprocal class is a set of *reciprocal characteristics*, which are space-time functions that determine the reciprocal class. We compute explicitly these characteristics, and divide them into two main families: *arc characteristics* and *cycle characteristics*. As a byproduct, we obtain an explicit criterion to check when two different Markov chains share their bridges.

Starting from the characteristics we offer two different descriptions of the reciprocal class, including its non-Markov probabilities.

The first one is based on a pathwise approach and the second one on short time asymptotic. With the first approach one produces a family of functional equations whose only solutions are precisely the elements of the reciprocal class. These equations are integration by parts on path space associated with derivative operators which perturb the paths by mean of the addition of random loops. Several geometrical tools are employed to construct such formulas. The problem of obtaining sharp characterizations is also considered, showing some interesting connections with discrete geometry. Examples of such formulas are given in the framework of counting processes and random walks on Abelian groups, where the set of loops has a group structure.

In addition to this global description, we propose a second approach by looking at the short time behavior of a reciprocal process. In the same way as the Markov property and short time expansions of transition probabilities characterize Markov chains, we show that a reciprocal class is characterized by imposing the reciprocal property and two families of short time expansions for the bridges. Such local approach is suitable to study reciprocal processes on general countable graphs. As application of our characterization, we considered several interesting graphs, such as lattices, planar graphs, the complete graph, and the hypercube.

Finally, we obtain some first results about concentration of measure implied by lower bounds on the reciprocal characteristics.

Zusammenfassung

Diese Dissertation behandelt die reziproke zufällige Prozesse mit Sprüngen. Gegeben eine zeitkontinuierliche Markovkette als Referenzdynamik, ist die assoziierte reziproke Klasse die Menge aller Wahrscheinlichkeiten auf dem Pfadraum, die als eine Mischung ihrer Brücken geschrieben werden kann. Reziproke Prozesse zeichnen sich durch eine Form der bedingten Unabhängigkeit aus, die die Markoveigenschaft verallgemeinert. Ursprünglich ist diese Idee auf Schrödinger zurückzuführen, der nach einer probabilistischen Interpretation für die Quantenmechanik suchte. Einer reziproken Klasse wird eine Familie reziproker Charakteristiken assoziiert. Dies sind Raum-Zeit Abbildungen, die die reziproke Klasse eindeutig definieren. Wir berechnen diese Charakteristiken explizit und unterteilen sie in zwei Typen: Bogen-Charakteristiken und Kreis-Charakteris-

tiken. Zusätzlich erhalten wir ein klares Kriterium zur Prüfung wann die Brücken von zwei verschiedenen Markovketten übereinstimmen.

Wir beschreiben auf zwei verschiedene Arten reziproken Klasse und berücksichtigen auch ihre nicht-Markov Elemente. Die erste Charakterisierung basiert auf einem pfadweisen Ansatz, während die zweite kurzzeit Asymptotik benutzt. Der erste Ansatz liefert eine Familie funktionaler Gleichungen deren einzige Lösungen die Elemente der reziproken Klasse sind. Die Gleichungen können als partielle Integration auf dem Pfadraum mit einem Ableitungsoperator, der eine Störung der Pfade durch zusätzliche zufällige Kreise hervorruft, interpretiert werden. Die Konstruktion dieser Gleichungen benötigt eine geometrische Analyse des Problems. Wir behandeln außerdem die Fragestellung einer scharfen Charakterisierung und zeigen interessante Verbindungen zur diskreten Geometrie. Beispiele, für die wir eine solche Formel finden konnten, sind für Zählprozesse und für Irrfahrten auf abelschen Gruppen, in denen die Menge der Kreise eine Gruppenstruktur erweist.

Zusätzlich zu diesem globalen Zugang, erforschen wir eine lokale Beschreibung durch die Analyse des kurzfristigen Verhaltens eines reziproken Prozesses. Analog zur Markoveigenschaft und kurzzeit Entwicklung ihrer Übergangswahrscheinlichkeit Markovketten charakterisieren, zeigen wir, dass eine reziproke Klasse charakterisiert werden kann indem wir ihre reziproke Eigenschaft und zwei Familien von Kurzzeit Entwicklungen der Brücken voraussetzen. Solche lokalen Ansatz ist geeignet, um Sprungprozesse auf allgemeine zählbaren Graphen zu studieren. Als Beispiele unserer Charakterisierung, betrachten wir Gitter, planare Graphen, komplette Graphen und die Hyperwürfel.

Zusätzlich präsentieren wir erste Ergebnisse über Maßenkonzentration eines reziproken Prozesses, als Konsequenz unterer Schranken seiner Charakteristiken.

Riassunto

In questa tesi si studiano le *classi reciproche* delle catene di Markov. Data una catena di Markov a tempo continuo su uno spazio numerabile, che svolge il ruolo di dinamica di riferimento, la sua classe reciproca è costituita da tutte le leggi sullo spazio dei cammini che si possono scrivere come un miscuglio dei ponti della legge di riferimento. Questi processi stocastici godono di una proprietà di indipendenza condizionale che generalizza la proprietà di Markov ed è ispirata ad un'idea avuta da Schrödinger nel tentativo di derivare un'interpretazione stocastica della meccanica quantistica.

A ciascuna classe reciproca è associato un insieme di *caratteristiche reciproche*. Una caratteristica reciproca è una proprietà della dinamica di riferimento che viene trasmessa a tutti gli elementi della classe, e viene espressa matematicamente da un'opportuna combinazione di funzionali del generatore della catena di riferimento. Nella tesi, le caratteristiche vengono calcolate esplicitamente e suddivise in due famiglie principali: le *caratteristiche di arco* e le *caratteristiche di ciclo*. Come sottoprodotto, otteniamo un criterio esplicito per decidere quando due catene di Markov hanno gli stessi ponti.

A partire dalle caratteristiche reciproche, vengono proposte due caratterizzazioni della classe reciproca, compresi i suoi elementi non Markoviani. La prima è basata su un approccio traiettoriale, mentre la seconda si basa sul comportamento asintotico locale dei processi reciproci. Utilizzando il primo approccio, si ottiene una famiglia di equazioni funzionali che ammette come soluzioni tutti e soli gli elementi della classe reciproca. Queste equazioni sono integrazioni per parti sullo spazio dei cammini associate ad operatori differenziali che perturbano le traiettorie del processo canonico con l'aggiunta di loops casuali. Nella costruzione di queste equazioni si impiegano tecniche di geometria discreta, stabilendo un interessante collegamento con risultati recenti in questo campo. Le caratterizzazioni ottenute sono ottimali, in quanto impiegano un numero minimo di equazioni per descrivere la classe. Con questo metodo vengono studiate le classi reciproche di processi di conteggio, di camminate aleatorie su gruppi Abeliani, dove l'insieme dei cicli gode anch'esso di una struttura di gruppo. Il secondo approccio, di natura locale, si basa su stime asintotiche in tempo corto. È ben noto come una catena di Markov sia caratterizzata dal fatto di possedere la proprietà di Markov e dal comportamento in tempo corto delle probabilità di transizione. In questa tesi mostriamo che una classe reciproca è caratterizzata dalla proprietà reciproca, e da due famiglie di stime asintotiche per i ponti del processo. Questo approccio lo-

cale permette di analizzare le classi reciproche di passeggiate aleatorie su grafi generali. Come applicazione dei risultati teorici, consideriamo i lat-
tici, i grafi planari, il grafo completo, e l'ipercubo discreto.
Infine, otteniamo delle stime di concentrazione della misura e sul com-
portamento globale dei ponti, sotto l'ipotesi di un limite inferiore per le
caratteristiche reciproche.

Acknowledgments

My first thanks are for my parents. For my father, who silently inspired me his passion for mathematics. For my mother, for having always supported me with many loud words.

I thank my brother Zeno for all the interesting non-mathematical activities we did together.

I owe my gratitude and my scientific growth to my advisors: Prof.Sylvie Roelly taught me how to merge in a unique effort the interest for science, and the interest for persons, simply as they are. Her endless energy in dealing with any aspect of the scientific life probably comes out of all the good she sees in this. I will try to make this attitude mine.

Prof.Paolo Dai Pra introduced me to Probability some years ago, and has never lost interest in my progresses. Besides having read all my horrible drafts, he knows and understands my scientific interests and preferences more than anybody else, and always helped me in finding the way towards what I like.

Even though he was not my advisor, Prof.Christian Léonard gave me a personal guided tour of Probability theory, as he likes it. This was great.

I had the pleasure to share an office and many coffees in Potsdam with Sara. In Berlin I stayed with Adrian, Alberto, and Atul. I had great football conversations with Adrian, while Atul offered me lessons of Rough Path theory, and often Kebab. With Alberto we investigated in depth his character of "artista maledetto", without taking this matter too seriously. It was so much fun.

I thank Giuseppe and Matti. They have been just in front of my door for the past two years, always with lots of "grinta" to share with me.

I thank Paolo and Michele for all the serious and all the stupid discussions we had in these years over Skype and in our meetings in Paris.

I thank Niccolò, Federico e Igor. They have always been waiting for me in Verona. I will now be able to organize few of the many "grigliate" which I have been announcing for years.

Many thanks to my cousins and family. I am looking forward for staying with you in San Zeno.

Anna arrived in my life right before I would leave for Berlin. However, I had the easiest long-distance relationship one could imagine, simply because she was so unconditionally supportive. If anything beautiful is hidden behind the calculations and proofs of this thesis, I dedicate it to her.

Contents

Introduction	1
1 The Schrödinger Problem	17
1.1 Statement of the problem	17
1.1.1 A small thought experiment	17
1.1.2 Statement of the entropy minimization problem	18
1.2 Representation of the solution	19
1.2.1 Decomposition of the entropy	19
1.2.2 A generalized Schrödinger problem	23
2 Reciprocal processes and continuous time Markov chains	25
2.1 The reciprocal property	26
2.1.1 Definition	26
2.1.2 The relation with the Markov property	26
2.2 The concept of reciprocal class	28
2.2.1 Probabilities with the same bridges	28
2.2.2 A representation result	30
2.2.3 Markov probabilities of a reciprocal class	31
2.3 Our framework	33
2.3.1 Markov chains as walks on a graph	33
2.3.2 The reference Markov walk and its reciprocal class	35
2.3.3 Girsanov Theorem for random walks on a graph	37
3 Counting processes	41
3.1 Duality formula for a counting process	43
3.1.1 Nice counting processes	43
3.1.2 The derivative and divergence operators	44
3.1.3 Duality formula	47
3.2 The duality formula characterizes the reciprocal class	50
3.2.1 The reciprocal characteristic	50
3.2.2 The Poisson case	54

3.2.3	The general case	55
3.3	Lower bounds on the reciprocal characteristics	57
4	Random walks on lattices	63
4.1	The jump matrix and some other notation	66
4.2	Time and space transformations	69
4.2.1	Time changes	69
4.2.2	Space transformations	71
4.3	Characterization of the reciprocal class	76
4.3.1	Main result	76
4.3.2	Comparing random walks through characteristics	81
4.4	Lattices and their generating sets	84
4.4.1	Some heuristics	85
4.4.2	Existence of a finite generating set	87
4.5	Sharp characterizations of conditional laws	90
4.5.1	Refining the main theorem	92
4.6	Short-time asymptotics of cycles	92
4.7	Characteristics and concentration of measure	97
4.7.1	A simple question	97
4.7.2	Poissonian concentration revisited	100
4.7.3	A useful staircase	103
5	Random walks on Abelian groups	109
5.1	Random walks and random measures	112
5.2	An iterated formula	115
5.3	Loops and their skeletons	117
5.3.1	Main result: an integral characterization of the reciprocal class	119
5.4	Proof of the main theorem	120
5.5	Examples	126
5.5.1	The group is infinite	127
5.5.2	G is the cyclic group $\mathbb{Z}/N\mathbb{Z}$	129
5.5.3	The state space is a product group	131
6	Random walks on a general graph	133
6.1	Preliminaries	135
6.1.1	Directed subgraphs associated with an intensity	135
6.1.2	Directed subgraphs associated with a random walk	136
6.1.3	Gradients and generating sets of cycles	137
6.2	Main results	138
6.2.1	Reciprocal characteristics	138

6.2.2	The main results	140
6.3	Proofs of the main results	143
6.3.1	Proof of Theorem 6.2.1	148
6.3.2	Proof of Theorem 6.2.2	149
6.3.3	Proof of Corollary 6.2.1	151
6.3.4	Proof of Proposition 6.2.1	151
6.4	More results	152
6.4.1	Jump intensity	152
6.4.2	Characteristic equation	153
6.4.3	Homogeneous walks	156
6.5	Examples	157
7	Perspectives	173
	Bibliography	174

List of Notations

General notation

$(\mathcal{X}, \rightarrow)$ A directed graph

$(X_t)_{t \in [0,1]}$ The canonical process

\mathcal{A} The arcs of $(\mathcal{X}, \rightarrow)$

$\mathcal{B}(\cdot)$ Space of bounded measurable functions

$\mathcal{P}(\cdot)$ The space of probability measures

Ω The path space

$\mathfrak{R}(R)$ The reciprocal class of the reference walk R

\mathcal{X} A countable state space

$C_b(\cdot)$ Space of continuous bounded functions

P A probability

$P(\cdot|\mathcal{G})$ Conditional expectation given the σ algebra \mathcal{G}

$P(F)$ Expectation of the functional F under P , sometimes also denoted $E_P(F)$

P^x A random walk started in x

P^{xy} The xy bridge of P

P_0 The initial distribution

P_t The marginal at time $t \in [0, 1]$.

P_{01} The endpoint marginal

P_I The image measure of X_I under P

R The reference walk

X_I The collection of random variables $(X_t)_{t \in I}$, for $I \subseteq [0, 1]$

Chapter 3

Ξ_j The reciprocal characteristic

\mathbf{P} The standard Poisson process

G_j The density of the reference walk w.r.t. to the Poisson process

\mathcal{D}_u Derivative operator

Chapter 4

$\Xi_j(l, s, t)$ Arc characteristic

$\ker_{\mathbb{Z}}(\mathbf{A})$ The lattice $\{\mathbf{z} \in \mathbb{Z}^A : \mathbf{A}\mathbf{z} = 0\}$

\mathbb{A} The set of possible jumps

\mathbf{A} The jump matrix

\mathfrak{p}_λ A multidimensional Poisson law

Φ_j^c Cycle characteristic

θ_v Shift transformation in \mathbb{Z}^A

Chapter 5

$(G, +)$ A countable Abelian group

Γ The space $[0, 1] \times G$

\mathcal{S}_Γ The space of point measures over Γ

\mathcal{L}^+ The loop-skeletons

$\Phi_{\varphi^*}^\nu$ Reciprocal characteristics

Chapter 6

$\mathcal{A}_\rightarrow(k)$ Active arcs of the intensity k when $\mathcal{A}_{\leftrightarrow}(t, k)$ does not depend on t

$\mathcal{A}_\rightarrow(P)$ Active arcs of P

$\mathcal{A}_{\rightarrow}(t, k)$ Active arcs of the intensity k at time t

$\mathcal{A}_{\rightarrow}^R(x, \mathcal{Y})$ Set of arcs that supported by a bridge R^{xy} , where $y \in \mathcal{Y}$

$\mathcal{A}_{\leftrightarrow}(t, k)$ Symmetric extension of $\mathcal{A}_{\rightarrow}(t, k)$

$\chi_a[P](t, z \rightarrow z')$ Arc characteristic

$\chi_c[P](t, \mathbf{c})$ Cycle characteristic

$\mathcal{X}(P)$ Vertices visited by P

$\mathcal{X}^R(x, \mathcal{Y})$ Set of endpoints of the arcs in $\mathcal{A}_{\rightarrow}^R(x, \mathcal{Y})$

Introduction

Reciprocal probabilities evolved from an idea of Schrödinger, who wanted to derive a stochastic interpretation of quantum mechanics.

In two papers [73] and [74] entitled “*Über die Umkehrung der Naturgesetze*” and “*La théorie relativiste de l’électron et l’interprétation de la mécanique quantique*” he introduced what is nowadays known as the *Schrödinger problem*. He himself provides a neat statement of it. The following is taken from [74]:

Imaginez que vous observez un système de particules en diffusion, qui soient en équilibre thermodynamique. Admettons qu’ à un instant donné t_0 vous les ayez trouvées en répartition à peu près uniforme et qu’ à $t_1 > t_0$ vous ayez trouvé un écart spontané et considérable par rapport à cette uniformité. On vous demande de quelle manière cet écart s’est produit. Quelle en est la manière la plus probable ?

In mathematical terms the problem is formulated as a constrained entropy minimization problem. The entropy is taken with respect to a path measure which models the motion in equilibrium of the particles, and is called the *reference measure*. The constraint is that the marginal distributions at times t_0 and t_1 are prescribed by empirical observations, and it shapes what Schrödinger calls *un écart considérable*. One year after Schrödinger, Bernstein made the observation that the Markov property may be replaced by another dependence structure, in order to better describe the dynamical properties of the solutions of the Schrödinger problem. His idea was that a more time-symmetric notion should come into play. He writes in [3] that:

“[...] si l’on veut reconstituer cette symétrie entre le passé et le futur [...] il faut renoncer à l’ emploi des chaînes de type Markov et les remplacer par de schémas d’une nature différente. ”

He then introduced in [3] the *reciprocal property*, as a weaker version of the Markov property. It is a time Markov field property. At this point, it

should be said that Bernstein was very likely not aware that the solutions to the classical Schrödinger problem are indeed Markovian probabilities. Therefore it was probably not necessary to generalize the Markov property at that point. But the property he introduced is shaped to describe the dynamics of the solutions of a slightly modified version of the Schrödinger problem, which we discuss in some detail in Chapter 1, and led to many further fruitful mathematical developments.

The one which is of primary interest for this thesis is the study of reciprocal classes of Markov processes.

A mathematical rigorous study of reciprocal probabilities was initiated by Jamison in the articles [38],[39], and [40]. He noticed that the reciprocal property is strictly weaker than the Markov one. This observation led him to introduce reciprocal transition probabilities and to formulate a list of axioms that encode the reciprocal property: they are essentially the reciprocal analogous of the Chapman-Kolmogorov consistency equation. One of his results is that, given a reciprocal transition kernel satisfying these axioms, there exists a unique reciprocal process associated with it.

Furthermore, Jamison explicitly characterized the covariance structure of reciprocal Gaussian processes through some differential equation. The theory of reciprocal Gaussian processes was further developed by Chay [12], Carmichael, Mass and Theodorescu in [11], and extended to the multivariate case by Levy [50].

The concept of *reciprocal class* is a bit more recent, even though it appears in an implicit form in [39]: it is the set of all path measures sharing the bridges with a given reference probability, which is assumed to be Markovian. Many authors focused on the case when the reference probability is a Brownian diffusion process: Krener started the search for *reciprocal characteristics* (often called reciprocal invariants) in [41]. He conjectured, using short time expansion of conditional probabilities, that the reciprocal class of a Brownian diffusion is described by some special functionals of the drift of the reference process. Clark gave a positive answer to this question in [17, Thm 1]. He provided what he calls a “local” characterization of reciprocal diffusions. His result is a characterization of the reciprocal class which tells what form the *semimartingale characteristics* of a reciprocal process should take in order for it to be in the reciprocal class of a Brownian diffusion. Such requirements are expressed in a list of equations, and each equation defines one of the reciprocal characteristics. In the paper [42, Thm2.1], Krener gave a full probabilistic interpretation of the characteristics. Each reciprocal diffusion in a reciprocal class is shown to satisfy a family of short time expansions, using heat kernel asymp-

otics, where the coefficients of the leading terms are expressed through the characteristics. All these expansions are inspired by the goal of developing a 'second order differential calculus' for diffusion processes. In particular, the *conditional mean acceleration* of a Brownian diffusion contains all information about the reciprocal class (see equation 2.18 in [42]). In the same article, he also established that the most likely path (i.e. the minimizer of the *Onsager Machlup functional*) of a Brownian diffusion satisfies an ODE expressed in terms of the reciprocal characteristics. Some years later, Roelly and Thieullen succeeded in characterizing the whole reciprocal class of a Brownian diffusion in [67] and [68] including the non Markovian elements, using duality formulae related to Malliavin calculus. Both results are condensed in the short survey [66]. This approach is based on earlier work of Roelly and Zessin [69] who characterized the law Brownian diffusion through a duality formula, which relates the Malliavin derivative operator with a compensated stochastic integral operator. In contrast with Clark's characterization, this is a non-local characterization. The derivatives which are computed there are not in short time, but are Fréchet derivatives on path space. Therefore, it is a pathwise approach. The key idea Roelly and Thieullen had was to look for probabilities satisfying the duality only within a well chosen set of directions of differentiation, namely the loops. Indeed, imposing the duality with respect to *all* directions of differentiation is too restrictive, since the reference diffusion is then the only solution, up to its initial distribution. What we have described so far are the main mathematical steps that motivated the work of this thesis. They constitute the starting point of our investigations, together with Murr's phd thesis [56], who started to study reciprocal counting processes. However, many other fields of research have established a fruitful interaction with the theory of reciprocal processes. Let us give a very concise overview of what seem to be the most important ones.

Stochastic Mechanics The time symmetric features of the reciprocal property inspired many authors, who continued Schrödinger's original program in several different directions. Stochastic mechanics, which is roughly the program of explaining quantum mechanics by using the idea that particle trajectories are governed by diffusion processes, has a long history, dating back at least to Nelson's book [58]. However, Nelson's notion of stochastic acceleration of a diffusion as well as Cruzeiro-Zambrini one ([81, 24] in the context of Euclidean quantum mechanics) are not "reciprocal invariants". It is the theory developed by Krener and Thieullen (see [41, 51, 77, 52]) that connects reciprocal processes and stochastic mechan-

ics. Krener introduced a notion of acceleration for a diffusion process, which is different from Nelson's acceleration and is expressed in terms of the reciprocal characteristics. Such an acceleration is one of the postulates which define the notion of solution to a "second order" stochastic differential equation : indeed the development of a second order calculus (see [77, sec.4,5] and [42, sec3]), based on reciprocal characteristics, is one of the most relevant contributions of these studies: to each reciprocal class is shown to be associated an Euler-Lagrange equation [77, sec 6], and a family of conservation laws for the mass and the momentum [42, sec 5]. Finally let us mention that similar ideas stand behind the Stochastic Calculus of variations [78], where a stochastic version of Noether's Theorem is derived, and in [64].

Optimal transport Mikami established in [55] an interesting connection between the Schrödinger problem and the Monge-Kantorovich problem. He showed that one can construct a solution to the Monge-Kantorovich problem with quadratic cost by considering the zero-noise limit of a sequence of static Schrödinger problems, where the reference dynamics is a Brownian motion. Léonard in [46] extended this result to arbitrary cost functions, showing that the Monge-Kantorovich problem associated with a given cost function c is the Γ -limit of a sequence of Schrödinger problems where the reference dynamics obeys a Large Deviation Principle with rate function given by c . The fact that, in the small noise regimes, the trajectories of a diffusion process stay close to geodesics with very high probability suggest that the limit of solutions to the Schrödinger problem converges to the so called *displacement interpolation* in optimal transport, and this is indeed proven in [46]. This justifies the fact that sometimes solutions to the Schrödinger problem are called entropic interpolations.

There has been a very recent upsurge in the research around this connection, mostly motivated by application in control engineering, due to Chen, Pavon and Georgiu. In a series of papers, they look for implementable solution of the Schrödinger problem. They extended the Benamou-Brenier fluid dynamical formulation of the optimal transport problem ([16], [14]) to the Schrödinger problem, and perform explicit computations in the Gaussian cases, including some degenerate situations when the diffusion matrix of the reference process is singular.

Stochastic control Building on earlier works of Wakolbinger [79], and Dai Pra and Pavon [27], Dai Pra formulated in [25] the Schrödinger problem as a stochastic control problem. The control is represented by a cor-

rection term that can be added to the drift of the reference process, and it is said to be admissible if it steers the diffusion to the desired final law at time 1. The problem is to find the control such that the resulting diffusion minimizes the relative entropy with respect to the reference diffusion. It is tackled with PDE methods: the optimal control is shown to satisfy a second order Hamilton-Jacobi-Bellmann-type equation. Renewed interest in the direction of applications stems from [15], [14].

The contribution of this thesis

This thesis contains a systematic study of reciprocal classes of continuous time Markov random walks on countable state spaces. We give here an overview of the results. To fix ideas, we give some definitions, which are maybe not entirely precise at this stage, but immediate to understand. Precise statements are given in the main body of the thesis, which is independent from this short overview.

We consider a countable directed graph $(\mathcal{X}, \rightarrow)$. An arc from z to z' is denoted $z \rightarrow z'$. The whole arc set is denoted \mathcal{A} . The space of càdlàg piecewise constant functions on \mathcal{X} , whose jumps take place using only the arcs in \mathcal{A} is our path space, and we call it Ω . Probabilities on Ω are called walks, even when they are not Markovian. A continuous time Markov walk R on $(\mathcal{X}, \rightarrow)$ is specified uniquely through an intensity of jump $j : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_+$ and an initial distribution μ . This process is the *reference walk*. The xy bridge of R is denoted R^{xy} and the joint law at times 0 and 1 by R_{01} . We study its reciprocal class $\mathfrak{R}(R)$, that is, the set of random bridges of R .

$$\mathfrak{R}(R) := \left\{ P = \int_{\text{supp}(R_{01})} R^{xy}(\cdot) \pi(dx dy); \pi \in \mathcal{P}(\mathcal{X}^2) \right\}$$

where $\mathcal{P}(\mathcal{X}^2)$ is the space of probability measures on \mathcal{X}^2

Reciprocal characteristics As we will see, a reciprocal class is constituted by many Markov elements, such as R , all its bridges and all its Doob h-transforms, but the most of the class is made of non-Markov probabilities. The first step for understanding it is to give a criterion to decide when do two Markov process belong to the same reciprocal class.

Roughly all Markovian walks can be characterized via their *jump intensity*. That is, to each $P \in \mathcal{P}(\Omega)$ is associated a function $k(t, z \rightarrow z')$, such

that

$$P(X_t + h = y | X_{[0,t]}) \approx h k(t, z \rightarrow z'), \quad \text{as } h \downarrow 0$$

The above mentioned criterion, should be given in terms of the intensities. We call the set of regular intensities \mathcal{K} :

$$\mathcal{K} := \{k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_+, k(\cdot, z \rightarrow z') \in C_b^1([0, 1]) \forall z \rightarrow z' \in \mathcal{A}\}$$

It is natural to give the following informal definition:

Definition (Reciprocal characteristics). *A functional $\chi : \mathcal{K} \rightarrow \mathbb{R}$ is a **reciprocal characteristic** if and only if for any pair of Markov walks R and P of intensities j and k respectively:*

$$P \in \mathfrak{R}(R) \Rightarrow \chi(j) = \chi(k) \quad (1)$$

It is one of the contribution of this thesis to show existence of the characteristics and to compute them explicitly where the reference walk R is a random walk on a countable graph. We refer to Definition 3.2.1, Definition 4.2.2, Definition 4.3.1, Corollary 5.3.1, and Definition 6.2.1, which is the most general form. Definition 3.2.1 had already been given by R.Murr in [56].

The characteristics are divided into two main categories: the arc characteristics and the cycle characteristics, see the two figures below.

Definition (Arc and cycle characteristics: informal). (i) *Let $(z \rightarrow z')$ be an arc of the $(\mathcal{X}, \rightarrow)$. Then*

$$\chi_a[P](t, z \rightarrow z') = \partial_t \log k(t, z \rightarrow z') + \bar{k}(t, z') - \bar{k}(t, z)$$

*is the **arc characteristic** associated with $z \rightarrow z'$. We denote by $\bar{k}(t, z)$ the total intensity of jump: $\bar{k}(t, z) := \sum_{z': z \rightarrow z'} k(t, z \rightarrow z')$.*

(ii) *Let $\mathbf{c} := (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_{|\mathbf{c}|} = x_0)$ be a cycle in the graph $(\mathcal{X}, \rightarrow)$. Then*

$$\chi_c[P](t, \mathbf{c}) := \prod_{(x_i \rightarrow x_{i+1}) \in \mathbf{c}} k(t, x_i \rightarrow x_{i+1})$$

*is the **cycle characteristic** of P associated with \mathbf{c} .*

$$\chi_c[R](t, \mathbf{c}) = j(x_0 \rightarrow x_1)j(x_1 \rightarrow x_2)j(x_2 \rightarrow x_0)$$

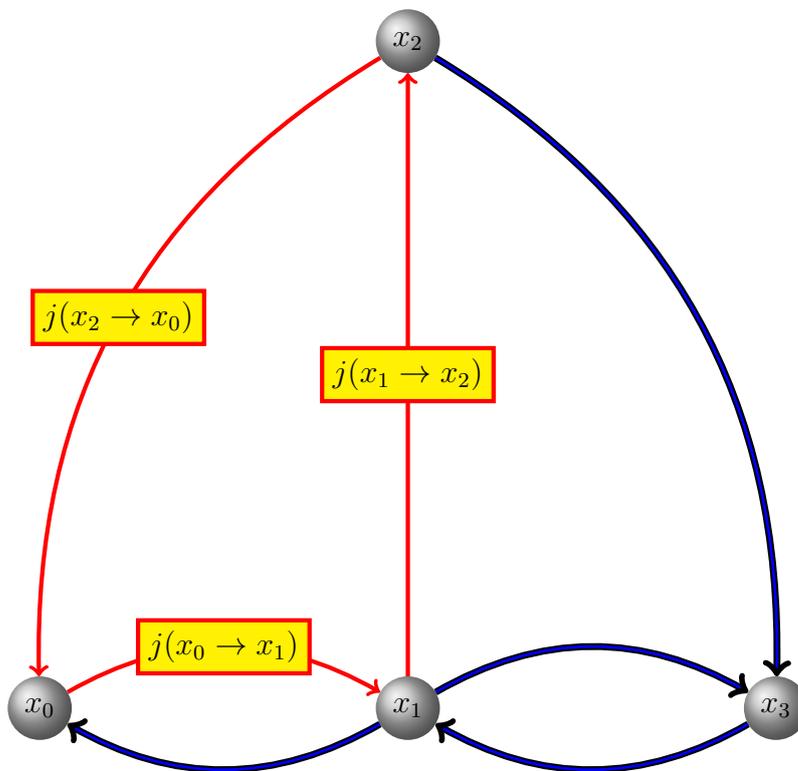


Figure 1: A random walk on the bridges of Königsberg: the reciprocal characteristic associated with the cycle $\mathbf{c} := (x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_0)$ is $\chi_c[R](t, \mathbf{c})$. In this picture, we have chosen a time-homogeneous reference intensity j .

$$\chi_a[R](t, x_1 \rightarrow x_2) = j(x_2 \rightarrow x_3) + j(x_2 \rightarrow x_0) - j(x_1 \rightarrow x_2) - j(x_1 \rightarrow x_3) - j(x_1 \rightarrow x_0)$$

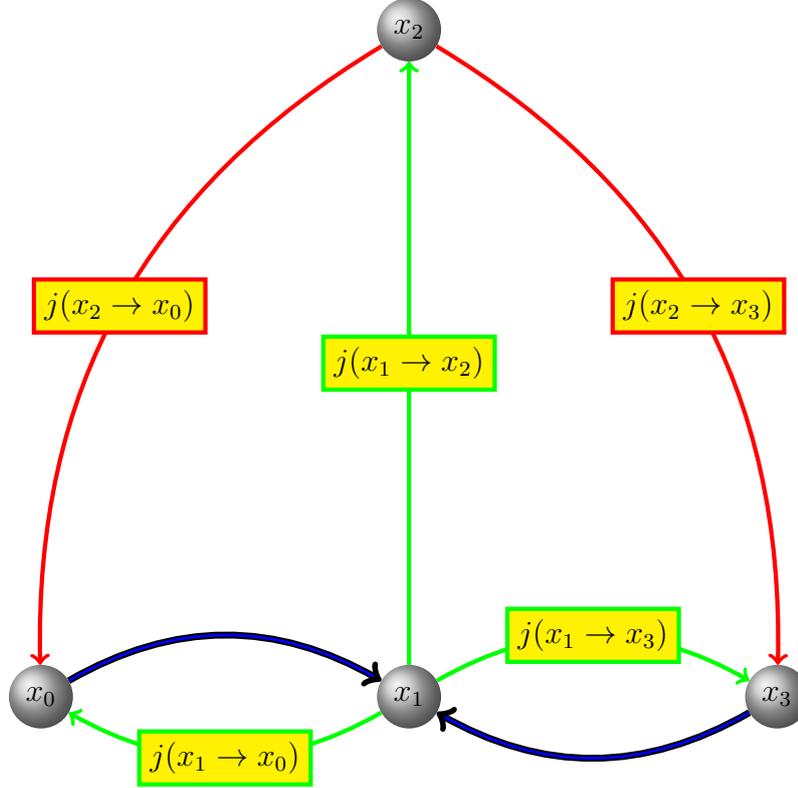


Figure 2: The reciprocal characteristic $\chi_a[R](t, x_1 \rightarrow x_2)$ associated with the arc $x_1 \rightarrow x_2$, assuming a time homogeneous intensity j .

Once characteristics have been identified, the most natural question is to find a minimal set of such χ that yields the reverse implication in equation (1). We are asking the question if it is possible to find a “basis” of the characteristics. Summarizing we look at the following:

Problem: Find a minimal \mathcal{X} such that

$$\chi(k) = \chi(j) \forall \chi \in \mathcal{X} \Rightarrow P \in \mathfrak{R}(R)$$

In this thesis, we answer this question in several interesting cases, such as Cayley graphs, lattices, the complete graph and planar graphs. The answers are very sensitive to the structure of $(\mathcal{X}, \rightarrow)$, and often are connected to its fine geometrical properties. For instance, in sections 4.4, and

4.5 we make use of some recent findings in discrete geometry to answer the above-mentioned problem.

Our results contain as a byproduct an efficient criterion for checking when Markov processes have the same bridges especially because it is explicit in terms of the jump intensities. Other criteria have been proposed, e.g. in [32], but they are implicit and not directly checkable. On the other hand, regarding diffusion processes, a similar result was proven by Clark [17], and in less generality by Benjamini and Lee [2].

Once one has understood the full picture concerning the Markov elements of $\mathfrak{R}(R)$, it is the time to look at the non Markov ones. The purpose is to employ the characteristics to go beyond the Markovian framework. We explored two ways of doing this: the duality formulae approach of Roelly and Thieullen [67, 68] and the short-time expansion of conditional probabilities.

Duality formulae: Chapters 3,4,5. With this approach one produces a family of functional equations whose only solutions are precisely the elements of $\mathfrak{R}(R)$. One of our contributions is a fairly robust scheme to construct such equations, inspired by the seminal works [67, 68]. We are going to describe it in the next lines. We use the word *duality formula* to equivalently refer to an *integration by parts on path space (IBPF)* or to a *change of measure formula*, and we shall see IBPFs as an infinitesimal version of change of measure.

What we mean by change of measure formula is the following:

Definition (Change of measure formula). *Let $\Psi : \Omega \rightarrow \Omega$ be a measurable map. Assume that the image measure $P \circ \Psi^{-1}$ is absolutely continuous with respect to P , and its density is G_{Ψ}^P . Then the relation*

$$P(F \circ \Psi) = P(F G_{\Psi}^P) \quad \forall F \in \mathcal{B}^+(\Omega)$$

*is called the **change of measure formula** associated with Ψ .*

The best known example of a change of measure formula on a path space is Girsanov's Theorem. In that case $\Omega = C([0, 1]; \mathbb{R})$, R is the Wiener measure and for ψ regular enough, $\Psi = \theta_{\psi}$ is the translation by ψ :

$$\theta_{\psi} : \Omega \longrightarrow \Omega, \quad \omega \mapsto \omega + \psi$$

We have that:

$$G_{\theta_{\psi}} := \exp \left(\int \psi_t d\omega_t - \frac{1}{2} \int_0^1 \psi_t^2 dt \right)$$

Assume now that for some Ψ you have the change of measure under the reference walk R :

$$R(F \circ \Psi) = R(FG_{\Psi}^R) \quad \forall F \in \mathcal{B}^+(\Omega). \quad (2)$$

We use the idea that, if Ψ leaves invariant the vector (X_0, X_1) , that is:

$$(X_0, X_1) \circ \Psi = X_0, X_1 \quad (3)$$

then the validity of the formula (2) extends to the whole class $\mathfrak{R}(R)$, keeping the same expression for the density.

$$\forall P \in \mathfrak{R}(R) \quad P(F \circ \Psi) = P(FG_{\Psi}^R) \quad \forall F \in \mathcal{B}^+(\Omega) \quad (4)$$

The reason for this, at least heuristically, is simple, and it goes back to the very general principle that conditioning a probability measure to a subset preserves the ratios between the probabilities of the elements of that subset. When considering bridge R^{xy} , we are conditioning $R \in \mathcal{P}(\Omega)$ to belong to the subset $\{X_0 = x, X_1 = y\}$. Therefore one believes that, whatever meaning the ‘‘probability of a path’’ $R(\omega)$ has, the following holds:

$$\frac{R(\omega)}{R(\tilde{\omega})} = \frac{R^{xy}(\omega)}{R^{xy}(\tilde{\omega})} \quad \forall \omega, \tilde{\omega} \text{ s.t. } \omega_0 = \tilde{\omega}_0 = x, \omega_1 = \tilde{\omega}_1 = y \quad (5)$$

But then, since morally:

$$G_{\Psi}^R(\omega) = \frac{R \circ \Psi^{-1}(\omega)}{R(\omega)},$$

combining the invariance property (3) of Ψ and the observation (5):

$$G_{\Psi}^R(\omega) = \frac{R^{X_0, X_1} \circ \Psi^{-1}(\omega)}{R^{X_0, X_1}(\omega)}$$

which indicates that $G_{\Psi}^R(\tilde{\omega})$ depends only on the bridges of R and therefore, if $P \in \mathfrak{R}(R)$, $G_{\Psi}^P = G_{\Psi}^R$. This motivates (4), and also indicates that the expression of G_{Ψ}^R should be expressed by the reciprocal characteristics. Imposing (4) for enough transformations Ψ , we arrive at the following prototype of result:

Theorem. *Let $P \in \mathcal{P}(\Omega)$. Then $P \in \mathfrak{R}(R)$ if and only if for enough transformations Ψ satisfying the invariance property (3):*

$$P(F \circ \Psi) = P(FG_{\Psi}^R) \quad \forall F \in \mathcal{B}^+(\Omega)$$

The construction of the Ψ , the decomposition of the density in terms of the reciprocal characteristics, and the fact that the formula is rich enough to characterize $\mathfrak{R}(R)$ are all *graph-dependent problems*, which have to be solved *ad hoc*. While it is clear how to shift paths on the Wiener space, this is far from obvious on path spaces built over graphs. Moreover, one has to design the transformations in such a way to respect the initial and final state. When $(\mathcal{X}, \rightarrow)$ has some translation invariant structure, as it is the case for lattices or more general Cayley graphs, we found a natural way of doing this, and devised the geometrical objects which allow to handle the algebraic expressions in a canonical way. We applied this strategy in Theorem 3.2.2, Theorem 4.3.1, and Theorem 5.3.1. In Chapter 3 we look at counting processes, in Chapter 4 at lattices and in Chapter 5 at Abelian groups. In going from Chapter 4 to Chapter 5, we put a geometrical assumption that allows for a factorization of the cycle space of the graph. This assumption is crucial to obtain true pathwise formulas and is satisfied in most of the cases of interest. When this hypothesis fails, several geometrical problems arise. They are discussed in sections 4.4 and 4.5. Therefore, the results of Chapter 5, which are obtained under this hypothesis, when applicable to the lattice case (recall that lattices are special instances of Cayley graphs) not only cover the results of Chapter 4, but improve them considerably. However, the results of Chapter 4, hold for a more general class of graphs. The resulting change of measure formulae are often new generalizations of other well known formulae, such as Slivnjak-Mecke identities [75, 54] or Chen's characterization of the Poisson distribution [13].

Duality formulae correspond to a pathwise viewpoint on reciprocal processes, in the sense that they tell by which amount the "probability of a path" changes when the path is subject to a perturbation which leaves invariant its endpoints.

In contrast with this global description, we have a corresponding local picture, which we obtain by looking at the short time behavior of a reciprocal process, and is illustrated in the next paragraph.

Short time expansions: Chapter 6 This approach leads to a characterization of the reciprocal class through the local (in time) behavior of its elements. With respect to the previous one, it has the advantage to hold for general graphs, even if they do not possess any symmetry. Take a graph $(\mathcal{X}, \rightarrow)$. It is well known that, modulo technical conditions, the reference Markov random walk R is characterized by the following two properties:

(i) The Markov property: for any $s < t \in [0, 1]$, $A \subseteq \mathcal{X}$:

$$R(X_t \in A | X_{[0,s]}) = R(X_t \in A | X_s) \quad R - a.s.$$

(ii) The jump intensity: for any $z \rightarrow z' \in \mathcal{A}$:

$$R(X_{t+h} = z' | X_t = z) = j(t, z \rightarrow z')h + o(h), \quad as \ h \downarrow 0 \quad (6)$$

Is there an analogous theorem for reciprocal processes? We answer affirmatively, by substituting to the Markov property the reciprocal property, which is a time Markov-field property, and to the expansion (6) some other expansions based on the reciprocal characteristics. In the context of diffusion processes, the characteristics were recovered in the short time expansions used to compute stochastic acceleration terms, see [42]. This result fits in the larger program to develop second order calculus for diffusion processes. Without relying on any physical interpretation, we obtained a very natural probabilistic interpretations of the reciprocal characteristics, which is turned into a characterization of the reciprocal class. No second order process is used, we simply compute the conditional probabilities of some suitably chosen events. Indeed, if some cycles of $(\mathcal{X}, \rightarrow)$ are longer than two, which is almost always the case, we need to expand conditional probabilities at higher order, so Krener's approach based on second order expansion is not suitable for graphs. We report a simplified form of Theorem 6.2.1, where we dropped most of the technical assumptions. There, we denote by T_j the j -th jump time of a walk, and by T_j^t the j -th jump time after t .

Theorem (Short-time expansions characterize $\mathfrak{R}(R)$). $P \in \mathcal{P}(\Omega)$ belongs to $\mathfrak{R}(R)$ if and only if the following assertions hold:

(i) P is reciprocal: for any times $s \leq u$ in $[0, 1]$ and for any event, $B \in \mathcal{X}_{[s,u]}$:

$$P(B | X_{[0,s]} \vee X_{[u,1]}) = P(B | X_s, X_u) \quad P - a.s.$$

(ii) For any $t \in (0, 1)$, any $(z \rightarrow z') \in \mathcal{A}$ and any $r \in [0, 1]$:

$$\begin{aligned} P(T_1^t = t+hdr \mid X_t = z, X_{t+h} = z', T_2^t > t+h) \\ = dr + h\chi_a[j](t, z \rightarrow z')(r - 1/2) dr + o(h) \quad ash \downarrow 0. \end{aligned}$$

(iii) For any $t \in (0, 1)$ and any cycle $\mathbf{c} = (x_0 \rightarrow x_1 \dots \rightarrow x_{|\mathbf{c}|} = x_0)$:

$$\begin{aligned} P\left((X_t \rightarrow X_{T_1^t} \rightarrow \dots \rightarrow X_{T_{|\mathbf{c}|}^t} = X_t) = \mathbf{c}, T_{|\mathbf{c}|}^t < t+h < T_{|\mathbf{c}|+1}^t \mid X_t = X_{t+h} = x_0\right) \\ = \chi_c[j](t, \mathbf{c}) \frac{h^{|\mathbf{c}|}}{|\mathbf{c}|!} + o(h^{|\mathbf{c}|}) \quad ash \downarrow 0. \end{aligned}$$

Let us comment on (ii) and (iii):

- (ii) Assume that you observe a reciprocal walk of $\mathfrak{R}(R)$ sitting in z at time t and after a short time interval you see it in z' , where $z \rightarrow z'$ is an arc of $(\mathcal{X}, \rightarrow)$. Then this has essentially happened through a single jump along $z \rightarrow z'$. The arc characteristics $\chi_a[j](t, z \rightarrow z')$ accounts for the distribution of the jump time. A positive arc characteristic implies that this distribution is concentrated around the end of the time interval, whereas a negative characteristic implies that the distribution is concentrated around the beginning of such interval.
- (iii) Assume that you observe a reciprocal walk of $\mathfrak{R}(R)$ sitting in a state x_0 at time t and you observe it there again after a short time interval h . Given this, the probability that in the time-window $[t, t + h]$ the walk has traveled along the cycle $\mathbf{c} = (x_0 \rightarrow x_1 \dots \rightarrow x_{|\mathbf{c}|})$ is proportional to the reciprocal characteristic $\chi_{\mathbf{c}}[j](t, \mathbf{c})$ of \mathbf{c} and to $h^{|\mathbf{c}|}$, where $|\mathbf{c}|$ is the length of the cycle.

Quantitative estimates for bridges: Sections 3.3 and 4.7 A last contribution of this thesis is to obtain, in some special models, quantitative estimates on the behavior of the bridges of the reference walk. The main point about these results is that they hold in non-asymptotic regimes, in contrast with the short time estimates used to characterize the reciprocal class, and that such estimates are expressed through the reciprocal characteristics, which are the natural parameters for reciprocal classes. This is, to the best of our knowledge, the first time when the role of reciprocal characteristics is made explicit in global estimates concerning bridges.

Our first result concerns counting processes, that is random walks on the graph $(\mathbb{Z}, \rightarrow)$, where $z \rightarrow z' \Leftrightarrow z' = z + 1$. In this type of graphs, there are no cycles, and therefore only the arc characteristics matter. A lower bound on the arc characteristics Ξ_j is shown to imply an estimate on the last jump time of the bridges of the reference walk. In particular, a positive bound implies that the bridge of the reference walk is *slower* than the bridge of a Poisson process, in the sense that it tends to reach its final state later than the Poisson bridge, and we have an accumulation of the jump times around time one. The following statement formalize this. It is Proposition 3.3.1.

Proposition. *Let R^{0n} be the bridge between 0 and n of R . Assume that*

$$\inf_{t \in [0,1], 0 \leq i \leq n-1} \Xi_j(i, t) \geq c \in \mathbb{R}$$

Then:

$$R^{0n}(T_n \leq t) \leq \left(\frac{\exp(ct) - 1}{\exp(c) - 1} \right)^n$$

Our second result is a concentration inequality for the number of jumps of the bridge of a continuous time random walk on $(\mathbb{Z}, \rightarrow)$, where the only allowed jumps are either of size -1 or k , where $k \in \mathbb{N}$. That is, $z \rightarrow z' \Leftrightarrow z - 1z' = z + k$. Under the reference walk the number of jumps of height k simply follows a Poisson law. This is not true under any bridge. We obtain in Chapter 4 (see Corollary 4.7.1) a characterization of this conditional distribution with a change of measure formula, where the role of the cycle characteristic Φ_j is highlighted. What we obtain is a formula that generalizes Chen's characterization of Poisson law, see [13]. Relying on a geometrical interpolation argument (Proposition 4.7.3) and on refinements of previously established concentration of measure results for the Poisson law (Proposition 4.7.1), we establish the following result. Here, by $o(R)$ we denote a function which grows sublinearly as $R \rightarrow +\infty$.

Theorem (Theorem 4.7.1:informal version). *Let $\rho \in \mathcal{P}(\mathbb{N})$ be the distribution of the number of jumps of size k under the 00 bridge of R , R^{00} . Then there exist $C_0 > 0$ such that for all f which are 1-Lipschitz and for all $R > C_0$:*

$$\rho(f \geq \rho(f) + R) \leq \exp(-(k+1)R \log R + (\log(\Phi_j) + C_1)R + o(R))$$

The constant C_1 does not depend on Φ_j . C_0 might depend on it.

Let us comment very briefly on the form of the concentration rate: the leading term is governed by the geometry of the jump set, since it only depends on k whereas the reciprocal characteristic Φ_j drives the exponential correction terms. Such concentration rates are not implied by any of the well known functional inequalities, such as the family of Modified Logarithmic Sobolev inequality studied, among the others, in [5],[26].

The reasoning we made to obtain Theorem 4.7.1 is likely to carry over to the treatment of a more general class of models.

The concentration inequality derived here gains his interest also outside the study of bridges of continuous time random walks. Let us clarify why: Chen's characterization of Poisson random variable is the fact that the Poisson law of mean λ is the only law satisfying

$$\forall f \quad \lambda \rho(f(n+1)) = \rho(f(n)n)$$

The measure ρ studied in Theorem 4.7.1 is shown to be the only solution to the change of measure formula:

$$\forall f \quad \Phi_j \rho(f(n+1)) = \rho(f(n)\gamma(n))$$

where $\gamma(n)$ is a polynomial of degree $k + 1$ (recall that k is the size of the large jump). What is known is that to a linear coefficient on the right hand side of the change of measure formula, as in Chen's formula, corresponds a concentration inequality with rate $-R \log R + (\log(\lambda) + C)R + o(R)$, where C is a numerical constant. Our result shows that to a polynomial coefficient on the right hand side of the change of measure formula corresponds a concentration inequality with rate $-(k+1)R \log R + (\log(\Phi_j) + C)R + o(R)$, where C is a numerical constant. Therefore we establish a clear relation between the form of the density in change of measure formulae and the rate of concentration.

Chapter 1

The Schrödinger Problem

Outline of the chapter This short chapter is meant as an introduction to the Schrödinger problem, which shall motivate the study of reciprocal classes. We give some heuristics that explain its formulation, and prove some structural results for its solution.

1.1 Statement of the problem

1.1.1 A small thought experiment

At time $t = 0$, we are given a large number Y_0^1, \dots, Y_0^N of independent indistinguishable particles. As $N \rightarrow +\infty$, their empirical distribution approaches a smooth profile μ_0 .

$$\frac{1}{N} \sum_{i=1}^N \delta_{Y_0^i} \rightarrow \mu_0, \quad \text{as } N \rightarrow +\infty$$

We let each particle travel independently from all the others with a Brownian motion for a unit of time. The law of large numbers tells that, as $N \rightarrow +\infty$ the empirical measure at time 1, which we call μ_1 , approaches $\tilde{\mu}_1$, defined by:

$$\tilde{\mu}_1(dy) := \int_{\mathbb{R}} r(y|x) \mu_0(dx),$$

where r is the Gaussian kernel. We are allowed to observe the empirical measure μ_1 at time 1. Schrödinger question is the following:

Given that N is very large and μ_1 is significantly different from $\tilde{\mu}_1$, what is the most likely behavior of the whole random system?

We can sketch an heuristic based on the theory of large deviations which explains the mathematical formulation of this question. Such argument was made rigorous by Föllmer in [34].

We call L_N the empirical measure associated to the particle system. Note that such a measure is defined over the space Ω of continuous trajectories, rather than on \mathbb{R} , as it was the case for μ_0 and μ_1 .

$$L^N := \frac{1}{N} \sum_{i=1}^N \delta_{((Y_t^i)_{t \in [0,1]})}$$

We denote by $Prob$ the distribution of L_n . The law of large number tells that L^N converges to a Brownian motion started in μ_0 , whose law R is called the *reference dynamics*. Our observations concerning the initial and final configurations of the particles tell us that:

$$L_N \in \{P : P_0 = \mu_0, P_1 = \mu_1\} \quad (1.1)$$

Using informally Sanov's Theorem (see [29, sec 6.2]) we have that, when N is very large the distribution of L_N is governed by the relative entropy $H(\cdot|R)$:

$$Prob(L_N \in A) \approx \exp(-N \inf_{P \in A} H(P|R)) \quad \forall A \subseteq \mathcal{P}(\Omega)$$

In this interpretation, the "most likely" evolution is clearly given by the minimizer of $H(\cdot|R)$ within the set of measures matching our observation, described in (1.1). We arrive at :

$$H(P|R) \rightarrow \min \quad P \in \mathcal{P}(\Omega), \quad P_0 = \mu_0, P_1 = \mu_1$$

This is the Schrödinger problem.

1.1.2 Statement of the entropy minimization problem

In this section we state rigorously the problem we have just discussed. Although in the presentation above particles were moving according to a Brownian motion, the same questions can be asked in a much more general setting, replacing the Brownian motion with another Markov process. Indeed, in this thesis, we will be concerned with random walks on graphs. We consider a Polish state space \mathcal{X} . The càdlàg space over it is denoted Ω , and the canonical process $(X_t)_{t \in [0,1]}$. All the standard conventions for sigma algebras and filtrations can be read in the table of notation.

Definition 1.1.1. *The dynamic Schrödinger problem associated with $R \in \mathcal{P}(\Omega)$, $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ is the following entropy minimization problem:*

$$H(P|R) \rightarrow \min \quad P \in \mathcal{P}(\Omega), \quad P_0 = \mu_0, P_1 = \mu_1 \quad (1.2)$$

where $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ are the prescribed marginals.

Projecting this problem onto the marginals at times $t = \{0, 1\}$ gives the, apparently simpler, *static Schrödinger problem*. For $\pi \in \mathcal{P}(\mathcal{X}^2)$, $i \in \{0, 1\}$, the image measure $\pi \circ (X_i)^{-1}$ is denoted by π_i .

Definition 1.1.2. *The static Schrödinger problem associated with $R \in \mathcal{P}(\Omega)$, $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X}^2)$ is the following entropy minimization problem:*

$$H(\pi|R_{01}) \rightarrow \min \quad P \in \mathcal{P}(\mathcal{X}^2), \quad \pi_0 = \mu_0, \pi_1 = \mu_1 \quad (1.3)$$

As it is clear from the formulation, there is more than an analogy with an optimal transport problem. Indeed it is shown in [55] and [46] that the classical Monge-Kantorovich problem can be obtained as the limit in a suitable sense of a sequence of (static) Schrödinger problems through a "slowing down" procedure.

1.2 Representation of the solution

1.2.1 Decomposition of the entropy

The first result of this subsection is that is Proposition (1.2.1), which says that the two problems are indeed equivalent. If one can solve the dynamical problem the solution to the static problem is given by a simple projection. The converse is also true. Given a solution to the static problem one obtains a solution to the dynamical problem by mixing bridges of the reference measure according to the solution of the static problem. This was first proven by Föllmer [34], although in a less general setting.

Before presenting the result, we recall that under the current hypotheses, both Ω and \mathcal{X}^2 are Polish spaces, and the projection $(X_0, X_1) : \Omega \rightarrow \mathcal{X}^2$ is measurable. Therefore there exist a regular conditional probability associated with it, that is, there exist a measurable map $P^{xy} : \mathcal{X}^2 \rightarrow \mathcal{P}(\Omega)$ such that for all $A \in \mathcal{F}$:

$$P(A) = \int_{\mathcal{X}^2} P^{xy}(A) P_{01}(dxdy)$$

The $P - a.s.$ well defined measure P^{xy} is called the *xy bridge*.

Under the current assumptions for any $P \in \mathcal{P}(\Omega)$, $\pi \in \mathcal{P}(\mathcal{X}^2)$, $\pi \ll P_{01}$ the probability measure

$$\int_{\mathcal{X}^2} P^{xy}(\cdot) \pi(dxdy)$$

is well defined. Having said this, we can prove the equivalence between the two problems, following [49].

Proposition 1.2.1. *Let $\mu_0, \mu_1 \in \mathcal{P}(\mathcal{X})$ be fixed. The dynamical and static Schrödinger problems both admit at most one solution. If \hat{P} solves the dynamical problem then $\hat{\pi} := \hat{P}_{01}$ solves the static problem associated with R . Conversely, if $\hat{\pi}$ solves the static problem, then*

$$\hat{P} = \int_{\mathcal{X}^2} R^{xy}(\cdot) \hat{\pi}(dxdy) \quad (1.4)$$

solves the dynamical problem.

Proof. Since the admissible region for both problem is a convex subset of either $\mathcal{P}(\Omega)$ or $\mathcal{P}(\mathcal{X}^2)$ and the relative entropy is a strictly convex function, both problems admit at most one solution. Assume now that \hat{P} solves (1.2). Using the well known disintegration formulas for the relative entropy:

$$H(\hat{P}|R) = \int H(\hat{P}^{xy}|R^{xy}) \hat{P}_{01}(dxdy) + H(\hat{P}_{01}|R_{01})$$

we deduce that $P^{xy} = R^{xy} \hat{P}_{01} - a.s.$, for otherwise the probability

$$\tilde{P}(\cdot) = \int_{\mathcal{X}^2} R^{xy}(\cdot) \hat{P}_{01}(dxdy)$$

would be such that $H(\tilde{P}|R) < H(\hat{P}|R)$, which contradicts the optimality of \hat{P} .

Consider now any other π in the admissible region of (1.3). Then the measure

$$Q_\pi(\cdot) = \int_{\mathcal{X}^2} R^{xy}(\cdot) \pi(dxdy)$$

is well defined.

Using again the disintegration formula for the relative entropy we have that $H(Q_\pi|R) = H(\pi|R_{01})$ and $H(\hat{P}|R) = H(P_{01}|R_{01})$. But since \hat{P} solves the dynamic problem, then $H(\pi|R_{01}) > H(\hat{P}_{01}|R_{01})$. This proves that $\hat{\pi}$ solves the static problem.

Conversely, let $\hat{\pi}$ a solution of the static problem and consider \hat{P} as in (1.4). The disintegration formula for the relative entropy tells us that

$H(\hat{P}|R) = H(\hat{\pi}|R_{01})$. Let us remark that \hat{P} is well defined under the current hypothesis. Consider now any Q in the admissible region of the dynamical problem. Then Q_{01} is clearly in the admissible region of (1.3). Using the disintegration formula for the entropy we have:

$$H(Q|R) = \int H(Q^{xy}|R^{xy})Q_{01}(dxdy) + H(Q_{01}|R_{01})$$

As the relative entropy is always non negative, and by assumption $H(Q_{01}|R_{01}) \geq H(\hat{\pi}|R_{01}) = H(\hat{P}|R)$, we conclude that $H(Q|R) \geq H(\hat{P}|R)$ and the conclusion follows. \square

The following proposition gives some information on the shape of the minimizers. It tells that the density of the solution decouples in a product of two functions $f(X_0)g(X_1)$.

f, g can also be found as solutions to the so called Schrodinger system, for which Fortet [35] and Beurling [4] proved the first existence results. A more general statement can be found in Section 2 of [48], where fine questions concerning the support of R_{01} are discussed. For the sake of simplicity, we present a simpler version Theorem 2.8 of [48] under the slightly more restrictive assumption as considered by Ruschendorff and Thomsen [71, Thm 3].

Proposition 1.2.2. *Assume that $R_{01} \ll R_0 \otimes R_1$, and that for some π in the admissible region, $H(\pi|R_{01}) < +\infty$. Then the static problem admits a unique solution $\hat{\pi}$ and there exist two measurable functions $f, g : \mathcal{X} \rightarrow \mathbb{R}_+$ such that*

$$\hat{\pi} = f(X_0)g(X_1)R_{01} \tag{1.5}$$

The functions f, g are R_{01} - a.s. solutions to the Schrödinger system:

$$\begin{cases} \frac{d\mu_0}{dR_0}(x) = f(x)R(g(X_1)|X_0 = x) \\ \frac{d\mu_1}{dR_1}(y) = g(y)R(f(X_0)|X_1 = y) \end{cases} \tag{1.6}$$

We do not give the proof of this theorem here, since the measure-theoretical arguments needed to show existence part are quite technical, and not strictly related to the content of this thesis. We shall rather give an intuition on why convex optimization techniques can be used to prove the factorization (1.5). The same ideas provide an informal derivation of the Schrödinger system.

Consider any π in the admissible region of the static problem. The following representation of the relative entropy is well known:

$$H(\pi|R_{01}) = \sup \left\{ \int_{\mathcal{X}^2} u d\pi : u \in C_b(\mathcal{X}^2), \int_{\mathcal{X}^2} \exp(u) dR_{01} = 1 \right\} \quad (1.7)$$

Consider now $\varphi, \psi \in C_b(\mathcal{X})$ and define $\varphi \oplus \psi \in C_b(\mathcal{X}^2)$ as follows:

$$\varphi \oplus \psi(x, y) := \varphi(x) + \psi(y)$$

Choosing $u = \varphi \oplus \psi$ in (1.7), and using the fact that $\pi_0 = \mu_0, \pi_1 = \mu_1$ we obtain:

$$H(\pi|R_{01}) \geq \int_{\mathcal{X}} \varphi d\mu_0 + \int_{\mathcal{X}} \psi d\mu_1, \quad \forall \varphi, \psi \text{ s.t. } \int_{\mathcal{X}^2} \exp(\varphi \oplus \psi) dR_{01} = 1$$

Let us note that the right hand side of the last identity is independent from the choice of π . Therefore, the optimal value of

$$\int_{\mathcal{X}} \varphi d\mu_0 + \int_{\mathcal{X}} \psi d\mu_1 \rightarrow \max, \quad \varphi, \psi \in C_b(\mathcal{X}), \int_{\mathcal{X}^2} \exp(\varphi \oplus \psi) dR_{01} \quad (1.8)$$

is a lower bound for the optimal value of the Schrödinger problem. Indeed, (1.8) is the dual problem of (1.4). In [45] it is proven that the optimal values of the two problems in most of the cases coincide, and are both attained. Now, assume that we are in one of these cases. If $\hat{\pi}$ is the solution to the static Schrödinger problem and $\hat{\varphi}, \hat{\psi}$ is an optimal pair for the dual problem (1.8) we have:

$$H(\hat{\pi}|R_{01}) = \int_{\mathcal{X}} \hat{\varphi} d\mu_0 + \int_{\mathcal{X}} \hat{\psi} d\mu_1 = \int_{\mathcal{X}^2} \hat{\varphi} \oplus \hat{\psi} d\hat{\pi}$$

one gets that:

$$\frac{d\hat{\pi}}{dR_{01}} = \exp(\hat{\varphi} \oplus \hat{\psi})$$

which partially explains (1.5), with $f = \exp(\hat{\varphi}), g = \exp(\hat{\psi})$. Considering the marginals, we obtain:

$$\begin{cases} \frac{d\hat{\pi}_0}{dR_0}(x) = f(x)R_{01}(g(X_1)|X_0 = x) \\ \frac{d\hat{\pi}_1}{dR_1}(y) = g(y)R_{01}(f(X_0)|X_1 = y) \end{cases}$$

But, since $\hat{\pi}$ is in the admissible region, $\hat{\pi}_0 = \mu_0, \hat{\pi}_1 = \mu_1$, and this gives the system (1.6). For more details, we redirect the interested reader to the proof of Theorem 2.8 in [46]. An interesting consequence Theorem 1.2.2 is that the solution of (1.2) inherits the Markov property from R . We only sketch the proof, since it will follow as a special case of a more general statement, which we prove in Proposition 2.2.3.

Proposition 1.2.3. *Assume that $R_{01} \ll R_0 \otimes R_1$ and that R is a Markov measure. Then the solution to the dynamical problem exists and it is also a Markov probability.*

Proof. Thanks to Proposition 2.2.2 the static problem admits a solution π which takes the form 1.5. Applying Proposition 1.2.1 the dynamical problem also admits a solution, which is:

$$P = \int_{\mathcal{X}^2} R^{xy} \pi(dxdy) = \int_{\mathcal{X}^2} R^{xy} f(x)g(y)R_{01}(dxdy)$$

which is equivalent to say that

$$P \ll R, \quad \text{and} \quad \frac{dP}{dR} = f(X_0)g(X_1) \quad R - a.s.$$

It will be proven in Proposition 2.2.3 that if R is Markov and P takes the form above, then P is Markov as well. \square

It is curious that, at this stage, there doesn't seem to be any need for a generalization of the Markov property to study the solution Schrödinger problem. Indeed Bernstein, in his 1932 paper was not aware that solutions of the Schrödinger problems are Markov. It seems that this has been first been pointed out by Jamison in [39]. However, the reciprocal property, which Bernstein introduced in the same work, is shown to describe exactly the dependence structure of solutions of a very natural generalization of the problem discussed above. That is, we impose a constraint not only on the endpoint marginals separately, but we also prescribe their dependence structure.

The constraint then changes from

$$P_0 = \mu_0, P_1 = \mu_1$$

to

$$P_{01} = \mu \in \mathcal{P}(\mathcal{X}^2)$$

1.2.2 A generalized Schrödinger problem

We now turn the attention to the generalized Schrödinger problem:

Definition 1.2.1. *We define the following entropy minimization problem, associated with $R \in \mathcal{P}(\Omega)$, $\mu \in \mathcal{P}(\mathcal{X}^2)$:*

$$H(P|R) \rightarrow \min \quad P \in \mathcal{P}(\Omega), \quad P_{01} = \mu$$

Let us note that there is not a static problem associated to this problem, as P_{01} is fixed within the admissible region. As in Proposition 1.2.1 we have a nice constructive result for the solution to (1.2.1): It says that the solution is a random bridge of R , where the mixing measure is given precisely by μ , rather than the solution of the static problem (1.3). We skip the proof, as it is completely analogous to that of Proposition 1.2.1.

Proposition 1.2.4. *The problem (1.2.1) admits a solution if and only if $H(\mu|R_{01}) < +\infty$. In this case, the solution is:*

$$\hat{P} = \int_{\mathcal{X}^2} R^{xy} \mu(dxdy) \quad (1.9)$$

Solutions to this last problem are truly reciprocal probabilities. The goal of the next section is to introduce the reciprocal property and give some very general notion about reciprocal probabilities. As a by product, we will obtain that \hat{P} defined in (1.9) is indeed reciprocal.

Chapter 2

Reciprocal processes and continuous time Markov chains

Outline of the chapter The aim of this chapter is to lay the foundations for the study of reciprocal probabilities on discrete structures. We review some basic general results and introduce the concept of reciprocal class of a Markov probability. It is shown to be the set of solution to the generalized Schrödinger problem introduced in the first chapter. We define the main object of study for this thesis: the reciprocal class of a Markov Chain. As a technical tool, which will be used systematically later on, a Girsanov Theorem for continuous time Markov chains is presented at the level of generality needed in this thesis.

The recent survey [49] introduces a measure-theoretical viewpoint on reciprocal processes, in contrast with Jamison reciprocal transition probabilities, and collects many basic results. It serves as a guideline for the first two sections of this chapter.

Organization of the chapter Section 2.1 is a self-contained introduction to reciprocal probabilities. Reciprocal classes are studied in Section 2.2. A first representation results for reciprocal classes is shown at Proposition 2.2.2. In Section 2.3 we specify our notations about continuous time Markov chains, and the assumptions on the reference measure. As a useful tool for the next chapters, a Girsanov theorem is proved.

2.1 The reciprocal property

2.1.1 Definition

A simple description of the Markov property of a probability is that, given the current state X_u , the sigma algebras $X_{[0,u]}$ and $X_{[u,1]}$ are independent. That is, $X_{[0,u]}$ and $X_{[u,1]}$ are independent under $P(\cdot|X_u)$.

The reciprocal property is the fact that for any pair of times $s < u$, given the position (X_s, X_u) of the process at these two times, the behavior of the process in $[s, u]$ is independent from the past up to s and the future from u on. Speaking about sigma algebras, we ask that $X_{[s,u]}$ is independent from $\sigma(X_{[0,s]} \vee X_{[u,1]})$ given (X_s, X_u) . The property we have just stated coincide with that of a Markov field, indexed by time, and reciprocal probabilities can also be seen from this point of view.

Definition 2.1.1. *A probability measure P on Ω is called **reciprocal** if for any times $s \leq u$ in $[0, 1]$ and for any event, $B \in X_{[s,u]}$:*

$$P(B|X_{[0,s]} \vee X_{[u,1]}) = P(B|\sigma(X_s, X_u)) \quad P - a.s. \quad (2.1)$$

Remark 2.1.1. *For any sigma algebra \mathcal{G} , $P(B|\mathcal{G})$ is an equivalent notation for the random variable $P(\mathbf{1}_{X \in B}|\mathcal{G})$. We shall use both expressions, depending on the context.*

From the very definition of the reciprocal property, one immediately sees a nice time-simmetry, where future and past are somehow exchangeable: a probability is reciprocal if and only if the time-reversed probability is so (see Theorem 2.2 of [49]). This is also true for Markov probabilities, but maybe less transparent from the definition.

2.1.2 The relation with the Markov property

Here, we show some of the most interesting properties of reciprocal probabilities. We follow the guidelines of [49], which relies on Jamison's original presentation.

At first, let us show that the reciprocal property is indeed a weakening of the Markov property.

Proposition 2.1.1. *Any Markov probability is reciprocal.*

Proof. We check directly Definition 2.1.1. The proof consists of two chains of identities. These identities are obtained one from the other using either the Markov property (in this case we mark the equality with (M)) or some of the properties of conditional expectation, (if so, we mark the equality with (E)). First we show that for any $A \in X_{[0,s]}, C \in X_{[u,1]}$:

$$P(\mathbf{1}_A \mathbf{1}_C | X_s, X_u) = P(\mathbf{1}_A | X_s) P(\mathbf{1}_C | X_u). \quad (2.2)$$

For this purpose, let us pick any pair of measurable sets $D, F \subseteq \mathcal{X}$. We have

$$\begin{aligned} P(\mathbf{1}_A \mathbf{1}_D(X_s) \mathbf{1}_F(X_u) \mathbf{1}_C) &\stackrel{(E)}{=} P(\mathbf{1}_A \mathbf{1}_D(X_s) \mathbf{1}_F(X_u) P(\mathbf{1}_C | X_{[0,u]})) \\ &\stackrel{(M)}{=} P(\mathbf{1}_A \mathbf{1}_D(X_s) \mathbf{1}_F(X_u) P(\mathbf{1}_C | X_u)) \\ &\stackrel{(E)}{=} P(P(\mathbf{1}_A | X_{[s,1]}) \mathbf{1}_D(X_s) \mathbf{1}_F(X_u) P(\mathbf{1}_C | X_u)) \\ &\stackrel{(M)}{=} P(P(\mathbf{1}_A | X_s) \mathbf{1}_D(X_s) \mathbf{1}_F(X_u) P(\mathbf{1}_C | X_u)), \end{aligned}$$

from which (2.2) follows by the very definition of conditional expectation.

Consider now any triplet of events A, B, C such that $A \in X_{[0,s]}, B \in X_{[s,u]}, C \in X_{[u,1]}$. We have:

$$\begin{aligned} P(\mathbf{1}_A \mathbf{1}_B \mathbf{1}_C) &\stackrel{(E)}{=} P(\mathbf{1}_A \mathbf{1}_B P(\mathbf{1}_C | X_{[0,u]})) \\ &\stackrel{(M)}{=} P(\mathbf{1}_A \mathbf{1}_B P(\mathbf{1}_C | X_u)) \\ &\stackrel{(E)}{=} P(P(\mathbf{1}_A | X_{[s,1]}) \mathbf{1}_B P(\mathbf{1}_C | X_u)) \\ &\stackrel{(M)}{=} P(P(\mathbf{1}_A | X_s) \mathbf{1}_B P(\mathbf{1}_C | X_u)) \\ &\stackrel{(E)}{=} P(P(\mathbf{1}_A | X_s) P(\mathbf{1}_B | X_s, X_u) P(\mathbf{1}_C | X_u)) \\ &\stackrel{(2.2)}{=} P(P(\mathbf{1}_A \mathbf{1}_C | X_s, X_u) P(\mathbf{1}_B | X_s, X_u)) \\ &\stackrel{(E)}{=} P(\mathbf{1}_A \mathbf{1}_C P(\mathbf{1}_B | X_s, X_u)). \end{aligned}$$

Since A, C were arbitrarily chosen in $X_{[0,s]}, X_{[u,1]}$, we have shown that

$$\forall B \in X_{[s,u]}, \quad P(\mathbf{1}_B | X_{[0,s]}, X_{[u,1]}) = P(\mathbf{1}_B | X_s, X_u) \quad P - a.s.$$

This shows that P is reciprocal. \square

The reciprocal property is not equivalent to the Markov property. We construct here a simple counterexample, based on the Poisson process.

Example 2.1.1. Let $\mathcal{X} = \mathbb{N}$ and R be the Poisson process with initial distribution $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$, where δ denotes the Dirac measure. We consider $P = \frac{1}{2}R^{01}(\cdot) + \frac{1}{2}R^{12}(\cdot)$, where R^{01} is the Poisson bridge from 0 to 1 and R^{12} is the Poisson bridge from 1 to 2. It is easy to see that one has:

$$P(X_1 = 1 | X_{\frac{1}{2}} = 1) = \frac{R^{01}(X_{\frac{1}{2}} = 1)}{R^{01}(X_{\frac{1}{2}} = 1) + R^{12}(X_{\frac{1}{2}} = 1)} < 1$$

because $R^{12}(X_{\frac{1}{2}} = 1) > 0$.

However:

$$P(X_1 = 1 | X_{\frac{1}{2}} = 1, X_0 = 0) = 1$$

This shows that P is not a Markov probability. But, thanks to Proposition 2.2.2, which we will prove below, P is reciprocal. Indeed the density $\frac{dP}{dR}$ is measurable with respect to the initial and final state. One can check that:

$$\frac{dP}{dR} = \frac{1}{e} \mathbf{1}_{\{(0,1),(1,2)\}}(X_0, X_1)$$

The next result is a sufficient condition for a reciprocal probability to be Markov.

Proposition 2.1.2. Let $P \in \mathcal{P}(\Omega)$ be reciprocal. If either X_0 or X_1 is almost surely constant, then P has the Markov property.

Proof. Assume, w.l.o.g. that X_1 is a.s. constant and take any $f \in \mathcal{B}(\mathcal{X})$. Then we have that, for any $s \leq u$:

$$P(f(X_u) | X_{[0,s]}) = P(f(X_u) | X_{[0,s]}, X_1).$$

Using the reciprocal property and the hypothesis:

$$P(f(X_u) | X_{[0,s]}, X_1) = P(f(X_u) | X_s, X_1) = P(f(X_u) | X_s),$$

which gives the conclusion. □

2.2 The concept of reciprocal class

2.2.1 Probabilities with the same bridges

Given a reference Markov probability R , (which plays the role of the reference dynamics in the Schrödinger Problem), the associated reciprocal class is the set of all bridge mixtures of R . In this sense, it can be seen as

the set of "random bridges" of R . Using Proposition 1.2.1, one sees that as the constraint π varies, the set of solutions to the modified Schrödinger problem (1.2.1) forms a reciprocal class.

In the rest of the thesis, we make the assumption that the state space \mathcal{X} is countable. When \mathcal{X} is not countable, one has to make a distinction between reciprocal *family* and reciprocal *class* because the bridges of the reference process may not be everywhere well defined, but only R -almost surely (see Section 2 of [49]). But since \mathcal{X} is assumed to be countable the xy bridge $R^{xy} \in \mathcal{P}(\Omega)$ is everywhere well defined on the set support or R_{01} , and there is no need to distinguish here.

Definition 2.2.1. (*Reciprocal Class*) Let R be a Markov probability. We define the following subset of probability measures:

$$\mathfrak{R}(R) := \left\{ P = \int_{\text{supp}(R_{01})} R^{xy}(\cdot) \pi(dx dy); \pi \in \mathcal{P}(\mathcal{X}^2), \text{supp}(\pi) \subseteq \text{supp}(R_{01}) \right\}$$

as the reciprocal class of R .

The next proposition is a general recipe to construct reciprocal probabilities as mixtures of the bridges of a reciprocal reference measure. Since any Markov probability is also reciprocal, as a by product we obtain that the elements of $\mathfrak{R}(R)$ are indeed reciprocal probabilities in the sense of Definition 2.1.1.

Proposition 2.2.1. Let R be a reciprocal probability. Then, for any $\pi \in \mathcal{P}(\mathcal{X}^2)$ such that $\text{supp}\pi \subseteq \text{supp}R_{01}$ the measure P defined by:

$$P(\cdot) = \int_{\mathcal{X}^2} R^{xy}(\cdot) \pi(dx dy) \quad (2.3)$$

is a reciprocal probability. Moreover, P also satisfies

i) For all $(x, y) \in \text{supp}P_{01}$:

$$P^{xy} = R^{xy} \quad P_{01} - a.s. \quad (2.4)$$

ii) For all $s \leq u$

$$P(\cdot | X_s, X_u) = R(\cdot | X_s, X_u) \quad P - a.s. \quad (2.5)$$

Proof. We check directly Definition 2.1.1. Consider $s \leq u$ and $A \in X_{[0,s]}$, $B \in X_{[s,u]}$, $C \in X_{[u,1]}$. In the same spirit as the proof of Proposition 2.1.1, whenever an equality is obtained with an application of the reciprocal property

we mark it with (R). We have:

$$\begin{aligned}
P(\mathbf{1}_A \mathbf{1}_B \mathbf{1}_C) &= \int_{\text{supp}(\pi)} \frac{1}{R_{01}(x, y)} R(\mathbf{1}_A \mathbf{1}_B \mathbf{1}_C \mathbf{1}_{\{X_0, X_1=(x, y)\}}) \pi(dxdy) \\
&\stackrel{(R)}{=} \int_{\text{supp}(\pi)} \frac{1}{R_{01}(x, y)} R(\mathbf{1}_A R(\mathbf{1}_B | X_s, X_t) \mathbf{1}_C \mathbf{1}_{\{X_0, X_1=(x, y)\}}) \pi(dxdy) \\
&= P(\mathbf{1}_A R(\mathbf{1}_B | X_s, X_t) \mathbf{1}_C)
\end{aligned}$$

By the very definition of conditional expectation, we conclude that

$$P(\mathbf{1}_B | X_{[0, s]}, X_{[u, 1]}) = R(\mathbf{1}_B | X_s, X_u) \quad \forall B \in \mathcal{X}_{[s, u]}.$$

But then:

$$\begin{aligned}
P(\mathbf{1}_B | X_s, X_u) &= P(P(\mathbf{1}_B | X_{[0, s]}, X_{[u, 1]}) | X_s, X_u) \\
&= P(R(\mathbf{1}_B | X_s, X_u) | X_s, X_u) = R(\mathbf{1}_B | X_s, X_u)
\end{aligned}$$

from which both the fact that P is reciprocal and ii) follow.

Claim i) on the equality of the bridges follows by ii) considering $s = 0, u = 1$. \square

Using this last proposition, we can see the announced fact that solutions to the generalized Schrödinger problem are indeed reciprocal probabilities. The last proposition has the following interesting corollary:

Corollary 2.2.1. *If a solution to the entropy minimization problem (1.2.1) exists, then it is a reciprocal probability.*

Moreover, combining Proposition 2.2.1 with 2.1.2 we get:

Corollary 2.2.2. *Bridges of reciprocal probabilities are Markov probabilities.*

2.2.2 A representation result

The next result, which will be very useful later on, says that a probability belongs to the reciprocal class if and only if its density w.r.t. to the reference measure R is of a particular form. It has to be compared with Proposition 2.2.3.

Proposition 2.2.2. *Let $P \in \mathcal{P}(\Omega)$. Then $P \in \mathfrak{R}(R)$ if and only if $P \ll R$ and $\frac{dP}{dR}$ is (X_0, X_1) -measurable.*

Remark 2.2.1. *If the state space is not countable, one cannot expect that members of the reciprocal class are dominated by the reference measure, because bridges are not.*

Proof. (\Rightarrow) Let $P \in \mathfrak{R}(R)$. Then, since \mathcal{X} is countable, then for any $(x, y) \in \text{supp}R_{01}$, $R^{xy} \ll R^x$. By mixing we obtain that $P \ll R$. Let us denote its density by M . We have, for all $F \in \mathcal{B}^+(\Omega)$, using repeatedly the properties of conditional expectation and the definition of $\mathfrak{R}(R)$:

$$\begin{aligned}
R(R(M|X_0, X_1)F) &= R(R(M|X_0, X_1)R(F|X_0, X_1)) \\
&= R(MR(F|X_0, X_1)) \\
&\stackrel{\underbrace{\phantom{P \in \mathfrak{R}(R)}}}{=} R(MP(F|X_0, X_1)) \\
&= P(P(F|X_0, X_1)) \\
&= P(F) \\
&= R(MF)
\end{aligned}$$

From this it follows that $M = R(M|X_0, X_1)$, which gives the conclusion.

(\Leftarrow) Assume that $P \ll \mathfrak{R}(R)$ and the density (which again we denote by M) is (X_0, X_1) measurable. Then $M = R(M|X_0, X_1)$ $R - a.s.$. Let $F \in \mathcal{B}^+(\Omega)$, and Z be (X_0, X_1) -measurable. We have, again by the very definition of conditional expectation:

$$\begin{aligned}
P(R(F|X_0, X_1)Z) &= R(R(F|X_0, X_1)) \underbrace{(MZ)}_{(X_0, X_1)\text{-measurable}} \\
&= R(FMZ) \\
&= P(FZ) \\
&= P(P(F|X_0, X_1)Z)
\end{aligned}$$

From which it follows that $R(F|X_0, X_1) = P(F|X_0, X_1)$ $P - a.s.$, and hence the conclusion. \square

2.2.3 Markov probabilities of a reciprocal class

By proving the next proposition we also complete the proof of Proposition 1.2.3 about the Markovianity of solutions to the Schrödinger Problem. We show that transforming a reference Markov probability R with a density enjoying the multiplicative decomposition (2.6) preserves Markovianity. Such a measure transformation generalizes the Doob h-transform [30]. More precisely, it is a time symmetric version of it.

Proposition 2.2.3. *Let R be Markov and $P \in \mathcal{P}(\Omega)$. Assume that there exist $f, g : \mathcal{X} \rightarrow \mathbb{R}_+$ such that:*

$$P = f(X_0)g(X_1)R \quad R - a.s. \quad (2.6)$$

Then P is also Markov.

The proof is based on the following well known lemma.

Lemma 2.2.1. *Let $P \ll R$ and $M = \frac{dP}{dR}$. Then, for every $F \in \mathcal{B}^+(\Omega)$:*

$$P(F|X_t) = \frac{R(MF|X_t)}{R(M|X_t)} \quad P - a.s. \quad (2.7)$$

Proof. First note that no division by zero on the right hand side occurs $P - a.s.$. We have, with the basic properties of conditional expectation:

$$\begin{aligned} P(F \mathbf{1}_A(X_t)) &= R(MF \mathbf{1}_A(X_t)) \\ &= R\left(R(M|X_t) \frac{MF \mathbf{1}_A(X_t)}{R(M|X_t)}\right) \\ &= R\left(R(M|X_t) \frac{R(MF|X_t) \mathbf{1}_A(X_t)}{R(M|X_t)}\right) \\ &= R\left(M \frac{R(MF|X_t) \mathbf{1}_A(X_t)}{R(M|X_t)}\right) \\ &= P\left(\frac{R(MF|X_t)}{R(M|X_t)} \mathbf{1}_A(X_t)\right) \end{aligned}$$

The conclusion follows by the definition of conditional expectation. \square

Proof. We have to show that for any $A \in X_{[0,t]}$, any $B \in X_{[t,1]}$:

$$P(\mathbf{1}_A \mathbf{1}_B | X_t) = P(\mathbf{1}_A | X_t) P(\mathbf{1}_B | X_t). \quad (2.8)$$

Using Lemma 1 and the Markov property of R :

$$\begin{aligned} P(\mathbf{1}_A \mathbf{1}_B | X_t) &= \frac{R(f(X_0) \mathbf{1}_A \mathbf{1}_B g(X_1) | X_t)}{R(f(X_0) g(X_1) | X_t)} \\ &= \frac{R(f(X_0) \mathbf{1}_A | X_t)}{R(f(X_0) | X_t)} \frac{R(\mathbf{1}_B g(X_1) | X_t)}{R(g(X_1) | X_t)} \end{aligned} \quad (2.9)$$

Applying twice Lemma 2.2.1 and the Markov property we obtain that $P(\mathbf{1}_A | X_t) P(\mathbf{1}_B | X_t)$ coincides with the expression in (2.9). This concludes the proof. \square

We refer to [44] for more details about the infinitesimal generator associated with P defined as in (2.6). It is expressed in terms of the solution of the Kolmogorov backward PDE associated with the generator of R and the ‘‘carré du champ’’ operator.

2.3 Our framework

This section is devoted to a precise definition of our main object of study: the reciprocal class $\mathfrak{R}(R)$ of a continuous time Markov Chain R , which is called the *reference walk*. Markov chains are essentially Markov processes on countable state spaces, and are among the most studied class of processes in Probability theory. Some very general references are the books [60],[8]. In this thesis only continuous time Markov chains are considered.

Our main goals in the next chapters will be to compute the *reciprocal characteristics* associated with a reciprocal class $\mathfrak{R}(R)$, give their probabilistic interpretation, and characterize $\mathfrak{R}(R)$ by means of the characteristics.

Since the Markov chains we will consider in the next chapter are of quite different nature, we need to establish a common framework and notation to treat them: this is done in 2.3.1. Even more notation on graphs will be required in Chapter 6, see section 6.1. In the first two subsections, we specify the main assumptions on the reference walk, ensuring its existence, and define the reciprocal class associated with it. Section 2.3.3 is used to discuss a Girsanov Theorem for Markov chains. It does not contain new results, but it is a translation in our setting of known results.

2.3.1 Markov chains as walks on a graph

In this subsection, we introduce some general notation and state an existence result for the reference measure R . We will view Markov chains as random walks on graphs. There is no loss of generality in this, it is simply the language we believe to be the most appropriate to present our results, and will cover all the processes studied in this thesis.

Therefore the words Markov chain, Markov walk, and Markov walk on a graph are used as synonymous.

In absence of further specification, the term random walk is used for general probabilities, which may also be non markovian.

Probabilities enjoying the reciprocal property are called reciprocal walks.

Our state space is a countable set \mathcal{X} of vertices. \mathcal{X} is equipped with the discrete topology, and the limits appearing in the next definitions are to be understood with respect to this topology. Any subset $\mathcal{A} \subseteq \mathcal{X}^2$ defines a *directed graph* on \mathcal{X} through the relation \rightarrow , which is defined for all $z, z' \in \mathcal{X}$ by

$$z \rightarrow z' \text{ if and only if } z, z' \in \mathcal{A}.$$

We denote by $(\mathcal{X}, \rightarrow)$ this directed graph. We call a pair z, z' such that $(z \rightarrow z')$ an *arc* of the graph $(\mathcal{X}, \rightarrow)$.

Cycles play a crucial role in the study of reciprocal classes. Let us give some definitions.

Definition 2.3.1 (path and cycles). *Let $\mathcal{A} \subset \mathcal{X}^2$ specify a directed graph $(\mathcal{X}, \rightarrow)$ on \mathcal{X} .*

- i) *For any $n \geq 1$ and $x_0, \dots, x_n \in \mathcal{X}$ such that $x_0 \rightarrow x_1, \dots, x_{n-1} \rightarrow x_n$, the ordered sequence of vertices $\mathbf{w} := (x_0, x_1, \dots, x_n)$ is called an \mathcal{A} -path, or shortly a walk. We adopt the more appealing notation $(x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n)$. The length n of \mathbf{w} is denoted by $|\mathbf{w}|$.*
- ii) *When $x_n = x_0$, the walk $(x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n = x_0)$ is a **cycle**.*
- iii) *A cycle $(x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_n = x_0)$ is said to be simple if the cardinal of the visited vertices $\{x_0, x_1, \dots, x_{n-1}\}$ is equal to the length n of the cycle. This means that a simple cycle cannot be further decomposed in cycles.*

Remark that in our definition, cycles come with an orientation: the cycles $(x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_{n-1} \rightarrow x_n = x_0)$ is different from the cycle $(x_n = x_0 \rightarrow x_{n-1} \rightarrow \dots \rightarrow x_1 \rightarrow x_0)$. Moreover, \mathcal{A} -path are not trajectories: they are simply path on the graph $(\mathcal{X}, \rightarrow)$.

For a given $\mathcal{A} \subseteq \mathcal{X}^2$, we will consider random walks on \mathcal{X} where only transition on the arcs of $(\mathcal{X}, \rightarrow)$ are allowed.

The left limit at t of a function $\omega \in \mathcal{X}^{[0,1]}$ is denoted by ω_{t-} , and the right limit by ω_{t+} .

The path space $\Omega \subseteq \mathcal{X}^{[0,1]}$ which describes the trajectories of the processes is the set of all càdlàg piecewise constant paths $\omega = (\omega_t)_{t \in [0,1]}$ on \mathcal{X} with *finitely many jumps* such that there are no jumps at time one, and transitions between vertices can happen only along the arcs in \mathcal{A} . Summarizing:

$$\Omega := \left\{ \omega : [0, 1] \rightarrow \mathcal{X}, \#\{t : \omega_t \neq \omega_{t-}\} < +\infty, \omega_{t+} = \omega_t \forall t \in [0, 1], \right. \\ \left. \omega_{t-} \neq \omega_t \Rightarrow (\omega_{t-} \rightarrow \omega_t), \omega_{1-} = \omega_1 \right\}$$

Ω is equipped with the canonical σ algebra generated by the canonical process $X = (X_t)_{t \in [0,1]}$. defined for each $t \in [0, 1]$ and $\omega \in \Omega$ by $X_t(\omega) = \omega_t$.

We define recursively for any $n \in \mathbb{N}$, the stopping time T_n , which is the time when the n th jump happened.

$$T_1 = \inf\{t \in [0, 1] \text{ s.t. } X_t \neq X_{t-}\}, \quad T_n := \inf\{t > T_{n-1} \text{ s.t. } X_t \neq X_{t-}\} \quad (2.10)$$

with the convention that $\inf \emptyset = +\infty$. We adopt a measure theoretical viewpoint. That is, we identify random processes with paths in Ω and their laws on Ω . We call any $P \in \mathcal{P}(\Omega)$ a random walk, or simply a walk, regardless if it is Markov or not.

2.3.2 The reference Markov walk and its reciprocal class

Prior to our choice of a reference walk, we fix a set $\mathcal{A} \subseteq \mathcal{X}^2$ and consider the directed graph $(\mathcal{X}, \rightarrow)$.

Our reference walk R is always Markov. It is specified through an *intensity of jump* along the arcs $j : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$. We always reserve R for the reference walk, and j for its intensity. No other probability or intensity will be labeled in the same way.

When j is not time dependent the dynamics of R has the following simple description: if the walker sits in z , it waits for a random time which is exponentially distributed with parameter $\sum_{z': z \rightarrow z'} j(z \rightarrow z')$. Then it chooses a neighbor z' of z with probability proportional to $j(z \rightarrow z')$ and jumps there. All these events are mutually independent

Assumption 2.3.1 ((graph and reference intensity)). $(\mathcal{X}, \rightarrow)$ and j satisfy the following assumptions:

i) $(\mathcal{X}, \rightarrow)$ has bounded degree:

$$\exists C < +\infty, \#\{z' \in \mathcal{X} : z \rightarrow z'\} \leq C \quad \forall z \in \mathcal{X}$$

.

ii) $(\mathcal{X}, \rightarrow)$ has no cycles of length one. That is, for all $z \in \mathcal{X}$, $(z, z) \notin \mathcal{A}$

iii) The intensity j is uniformly bounded from above:

$$\sup_{t \in [0, 1], z \in \mathcal{X}} \bar{j}(t, z) < +\infty \quad (2.11)$$

where \bar{j} is the total intensity of jump in z ,

$$\bar{j}(t, z) := \sum_{z' \in \mathcal{X} : z \rightarrow z'} j(t, z \rightarrow z').$$

iv) There is some subset $\mathcal{A}_{\rightarrow}(j) \subseteq \mathcal{A}$ such that:

$$(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(j) \Rightarrow j(t, z \rightarrow z') > 0 \quad \forall t \in (0, 1),$$

and:

$$(z \rightarrow z') \notin \mathcal{A}_\rightarrow(j) \Rightarrow j(t, z \rightarrow z') = 0 \quad \forall t \in [0, 1].$$

We call $\mathcal{A}_\rightarrow(j)$ the active arcs of j . Furthermore, we assume that j has a uniform positive lower bound on $[0, 1] \times \mathcal{A}_\rightarrow(j)$.

v) The intensity j is continuously t -differentiable, i.e. for any $z \rightarrow z' \in \mathcal{A}_\rightarrow(j)$, $t \mapsto j(t, z \rightarrow z')$ is continuously differentiable.

Point iv) of Assumption 2.3.1 simply means that if an arc $(z \rightarrow z')$ is a possible choice for the walker at some time $t \in [0, 1]$, then it can also be chosen at any other time, provided that the walker sits in z . It ensures that the support of R_t does not change with time.

Associated to any intensity $k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ and $t > 0$ there is a formal generator \mathcal{K}_t , which acts on functions $u : \mathcal{X} \rightarrow \mathbb{R}$ of finite support as follows:

$$\mathcal{K}_t u(z) = \sum_{z' : z \rightarrow z'} j(t, z \rightarrow z')(u(z') - u(z)) \quad (2.12)$$

The generator associated with the reference intensity is denoted \mathcal{G}_t .

In the next definition the intensity k does not necessarily satisfy iii) and iv) of Assumption 2.3.1.

Definition 2.3.2. We say that a law $P \in \mathcal{P}(\Omega)$ is a Markov walk of intensity $k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ if for all $u : \mathcal{X} \rightarrow \mathbb{R}$ with finite support

$$u(X_t) - \int_0^t \mathcal{K}_s u(X_s) ds \quad (2.13)$$

is a local P -martingale. $(\mathcal{K}_s)_{s \in [0, 1]}$ is the **generator** of P

Adapting the much more general Theorem 3.6 in [36], (or Theorem 6.7 of [53],) it follows that for any j satisfying Assumption 2.3.1, and $x \in \mathcal{X}$, there exists a unique Markov walk R^x of intensity j and initial distribution δ_x .

Clearly Assumption 2.3.1 can be strongly relaxed to ensure the existence of the process. However, it will turn out to be very convenient in view of the results of the next chapters.

In all what follows a graph $(\mathcal{X}, \rightarrow)$ is given. On it, an intensity j satisfying Assumption 2.3.1 is defined, and we consider a Markov walk R of intensity j with initial measure of full support. They are the data of the problem, which is to study the reciprocal class $\mathfrak{R}(R)$:

$$\mathfrak{R}(R) := \left\{ P = \int_{\text{supp}(R_{01})} R^{xy}(\cdot) \pi(dx dy); \pi \in \mathcal{P}(\mathcal{X}^2), \text{supp}(\pi) \subseteq \text{supp}(R_{01}) \right\}$$

Since \mathcal{X} is countable, the bridge R^{xy} is always well defined for $x, y \in \text{supp}R_{01}$. The reciprocal class is well defined too.

2.3.3 Girsanov Theorem for random walks on a graph

The following Girsanov Theorem is a translation of the abstract results of [36], which are written for *multivariate Point processes*. We are dealing with random walks on graphs. But there is a natural way to see a walk as a multivariate point process, by associating to each path the sequence $(T_n, A_n)(\omega)$ where T_n is the n th jump time and A_n is the arc along which the walk jumps at T_n . Conversely, a random walk is naturally associated to a multivariate point processes, by inverting the above construction.

Girsanov Theorem is a standard result for SDEs driven by the Brownian motion, but it is less studied for jump processes. Very general statements are in [36],[37] but it is not straightforward to specify them to our situation, which is much less general.

An exposition of Girsanov theory for jump processes in \mathbb{R}^d under a finite entropy condition can be found in [47], and the Lévy processes case is treated in [72]. But for countable state spaces without a vector space structure we found no reference, and therefore it is worth to spend some words on this, since we will use these results at different times in this thesis.

Theorem 2.3.1 ((Girsanov Theorem)). *Let $k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ be such that*

$$\{(t, z \rightarrow z') \text{ s.t. } k(t, z \rightarrow z') > 0\} \subseteq [0, 1] \times \mathcal{A}_{\rightarrow}(j). \quad (2.14)$$

We define:

$$\tau := \inf \left\{ t \in [0, 1] : k(t, X_{t-} \rightarrow X_t) = 0 \text{ or } \int_0^t \bar{k}(s, X_{s-}) ds = +\infty \right\}$$

with the convention $\inf \emptyset = +\infty$. We define the random variable:

$$Z_\infty := \mathbf{1}_{\{\tau=+\infty\}} \exp \left(- \int_0^1 (\bar{k} - \bar{j})(s, X_{s-}) ds \right) \prod_{i:T_i < 1} \frac{k}{j}(T_i, X_{T_i-} \rightarrow X_{T_i}) \quad (2.15)$$

If $R(Z_\infty) = 1$, then the measure $P \in \mathcal{P}(\Omega)$ defined by

$$P = Z_\infty R \quad (2.16)$$

is the Markov random walk on $(\mathcal{X}, \rightarrow)$ with intensity k and initial distribution R_0 .

We have the following important corollary:

Corollary 2.3.1. *Let k be a jump intensity satisfying Assumption 2.3.1 and $\mathcal{A}_{\rightarrow}(k) \subseteq \mathcal{A}_{\rightarrow}(j)$. If P is a Markov walk with intensity k such that $P_0 = R_0$, then $P \ll R$ and $\frac{dP}{dR}$ is given by (2.15).*

$$Z_{\infty} := \mathbf{1}_{\{\tau=+\infty\}} \exp\left(-\int_0^1 (\bar{k}-\bar{j})(s, X_{s-}) ds\right) \prod_{i:T_i < 1} \frac{k}{j}(T_i, X_{T_i^-} \rightarrow X_{T_i}) \quad (2.17)$$

Let us first prove Theorem 2.3.1.

Proof. Let the sequence of jump times $\{T_n\}_{n \geq 1}$ be defined by (2.10). We also define the sequence of the arcs along which the walk jumps at T_n :

$$A_n : \Omega \rightarrow \mathcal{A}, \quad A_n = X_{T_n^-} \rightarrow X_{T_n} \in \mathcal{A}$$

Then the sequence (T_n, A_n) is a *multivariate point process* in the sense of paragraph 3, page 238 of [36]. We define the random measure $\mu \in \mathcal{M}^+([0, 1] \times \mathcal{A})$ as follows:

$$\mu(dt da) = \sum_{n \geq 1} \delta_{(T_n, A_n)} \quad (2.18)$$

where $\delta_{t,a}$ stands for the Dirac measure on $[0, 1] \times \mathcal{A}$. Using Proposition 3.1 of [36] we obtain, as a general fact, that $Q \in P(\Omega)$ is a Markov random walk with intensity $k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ if and only if a version of the *predictable projection* of μ in the sense of Theorem 2.1 in [36] is the following (predictable) random measure $\mu'(dt da) \in \mathcal{M}^+([0, 1] \times \mathcal{A})$:

$$\mu'(dt da) = \sum_{z': X_{t-} \rightarrow z'} k(t, X_{t-} \rightarrow z') \delta_{(X_{t-} \rightarrow z')} (da) dt \quad (2.19)$$

Now, let us consider our reference intensity j , and another intensity $k : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ satisfying (2.14). We define $Y : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ by:

$$Y(t, z \rightarrow z') = \frac{k}{j}(t, z \rightarrow z') \quad (2.20)$$

Because of (2.14), Y is well defined: no division by zero occurs.

Then we can see that the random variable Z_{∞} defined at (2.15) coincides with the random variable obtained by setting $S = \infty$ at equation (14), p.247 of [36], when Y is given by (2.20).

Since $R(Z_{\infty}) = 1$, Theorem 4.5 of the same article applies: under the measure P defined by (2.16) the random measure μ admits as predictable projection the measure

$$\begin{aligned}
\mu'(dt da) &= \sum_{z': X_{t-} \rightarrow z'} Y(t, X_{t-} \rightarrow z') j(t, X_{t-} \rightarrow z') \delta_{(X_{t-} \rightarrow z')}(da) dt \\
&= \sum_{z': X_{t-} \rightarrow z'} k(t, X_{t-} \rightarrow z') \delta_{(X_{t-} \rightarrow z')}(da) dt
\end{aligned}$$

Therefore, from the discussion above, P is a Markov random walk of intensity k . \square

We can get back to the proof of the corollary:

Proof. All what we need to show is that if Z_∞ is given by (2.15), then $R(Z_\infty) = 1$. The conclusion then follows by Theorem 2.3.1. Using Proposition 4.3 in [36] we have that $(Z_t)_{t \in [0,1]}$ defined by

$$Z_t := \mathbf{1}_{\{\tau \geq t\}} \exp\left(-\int_0^t (\bar{k} - \bar{j})(s, X_{s-}) ds\right) \prod_{i: T_i < t} \frac{k}{j}(T_i, X_{T_i-} \rightarrow X_{T_i})$$

is a local R-martingale. Therefore it is sufficient to prove that $(Z_t)_{t \in [0,1]}$ is a true martingale to prove the corollary. A known criterion for this to happen is that:

$$R\left(\sup_{t \in [0,1]} Z_t\right) < +\infty$$

But if both j and k satisfy Assumption 3, and $\mathcal{A}_\rightarrow(k) \subseteq \mathcal{A}_\rightarrow(j)$, then it is easy to see that there exist constants a, b such that

$$\sup_{t \in [0,1]} Z_t \leq a \exp(bN_1) \quad R - a.s.$$

where

$$N_1 = \#\{t \in [0, 1] : X_t \neq X_{t-}\}$$

is the total number of jumps up to time 1. The assumptions that j is upper bounded and the graph is of bounded degree imply that N_1 has all exponential moments. This completes the proof. \square

Chapter 3

Counting processes

Outline of the chapter This chapter is an exposition of the results contained in the published paper [22], where a duality formula between a derivative operator, based on time-changes, and a stochastic integral operator is shown to characterize the reciprocal class of a counting process. The duality formula is expressed in terms of the reciprocal characteristic. Effects of lower bounds on the reciprocal characteristics are studied as well.

Counting processes are jump processes on \mathbb{Z} which can only make positive jumps of size one and are used to model a stream of random events. They gain their interest even outside Markovian framework, as it is the case for renewal processes. The range of application is very broad, covering survival analysis [33], statistics [1], control and engineering [7] and, more recently, environmental sciences [65]. In our work we will profit from the martingale approach to study their dynamics. This was first put explored by Watanabe in [80], who proved a characterization of the Poisson process as the unique solution of a Martingale problem, in the spirit of Lévy characterization of the Brownian motion. Brémaud extended in [6] this approach to general counting processes. He proved that a counting process is characterized by a predictable process called the compensator. The compensator is defined as the unique process such that, when subtracted to the canonical process, turns it into a martingale. Jacod [36] extended this notion to a much larger class of jump processes.

Bridge mixtures of Markov counting processes provide the simplest examples of reciprocal jump processes. As we shall see, this is because

the underlying graph is *acyclic* and has a translation invariance property. Moreover, it is natural to start a study of reciprocal processes over discrete structures by considering the Poisson process, which is in many respects the discrete-space analogous of the Brownian motion. The *reciprocal characteristic* of a counting process is computed explicitly: it is a space time function associated with the intensity of the reference process which describes the dynamics of a reciprocal class. This is made clear in our main result: an integration by parts (duality) formula between a derivative operator and a stochastic integral which characterizes reciprocal class. This is Theorem 3.2.2. Integration by parts formulae are at the core of the characterization of the reciprocal class of a Brownian diffusion, see [67],[68], and our result has to be interpreted in the same spirit. Quite remarkably, there already exist more than one type of integration by parts formula for the Poisson process. They reflect two different approaches to stochastic calculus for jump process, the one based on perturbation of jump times(differential calculus), and the one based on addition of jumps to the trajectories(difference calculus). Only the first approach will turn out to be useful in the study of reciprocal probabilities. This is in contrast with the diffusion case, where essentially there exist only one duality formula which stands at the basis of Malliavin calculus.

In section 3.3 we shall be concerned with lower bounds on the reciprocal characteristic. A lower bound is shown to imply an estimate on the last jump time of the bridges of the reference walk, see Proposition 3.3.1. In particular, a positive bound implies that the bridge of the reference walk is *slower* than the bridge of a Poisson process, in the sense that it tends to reach its final state later than the Poisson bridge, and we have an accumulation of the jump times around time one.

Studying the reciprocal class of a counting process has to be seen as a preliminary step of a larger program. Indeed the simple graph structure hides many difficulties, and, most importantly, does not allow to see the role of cycles. This will emerge clearly from Chapter 4 on. However, the reciprocal characteristic we compute here, serves as a prototype for the *arc characteristic* in a general graph, which we compute in Chapter 6.

Organization of the chapter In section 3.1 we recall some basic facts about counting process. Then we define all the objects which appear in the duality formula and prove, as a new result, a duality formula for the law of a nice Markov counting process at Proposition 3.1.5. Section 3.2 is devoted to the study of reciprocal class of a counting process. It contains two new results: In Theorem 3.2.1 we compute the reciprocal characteris-

tic associated with a counting process and in Theorem 3.2.2 we show that the duality formula proved in the former section is expressed in terms of the reciprocal characteristic and indeed characterizes the reciprocal class. Section 3.3 is devoted to study the quantitative effects on the dynamics of a bridge of lower bounds for the reciprocal characteristic.

The first section of this chapter is based on Murr's Phd thesis [56]. In particular Proposition 3.2.1 is a rewriting of its Theorem 6.58. From section 3.2.2 on, the results contained are original. The following chart helps in connecting the notation of this chapter with the general framework we set up in Chapter 2.

	General def	Local def
State space	\mathcal{X}	\mathbb{Z}
Arcs	\mathcal{A}	$(z \rightarrow z + 1), z \in \mathbb{Z}$
Active arcs	$\mathcal{A}_{\rightarrow}(j)$	$(z \rightarrow z + 1), z \in \mathbb{Z}$
Reference intensity	$j(t, z \rightarrow z + 1)$	$j(t, z)$

3.1 Duality formula for a counting process

3.1.1 Nice counting processes

Using the notation and language of Chapter 2, a counting process is a continuous time markov chain on the graph specified $\mathcal{X} = \mathbb{Z}$ and $\mathcal{A} = \{(z \rightarrow z + 1), z \in \mathbb{Z}\}$.

Then Ω consists of all cadl g step functions, with finitely many jumps of amplitude +1 and any initial value in \mathbb{Z} .

Any path is described by the collection (x, t_1, \dots, t_n) of its initial position $x \in \mathbb{Z}$ and its $n = X_1 - X_0$ instants of jumps $0 < t_1 < \dots < t_n < 1$. It is practical to set $t_i = 1$ for $i > n$. We denote $T_i(\omega) = t_i$ the i -th instant of jump.

Any $Q \in \mathcal{P}(\Omega)$ admits an increasing predictable process $A : [0, 1] \times \Omega \rightarrow \mathbb{R}_{\geq 0}$ such that $Q(A(0) = 0) = 1$ and

$$X_t - X_0 - A(t)$$

is a local Q -martingale. A characterizes the dynamics of Q (see, for instance Theorem 2.1 in Jacod [36] for a general result). A is called the *compensator* and its time derivative j (if it exists) is the intensity of jump of the process.

We translate our Assumption 2.3.1 in the present framework: since the structure of the graph is very simple, there is no ambiguity in identifying the arc $(z \rightarrow z + 1)$ with z . Our intensity is then a function $j : [0, 1] \times \mathbb{Z} \rightarrow \mathbb{R}_+$. Assumption 2.3.1 is seen to be equal to the following:

Assumption 3.1.1. *The reference process R and its jump intensity j satisfy:*

1. $j : [0, 1] \times \mathbb{Z} \rightarrow \mathbb{R}_+$ is strictly positive and upper bounded, and for all $z \in \mathbb{Z}$ the function $t \mapsto j(t, z)$ is of class C^1
2. R_0 has full support.

Definition 3.1.1 (Nice Markov counting processes). *i) Any $P \in \mathcal{P}(\Omega)$ which is Markov and whose jump intensity exists and satisfies Hypothesis (3.1.1) is called a **nice Markov counting process**, NMC for short.*

- ii) We call a **Poisson process** any NMC that satisfies $j(t, z) \equiv 1$, no matter what it is the initial distribution. Similarly a Poisson process of parameter λ is a NMC with $j(t, z) \equiv \lambda$.*

Poisson processes play a distinguished role, and, when they do not coincide with the reference measure R , they are denoted by \mathbf{P} .

The Girsanov formula reads as follows:

Proposition 3.1.1. *Let $P \in \mathcal{P}(\Omega)$ be any nice counting process of intensity k such that $P_0 = R_0$. Then $P \ll R$ and*

$$\frac{dP}{dR} = \exp \left(- \int_0^1 (k(s, X_{s-}) - j(s, X_{s-})) ds \right) \prod_{i: T_i < 1} \frac{k(T_i, X_{T_i-})}{j(T_i, X_{T_i-})}$$

In particular, if \mathbf{P} is a Poisson process and R a NMC with of intensity j , and $\mathbf{P}_0 = R_0$, then the density $\frac{dR}{d\mathbf{P}}$ is denoted by G_j . We have :

$$G_j := \exp \left(\int_0^1 1 - j(s, X_{s-}) ds \right) \prod_{i: T_i < 1} j(T_i, X_{T_i-}) \quad (3.1)$$

3.1.2 The derivative and divergence operators

We present an approach to Malliavin calculus on Ω first developed by Carlen and Pardoux in [10]. It is the core of our characterization of $\mathfrak{A}(R)$. Despite other approaches (see e.g. Chapter 7,8 of [63]) where the derivative operator is actually a difference operator, this one has the advantage that the operator introduced is a true derivative, in the sense that it enjoys a product rule and a chain rule. These two properties will turn out to be crucial in the characterization of the reciprocal class.

As for the classical Malliavin calculus we establish an integration by parts formula which puts in duality a derivative operator with a so-called

divergence operator. In the case of the Wiener measure, (Lemma 1.2 of [61]) it is the Malliavin derivative that is in duality with the Skorhorod integral.

In that case, the derivative operator can be interpreted as a directional derivative, where the paths are perturbed by a small shift.

Here, we consider a directional derivative on Ω by considering infinitesimal changes of the time parametrization (see also Elliott & Tsoi [31]), and establish a duality with a stochastic integral operator.

We use the following scheme: at first we introduce a time-perturbation, and then use it to define a derivative operator acting on a suitable space of test functions. Eventually, we define our divergence operator and establish the duality formula.

Definition 3.1.2 (The set \mathcal{U} of perturbation functions). *The set \mathcal{U} of perturbation functions consists of all C^1 -functions $u : [0, 1] \rightarrow \mathbb{R}$ such that $u(0) = u(1) = 0$.*

For any function $u \in \mathcal{U}$ and $\varepsilon > 0$ small enough, we define the change of time $\theta_u^\varepsilon : [0, 1] \rightarrow [0, 1]$ by

$$\theta_u^\varepsilon(t) = t + \varepsilon u(t).$$

The boundedness of the derivative $\partial_t u$ of u and the property $u(0) = u(1) = 0$ ensure that for any ε small enough, θ_u^ε is indeed a change of time with $\theta_u^\varepsilon(0) = 0$ and $\theta_u^\varepsilon(1) = 1$.

The perturbation is defined for any path $\omega \in \Omega$ by:

$$\Theta_u^\varepsilon : \Omega \rightarrow \Omega, \quad X(\Theta_u^\varepsilon(\omega))_t = X(\omega)_{\theta_u^\varepsilon(t)} \quad \forall t \in [0, 1] \quad (3.2)$$

Note that the operator Θ_u^ε keeps the initial and final values of the path unchanged. It modifies the jump times according to the following rule:

$$T_i \circ \Theta_u^\varepsilon + \varepsilon u(T_i \circ \Theta_u^\varepsilon) = T_i \quad \forall i \in \mathbb{N} \quad (3.3)$$

We define a derivative in the direction of the elements of \mathcal{U} ,

Definition 3.1.3 (The derivative $\mathcal{D}_u \Phi$). *Let Φ be a measurable real function on Ω and $u \in \mathcal{U}$ a perturbation function. We define*

$$\mathcal{D}_u \Phi := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (\Phi \circ \Theta_u^\varepsilon - \Phi), \quad (3.4)$$

provided that this limit exists

We do not ask the limit to exist in L^2 but pointwise. Indeed, we have not specified a reference measure at the moment. This is unusual for Malliavin calculus, but it turns out to be useful here. As a by-product of Definition (3.1.3) we deduce that \mathcal{D}_u satisfies both the chain rule and the product rule. We do not give the proof, because it uses the same arguments used to prove the standard rules of calculus for real functions.

Proposition 3.1.2. *For any differentiable pair of functions $\Phi, \Psi, u \in \mathcal{U}$ we have:*

$$\mathcal{D}_u(\Phi\Psi) = \Phi \mathcal{D}_u\Psi + \Psi\mathcal{D}_u\Phi \quad (3.5)$$

If $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ is C^∞ function and Φ is differentiable then:

$$\mathcal{D}_u(\alpha(\Phi)) = \alpha'(\Phi)\mathcal{D}_u\Phi \quad (3.6)$$

Let us remark that we slightly changed the notations introduced by Carlen and Pardoux in [10]. We write \mathcal{D}_u instead of \mathcal{D}_u . In the next proposition we show that for a large class of functions, the derivative exists and can be computed. They are smooth functionals of the jump times. On the contrary, the derivative of a function $\Phi(X_t)$ of the position of the process is zero except for those paths which have a jump exactly at t . Therefore this class of functionals is not meaningful for the derivative operator we have just defined.

Definition 3.1.4 (The set \mathcal{S} of smooth functions). *We say that $\Phi : \Omega \rightarrow \mathbb{R}$ belongs to the set \mathcal{S} of smooth functions if there exists $m \geq 1$ such that $\Phi = \varphi(X_0; T_1, \dots, T_m)$ for some bounded $\varphi : \mathbb{Z} \times [0, 1]^m \rightarrow \mathbb{R}$ such that for all $x \in \mathbb{Z}$, the partial functions $\varphi(x; \cdot)$ are C^1 -differentiable on the simplex $(0 < t_1 < \dots < t_m < 1)$, and have a C^1 extension to the closure. Finally, the partial derivatives are bounded as well:*

$$\sup_{i \leq m, x \in \mathbb{Z}} \partial_{t_i} \varphi(x, t_1, \dots, t_m) < +\infty.$$

These functions are differentiable on the Poisson space in a natural way, see also Thm. 1.3 in [10].

Proposition 3.1.3. *Let $\Phi \in \mathcal{S}$ be a simple function. It is differentiable in the direction of any $u \in \mathcal{U}$ and one has:*

$$\begin{aligned} \mathcal{D}_u\Phi &= \mathcal{D}_u\varphi(X_0; T_1, \dots, T_m) = - \sum_{j=1}^m \partial_{t_j} \varphi(X_0; T_1, \dots, T_m) u(T_j) \\ &= - \int_0^1 \left(\sum_{j=1}^m \partial_{t_j} \varphi(X_0; T_1, \dots, T_m) \mathbf{1}_{[0, T_j]}(t) \right) \dot{u}(t) dt. \end{aligned} \quad (3.7)$$

Proof. Let $C = \sup_{t \in [0,1]} \partial_t u(t) < +\infty$. From (3.3) we have that, for all i :

$$\begin{aligned} |T_i \circ \Theta_u^\varepsilon - T_i + \varepsilon u(T_i)| &= \varepsilon |u(T_i) - u(T_i \circ \Theta_u^\varepsilon)| \\ &\leq C\varepsilon |T_i - T_i \circ \Theta_u^\varepsilon| \\ &\leq C\varepsilon^2 |u(T_i \circ \Theta_u^\varepsilon)| \\ &\leq C^2 \varepsilon^2 \end{aligned} \quad (3.8)$$

To finish the proof, we just need to apply the definition of \mathcal{D}_u , the regularity of φ , and the standard rules of calculus. \square

The last ingredient of the duality formula is the compensated stochastic integral, playing the role of the divergence operator.

Definition 3.1.5. Let $u \in \mathcal{U}$. We define $\delta(u) : \Omega \rightarrow \mathbb{R}$ as:

$$\delta(u) := \int_0^1 \dot{u}(t) d(X_t - t)$$

As the space Ω is very simple, the stochastic integral is clearly well defined for any $\omega \in \Omega$ as:

$$\int_0^1 \dot{u}(t) dX_t := \sum_{i: T_i < 1} \dot{u}(T_i^-) \quad (3.9)$$

Moreover, let us note that since $u(1) = u(0) = 0$, $\delta(u)$ coincides with (3.9).

3.1.3 Duality formula

As a first step, we prove a duality under the Poisson process.

Proposition 3.1.4. Let R be a Poisson process, and $u \in \mathcal{U}$. Then for all $\Phi \in \mathcal{S}$:

$$R(\mathcal{D}_u \Phi) = R(\Phi \delta(u)) \quad (3.10)$$

Proof. Let $\Phi \in \mathcal{S}$. Then, by definition:

$$R(\mathcal{D}_u \Phi) = R\left(\lim_{\varepsilon \rightarrow 0} \frac{\Phi \circ \Theta_u^\varepsilon - \Phi}{\varepsilon}\right)$$

Because of the smoothness of Φ and the boundedness of $\mathcal{D}_u \Phi$, (see (3.7)) we can exchange limit and expectation, so that:

$$R(\mathcal{D}_u \Phi) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} R(\Phi \circ \Theta_u^\varepsilon - \Phi) \quad (3.11)$$

We prove that the measure $R^\varepsilon := R(\Theta_u^\varepsilon \in \cdot)$ is a nice counting process with intensity

$$k^\varepsilon(t, z) := 1 + \varepsilon \dot{u}(t)$$

This follows from the fact that clearly R^ε is Markov, and

$$R^\varepsilon(X_t - X_s | X_s) = R(X_{t+\varepsilon u(t)} - X_{s+\varepsilon u(s)} | X_{s+\varepsilon u(s)}) = t - s + \varepsilon \int_s^t \dot{u}_l dr$$

Therefore $X_t - \int_0^t (1 + \varepsilon \dot{u}(s)) ds$ is a R^ε martingale. We have thus identified the compensator of R^ε . It follows that R^ε is a nice counting process with the desired intensity. An application of Girsanov Theorem at Proposition 3.1.1, combined with the assumption $u(1) = u(0) = 0$ tells that:

$$\frac{dR^\varepsilon}{dR} = \prod_{i: T_i < 1} (1 + \varepsilon \dot{u}(T_i^-))$$

and therefore, by (3.11):

$$R(\Phi \circ \Theta_u^\varepsilon - \Phi) = R\left(\Phi\left(\prod_{i: T_i < 1} (1 + \varepsilon \dot{u}(T_i^-) - 1)\right)\right)$$

A simple second order expansion in ε , together with the boundedness of \dot{u} yields the conclusion. \square

Having in mind the program to carry out a Malliavin calculus to study the laws of Poisson driven equations, the validity of that formula extends to a much larger class of test functions and directions of differentiation. For our purposes Proposition 3.1.4 is general enough to provide a characterization of the reciprocal class, and we will not extend this duality. However, let us just mention that \mathcal{D}_u is an unbounded closable densely defined operator from $L^2(\Omega)$ into $L^2(\Omega)$ (see Theorem 1.5 of [10]). This gives the opportunity to define the derivative of functionals which are not in \mathcal{S} , but can be approximated in a suitable sense. As for the Malliavin calculus on the Wiener space, this is crucial to study properties of the marginals of stochastic differential equations. We rather move to our next goal: prove a duality formula for the law of a NMC. Here, we exploit the fact that \mathcal{D}_u is a true derivative:

Proposition 3.1.5 below states a duality relation between the derivative operator \mathcal{D} and some stochastic integral. R is no longer assumed to be a Poisson process.

Proposition 3.1.5. *Let R be a nice counting process of intensity j . Then it satisfies the following duality formula. For all $\Phi \in \mathcal{S}$ and $u \in \mathcal{U}$:*

$$R(\mathcal{D}_u \Phi) = R\left(\Phi \int_0^1 [\dot{u}(t) + (\partial_t \log j(t, X_{t-}) + j(t, X_{t-} + 1) - j(t, X_{t-}))u(t)] dX_t\right) \quad (3.12)$$

We do not make the initial distribution μ precise since it does not play any role.

Proof. Let \mathbf{P} be the Poisson process with $\mathbf{P}_0 = R_0$. Recalling Girsanov's Theorem we see that $R \ll \mathbf{P}$. We call G_j the density $\frac{dR}{d\mathbf{P}}$ whose expression is given at (3.1). The first part of the proof is to prove the differentiability of the density and to compute the directional derivative of $\log G_j$. To this aim, we define the event $\Omega_m := \{X_1 - X_0 = m\}$. We observe that G_j coincides on the event Ω_m with G_j^m , defined as follows:

$$G_j^m := \exp\left(-\sum_{i=1}^{m+1} \int_{T_{i-1}}^{T_i} (j(s, X_0 + (i-1)) - 1) ds\right) \prod_{i=1}^m j(T_i, X_0 + i) \quad (3.13)$$

where we adopted the convention $T_{m+1} = 1$. and therefore we can reduce ourselves to compute each of the $\mathcal{D}_u \log G_j^m$. The directional derivative of $\log G_j$ is then given by:

$$\mathcal{D}_u \log G_j = \sum_{m=1}^{+\infty} \mathcal{D}_u \log G_j^m \mathbf{1}_{\Omega_m}$$

But for any m , $G_j^m \in \mathcal{S}$, thanks to Assumption 3.1.1. In Proposition 3.1.3 we computed the derivative of a smooth functional. Using that result we obtain:

$$\begin{aligned} \mathcal{D}_u \log G_j^m &= j(T_1, X_0)u(T_1) + \sum_{i=2}^m [j(T_i, X_{T_i^-})u(T_i) - j(T_{i-1}, X_{T_i^-})u(T_{i-1})] \\ &\quad - j(T_m, X_1)u(T_m) - \sum_{i=1}^m \partial_t \log j(T_i, X_{T_i})u(T_i) \\ &= -\sum_{i=1}^m u(T_i) [\partial_t \log j(T_i, X_{T_i^-}) + j(T_i, X_{T_i}) - j(T_i, X_{T_i^-})] \\ &= -\int_0^1 [\partial_t \log j(t, X_{t-}) + j(t, X_{t-} + 1) - j(t, X_{t-})] u(t) dX_t \end{aligned}$$

This proves that $\log G_j$ is differentiable with derivative:

$$\mathcal{D}_u \log G_j = - \int_0^1 [\partial_t \log j(t, X_{t-}) + j(t, X_{t-} + 1) - j(t, X_{t-})] u(t) dX_t. \quad (3.14)$$

Applying Girsanov's theorem, the product rule and the duality formula under the Poisson process we obtain:

$$\begin{aligned} R(\mathcal{D}_u \Phi) &= \mathbf{P}(G_j \mathcal{D}_u \Phi) \\ &= \mathbf{P}(\mathcal{D}_u(G_j \Phi)) - \mathbf{P}(\Phi \mathcal{D}_u G_j) \\ &= \mathbf{P}(G_j \Phi \delta(u)) - \mathbf{P}(\Phi \mathcal{D}_u G_j) \\ &= R(\Phi \delta(u)) - R(\Phi \mathcal{D}_u \log G_j) \end{aligned} \quad (3.15)$$

where to obtain the last equality we applied the chain rule. Using the expression derived in (3.14) for the logarithmic derivative of G_j , the conclusion follows. \square

3.2 The duality formula characterizes the reciprocal class

3.2.1 The reciprocal characteristic

We compute the reciprocal characteristic associated with a jump process. As it is clear from the introduction, the reciprocal characteristic is the natural parameter to describe reciprocal classes, in the same way as the intensity is the natural way to characterize a random walk.

Definition 3.2.1 (Reciprocal characteristic). *Let j be the intensity of a nice Markov counting process. We define the map $\Xi_j : [0, 1] \times \mathbb{Z} \rightarrow \mathbb{R}_+$*

$$\Xi_j(t, z) := \partial_t \log j(t, z) + j(t, z + 1) - j(t, z) \quad (3.16)$$

and call it the reciprocal characteristic associated with j .

We can see from the definition that the duality formula (3.12) can be rewritten as follows:

$$R(\mathcal{D}_u \Phi) = R\left(\Phi \int_0^1 [\dot{u}(t) + \Xi_j(t, X_{t-})u(t)] dX_t\right) \quad (3.17)$$

As a first result, let us show the reciprocal characteristic identifies all Markov processes having the same bridges:

3.2. THE DUALITY FORMULA CHARACTERIZES THE RECIPROCAL CLASS51

Theorem 3.2.1. *Let R, P be two nice Markov counting process and j, k be their intensities. Then $P \in \mathfrak{R}(R)$ if and only if $\Xi_j = \Xi_k$*

The proof of this theorem relies on the explicit computation of the intensity of a Doob h-transform of a NMC.

Proposition 3.2.1 (h transform). *Let R be the law of a NMC of intensity j . Let $h : \mathbb{Z} \rightarrow \mathbb{R}_+$ be such that $R(h(X_1)) = 1$. Then the process whose law P is given by:*

$$P = h(X_1)R$$

is a Markov counting process and its intensity $k : [0, 1] \times \mathbb{Z} \rightarrow \mathbb{R}$ satisfies:

$$k(t, X_{t-}) = \frac{h(t, X_{t-} + 1)}{h(t, X_{t-})} j(t, X_{t-}) \quad (3.18)$$

where $h(t, z) = R(h(X_1) | X_t = z)$.

Proof. First note that $h(t, z)$ is time-differentiable (a standard semigroup argument) and space time harmonic, that is, a solution of the Kolmogorov Backward PDE:

$$\partial_t h(t, z) + j(t, z)[h(t, z + 1) - h(t, z)] = 0 \quad R_t - a.s. \forall t \in [0, 1]$$

Define $\psi(t, z) := \log h(t, z)$. It satisfies:

$$\partial_t \psi(t, z) = j(t, z) [1 - \exp(\psi(t, z + 1) - \psi(t, z))] \quad (3.19)$$

By Itô formula:

$$\begin{aligned} \psi(t, X_{t-}) &= \int_0^t \partial_t \psi(s, X_{s-}) ds + \sum_{i: T_i \leq t} \psi(T_i, X_{T_i}) - \psi(T_i, X_{T_i-}) \\ &= \int_0^t \left(1 - \frac{h(s, X_{s-} + 1)}{h(s, X_{s-})}\right) j(s, X_{s-}) ds + \sum_{i: T_i \leq t} \log\left(\frac{h(T_i, X_{T_i})}{h(T_i, X_{T_i-})}\right) \end{aligned}$$

Since $h = \exp(\psi)$, this implies that

$$h(X_1) = \exp\left(-\int_0^1 \left(\frac{h(s, X_{s-} + 1)}{h(s, X_{s-})} - 1\right) j(s, X_{s-}) ds\right) \prod_{i: T_i < 1} \frac{h(T_i, X_{T_i})}{h(T_i, X_{T_i-})}$$

and

$$h(X_1)G_j = \exp\left(-\int_0^1 j(s, X_{s-}) \frac{h(s, X_{s-} + 1)}{h(s, X_{s-})} - 1 ds\right) \prod_{i: T_i < 1} j(T_i, X_{T_i-}) \frac{h(T_i, X_{T_i})}{h(T_i, X_{T_i-})}$$

where we recall the definition of G_j at (3.1). With Girsanov's Theorem we see that P admits the intensity k defined in (3.18). \square

We can get back to the proof of Theorem 3.2.1.

Proof. (\Rightarrow) Assume that $P \in \mathfrak{R}(R)$. Then by Proposition 2.2.2, for any $x \in \text{supp}P_0$ there exists $h : \mathbb{Z} \rightarrow \mathbb{R}$ such that $P^x = h(X_1)R^x$. This implies, by Theorem 3.2.2 that k and j are related through the relation (3.18):

$$k(t, z) = \exp(\psi(t, z + 1) - \psi(t, z))j(t, z) \quad (3.20)$$

where ψ solves the equation (3.19): This leads to:

$$\begin{aligned} \partial_t(\log k(t, z) - \log j(t, z)) &= \partial_t(\psi(t, z + 1) - \psi(t, z)) \\ &\stackrel{(3.19)}{=} j(t, z + 1)[1 - \exp(\psi(t, z + 2) - \psi(t, z + 1))] \\ &\quad - j(t, z)[1 - \exp(\psi(t, z + 2) - \psi(t, z + 1))] \\ &\stackrel{(3.20)}{=} (j(t, z + 1) - j(t, z)) - (k(t, z + 1) - k(t, z)) \end{aligned}$$

which implies the equality of Ξ_k and Ξ_j .

(\Leftarrow) Assume that $\Xi_k = \Xi_j$, i.e.:

$$\partial_t \log j(t, z) + j(t, z + 1) - j(t, z) = \partial_t \log k(t, z) + k(t, z + 1) - k(t, z) \quad (3.21)$$

Integrating with respect to time (3.21), we obtain that there exists a space dependent function $c : \mathbb{Z} \rightarrow \mathbb{R}$ such that:

$$\log\left(\frac{k}{j}(t, z)\right) = \phi(t, z + 1) - \phi(t, z) + c(z)$$

where

$$\phi(t, z) = \int_0^t j(s, z) - k(s, z) ds$$

By applying Girsanov's Theorem:

$$\frac{dP}{dR} = \exp\left(\int_0^1 (k(s, X_{s-}) - j(s, X_{s-})) ds\right) \prod_{i:T_i < 1} \frac{k}{j}(T_i, X_{T_i^-})$$

We can rewrite the density using in terms of ψ, c :

$$\begin{aligned} \frac{dP}{dR} &= \exp\left(\int_0^1 k(s, X_{s-}) - j(s, X_{s-}) ds\right) \prod_{i:T_i < 1} \frac{k}{j}(T_i, X_{T_i^-}) \\ &= \exp\left(\int_0^1 \partial_t \phi(t, X_{s-}) ds + \sum_{i:T_i < 1} (\log k - \log j)(T_i, X_{T_i^-})\right) \\ &= \exp\left(\int_0^1 \partial_t \phi(t, X_{s-}) ds + \sum_{i:T_i < 1} \phi(T_i, X_{T_i}) - \phi(T_i, X_{T_i^-}) + \sum_{i:T_i < 1} c(X_{T_i^-})\right) \\ &= \exp(\phi(1, X_1) - \phi(0, X_0) + \sum_{i:T_i < 1} c(X_{T_i^-})) \end{aligned}$$

3.2. THE DUALITY FORMULA CHARACTERIZES THE RECIPROCAL CLASS 53

where to derive the last equality we applied the Itô formula. But $\sum_{i:T_i < 1} c(X_{T_i^-}) = c(X_0 + 1) + \dots + c(X_1 - 1)$ depends only on the initial and final state X_0 and X_1 . Proposition 2.2.2 gives the conclusion. \square

Example 3.2.1. 1. A Poisson process of parameter λ always belongs to the reciprocal class of a standard Poisson process (i.e. $\lambda = 1$): it is the well known fact that Poisson processes with different intensities have the same bridges.

2. A Poisson bridge \mathbf{P}^{xy} belongs, by definition, to the reciprocal class $\mathfrak{R}(\mathbf{P})$ of a Poisson process. Theorem 3.2.1 gives another proof of this fact. Indeed, it is well known that the intensity of R^{xy} is:

$$j^{xy}(t, z \rightarrow z + 1) = \frac{y - z}{(1 - t)}$$

Therefore the reciprocal characteristic $\Xi_{j^{xy}}(t, z)$ is:

$$\partial_t \log \left(\frac{y - z}{(1 - t)} \right) + \frac{y - (z + 1)}{(1 - t)} - \frac{y - z}{(1 - t)} = \frac{1}{(1 - t)} - \frac{1}{(1 - t)} = 0$$

The same computation can be repeated for any reference intensity j , using (3.16) and (3.21).

3. Let $j(z)$ and $k(z)$ be two time homogeneous intensities. Then Theorem 3.2.1 tells that the associated processes share their bridges if and only if there exist a constant λ such that:

$$k(z) = j(z) + \lambda \quad \forall z \in \mathbb{Z}$$

This means that one process can be obtained from the other by superposition with a Poisson process of intensity λ

4. Consider two intensities which are only time-dependent, say $j(t)$ and $k(t)$. Then the associated walks share the same bridges if and only if there exist a constant c such that

$$j(t) = c k(t) \quad \forall t \in [0, 1].$$

This means that the associated counting processes can be obtained one from the other by thinning.

Theorem 3.2.1 is useful to confront the reciprocal classes of two Markov processes, and therefore it does not concern reciprocal probabilities in the strict sense. But the Duality formula we have derived above does, applies in this more general framework. We prove our main result of this section in two steps. First, we work out the Poisson case and then consider general intensities.

3.2.2 The Poisson case

The first interesting result is that the duality formula characterizes the Poisson bridge, under the obvious endpoint marginal constraints.

Proposition 3.2.2. *Let $P \in \mathcal{P}(\Omega)$ admit the endpoint marginal $P_{01} = \delta_{(x,y)}$ with $x \leq y \in \mathbb{Z}$. The process P is the Poisson bridge between x and y if and only if*

$$P(\mathcal{D}_u \Phi) = P\left(\Phi \int_0^1 \dot{u}(t) dX_t\right) \quad (3.22)$$

holds for any $\Phi \in \mathcal{S}$ and any $u \in \mathcal{U}$.

Proof. (\Rightarrow) Let $P = R^{xy}$ be the Poisson bridge from x to y . Then it is easy to see that $R^{xy} \ll R^x$ and that a version of the density is given by

$$\frac{dR^{xy}}{dR^x} = \frac{1}{R^x(X_1 = y)} \mathbf{1}_{\{X_1 = y\}} := h(X_1)$$

Since $X_1 \circ \Theta_u^\varepsilon = X_1$ for all $u \in \mathcal{U}$, we have that $\mathcal{D}_u h(X_1) = 0$ almost surely. But then, using the duality formula of Proposition 3.1.4 for R^x :

$$\begin{aligned} R^{xy}(\mathcal{D}_u \Phi) &= R^x(h(X_1) \mathcal{D}_u \Phi) \\ &= R^x(\mathcal{D}_u(h(X_1) \Phi)) \\ &= R^x(h(X_1) \Phi \delta(u)) \\ &= R^{xy}(\Phi \delta(u)) \end{aligned}$$

which shows that R^{xy} satisfies (3.22).

(\Leftarrow) Let us prove the converse statement by computing the jump intensity of Q by means of a Nelson stochastic derivative.

Fix $t \in [0, 1)$. We are going to apply (3.10) with Φ an $X_{[0,t]}$ -measurable function and $u \in \mathcal{U}$ such that $\dot{u} = \frac{1}{\varepsilon} \mathbf{1}_{[t, t+\varepsilon]} - \frac{1}{1-(t+\varepsilon')} \mathbf{1}_{[t+\varepsilon', 1]}$ where $0 < \varepsilon < \varepsilon'$. With these choices, we see that the left hand side of (3.22) is worth zero. The right hand side is worth

$$\frac{1}{\varepsilon} P\left(\Phi \int_t^{t+\varepsilon} dX_r\right) - \frac{1}{1-(t+\varepsilon')} P\left(\Phi \int_{t+\varepsilon'}^1 dX_r\right)$$

By taking the conditional expectation, since Φ is an arbitrary $X_{[0,t]}$ measurable functional, we obtain:

$$\frac{1}{\varepsilon} P\left(\int_t^{t+\varepsilon} dX_r \mid X_{[0,t]}\right) = \frac{1}{1-(t+\varepsilon')} P\left(\int_{t+\varepsilon'}^1 dX_r \mid X_{[0,t]}\right)$$

3.2. THE DUALITY FORMULA CHARACTERIZES THE RECIPROCAL CLASS 55

for every small enough ε' . Remark that both sides of the equality are constant as functions of ε and ε' . In particular, for almost every t the stochastic derivative

$$a(t) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} P\left(\int_t^{t+\varepsilon} dX_r \mid X_{[0,t]}\right)$$

exists (and is equal to the right hand side). This shows that P admits the $dt P(d\omega)$ -almost everywhere defined process $(t, \omega) \mapsto a(t, \omega)$ as its intensity. Letting ε' tend to zero gives

$$a(t) = P\left(\frac{1}{1-t}(X_1 - X_t) \mid \mathcal{F}_{[0,t]}\right) = \frac{y - X_t}{1-t}.$$

We recognize the intensity of a Poisson bridge at time t with final condition y . \square

By randomizing the endpoint marginal of P in (3.22), we obtain the following characterization of the reciprocal class of a Poisson process:

Corollary 3.2.1. *If for any $\Phi \in \mathcal{S}$ and any $u \in \mathcal{U}$ the duality formula*

$$P(\mathcal{D}_u \Phi) = P\left(\Phi \int_0^1 \dot{u}(t) dX_t\right)$$

holds under $P \in \mathcal{P}(\Omega)$ where $P(X_1 - X_0) < +\infty$, then P belongs to the reciprocal class $\mathfrak{R}(R)$ of the Poisson process.

Note that we added the requirement $P(X_1 - X_0) < +\infty$ for the right hand side of the equation to be well defined.

3.2.3 The general case

Next result emphasizes that the duality formula (3.12) characterizes the reciprocal class of any NMC intensity j . A natural idea would be to follow the guideline of the proof of Proposition 3.2.2. Unfortunately, this leads to an implicit equation for the intensity, in contrast with the special Poissonian case where the reciprocal invariant $\Xi_{j=1}$ vanishes and leads to (3.22). However, a fruitful method consists in relying on the last corollary and pathwise techniques.

Theorem 3.2.2. *Let $P \in \mathcal{P}(\Omega)$ be such that $P(X_1 - X_0) < +\infty$. Then P is in $\mathfrak{R}(R)$ if and only if the duality formula*

$$P(\mathcal{D}_u \Phi) = P\left(\Phi \int_0^1 [\dot{u}(t) + \Xi_j(t, X_{t-})u(t)] dX_t\right) \quad (3.23)$$

holds for any $\Phi \in \mathcal{S}$ and any $u \in \mathcal{U}$.

Note that each term in (3.23) is meaningful since Ξ_j is bounded and $X_1 - X_0 \in L^1(P)$.

Proof. The direct statement is analogous to the Poisson case. Assume that $P \in \mathfrak{R}(R)$. Then, by Proposition 2.2.2, $P = h(X_0, X_1)R$ for some h . Since, clearly, for all $u \in \mathcal{U}$, $\mathcal{D}_u h(X_0, X_1) = 0$ we have:

$$\begin{aligned} P(\mathcal{D}_u \Phi) &= R(h(X_0, X_1)\mathcal{D}_u \Phi) \\ &\stackrel{\mathcal{D}_u h=0}{=} R(\mathcal{D}_u(h(X_0, X_1)\Phi)) \\ &= R((h(X_0, X_1)\Phi) \int_0^1 [\dot{u}(t) + \Xi_j(t, X_{t-})u(t)] dX_t) \\ &= P((h(X_0, X_1)\Phi) \int_0^1 [\dot{u}(t) + \Xi_j(t, X_{t-})u(t)] dX_t) \end{aligned}$$

The converse statement is based on the former Proposition 3.2.2. Let $\Omega_m = \{X_1 - X_0 = m\}$ and consider m such that $P(\Omega_m) > 0$. We note that (3.23) is satisfied by the measure

$$P^m := \frac{\mathbf{1}_{\Omega_m}}{P(\Omega_m)} P$$

as well. This can be shown by observing that the density of P^m w.r.t. to P depends only on the vector (X_0, X_1) and using the same argument we used to prove the direct statement.

Let G_j be as in equation (3.1). We define the probability measure \tilde{P}^m as follows:

$$d\tilde{P}^m := c \frac{1}{G_j} dP^m$$

where c is the normalizing constant. Since G_j is uniformly bounded from above and below on Ω_m , \tilde{P}^m is well defined. With (3.14), the identity (3.23) leads us to:

$$\begin{aligned} P^m(\mathcal{D}_u(G_j^{-1}\Phi)) &= P^m\left(G_j^{-1}\Phi \int_0^1 [\dot{u}_t + \Xi_j(t, X_{t-})u_t] dX_t\right) \\ &= P^m\left(G_j^{-1}\Phi \int_0^1 \dot{u}_t dX_t\right) - P^m\left(G_j^{-1}\mathcal{D}_u \log(G_j)\Phi\right) \\ &= P^m\left(G_j^{-1}\Phi \int_0^1 \dot{u}_t dX_t\right) + P^m\left(\mathcal{D}_u(G_j^{-1})\Phi\right). \end{aligned} \quad (3.24)$$

Hence,

$$\begin{aligned} \tilde{P}^m(\mathcal{D}_u\Phi) &= c P^m(G_j^{-1}\mathcal{D}_u\Phi) = -c P^m(\mathcal{D}_u(G_j^{-1}\Phi)) + c P^m(\mathcal{D}_u(G_j^{-1}\Phi)) \\ &\stackrel{(3.24)}{=} c P^m\left(G_j^{-1}\Phi \int_0^1 \dot{u}(t)dX_t\right) = \tilde{P}^m\left(\Phi \int_0^1 \dot{u}(t)dX_t\right). \end{aligned}$$

It follows from Corollary 3.2.1 that \tilde{P}^m is in the reciprocal class of a Poisson process, say \mathbf{P} . By Theorem 2.2.2 there exists h such that $d\tilde{P}^m = h(X_0, X_1) \mathbf{P}$. But this implies that

$$dP^m = c^{-1} G_j d\tilde{P}^m = c^{-1} G_j h(X_0, X_1) d\mathbf{P} = c^{-1} h(X_0, X_1) dR$$

and therefore $P^m \in \mathfrak{R}(R)$. By integrating with respect to m , we obtain that $P \in \mathfrak{R}(R)$, which is the desired result. \square

Theorem 3.2.2 improves Theorem 3.2.1 significantly because (i) it is not required a priori that the process which stands in the reciprocal class is an NMC process and (ii) no explicit expression of its intensity of jump is required. At this point the characterization of the reciprocal class of a nice counting process is complete. However, one may still wonder about the probabilistic meaning of the reciprocal characteristic. We will come back to this question in a more general framework in Chapter 6.

3.3 Lower bounds on the reciprocal characteristics

In this section we are going to investigate what are the consequences of a lower bound on the reciprocal characteristics for the dynamics of the bridge of a counting process. We show that this implies an estimate on the distribution of the last jump time, which we can be viewed as a measure of the speed at which a bridge converges to its final state. The estimate is sharp in the sense it is an exact computation when the reciprocal characteristic is constant as a function of space and time. It establishes an ordering between the reciprocal classes: the higher the lower bound, the lazier the bridges. Intuitively, can say that zero characteristics imply a uniform distribution of the jump times, which is the Poisson bridge, a positive characteristics imply that the jump times are concentrated around time one, while negative characteristics imply that they are concentrated around zero. The figure below illustrates these two opposite behaviors.

Proposition 3.3.1. *Let R^{0n} be the bridge between 0 and n of R . Assume that*

$$\inf_{t \in [0,1], 0 \leq i \leq n-1} \Xi_j(i, t) \geq c \in \mathbb{R} \quad (3.25)$$

Then:

$$R^{0n}(T_n \leq t) \leq \left(\frac{\exp(ct) - 1}{\exp(c) - 1} \right)^n \quad (3.26)$$

Remark 3.3.1. (i) *If we have a positive lower bound on the characteristic, jumps accumulate around time 1. In case of a negative upper bound, jump accumulate around time 0.*

(ii) *In (3.26) only considered the last jump time. However, a more general result might hold. That is, under the hypothesis (3.25) we can construct a coupling between the $0n$ bridge of R and the $0n$ bridge of a counting process \tilde{R} whose characteristics is constantly c in such a way that, at any time, the bridge of R is below that of \tilde{R} .*

(iii) *The estimate holds at equality if the reciprocal characteristic is constantly c*

(iv) *Yet another meaningful interpretation of (3.26) would be that it gives quantitative information on when did the last customer in a line arrived, knowing that n customers arrived in a unit of time. This fact may be of interest in queuing theory.*

Before the proof Proposition (3.3.1) we introduce some notation we are going to use and prove an auxiliary result. A typical element of $[0, 1]^n$ is denoted \mathbf{s} and its i -th coordinate s_i . For an element $\mathbf{s} \in [0, 1]^n$, and for $i \leq n$ we adopt the conventions

$$[0, 1]^{n-1} \in \mathbf{s}^i := (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n), \quad [0, 1]^n \ni \mathbf{s}_t^i := (s_1, \dots, s_{i-1}, t, s_{i+1}, \dots, s_n)$$

We write \cdot for the scalar product. In particular, we write $\mathbf{1} \cdot \mathbf{s}$ for $\sum_{i=1}^n s_i$. We can state and prove the following auxiliary result.

Proposition 3.3.2. *Let $\lambda : [0, 1]^n \rightarrow \mathbb{R}_+$ be continuous and almost everywhere continuously differentiable. If*

$$\inf_{\mathbf{s} \in [0,1]^n, 1 \leq i \leq n} \partial_{s_i} \lambda(\mathbf{s}) \geq c \quad (3.27)$$

then the function $\rho(t)$ defined by:

$$t \mapsto \rho(t) := \log \left(\int_{[0,t]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s} \right) - \log \left(\int_{[0,t]^n} \exp(c \mathbf{1} \cdot \mathbf{s}) d\mathbf{s} \right) \quad (3.28)$$

is non-decreasing.

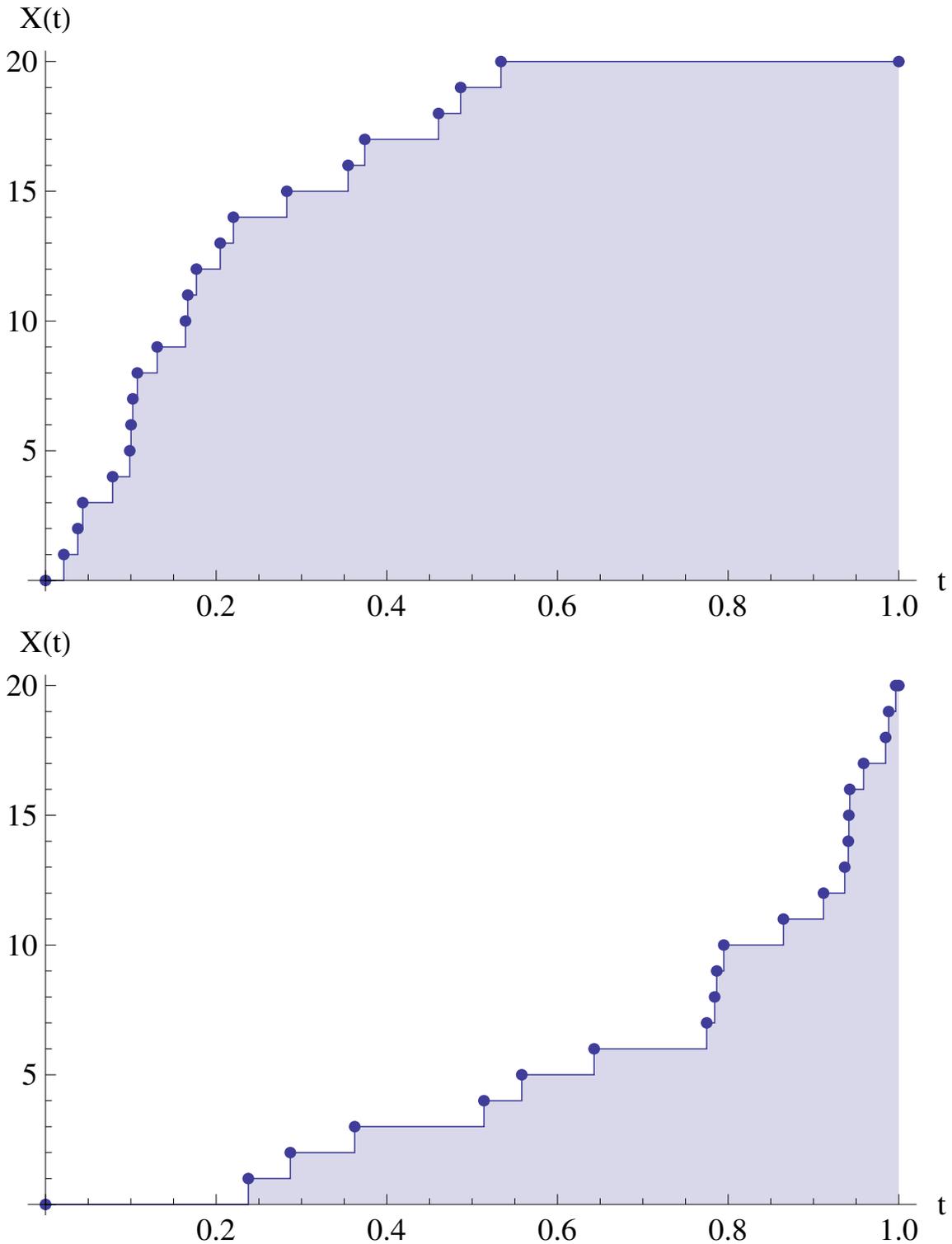


Figure 3.1: A fast and a lazy bridge. In the upper picture we simulated a trajectory of the bridge from 0 to 20 of a counting process whose reciprocal characteristic is constantly equal to -3 , whereas in the bottom picture we simulated the trajectory of the same bridge for a random walk whose reciprocal characteristic is constantly equal to 3 . The simulation clearly shows how the trajectory of the process with positive lower bound on the reciprocal characteristics has a concave profile, while the other one has a convex profile, reflecting the two different speeds at which they travel towards their final state

Proof. We can write ρ as

$$\rho(t) = \log(f(t)) - \log(g(t))$$

with the obvious identifications given by (3.28):

$$f(t) := \int_{[0,t]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s}, \quad g(t) := \int_{[0,t]^n} \exp(c\mathbf{1} \cdot \mathbf{s}) d\mathbf{s}$$

An application of the chain rule tells that ρ is non decreasing if and only if $(\partial_t f g - \partial_t g f)(t) \geq 0$ for all $t \in [0, 1]$. We can compute explicitly $\partial_t f(t)$:

$$\partial_t f(t) = \sum_{i=1}^n \int_{[0,t]^{n-1}} \exp(\lambda(\mathbf{s}_t^i)) ds^i$$

and similarly:

$$\partial_t g(t) = \sum_{i=1}^n \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}_t^i) ds^i$$

Therefore the conclusion would follow if we could prove that for all $t \in [0, 1]$, and for all $i \leq n$:

$$g(t) \int_{[0,t]^{n-1}} \exp(-\lambda(\mathbf{s}_t^i)) ds^i \geq f(t) \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}_t^i) ds^i \quad (3.29)$$

Using the elementary fact that

$$g(t) = \int_{[0,t]^n} \exp(c\mathbf{1} \cdot \mathbf{s}) d\mathbf{s} = \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}^i) ds^i \int_0^t \exp(cs_i) ds_i$$

we can rewrite the left hand side of (3.29) as follows :

$$\begin{aligned} g(t) \int_{[0,t]^{n-1}} \exp(\lambda(\mathbf{s}_t^i)) ds^i &= \int_{[0,t]^n} \exp(c\mathbf{1} \cdot \mathbf{s}) d\mathbf{s} \int_{[0,t]^{n-1}} \exp(\lambda(\mathbf{s}_t^i)) ds^i \\ &= \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}^i) ds^i \int_{[0,t]^n} \exp(\lambda(\mathbf{s}_t^i) cs_i) ds \end{aligned}$$

Using the hypothesis (3.27) we obtain for all $\mathbf{s} \in [0, t]^n$:

$$\lambda(\mathbf{s}_t^i) + cs_i - (\lambda(\mathbf{s})ct) = \left(\int_{s_i}^t \underbrace{\partial_{s_i} \lambda(\mathbf{s} + (h - s_i)\mathbf{e}_i)}_{\geq c} dh \right) - c(t - s_i) \geq 0$$

This implies that

$$\begin{aligned}
 & \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}^i) d\mathbf{s}^i \int_{[0,t]^n} \exp(\lambda(\mathbf{s}_t^i) c s_j) ds \\
 & \geq \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}^i) d\mathbf{s}^i \int_{[0,t]^n} \exp(\lambda(\mathbf{s}) ct) ds \\
 & = \int_{[0,t]^{n-1}} \exp(c\mathbf{1} \cdot \mathbf{s}^i) d\mathbf{s}^i \int_{[0,t]^n} \exp(\lambda(\mathbf{s})) ds = \partial_t g f(t)
 \end{aligned}$$

from which (3.29) follows. \square

We can go back to the proof of Proposition 3.3.1.

Proof. We first identify the density G_j^{0n} of R^{0n} with respect to the bridge of the standard Poisson process \mathbf{P}^{0n} . We have that

$$G_j^{0n} = \frac{1}{Z} \exp\left(\sum_{i=1}^n \xi^i(T_i)\right) \mathbf{P}^{0n} - a.s. \quad (3.30)$$

where $\xi^i(\cdot)$ is the primitive of $\Xi_j(i-1, \cdot)$ over $[0, 1]$ and Z is a normalization constant. The validity of this statement can be checked directly by showing that $\tilde{R}^{0n} := G_j^{0n} \mathbf{P}^{0n}$ satisfies the duality formula (3.23), using the properties of the derivative operator \mathcal{D}_u and the duality under the Poisson process. The computations are analogous to the ones we did above. Since under the Poisson bridge the vector (T_1, \dots, T_n) has the distribution of n ordered points in $[0, 1]$, we find that, if we define $\mathcal{P}([0, 1]^n) \ni r := R^{0n} \circ (T_1, \dots, T_n)^{-1}$, then :

$$r(ds) := \frac{1}{Z} \mathbf{1}_{s \in O} \exp\left(\sum_{i=1}^n \xi^i(s_i)\right) ds$$

where dt stands for the Lebesgue measure $[0, 1]^n$ and

$$O := \{s \in [0, 1]^n : 0 < s_1 < s_2 < \dots < s_n < 1\} \quad (3.31)$$

Therefore, making explicit the normalization constant:

$$r(t_n \leq t) = \frac{\int_{[0,t]^n \cap O} \exp(-\sum_{i=1}^n \xi^i(s_i)) ds}{\int_O \exp(-\sum_{i=1}^n \xi^i(s_i)) ds} \quad (3.32)$$

By symmetrizing the density we rewrite the latter as an integral on cubes, in view of an application of Proposition 3.3.1. To this aim, we define for every permutation $\sigma \in S_n$:

$$O^\sigma := \{s \in [0, 1]^n : 0 < s_{\sigma(1)} < s_{\sigma(2)} < \dots < s_{\sigma(n)} < 1\}$$

Note that O in (3.31) corresponds to O^{id} in this last definition, where id is the identity of S_n . We shall also the function $\lambda : [0, 1]^n \rightarrow \mathbb{R}_+$:

$$\lambda(\mathbf{s}) = \sum_{\sigma \in S_n} \mathbf{1}_{\mathbf{s} \in O^\sigma} \exp \left(\sum_{i=1}^n \xi^i(s_{\sigma(i)}) \right) \forall \mathbf{s} \in [0, 1]^n$$

This allows us to rewrite (3.32) as:

$$r(T_n \leq t) = \frac{\int_{[0,t]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s}}{\int_{[0,1]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s}} \quad (3.33)$$

Since almost everywhere (even though not everywhere, as at the boundaries between the O^σ λ is not differentiable) we have:

$$\partial_{s_i} \lambda(\mathbf{s}) = \sum_{\sigma \in S_n} \mathbf{1}_{\mathbf{s} \in O^\sigma} (\partial_s \xi^{\sigma^{-1}(i)})(s_i) \geq c$$

and it can be checked that λ is continuous we have that we can apply the former Proposition 3.3.1. Namely, $\rho(t) \leq \rho(1)$ (see (3.28)) tells that:

$$\frac{\int_{[0,t]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s}}{\int_{[0,t]^n} \exp(c \mathbf{1} \cdot \mathbf{s}) d\mathbf{s}} \leq \frac{\int_{[0,1]^n} \exp(\lambda(\mathbf{s})) d\mathbf{s}}{\int_{[0,1]^n} \exp(c \mathbf{1} \cdot \mathbf{s}) d\mathbf{s}}$$

which, by rearranging the terms to get the right hand side of 3.33, and evaluating the other integrals to their exact value, yields the conclusion. \square

Chapter 4

Random walks on lattices

Outline of the chapter In this chapter we study the reciprocal class of a random walk on a lattice. At each site of the lattice, the walker can choose among a set $\mathbb{A} = \{a^1, \dots, a^A\}$ of possible jumps. As A may be a large number, the cycles of the underlying graph have a very rich structure. The key tool to study them, is through the analysis of the geometrical properties of a specific sublattice of \mathbb{Z}^A , different from the one where the walk takes place. This sublattice is proven to be helpful to identify reciprocal characteristics and to construct a functional equation based on the characteristics, which characterizes $\mathfrak{R}(R)$. To achieve a sharp characterization we use generating sets of lattices, a tool from discrete geometry. This chapter is based on the submitted work [20], and on the forthcoming work [19].

Random walks on the square lattice or on the triangular lattice are among the most basic examples of stochastic processes, and a continuous source of interesting questions. Here we consider general lattices as state space, and graph structures which are *translation invariant*, meaning that there exist a universal set $\mathbb{A} = \{a^1, \dots, a^A\}$ such that at each site x the neighbors are the points $x + a^1, \dots, x + a^A$. The jump set \mathbb{A} shapes the structure of cycles of the graph. We see here for the first time how cycles play a crucial role in the study of reciprocal classes. We compute their characteristics: they are the product of the values of the reference intensity along them, and we will show how they determine the distribution of the jump types of the elements of $\mathfrak{R}(R)$.

The reference Markov walk is a walk on the lattice whose jump intensity is also translation invariant. This choice is motivated by the fact that

we want to achieve a characterization of the class in the most canonical way, as the set of probability measures satisfying a set of change of measure formulae. As we shall see, the formulae splits into two families. In one case, they arise from time changes, following the ideas of Chapter 3. The second family deals with the random vector $\mathbf{N}_1 \in \mathbb{N}^A$, which counts, for $1 \leq l \leq A$, how many times the walker moved along an arc of the form $z \rightarrow z + a^l$ over the time interval $[0, 1]$.

$$\mathbf{N}_1 = (N_1^l)_{1 \leq l \leq A}, \quad N_1^l(\omega) = \sum_{s \leq 1} \mathbf{1}_{\{\omega_s - \omega_{s-} = a^l\}}.$$

Because the reference jump intensity is space homogeneous, under the reference walk R , \mathbf{N}_1 follows a multivariate Poisson law.

That is, for some $(\lambda^1, \dots, \lambda^A)$, the l -th coordinate is a Poisson random variable of parameter λ^l and all coordinates are independent from each other.

This fails to be true under any bridge. What we are going to see is that *some* identities, which are valid under the multivariate Poisson distribution $R(\mathbf{N}_1 \in \cdot) \in \mathcal{P}(\mathbb{N}^A)$ are true under the law $R^{xy}(\mathbf{N}_1 \in \cdot)$ as well. Together with the obvious constraint

$$R^{xy} \left(\sum_{l=1}^A a^l \mathbf{N}_1^l = y - x \right) = 1$$

they are shown to characterize it. The identities are derived from the study of the law of the shifted vector $\mathbf{N}_1 + \mathbf{v}$, for $\mathbf{v} \in \mathbb{Z}^A$. In the one dimensional case Chen's characterization does this. It is known [13] that the Poisson law on \mathbb{N} of parameter λ , denoted by \mathfrak{p}_λ , is characterized by:

$$\mathfrak{p}_\lambda(f(n+1)) = \lambda^{-1} \mathfrak{p}_\lambda(f(n)n)$$

It is simple to extend this to the multivariate Poisson law $R(\mathbf{N}_1 \in \cdot)$. For any $\mathbf{v} \in \mathbb{Z}^A$, one can compute an explicit function¹ $G_{\mathbf{v}}^\lambda : \mathbb{N}^A \rightarrow \mathbb{R}_{\geq 0}$ such that:

$$R(f(\mathbf{N}_1 + \mathbf{v})) = R(f(\mathbf{N}_1)G_{\mathbf{v}}^\lambda(\mathbf{N}_1)) \quad \forall f \in \mathcal{B}(\mathbb{N}^A)$$

The crucial observation is that, if \mathbf{v} is chosen in a suitable set, then this identity extends to any bridge. The right set to look at is the set of those vectors $\mathbf{c} \in \mathbb{Z}^A$ with the property that:

¹The superscript λ emphasizes the dependence on the parameters of the Poisson multivariate law

$$\sum_{l=1}^A a^l c^l = 0 \quad (4.1)$$

This is clearly a *loop* condition, reminiscent of that of [67],[68] in the diffusion case. The vectors fulfilling it define another sublattice of \mathbb{Z}^A , that we call $\ker_{\mathbb{Z}}(\mathbf{A})$.

Summing up, we obtain that for any pair of initial and final states x, y :

$$R^{xy}(f(\mathbf{N}_1 + \mathbf{c})) = R^{xy}(f(\mathbf{N}_1)G_{\mathbf{c}}^{\lambda}(\mathbf{N}_1)) \quad \forall f \in \mathcal{B}(\mathbb{N}^A), \mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$$

The interpretation of condition 4.1 is that $\ker_{\mathbb{Z}}(\mathbf{A})$ gives an efficient way of looking at cycles. Indeed, if ω and $\tilde{\omega}$ are two paths in Ω with the same initial and final state, then it is natural to guess that be obtained from the other by adding a sequence of cycles, properly embedded in the time interval $[0, 1]$. This can be expressed by saying that $\mathbf{N}_1(\omega) - \mathbf{N}_1(\tilde{\omega}) \in \ker_{\mathbb{Z}}(\mathbf{A})$. The detailed expression of $G_{\mathbf{c}}^{\lambda}$ will also include the exact expression for the cycle characteristics associated with the random walk.

We have thus described the two main steps in the characterization of $\mathfrak{R}(R)$, which is given at Theorem 4.3.1.

We are also interested in achieving a *sharp* characterization, where the number of equations used is the minimal one. What seems natural, since lattices admit basis, is to work with a basis of $\ker_{\mathbb{Z}}(\mathbf{A})$ and the associated shifts, rather than *all* vectors of $\ker_{\mathbb{Z}}(\mathbf{A})$. We show that this is not in general possible. It is a quite technical problem, essentially due the fact that $\ker_{\mathbb{Z}}(\mathbf{A})$ may not have a basis of non negative vectors. The search for a sharp characterization leads to the study of *generating sets*, a useful generalization of the concept of lattice basis.

In the delicate situation when $\ker_{\mathbb{Z}}(\mathbf{A})$ has no basis with non negative vectors, we also need to be very careful in the parallelism between shifting \mathbf{N}_1 by $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$, and the addition a cycle to the canonical process. As we have said above, this certainly gives the right intuition. However, if $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$ has a strictly negative coordinate, the output of the shift $\mathbf{N}_1 + \mathbf{c}$ cannot directly be interpreted as an element of Ω , since the resulting path would have a negative number of jumps of some type. This is the reason why we need to keep the two families of formulae we discussed above distinct.

In Chapter 5 we present, under the assumption that $\ker_{\mathbb{Z}}(\mathbf{A})$ has a non negative basis, a unique formula characterizing the reciprocal class, where the shifts can effectively be seen as addition paths whose trace is a cycle (which we will call *loops*) to the paths of the canonical process.

In the last section 4.7 we prove a concentration of measure inequality for the distribution of the vector \mathbf{N}_1 under the bridge of the reference walk in a specific class of models, see Theorem 4.7.1. The concentration rates are determined by the reciprocal characteristic and the geometrical properties of the underlying lattice. This preliminary result make use of various observations, which are likely to be valid in more generality.

Organization of the chapter We provide the necessary setup in 4.1. Next, we present the two families of identities which will be used in the characterization of the reciprocal class we will give in Theorem 4.3.1. Section 4.4 is devoted to generating sets. In 4.5 we show how to use them to obtain refined versions of the results of the former sections. The last section contains a first result about the probabilistic interpretation of the reciprocal characteristics. Section 4.7 is dedicated to establish concentration of measure inequalities. The following chart helps in connecting the notation of this chapter with the general framework we set up in Chapter 2.

	General def	Local def
State space	\mathcal{X}	Lattice spanned by $\{a^1, \dots, a^A\}$
Arcs	\mathcal{A}	$(z \rightarrow z + a^l), 1 \leq l \leq A$
Active arcs	$\mathcal{A}_{\rightarrow}(j)$	$(z \rightarrow z + a^l), 1 \leq l \leq A$
Reference intensity	$j(t, z \rightarrow z + a^l)$	$j^l(t)$

4.1 The jump matrix and some other notation

Let us recall what a lattice of \mathbb{R}^d is:

Definition 4.1.1 (Lattice). *Let V be a d dimensional vector space on \mathbb{R} .*

1. *A lattice in V is a subgroup of the form:*

$$\mathcal{L} = \mathbb{Z}v^1 + \dots + \mathbb{Z}v^m$$

where the v^1, \dots, v^m are linearly independent elements of V . The m -uple (v^1, \dots, v^m) is called a basis of \mathcal{L} .

2. *Let $\mathbb{A} = \{a^1, \dots, a^A\}$ be finitely many vectors. Then the lattice generated by \mathbb{A} is the set:*

$$\mathcal{L} = \left\{ \sum_{l=1}^A z_l a^l, z_l \in \mathbb{Z} \forall 1 \leq l \leq A \right\}$$

Note that \mathbb{A} might not be a basis for the lattice it generates.

Definition 4.1.2 (graph and reference intensity). *Let $\mathbb{A} := \{a^1, \dots, a^A\}$ be a subset of \mathbb{R}^d . We take \mathcal{X} as the lattice generated by \mathbb{A} :*

$$\mathcal{X} = \left\{ \sum_{l=1}^A z_l a^l, z_l \in \mathbb{Z} \right\}$$

The set of arcs is given by

$$\mathcal{A} := \{(x \rightarrow x + a^l), x \in \mathcal{X}, 1 \leq l \leq A\}$$

We consider the reference intensity to be given by:

$$j(t, z \rightarrow z + a^l) = j^l(t), \quad \forall z \in \mathcal{X}, 1 \leq l \leq A \quad (4.2)$$

for some strictly positive C^1 function $j : [0, 1] \times \{1, \dots, A\} \rightarrow \mathbb{R}_+$.

With these choice, the reference walk R is well defined. Let us observe that, when the walk starts at $x \in \mathcal{X}$, then:

$$\text{supp}R_t^x = x + \left\{ \sum_{l=1}^A a^l n_l, n^l \in \mathbb{N} \right\}$$

which may be strict subset of \mathcal{X} . The graph $(\mathcal{X}, \rightarrow)$ is space homogeneous: the neighborhood of any vertex always looks the same, this is the lattice property. The intensity is space homogeneous as well. At any time t , the walker moves along an arc parallel to a^l at rate $j^l(t)$, no matter what is his current position.

As $\mathcal{X} \subseteq \mathbb{R}^d$ the random variable $X_t - X_{t-}$ is well defined for all $t \in [0, 1]$ and takes values in $\{0\} \cup \mathbb{A}$. As a consequence of our choices for the reference intensity, R has independent increments. The distribution of $X_t - X_s$ can be characterized through the Lévy Khintchine formula:

$$R\left(\exp(i\lambda \cdot (X_t - X_s))\right) = \exp\left(\sum_{l=1}^A (e^{i\lambda \cdot a^l} - 1) \int_s^t j^l(u) du\right), \quad \lambda \in \mathbb{R}^d \quad (4.3)$$

where we denoted by \cdot the scalar product.

Paths can be described by the jump processes corresponding to each element of \mathbb{A} . It is therefore natural to introduce the following random variables:

Definition 4.1.3. Let define $\mathbf{N} = (\mathbf{N}_t)_{0 \leq t \leq 1}$, where $\mathbf{N}_t := (N_t^1, \dots, N_t^A)$ and, for any $j \in \{1, \dots, A\}$, N_t^j counts how many times the jump a^j has occurred up to time t :

$$N_t^l(\omega) = \sum_{s \leq t} \mathbf{1}_{\{\omega_s - \omega_{s-} = a^l\}}.$$

The total amount of jumps up to time t , $|N|_t$, is given by the sum of the coordinates of \mathbf{N}_t , that is $|N|_t := \sum_{l=1}^A N_t^l$.
The i -th jump time of type a^l is:

$$T_i^l := \inf \{t \in [0, 1] : N_t^l = i\} \wedge 1.$$

Finally, the i^{th} jump time of the process is:

$$T_i := \inf \{t \in [0, 1] : |N|_t = i\} \wedge 1.$$

The jump matrix \mathbf{A} will play a crucial role through the lattice $\ker_{\mathbb{Z}}(\mathbf{A})$, which we introduce later on.

Definition 4.1.4 (Jump matrix). We define the matrix \mathbf{A} associated to \mathbb{A} by

$$\mathbf{A} = (a_i^l)_{1 \leq i \leq d, 1 \leq l \leq A} \in \mathbb{R}^{d \times A}$$

where a_i^l is the i -th coordinate of $a^l \in \mathbb{A}$.

We can express the canonical process in a compact form using the jump matrix as

$$X_t = X_0 + \sum_l a^l N_t^l = X_0 + \mathbf{A} \mathbf{N}_t.$$

The path space Ω , which we defined in Chapter 2, has the following simple description :

$$\Omega = \{\omega : |N|_1(\omega) < +\infty \text{ and } X_t(\omega) = X_0(\omega) + \mathbf{A} \mathbf{N}_t(\omega), 0 \leq t \leq 1, X_0 \in \mathcal{X}\} \\ \subseteq \mathbb{D}([0, 1], \mathbb{R}^d).$$

where $\mathbb{D}([0, 1], \mathbb{R}^d)$ is the usual càdlàg space over \mathbb{R}^d .

4.2 Time and space transformations

4.2.1 Time changes

We consider the set \mathcal{U} of all regular diffeomorphisms of the time interval $[0, 1]$, parametrized by the set \mathbb{A} :

$$\mathcal{U} = \left\{ u \in C^1(\{1, \dots, A\} \times [0, 1]; [0, 1]), u(\cdot, 0) \equiv 0, u(\cdot, 1) \equiv 1, \min_{l \leq A, t \in [0, 1]} \dot{u}(l, t) > 0 \right\}.$$

With the help of each $u \in \mathcal{U}$ we construct a transformation of the reference walk by time changes acting separately on each component process $N^l, l = 1, \dots, A$. This is similar to what we did in Chapter 3. However, we do not take derivatives here, and do not consider infinitesimal time changes.

Definition 4.2.1. Let $u \in \mathcal{U}$. We define the time-change transformation π_u by:

$$\begin{aligned} \pi_u : \Omega &\longrightarrow \mathbb{D}([0, 1], \mathbb{R}^d) \\ \pi_u(\omega)(t) &:= \omega(0) + \sum_{l=1}^A a^l N_{u(l,t)}^l(\omega), \quad 0 \leq t \leq 1. \end{aligned}$$

Remark 4.2.1. We cannot a priori be sure that π_u takes values in Ω since it may happen that jumps synchronize, i.e. $u^{-1}(l, T_i^l) = u^{-1}(l', T_i^{l'})$ for some l, l' . However it is easy to see that this happens with zero probability under R .

We now define a family of maps. They are the arc characteristics.

Definition 4.2.2. The *arc characteristic* associated to j is the function:

$$\Xi_j : \{1, \dots, A\} \times [0, 1]^2 \rightarrow \mathbb{R}_+, \quad \Xi_j(l, s, t) := \frac{j^l(t)}{j^l(s)} \quad (4.4)$$

Remark 4.2.2. In the time-homogeneous case $\Xi_j \equiv 1$.

Remark 4.2.3. Note that, for all $z \in \mathcal{X}, 1 \leq l \leq A, s, t \in [0, 1]$, we have that:

$$\Xi_j(l, s, t) = \frac{j(t, z \rightarrow z + a^l)}{j(s, z \rightarrow z + a^l)}$$

If we had directly generalized Definition 3.2.1, valid for counting processes ($\mathcal{X} = \mathbb{Z}, \mathbb{A} = \{1\}$) we would take the following expression as a definition for the arc characteristic:

$$\begin{aligned} \partial_t \log j(t, z \rightarrow z + a^l) &+ \sum_{l'=1}^A j(t, z + a^l \rightarrow z + a^l + a^{l'}) - \sum_{l'=1}^A j(t, z \rightarrow z + a^{l'}) \\ &= \partial_t \log j^l(t) \end{aligned}$$

We see the two definitions (3.16) and (4.4) are coherent, in the sense that one can be obtained from the other by some standard algebraic manipulations, and thus they depend on j through the same functionals. Indeed we have

$$\partial_t \log j(t, z \rightarrow z + a^l) = \lim_{\varepsilon \downarrow 0} \frac{\Xi_j(l, t, t + \varepsilon) - 1}{\varepsilon}$$

On the contrary, by integrating, we recover $\Xi_j(l, s, t)$ from the function $u \mapsto \partial_t \log(j(u, z \rightarrow z + a^l))$.

In the next proposition we shall prove that the image of R under the above time change π_u is absolutely continuous with respect to R , and that its density is indeed a function of the arc characteristic Ξ_j .

Proposition 4.2.1. *The following functional equation holds under R : For all $u \in \mathcal{U}$ and $F \in \mathcal{B}^+(\Omega)$,*

$$R(F \circ \pi_u) = R\left(F \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right). \quad (4.5)$$

Proof. We first observe that, for every fixed $l \in \{1, \dots, A\}$ the process

$$N_t^l \circ \pi_u - \int_0^t j^l(u(l, s)) \dot{u}(l, s) ds \quad (4.6)$$

is a R -martingale w.r.t. to its natural filtration $\tilde{\mathcal{F}}$. Indeed, for any $s \leq t$ and any $F \tilde{\mathcal{F}}_s$ -measurable, by applying the basic properties of processes with independent increments, we obtain:

$$\begin{aligned} R(F (N_t^l - N_s^l) \circ \pi_u) &= R(F) \int_{u(l, s)}^{u(l, t)} j^l(\tau) d\tau \\ &= R(F) \int_s^t j^l(u(l, \tau)) \dot{u}(l, \tau) d\tau. \end{aligned}$$

Therefore $N_t^l \circ \pi_u$ is a Markov counting process with intensity $(t, z) \mapsto j^l(u(l, t))\dot{u}(l, t)$. Moreover, if $l \neq l'$, $N_t^l \circ \pi_u$ and $N_t^{l'} \circ \pi_u$ are independent processes under R , because the processes N^l and $N^{l'}$ are independent and π_u acts separately on each component. This implies that the image of R under π_u , $R \circ \pi_u^{-1}$, is a random walk on \mathcal{X} whose intensity k is:

$$k(t, z \rightarrow z + a^l) = j^l(u(l, t))\dot{u}(l, t), \quad \forall z \in \mathcal{X}, t \in [0, 1], l \leq A.$$

We can now apply the Girsanov theorem to get the density of the push-forward measure $R \circ \pi_u^{-1}$ w.r.t. R :

$$\begin{aligned} \frac{dR \circ \pi_u^{-1}}{dR} &= \exp \left[\sum_{l=1}^A \left(\int_0^1 (j^l(u(l, t))\dot{u}(l, t) - j^l(t)) dt \right. \right. \\ &\quad \left. \left. + \int_0^1 \log(\Xi_j(l, t, u(l, t))\dot{u}(l, t) dN_t^l) \right) \right]. \end{aligned}$$

With the change of variable $t = u^{-1}(t')$ we have for any l :

$$\int_0^1 j^l(u(r, t))\dot{u}(l, t) dt = \int_0^1 j^l(t') dt'.$$

Therefore the first integral disappears and the conclusion follows. \square

4.2.2 Space transformations

The transformations π_u introduced in the previous section, when acting on a given path, change the jump times leaving unchanged the total number of jumps of each type. We now introduce transformations that modify the total number of jumps; these transformations act on the counting variable \mathbf{N}_1 taking its values in \mathbb{N}^A , which we embed into \mathbb{Z}^A to take advantage of the lattice structure. Here, we make a short deviation from the study of reciprocal classes to study a different, but tightly related problem, which is the problem of finding good characterizations of the conditional laws of Poisson random vectors. This subject is interesting in its own right, and we are going to prove generalizations of well known formulas and introduce some new objects whose interest goes beyond this specific problem. We make clear the connection between this problem and the characterization of $\mathfrak{R}(R)$ in Theorem 4.3.1.

Shifting a Poisson random vector

We consider a multivariate Poisson distribution $\mathbf{p}_\lambda \in \mathcal{P}(\mathbb{N}^A)$ where $\lambda = (\lambda^1, \dots, \lambda^A) \in \mathbb{R}_+^A$:

$$\forall \mathbf{n} \in \mathbb{N}^A, \quad \mathbf{p}_\lambda(\mathbf{n}) = \exp\left(-\sum_{l=1}^A \lambda^l\right) \frac{\lambda^{\mathbf{n}}}{\mathbf{n}!} . \quad (4.7)$$

where we use the notation:

$$\lambda^{\mathbf{n}} := \prod_{l=1}^A (\lambda^l)^{n^l}, \quad \mathbf{n}! = \prod_{l=1}^A n^l!$$

We recall Chen's characterization of the 1-dimensional Poisson random variable of parameter λ :

$$\mathbf{p}_\lambda(f(n+1)) = \frac{1}{\lambda} \mathbf{p}_\lambda(f(n)n)$$

Chen introduced it to estimate the rate of convergence of sum of dependent trials to the Poisson distribution. (see the original paper [13] and Chapter 9 in [76] for a complete account of Chen's method). Let us first give a straightforward multidimensional version of it.

Proposition 4.2.2. *Let $\lambda \in (\mathbb{R}_+)^A$. Then $\rho \in \mathcal{P}(\mathbb{N}^A)$ is the multivariate Poisson distribution \mathbf{p}_λ if and only if*

$$\forall \mathbf{e}^l, l = 1, \dots, A, \quad \rho(f(\mathbf{n} + \mathbf{e}^l)) = \frac{1}{\lambda^l} \rho(f(\mathbf{n})n^l), \quad \forall f \in \mathcal{B}(\mathbb{N}^A),$$

where \mathbf{e}^l denote the l -th vector of the canonical basis of \mathbb{Z}^A .

One can interpret this characterization as the computation of the density of the image measure by any shift along the canonical basis of \mathbb{N}^A .

Now we consider as more general transformations multiple left- and right-shifts, acting simultaneously on each coordinate, that is, we shift by vectors $\mathbf{v} \in \mathbb{Z}^A$.

Definition 4.2.3. *Let $\mathbf{v} \in \mathbb{Z}^A$. We define the \mathbf{v} -shift by*

$$\begin{aligned} \theta_{\mathbf{v}} : \mathbb{Z}^A &\longrightarrow \mathbb{Z}^A \\ \mathbf{z} &\longmapsto \theta_{\mathbf{v}}(\mathbf{z}) = \mathbf{z} + \mathbf{v}. \end{aligned}$$

Consider the image of \mathfrak{p}_λ under $\theta_{\mathbf{v}}$. It is a probability measure whose support is no more included in \mathbb{N}^A since there may be $\mathbf{z} \in \mathbb{N}^A$ such that $\theta_{\mathbf{v}}(\mathbf{z}) \notin \mathbb{N}^A$. Therefore we only compute the density of its absolutely continuous component, appearing in the Radon-Nykodim decomposition:

$$\mathfrak{p}_\lambda \circ \theta_{\mathbf{v}}^{-1} = \mathfrak{p}_\lambda^{\mathbf{v},ac} + \mathfrak{p}_\lambda^{\mathbf{v},sing}. \quad (4.8)$$

A version of the density of the absolutely continuous component is given by

$$\frac{d\mathfrak{p}_\lambda^{\mathbf{v},ac}}{d\mathfrak{p}_\lambda}(\mathbf{n}) = \lambda^{-\mathbf{v}} \frac{\mathbf{n}!}{(\mathbf{n} - \mathbf{v})!} \mathbf{1}_{\{\mathbf{n} - \mathbf{v} \in \mathbb{N}^A\}}$$

In view of obtaining a change of measure formula as in Proposition 4.2.2 we define

$$G_{\mathbf{v}}(\mathbf{n}) := \frac{\mathbf{n}!}{(\mathbf{n} - \mathbf{v})!} \mathbf{1}_{\{\mathbf{n} - \mathbf{v} \in \mathbb{N}^A\}} \quad (4.9)$$

Let us now consider the space $\mathcal{B}^\sharp(\mathbb{Z}^A) \subseteq \mathcal{B}(\mathbb{Z}^A)$ consisting of test functions with support in \mathbb{N}^A :

$$\mathcal{B}^\sharp(\mathbb{Z}^A) := \{f \in \mathcal{B}(\mathbb{Z}^A) : f(\mathbf{z}) = 0 \forall \mathbf{z} \notin \mathbb{N}^A\}.$$

Then, the considerations above can be summarized in the following formula:

$$\mathfrak{p}_\lambda(f \circ \theta_{\mathbf{v}}) = \lambda^{-\mathbf{v}} \mathfrak{p}_\lambda(f G_{\mathbf{v}}), \quad \forall f \in \mathcal{B}^\sharp(\mathbb{Z}^A). \quad (4.10)$$

Note that the dependence on the parameters of the Poisson law is only on $\lambda^{-\mathbf{v}}$ and not on $G_{\mathbf{v}}$.

Example 4.2.1. Let $\mathbb{A} = \{-1, 1\}$. We call n^- (rather than n^1) and n^+ (rather than n^2) the counting variables for the jumps -1 and 1 respectively. The same convention is adopted for the intensity vector $\lambda = (\lambda^-, \lambda^+)$. Then, for $\mathbf{v} = (1, 1)$ (resp. $\mathbf{v} = (1, -1)$) and for any $f \in \mathcal{B}^\sharp(\mathbb{Z}^2)$,

$$\mathfrak{p}_\lambda\left(f(n^- + 1, n^+ + 1)\right) = \frac{1}{\lambda^- \lambda^+} \mathfrak{p}_\lambda\left(f(n^-, n^+) n^- n^+\right),$$

$$\mathfrak{p}_\lambda\left(f(n^- + 1, n^+ - 1)\right) = \frac{\lambda^+}{\lambda^-} \mathfrak{p}_\lambda\left(f(n^-, n^+) \frac{n^-}{n^+ + 1}\right).$$

Lattices and conditional distributions

We now consider, associated to a measure $\mu \in \mathcal{P}(\mathbb{N}^A)$, the following set of probability measures on \mathbb{N}^A :

$$\mathcal{R}_{\mathbf{A}}(\mu) := \left\{ \rho \in \mathcal{P}(\mathbb{N}^A) : \rho(\cdot) = \int \mu(\cdot | \sigma(\mathbf{A})) d\rho_{\sigma(\mathbf{A})} \right\}, \quad (4.11)$$

where the σ -algebra $\sigma(\mathbf{A})$ is generated by the application $\mathbf{z} \mapsto \mathbf{A}\mathbf{z}$ defined on \mathbb{Z}^A , and the measure $\rho_{\sigma(\mathbf{A})}$ is the projection of ρ on $\sigma(\mathbf{A})$.

The set $\mathcal{R}_{\mathbf{A}}(\mu)$ presents strong analogies with a reciprocal class. Indeed, one can prove an analogous of Proposition 2.2.2, that is:

Proposition 4.2.3. $\rho \in \mathcal{R}_{\mathbf{A}}(\mu)$ if and only if $\rho \ll \mu$ and $\frac{d\rho}{d\mu}$ is $\sigma(\mathbf{A})$ -measurable

Proof. If $\rho \in \mathcal{R}_{\mathbf{A}}(\mu)$ then we have that for all \mathbf{m}', \mathbf{m} such that $\mathbf{A}\mathbf{m} = \mathbf{A}\mathbf{m}'$:

$$\rho(\mathbf{n} = \mathbf{m} | \mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m}) = \mu(\mathbf{n} = \mathbf{m} | \mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m})$$

and

$$\rho(\mathbf{n} = \mathbf{m}' | \mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m}) = \mu(\mathbf{n} = \mathbf{m}' | \mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m})$$

These two condition imply that

$$\frac{\rho(\mathbf{m})}{\mu(\mathbf{m})} = \frac{\rho(\mathbf{m}')}{\mu(\mathbf{m}')} \quad \forall \mathbf{m}, \mathbf{m}' \text{ s.t. } \mathbf{A}\mathbf{m} = \mathbf{A}\mathbf{m}'$$

From this, it follows that the density of ρ w.r.t. to μ is $\sigma(\mathbf{A})$ -measurable. On the other hand, if the density is $\sigma(\mathbf{A})$ -measurable, then there exist a function h such that

$$\frac{\rho(\mathbf{m})}{\mu(\mathbf{m})} = h(\mathbf{A}\mathbf{m}) \quad \forall \mathbf{m} \in \mathbb{N}^A$$

But then, for all $\mathbf{v} \in \mathbb{R}^d$ and \mathbf{m} such that $\mathbf{A}\mathbf{m} = \mathbf{v}$, using the equation above:

$$\rho(\mathbf{n} = \mathbf{m} | \mathbf{A}\mathbf{n} = \mathbf{v}) = \frac{h(\mathbf{v})\mu(\mathbf{m})}{\sum_{\mathbf{m}': \mathbf{A}\mathbf{m}' = \mathbf{v}} h(\mathbf{v})\mu(\mathbf{m}')} = \mu(\mathbf{m}' | \mathbf{A}\mathbf{m}' = \mathbf{v})$$

□

Our first goal is to characterize $\mathcal{R}_{\mathbf{A}}(\mathfrak{p}_\lambda)$ using the formula (4.10) computed for a suitably chosen set of shift vectors \mathbf{v} . The right set will be the following sublattice of \mathbb{Z}^{A^2} :

$$\ker_{\mathbb{Z}}(\mathbf{A}) := \ker(\mathbf{A}) \cap \mathbb{Z}^A. \quad (4.12)$$

To see that $\ker_{\mathbb{Z}}(\mathbf{A})$ is a lattice, it is enough to check that it is discrete and closed under summation (see e.g. Proposition 4.2 [59]), which is certainly the case here. The next statement clarifies the role of $\ker_{\mathbb{Z}}(\mathbf{A})$.

²As usual, $\ker \mathbf{A} = \{\mathbf{z} \in \mathbb{R}^A \text{ s.t. } \mathbf{A}\mathbf{z} = 0\}$

Proposition 4.2.4. *Let $\rho \in \mathcal{P}(\mathbb{N}^A)$. Then $\rho \in \mathcal{R}_A(\mathfrak{p}_\lambda)$ if and only if*

$$\forall \mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A}), \quad \rho(f \circ \theta_{\mathbf{c}}) = \frac{1}{\lambda^{\mathbf{c}}} \rho(f G_{\mathbf{c}}) \quad \forall f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A), \quad (4.13)$$

where $G_{\mathbf{c}}$ is defined in (4.9).

Proof. (\Rightarrow) Let $f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A)$ and $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$. By definition of $\ker_{\mathbb{Z}}(\mathbf{A})$ and $\mathcal{R}_A(\mathfrak{p}_\lambda)$ we can choose a version of the density $h = \frac{d\rho}{d\mathfrak{p}_\lambda}$ such that $h \circ \theta_{\mathbf{c}} = h$. Applying the formula (4.10) we obtain:

$$\begin{aligned} \rho(f \circ \theta_{\mathbf{c}}) &= \mathfrak{p}_\lambda((f \circ \theta_{\mathbf{c}})h) = \rho((fh) \circ \theta_{\mathbf{c}}) \\ &= \lambda^{-\mathbf{c}} \mathfrak{p}_\lambda(fh) = \lambda^{-\mathbf{c}} \rho(fh) \end{aligned}$$

(\Leftarrow) Let $\mathbf{n}, \mathbf{m} \in \mathbb{N}^A$ be such that $\mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m}$. Set $f := \mathbf{1}_{\mathbf{n}}$, $\mathbf{c} := \mathbf{n} - \mathbf{m} \in \ker_{\mathbb{Z}}(\mathbf{A})$. Then (4.13) gives:

$$\rho(\mathbf{m}) = \rho(f \circ \theta_{\mathbf{c}}) = \lambda^{-\mathbf{c}} G_{\mathbf{c}}(\mathbf{n}) \rho(\mathbf{n}).$$

Since, by (4.10), the same relation holds under \mathfrak{p}_λ , we have

$$\frac{d\rho}{d\mathfrak{p}_\lambda}(\mathbf{m}) = \frac{d\rho}{d\mathfrak{p}_\lambda}(\mathbf{n}),$$

which completes the proof. \square

Remark 4.2.4. *We have denoted the elements of $\ker_{\mathbb{Z}}(\mathbf{A})$ by \mathbf{c} . The same notation was used to denote the cycles of $(\mathcal{X}, \rightarrow)$. Even though the two concepts are similar, they are not the same one, and we have to be careful in not identifying cycles with elements of $\ker_{\mathbb{Z}}(\mathbf{A})$. For a given cycle, we can naturally associate an element of $\ker_{\mathbb{Z}}(\mathbf{A})$ by counting how many times each type of arc occurred in the cycle. This will be done in section 4.6 at equation (4.38).*

Example 4.2.2. *Resuming Example 4.2.1, we verify that, in this case, $\ker_{\mathbb{Z}}(\mathbf{A}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mathbb{Z}$. Proposition 4.2.4 tells us that a probability distribution ρ on \mathbb{N}^2 satisfies*

$$\rho(\cdot \mid n^+ - n^- = x) = \mathfrak{p}_\lambda(\cdot \mid n^+ - n^- = x) \quad \forall x \in \mathbb{Z}$$

if and only if, for all k in \mathbb{N}^* and for all $f \in \mathcal{B}^{\sharp}(\mathbb{Z}^2)$,

$$\rho\left(f(n^- + k, n^+ + k)\right) = \frac{1}{(\lambda^+ \lambda^-)^k} \rho\left(f(n^-, n^+) \prod_{i=0}^{k-1} (n^- - i)(n^+ - i)\right)$$

and

$$\rho\left(f(n^- - k, n^+ - k)\right) = (\lambda^+ \lambda^-)^k \rho\left(f(n^-, n^+) \prod_{i=1}^k \frac{1}{(n^- + i)(n^+ + i)}\right).$$

In particular, consider ρ , the distribution of a two-dimensional Poisson random vector (n^-, n^+) conditioned by the event $\{n^+ - n^- = 0\}$. Plugging in the above formula a functional of the form $f(n^-, n^+) = \exp(rn^+)$, and using the fact that $n^+ - n^- = 0$, one gets:

$$\exp(r)\rho(\exp(rn^+)) = \frac{1}{\lambda^- \lambda^+} \rho(\exp(rn^+)(n^+)^2)$$

That is, the Laplace transform of ρ , $\phi(r) := \rho(\exp(rn^+))$, satisfies the ODE:

$$\exp(r)\phi(r) = \frac{1}{\lambda^- \lambda^+} \phi''(r).$$

Such equation is promising in view of deriving concentration properties for the conditional distribution, and generalizes the well known ODE satisfied by the Laplace transform of a Poisson distribution of parameter λ :

$$\exp(r)\phi(r) = \frac{1}{\lambda} \phi'(r).$$

Proposition 4.2.4 characterize $\mathcal{R}_{\mathbf{A}}(\mathfrak{p}_{\lambda})$ in terms of countably many equations, those in (4.13). It is natural to ask whether one can reduce to finitely many equations. In particular, since $\ker_{\mathbb{Z}}(\mathbf{A})$ is a sublattice of \mathbb{Z}^A , one wants to understand if restricting to a basis suffices. In general this is false, as we will see in Example 4.4.1 of Section 4.4. This problem is related to rather deep geometrical properties of $\ker_{\mathbb{Z}}(\mathbf{A})$, and will be discussed in Section 4.4.

4.3 Characterization of the reciprocal class

4.3.1 Main result

We present here one of the main results of this chapter: the reciprocal class $\mathfrak{R}(R)$ associated to random walk on a lattice is characterized as the set of all probabilities for which a family of transformations induces the same density, expressed in terms of the reciprocal characteristics. We have already introduced in the previous section the family of reciprocal arc characteristics. Let us now introduce the family of cycle characteristics.

Definition 4.3.1. Let j be a jump intensity as in (4.2). For any $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$ we call *cycle characteristic* the positive number $\Phi_j^{\mathbf{c}}$:

$$\Phi_j^{\mathbf{c}} := \prod_{l=1}^A \left(\int_0^1 j^l(t) dt \right)^{-c^l}.$$

Remark 4.3.1. In the time homogeneous case, $j^l(t) \equiv j^l$, $\Phi_j^{\mathbf{c}} = 1 / \prod_{l=1}^A (j^l)^{c^l}$.

We can now use the characteristics to characterize the reciprocal class.

Theorem 4.3.1. Let j be as in (4.2) and $P \in \mathcal{P}(\Omega)$. Then P belongs to the reciprocal class $\mathfrak{R}(R)$ if and only if:

i) For all $u \in \mathcal{U}$ and all $F \in \mathcal{B}(\Omega)$,

$$P(F \circ \pi_u) = P\left(F \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right). \quad (4.14)$$

ii) For every $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$ and every $f \in \mathcal{B}^{\#}(\mathbb{Z}^A)$, the following identity holds:

$$\rho(f \circ \theta_{\mathbf{c}}) = \Phi_j^{\mathbf{c}} \rho(f G_{\mathbf{c}}), \quad (4.15)$$

where $\rho := P \circ \mathbf{N}_1^{-1} \in \mathcal{P}(\mathbb{N}^A)$ is the law of \mathbf{N}_1 under P .

Remark 4.3.2. Note that identities similar to (4.15) hold for any $t \in]0, 1]$, $P \in \mathfrak{R}(R)$ satisfies (we assume a time homogeneous intensity, for simplicity):

$$P(f \circ \theta_{\mathbf{c}}(\mathbf{N}_t)) = \Phi_j^{\mathbf{c}} (1 - t)^{-|\mathbf{c}|} P((f G_{\mathbf{c}})(\mathbf{N}_t)), \quad \forall f \in \mathcal{B}^{\#}(\mathbb{Z}^A), 0 < t \leq 1, \quad (4.16)$$

where $|\mathbf{c}| := \sum_{l=1}^A c^l$. However, the identities (4.16) do not contain enough information to characterize the reciprocal class as the arc characteristics do not appear.

Proof. (\Rightarrow) Let $P \in \mathfrak{R}(R)$. An application of Proposition 2.2.2 tells that $P \ll R$, and $h := \frac{dP}{dR}$ is $\sigma(X_0, X_1)$ -measurable. Consider now $u \in \mathcal{U}$. By definition of u , for any l , $N_1^l \circ \pi_u = N_1^l$, so that $(X_0, X_1) \circ \pi_u = (X_0, X_1)$, R -a.s..

We then consider $F \in \mathcal{B}(\Omega)$ and apply Proposition 4.2.1 under the measure R , which leads to:

$$\begin{aligned} P(F \circ \pi_u) &= R\left((F \circ \pi_u)h(X_0, X_1)\right) = R\left((Fh(X_0, X_1)) \circ \pi_u\right) \\ &= R\left(Fh(X_0, X_1) \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right) \\ &= P\left(F \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right). \end{aligned}$$

In a similar way, if $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$, we have that $\mathbf{A}(\theta_{\mathbf{c}}\mathbf{N}_1) = \mathbf{A}\mathbf{N}_1$. We observe that $R(\mathbf{N}_1 \in \cdot | X_0 = x) = \mathfrak{p}_\lambda$, where

$$\lambda^l := \int_0^1 j^l(t) dt. \quad (4.17)$$

For $f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A)$ and $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$ we use Proposition 4.2.4, observing that \mathbf{N}_1 has law \mathfrak{p}_λ and is independent of X_0 , to obtain

$$\begin{aligned} \rho(f \circ \theta_{\mathbf{c}}) &= P(f \circ \theta_{\mathbf{c}}(\mathbf{N}_1)) \\ &= R(h(X_0, X_1) f \circ \theta_{\mathbf{c}} \circ \mathbf{N}_1) \\ &= R(h(X_0, X_0 + \mathbf{A}(\theta_{\mathbf{c}}\mathbf{N}_1)) f \circ \theta_{\mathbf{c}} \circ \mathbf{N}_1) \\ &= R(R^{X_0}(h(X_0, X_0 + \mathbf{A}(\theta_{\mathbf{c}}\mathbf{N}_1)) f \circ \theta_{\mathbf{c}} \circ \mathbf{N}_1)) \\ &= \Phi_j^{\mathbf{c}} R(h(X_0, X_1)(f G_{\mathbf{c}}) \circ \mathbf{N}_1) = \Phi_j^{\mathbf{c}} \rho(f G_{\mathbf{c}}) \end{aligned}$$

and *ii*) is now proven.

(\Leftarrow) We will show that $P = h(X_0, X_1)R$, which is equivalent to $P \in \mathfrak{A}(R)$ by Proposition 2.2.2. We divide the proof in three steps. In a first step, we prove the absolute continuity. In a second step we prove that the density is $\sigma(X_0, \mathbf{N}_1)$ -measurable and in a third one we prove that this density is indeed $\sigma(X_0, X_1)$ -measurable.

Step 1: Absolute continuity

We first observe that it is sufficient to prove that

$$P(\cdot | \mathbf{N}_1 = \mathbf{n}) \ll R(\cdot | \mathbf{N}_1 = \mathbf{n}) \text{ for all } \mathbf{n} \text{ such that } P(\mathbf{N}_1 = \mathbf{n}) > 0.$$

To this aim, we use an approximation argument.

Let us fix \mathbf{n} and construct a discrete (dyadic) approximation of the jump times. For $m \geq \max_{l=1, \dots, A} \log_2(n^l) + 1 := \bar{m}$, \mathcal{D}^m is composed by A ordered sequences of dyadic numbers, the l -th sequence having length n^l :

$$\mathcal{D}^m := \{\mathbf{k} = (k_i^l)_{l \leq A, i \leq n^l} : k_i^l \in 2^{-m}\mathbb{N}, 0 < k_{i-1}^l < k_i^l \leq 1, \forall l \leq A, \forall i \leq n^l\}$$

For $\mathbf{k} \in \mathcal{D}^m$ we define the subset of trajectories whose jump times are localized around \mathbf{k} :

$$O_{\mathbf{k}}^m = \{\mathbf{N}_1 = \mathbf{n}\} \cap \bigcap_{\substack{l \leq A \\ i \leq n^l}} \{0 \leq k_i^l - T_i^l < 2^{-m}\} \quad (4.18)$$

Moreover, as a final preparatory step, we observe for every $m \geq \bar{m}$, $\mathbf{k}, \mathbf{k}' \in \mathcal{D}^m$, one can easily construct $u \in \mathcal{U}$ such that:

$$u(l, t) = t + k_i^l - k_i^l, \quad \forall l \leq A, i \leq n^l \text{ and } t \text{ s.t. } 0 \leq k_i^l - t < 2^{-m}. \quad (4.19)$$

We can observe that (4.19) ensures $\dot{u}(l, T_i^l) = 1$ over $O_{\mathbf{k}}^m$, and that $O_{\mathbf{k}'}^m = \pi_u^{-1}(O_{\mathbf{k}}^m)$. We choose $F = \mathbf{1}_{O_{\mathbf{k}'}^m} \mathbf{1}_{\{\mathbf{N}_1 = \mathbf{n}\}} / P(\mathbf{N}_1 = \mathbf{n})$ and u as in (4.19) and apply (4.14) to obtain:

$$\begin{aligned} P(O_{\mathbf{k}'}^m | \mathbf{N}_1 = \mathbf{n}) &= P(\{\omega : \pi_u(\omega) \in O_{\mathbf{k}}^m\} | \mathbf{N}_1 = \mathbf{n}) \\ &= P\left(\mathbf{1}_{O_{\mathbf{k}}^m} \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right) \middle| \mathbf{N}_1 = \mathbf{n}\right) \\ &\geq C P(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}), \end{aligned}$$

where

$$C := \left(\inf_{s, t \in [0, 1], l \leq A} \Xi_j(l, s, t)\right)^{\sum_l \mathbf{n}^l} > 0 \quad (4.20)$$

since j satisfies Assumption 2.3.1. With a simple covering argument we obtain, for all $m \geq \bar{m}$ and $\mathbf{k} \in \mathcal{D}^m$,

$$\begin{aligned} \#\mathcal{D}^m \min\left\{1, \frac{1}{C}\right\} P(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}) \\ \leq P(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}) + \sum_{\substack{\mathbf{k}' \in \mathcal{D}^m \\ \mathbf{k}' \neq \mathbf{k}}} P(O_{\mathbf{k}'}^m | \mathbf{N}_1 = \mathbf{n}) \leq 1. \end{aligned}$$

It can be shown with a direct computation that $\frac{1}{|\mathcal{D}^m|} \leq C' R(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n})$ for some $C' > 0$ uniformly in $m, \mathbf{k} \in \mathcal{D}^m$ (the proof is given separately in Lemma 4.3.1). Therefore there exists a constant $C'' > 0$ such that:

$$P(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}) \leq C'' R(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}), \quad \forall m \geq \bar{m}, \mathbf{k} \in \mathcal{D}^m.$$

With a standard approximation argument we can extend the last bound to any measurable set. This completes the proof of the absolute continuity.

Step 2: The density $H := \frac{dP}{dR}$ is invariant under time change.

We show that, for any $u \in \mathcal{U}$, H is π_u -invariant, i.e. $H \circ \pi_u = H R - a.s.$. By definition, π_u is $R - a.s.$ invertible. Applying the identity (4.4) under R

and point i) of the hypothesis, we obtain, for any $F \in \mathcal{B}(\Omega)$:

$$\begin{aligned}
R(FH \circ \pi_u) &= R((F \circ \pi_u^{-1} H) \circ \pi_u) \\
&= R\left(F \circ \pi_u^{-1} H \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right) \\
&= P\left(F \circ \pi_u^{-1} \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right) \\
&= P(F) = R(FH)
\end{aligned}$$

which gives us the desired invariance, since F is arbitrary.

We claim that this implies that H is $\sigma(X_0, \mathbf{N}_1)$ -measurable, i.e. that there exists a function $h : \mathcal{X} \times \mathbb{N}^A \rightarrow \mathbb{R}^+$ such that

$$H = \frac{dP}{dR} = \frac{dP \circ (X_0, \mathbf{N}_1)^{-1}}{dR \circ (X_0, \mathbf{N}_1)^{-1}} = h(X_0, \mathbf{N}_1) \quad R - a.s.$$

This is true since, given any two $\omega, \omega' \in \Omega$ with the same initial state and the same number of jumps of each type, one can construct $u \in \mathcal{U}$ such that $\omega' = \pi_u(\omega)$.

Step 3: The density H is invariant under shifts in $\ker_{\mathbb{Z}}(\mathbf{A})$.

Let us recall that $R(\mathbf{N}_1 \in \cdot | X_0 = x) = \mathfrak{p}_\lambda$, where λ is given by (4.17). Under our assumption we might apply Proposition 4.2.4 to $\mathfrak{p}_\lambda = R(\mathbf{N}_1 \in \cdot | X_0 = x)$ and $\rho^x = P(\mathbf{N}_1 \in \cdot | X_0 = x)$. We obtain that the conditional density $\frac{d\rho^x}{d\mathfrak{p}_\lambda}$ is \mathbf{AN}_1 -measurable $P_0 - a.s.$ and, by mixing over the initial condition, that $\frac{dP \circ (X_0, \mathbf{N}_1)^{-1}}{dR \circ (X_0, \mathbf{N}_1)^{-1}} = \frac{dP}{dR}$ is $\sigma(X_0, \mathbf{AN}_1) = \sigma(X_0, X_1)$ -measurable. \square

Lemma 4.3.1. *Let \mathcal{D}^m and R as before. Then there exists a constant C' such that for m large enough,*

$$C' R(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}) \geq \frac{1}{\#\mathcal{D}^m}$$

Proof. We want to prove that, for $\mathbf{n} \in \mathbb{N}^A$:

$$\frac{1}{\#\mathcal{D}^m} \leq C' R(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}), \quad \forall m \geq \max_{l \leq A} \log(n^l) + 1, \quad \mathbf{k} \in \mathcal{D}^m \quad (4.21)$$

We can first compute explicitly $\#\mathcal{D}^m$ with a simple combinatorial argument: each $\mathbf{k} \in \mathcal{D}^m$ is constructed by choosing n^l dyadic intervals, $l \leq A$, and ordering them. Therefore

$$\#\mathcal{D}^m = \prod_{l=1}^A \binom{2^m}{n^l}. \quad (4.22)$$

On the other hand, we observe that defining

$$\tilde{j}(t, z \rightarrow z + a^l) = 1 \quad \forall t \in [0, 1], z \in \mathcal{X}, l \leq A$$

then the corresponding random walk, \tilde{R} , is equivalent to R . Therefore, we can prove (4.21) replacing R with \tilde{R} . To do this, for each $\mathbf{k} \in \mathcal{D}^m$ we define the function:

$$\delta : \{1, \dots, 2^m\} \times \{1, \dots, A\} \longrightarrow \{0, 1\}$$

$$\delta(i, l) := \begin{cases} 1, & \text{if } i \in \{2^m \mathbf{k}_1^l, \dots, 2^m \mathbf{k}_{n^l}^l\} \\ 0, & \text{otherwise.} \end{cases}$$

Then, using the explicit distribution of \tilde{R} ,

$$\begin{aligned} & \tilde{R}(O_{\mathbf{k}}^m | \mathbf{N}_1 = \mathbf{n}) \\ &= \tilde{R} \left(\bigcap_{(i,r) \in \{1, \dots, 2^m\} \times \{1, \dots, A\}} \{N_{\frac{i}{2^m}}^l - N_{\frac{i}{2^m}}^l = \delta(i, r)\} | \mathbf{N}_1 = \mathbf{n} \right) \\ &= \exp(A) \exp(-2^{-m})^{2^m A} (2^{-m})^{(\sum_l n^l)} \prod_{l=1}^A n^l! = \prod_{l=1}^A 2^{-mn^l} n^l! \end{aligned}$$

It is now easy to see that there exists a constant $C_0 > 0$ such that:

$$\binom{2^m}{n^l} \geq C_0 \frac{2^{mn^l}}{n^l!}, \quad \forall l \leq A, m \geq \max_{l \leq A} \log(n^l) + 1, \mathbf{k} \in \mathcal{D}^m$$

from which the conclusion follows using (4.22). \square

4.3.2 Comparing random walks through characteristics

In what follows and in the next subsections, we consider jump rates as in (4.2) which are time-homogeneous. In this case there is no ambiguity in identifying j with the vector $(j^1(0), \dots, j^A(0)) \in \mathbb{R}_+^A$.

Accordingly, for any $\tilde{j} \in \mathbb{R}_+^A$ we denote by \tilde{R} a Markov random walk on $(\mathcal{X}, \rightarrow)$ whose intensity is given by:

$$\tilde{j}(t, z \rightarrow z + a^l) = \tilde{j}^l \quad \forall z \in \mathcal{X}, t \in [0, 1], r \leq A$$

We do not specify the initial distribution of \tilde{R} , since it does not play any role in what follows. We present in Theorem 4.3.2 a set of explicit necessary and sufficient conditions for two random walks R and \tilde{R} to have the

same bridges, or equivalently, to belong to the same reciprocal class. In Chapter 3, we proved that two counting processes have the same bridges if and only if their reciprocal characteristic coincide. We offer here a significant generalization of that result.

We denote by $\ker_{\mathbb{Z}}(\mathbf{A})^{\perp}$ the orthogonal complement of the affine hull of $\ker_{\mathbb{Z}}(\mathbf{A})$, and the logarithm of the vector $j \in \mathbb{R}_+^A$, denoted by $\log(j)$, has to be understood componentwise.

Theorem 4.3.2. *Let $j, \tilde{j} \in \mathbb{R}_+^A$ and $\ker_{\mathbb{Z}}^*(\mathbf{A})$ be a lattice basis of $\ker_{\mathbb{Z}}(\mathbf{A})$. We let \tilde{R} be the walk associated with \tilde{j} . The following assertions are equivalent:*

- i) $\tilde{R} \in \mathfrak{R}(R)$.*
- ii) For every $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$ the equality $\Phi_j^{\mathbf{c}} = \Phi_{\tilde{j}}^{\mathbf{c}}$ holds.*
- iii) There exists $v \in \ker_{\mathbb{Z}}(\mathbf{A})^{\perp}$ such that $\log(\tilde{j}) = \log(j) + v$.*

Proof. *i) \Rightarrow ii)* By applying (4.15) and the trivial fact that $\tilde{R} \in \mathcal{R}(\tilde{R})$, we have

$$\Phi_{\tilde{j}}^{\mathbf{c}} \tilde{R}(fG_{\mathbf{v}} \circ \mathbf{N}_1) = \tilde{R}(f \circ \theta_{\mathbf{c}} \circ \mathbf{N}_1) = \Phi_j^{\mathbf{c}} \tilde{R}(fG_{\mathbf{v}} \circ \mathbf{N}_1), \quad \forall f \in \mathcal{B}^{\#}(\mathbb{Z}^A), \quad (4.23)$$

and *ii)* follows.

ii) \Rightarrow i) Observe that since $\ker_{\mathbb{Z}}(\mathbf{A})^*$ is a lattice basis, any $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$ can be written as an integer combination of the elements of $\ker_{\mathbb{Z}}^*(\mathbf{A})$, i.e. $\mathbf{c} = \sum_{\mathbf{c}^* \in \ker_{\mathbb{Z}}^*(\mathbf{A})} z_{\mathbf{c}^*} \mathbf{c}^*$, $z_{\mathbf{c}^*} \in \mathbb{Z}$. Therefore all the cycle characteristics coincide since:

$$\Phi_j^{\mathbf{c}} = \prod_{\mathbf{c}^* \in \ker_{\mathbb{Z}}^*(\mathbf{A})} (\Phi_j^{\mathbf{c}^*})^{z_{\mathbf{c}^*}} = \prod_{\mathbf{c}^* \in \ker_{\mathbb{Z}}^*(\mathbf{A})} (\Phi_{\tilde{j}}^{\mathbf{c}^*})^{z_{\mathbf{c}^*}} = \Phi_{\tilde{j}}^{\mathbf{c}}, \quad \forall \mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A}). \quad (4.24)$$

With a similar argument as above one proves that the identity (4.15) is satisfied under \tilde{R} . The functional equation (4.14) is trivially satisfied by \tilde{R} because $\Xi_j \equiv \Xi_{\tilde{j}} = 1$. The conclusion follows by applying Theorem 4.3.1.

ii) \Leftrightarrow iii) We just observe that the equality $\Phi_j^{\mathbf{c}} = \Phi_{\tilde{j}}^{\mathbf{c}}$ is equivalent to:

$$\sum_{l=1}^A \log(j^l) c^l = \sum_{l=1}^A \log(\tilde{j}^l) c^l.$$

Since a lattice basis $\ker_{\mathbb{Z}}^*(\mathbf{A})$ of $\ker_{\mathbb{Z}}(\mathbf{A})$ is a linear basis of the affine hull of $\ker_{\mathbb{Z}}(\mathbf{A})$ *ii)* is equivalent to the fact that $\log(v)$ and $\log(\tilde{v})$ have the same projection onto $\ker_{\mathbb{Z}}(\mathbf{A})$, which is equivalent to *iii)*. \square

Example 4.3.1. Continuing on Example 4.2.2, two time-homogeneous Markov walks with jumps in $\mathbb{A} = \{-1, 1\}$ and rate $j = (j^-, j^+)$ resp. $\tilde{j} = (\tilde{j}^-, \tilde{j}^+)$ have the same bridges if and only if

$$j^- j^+ = \tilde{j}^- \tilde{j}^+.$$

Example 4.3.2. Let $\mathbb{A} = \{-1, 3\}$ and define two time-homogeneous Markov walks with jumps in \mathbb{A} and rate $j = (j^-, j^+)$ resp. $\tilde{j} = (\tilde{j}^-, \tilde{j}^+)$. They have the same bridges if and only if

$$(j^-)^3 j^+ = (\tilde{j}^-)^3 \tilde{j}^+.$$

Example 4.3.3. Let $\mathbb{A} = \{a^1, \dots, a^6\}$ be the vertices of an hexagon, see the Figure 2:

$$a^i = \left(\cos\left(\frac{2\pi}{6}(i-1)\right), \sin\left(\frac{2\pi}{6}(i-1)\right) \right) \in \mathbb{R}^2, \quad i = 1, \dots, 6. \quad (4.25)$$

Then a basis of $\ker_{\mathbb{Z}}(\mathbf{A})$ is:

$$\ker_{\mathbb{Z}}^*(\mathbf{A}) = \{\mathbf{e}_1 + \mathbf{e}_4, \mathbf{e}_2 + \mathbf{e}_5, \mathbf{e}_1 + \mathbf{e}_3 + \mathbf{e}_5, \mathbf{e}_2 + \mathbf{e}_4 + \mathbf{e}_6\}. \quad (4.26)$$

By Theorem 4.3.2, \tilde{R} with jump rates $(\tilde{j}^1, \dots, \tilde{j}^6)$ belongs to $\mathfrak{R}(R)$ if and only if

$$\begin{cases} j^1 j^4 = \tilde{j}^1 \tilde{j}^4, \\ j^2 j^5 = \tilde{j}^2 \tilde{j}^5, \\ j^1 j^3 j^5 = \tilde{j}^1 \tilde{j}^3 \tilde{j}^5, \\ j^2 j^4 j^6 = \tilde{j}^2 \tilde{j}^4 \tilde{j}^6 \end{cases}$$

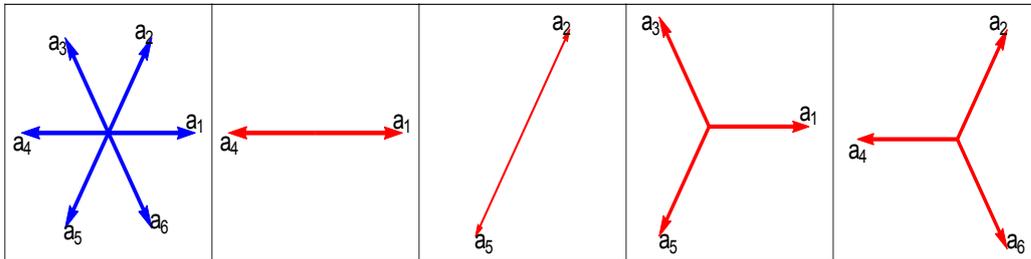


Figure 4.1: A representation of the vectors of \mathbb{A} and of the incidence vectors of $\ker_{\mathbb{Z}}^*(\mathbf{A})$

4.4 Lattices and their generating sets

The aim of this section is to reformulate Proposition 4.2.4 and Theorem 4.3.1 in terms of finitely many conditions. We recall, for example, that in Proposition 4.2.4, condition $\rho \in \mathcal{R}_{\mathbf{A}}(\mathfrak{p}_\lambda)$ was shown to be equivalent to the countably many equations in (4.13). The equations (4.13), for $\mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A})$, essentially tell us that, if $\mathbf{n} \in \mathbb{N}^A$ is such that $\mathbf{m} := \theta_{-\mathbf{c}}\mathbf{n}$ is also an element of \mathbb{N}^A , then $\rho(\mathbf{m}) = \rho(\mathbf{n})\mathfrak{p}_\lambda(\mathbf{m})/\mathfrak{p}_\lambda(\mathbf{n})$. We now show with a counterexample that the validity of this statement for \mathbf{c} varying in a basis of $\ker_{\mathbb{Z}}(\mathbf{A})$ is not enough to guarantee the validity in all $\ker_{\mathbb{Z}}(\mathbf{A})$. In the next paragraph, we will indeed reformulate this problem as a connectivity problem for a certain family of graphs, and propose a solution in this framework using *generating sets* of lattices.

Counterexample 4.4.1. : Let $\mathbb{A} = \{3, 4, 5\}$. Then

$$\ker_{\mathbb{Z}}(\mathbf{A}) = \{\mathbf{c} \in \mathbb{Z}^3 : 3c_1 + 4c_2 + 5c_3 = 0\}.$$

We define three vectors

$$f = (-3, 1, 1), \quad g = (1, -2, 1), \quad h = (2, 1, -2).$$

Note that $\{f, g, h\} \subseteq \ker_{\mathbb{Z}}(\mathbf{A})$. We also define

$$\mathbf{n}_f := (3, 0, 0), \quad \mathbf{n}_g := (0, 2, 0), \quad \mathbf{n}_h := (0, 0, 2).$$

Moreover, we observe that

$$\text{if, for some } \mathbf{c} \in \ker_{\mathbb{Z}}(\mathbf{A}), \theta_{\mathbf{c}}\mathbf{n}_f \in \mathbb{N}^3 \text{ then } \mathbf{c} = f. \quad (4.27)$$

This can be checked with a direct computation. The analogous statement also holds for g and h , i.e.

$$\theta_{\mathbf{c}}\mathbf{n}_g \in \mathbb{N}^3 \Rightarrow \mathbf{c} = g, \quad \theta_{\mathbf{c}}\mathbf{n}_h \in \mathbb{N}^3 \Rightarrow \mathbf{c} = h.$$

Let us now consider any basis $\ker_{\mathbb{Z}}^*(\mathbf{A})$ of $\ker_{\mathbb{Z}}(\mathbf{A})$. Since $\ker_{\mathbb{Z}}(\mathbf{A})$ is two dimensional, at least one vector, f or g or h , does not belong to $\ker_{\mathbb{Z}}^*(\mathbf{A})$. We assume w.l.o.g that $f \notin \ker_{\mathbb{Z}}^*(\mathbf{A})$. For any $0 < \varepsilon < 1$, $\lambda \in \mathbb{R}_+^3$, we define the probability measure $\rho \in \mathcal{P}(\mathbb{N}^3)$ as a mixture between the degenerate measure $\delta_{\mathbf{n}_f}$ and \mathfrak{p}_λ as follows:

$$\rho = \varepsilon\delta_{\mathbf{n}_f} + (1 - \varepsilon)\mathfrak{p}_\lambda. \quad (4.28)$$

Note that $\rho \notin \mathcal{R}_A(\mathfrak{p}_\lambda)$. Indeed any version of the density must be such that:

$$\frac{d\rho}{d\mathfrak{p}_\lambda}(\mathbf{n}_f) = \frac{\varepsilon}{\mathfrak{p}_\lambda(\mathbf{n}_f)} + (1 - \varepsilon), \quad \frac{d\rho}{d\mathfrak{p}_\lambda}(\theta_{\mathbf{c}_f} \mathbf{n}_f) = 1 - \varepsilon.$$

But, on the other hand, identity (4.13) is satisfied for any $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$. Let us pick any test function $f = \mathbf{1}_{\{\mathbf{z}=\bar{\mathbf{n}}\}}$, where $\bar{\mathbf{n}} \in \mathbb{N}^3$ and $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$. There are two possibilities:

- Either $\theta_{-\mathbf{c}}\bar{\mathbf{n}} \in \mathbb{Z}^3 \setminus \mathbb{N}^3$. In this case (4.13) is satisfied by ρ because both sides of the equality are zero, the left side because $\theta_{-\mathbf{c}}\bar{\mathbf{n}} \notin \mathbb{N}^A$, $\rho(\mathbb{N}^A) = 1$ and the right side because $G_{\mathbf{c}}(\bar{\mathbf{n}}) = 0$.
- Or $\theta_{-\mathbf{c}}\bar{\mathbf{n}} \in \mathbb{N}^A$. In this case, thanks to (4.27) and $f \notin \ker_{\mathbb{Z}}^*(\mathbf{A})$ we have $\bar{\mathbf{n}} \neq \mathbf{n}_f$ and $\theta_{-\mathbf{c}}\bar{\mathbf{n}} \neq \mathbf{n}_f$. Therefore, by (4.28),

$$\begin{aligned} \rho(\mathbf{1}_{\{\theta_{\mathbf{c}}\mathbf{z}=\bar{\mathbf{n}}\}}) &= \frac{\rho(\theta_{-\mathbf{c}}\bar{\mathbf{n}})}{\rho(\bar{\mathbf{n}})} \rho(\mathbf{1}_{\{\mathbf{z}=\bar{\mathbf{n}}\}}) = \frac{\mathfrak{p}_\lambda(\theta_{-\mathbf{c}}\bar{\mathbf{n}})}{\mathfrak{p}_\lambda(\bar{\mathbf{n}})} \rho(\mathbf{1}_{\{\mathbf{z}=\bar{\mathbf{n}}\}}) \\ &= \lambda^{-\mathbf{c}} \rho(\mathbf{1}_{\{\mathbf{z}=\bar{\mathbf{n}}\}} G_{\mathbf{c}}(\mathbf{z})) \end{aligned}$$

which is equivalent to (4.13).

We thus obtain an example of a set \mathcal{A} such that, for any $\lambda \in \mathbb{R}_+^3$ and any basis $\ker_{\mathbb{Z}}^*(\mathbf{A})$ of $\ker_{\mathbb{Z}}(\mathbf{A})$ we can construct a probability measure ρ which satisfies (4.13) for $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$ and $f \in \mathcal{B}^\#(\mathbb{Z}^A)$ but does not belong to $\mathcal{R}_A(\mathfrak{p}_\lambda)$.

4.4.1 Some heuristics

A generating set for a lattice is a set which has some special properties which generalize those of a lattice basis. Indeed, one can prove that any generating set contains at least a lattice basis. These geometrical objects find their application in both the areas of discrete geometry and computational algebra, in connection with the study of Gröbner basis. We will adopt the viewpoint of discrete geometry, which is closer to our objectives. However, the problem of computing a generating set for a lattice can be translated into the problem of computing a set of generators for a *lattice ideal*, which is an ideal of polynomials associated with \mathcal{L} . We shall use this parallelism in some of the results we are going to prove. Leaving all precise statements to the next section, let us make some heuristic considerations to gain some intuition.

We start with a lattice \mathcal{L} in \mathbb{Z}^A , a set $V \subseteq \mathbb{Z}^A$ and $S \subseteq \mathcal{L}$. The lattice \mathcal{L} induces naturally a foliation of V . A leaf can be defined as an equivalence class by saying that \mathbf{m}, \mathbf{m}' belong to the same leaf \mathcal{F} if and only if $\mathbf{m} - \mathbf{m}' \in \mathcal{L}$. We can construct an undirected graph on each leaf \mathcal{F} by drawing an edge between \mathbf{m} and \mathbf{m}' whenever $\mathbf{m} - \mathbf{m}' \in S$. The edge set is denoted

by \mathcal{E} . It is then an interesting problem to study the connectivity properties of the graph $(\mathcal{F}, \mathcal{E})$, which depend from both S and V . We compare the cases when $V = \mathbb{Z}^A$ and when $V = \mathbb{N}^A$.

It is easy to see that if $V = \mathbb{Z}^A$, it suffices that S contains a lattice basis \mathcal{B} for $(\mathcal{F}, \mathcal{E})$ to be connected. To convince ourselves of this, let us just take \mathbf{m}, \mathbf{m}' in the same leaf \mathcal{F} . Then, since S contains a basis, there exist s_1, \dots, s_M in \mathcal{B} such that

$$\mathbf{m} = \mathbf{m}' + s_1 + \dots + s_M$$

Then the sequence

$$v_0 = \mathbf{m}', \quad v_{k+1} = v_k + s_k$$

clearly defines a path from \mathbf{m} to \mathbf{m}' in $(\mathbb{Z}^A, \mathcal{E})$.

However, nothing ensures that this path touches only vertices of \mathbb{N}^A , so it is not a priori clear that, if S contains a lattice basis, $(\mathbb{N}^A, \mathcal{E})$ is connected. It can be that some leaves are connected, and some are not. Indeed, this fact is what stands behind Counterexample 4.4.1. Let us revisit this counterexample from a purely geometrical viewpoint: it is an example of lattice \mathcal{L} such that any lattice basis fails to be a generating set.

Example 4.4.1. Let \mathcal{L} be the following lattice of \mathbb{Z}^3 :

$$\mathcal{L} = \{ \mathbf{n} \in \mathbb{Z}^3 : 3\mathbf{n}_1 + 4\mathbf{n}_2 + 5\mathbf{n}_3 = 0 \}$$

We consider the following three vectors in \mathbb{Z}^3 :

$$f = (-3, 1, 1), \quad g = (1, -2, 1), \quad h = (2, 1, -2)$$

and the following three vectors in \mathbb{N}^3 :

$$\mathbf{n}_f = (3, 0, 0), \quad \mathbf{n}_g = (0, 2, 0), \quad \mathbf{n}_h = (0, 0, 2)$$

We take $V = \mathbb{N}^3$. We denote the leaf associated to f by \mathcal{F}_f , and we adopt the same convention for $\mathcal{F}_g, \mathcal{F}_h$. It is easy to see with a direct computation that $\mathcal{F}_f = \{ \mathbf{n}_f, \mathbf{n}_f + f \}$. In the same way, it is seen that $\mathcal{F}_g = \{ \mathbf{n}_g, \mathbf{n}_g + g \}$ and $\mathcal{F}_h = \{ \mathbf{n}_h, \mathbf{n}_h + h \}$. Therefore for any $S \subseteq \mathcal{L}$, the graph $(\mathcal{F}_f, \mathcal{E})$ constructed as above is connected if and only if S contains either f or $-f$. Repeating the same argument, we have that $(\mathcal{F}_g, \mathcal{E})$ is connected if and only if S contains either g or $-g$, and $(\mathcal{F}_h, \mathcal{E})$ is connected if and only if S contains either h or $-h$. We have thus shown that to make each leaf connected we need to include in S at least three vectors. Since the dimension of \mathcal{L} is two, this means that no lattice basis suffices to make all leaves connected graphs.

Given $V \subseteq \mathbb{Z}^A$ A *generating set* for \mathcal{L} is a set $S \subseteq \mathcal{L}$ such that all graphs associated to the leaves are connected. Unlike lattice basis, their cardinality can grow exponentially with the dimension of the lattice. Actually, from the viewpoint of complexity, it is a NP-hard problem to compute them. The fastest algorithms available at the moment are illustrated in Section 11.4 of [28]. The most important theoretical result which we will need is the fact that each lattice admits a *finite* generating set. The next subsection is devoted to the proof of this. It is not an original result of this thesis. We rather follow chapters 10 and 11 of the book [28], highlighting the main ideas which are important for our purposes. The interested reader can find there for additional material about generating sets.

4.4.2 Existence of a finite generating set

We formalize the heuristic considerations above. Since we will only be concerned with the case $V = \mathbb{N}^A$, we will adapt all the definitions to this situation.

We first define the *foliation* that the lattice \mathcal{L} induces on \mathbb{N}^A :

Definition 4.4.1. Given $\mathbf{n} \in \mathbb{N}^A$, the leaf $\mathcal{F}_{\mathbf{n}}$ containing \mathbf{n} is:

$$\mathcal{F}_{\mathbf{n}} := \{\mathbf{n} + \mathcal{L}\} \cap \mathbb{N}^A. \quad (4.29)$$

Fix now $S \subseteq \mathcal{L}$. S induces a graph structure on each leaf (see e.g. [59]):

Definition 4.4.2. For $S \subseteq \mathcal{L}$ and $\mathbf{n} \in \mathbb{N}^A$ we define $(\mathcal{F}_{\mathbf{n}}, \mathcal{E}_{\mathbf{n}})$ as the undirected graph whose vertex set is $\mathcal{F}_{\mathbf{n}}$ and whose edge set is given by

$$\mathcal{E}_{\mathbf{n}} := \{(\mathbf{m}, \mathbf{m}') \in \mathcal{F}_{\mathbf{n}} \times \mathcal{F}_{\mathbf{n}} : \exists \mathbf{s} \in S \text{ with } \mathbf{m} - \mathbf{m}' = \pm \mathbf{s}\}.$$

We are now ready to introduce the notion of generating set for \mathcal{L} .

Definition 4.4.3. The set S is a generating set for \mathcal{L} if, for all $\mathbf{n} \in \mathbb{N}^A$, $(\mathcal{F}_{\mathbf{n}}, \mathcal{E}_{\mathbf{n}})$ is a connected graph.

The following theorem states that each lattice admits a finite generating set.

Theorem 4.4.1. Let $\mathcal{L} \subseteq \mathbb{Z}^A$. Then \mathcal{L} admits a finite generating set.

The main idea of the proof is to study a certain ideal of polynomials, which is the so-called *lattice ideal* associated with \mathcal{L} . To do this, let us recall some algebraic notions. For A fixed, we denote by $\mathbb{K}[x_1, \dots, x_A]$ the ring of polynomials in A indeterminates. If $\mathbf{v} \in \mathbb{N}^A$ we use the notation

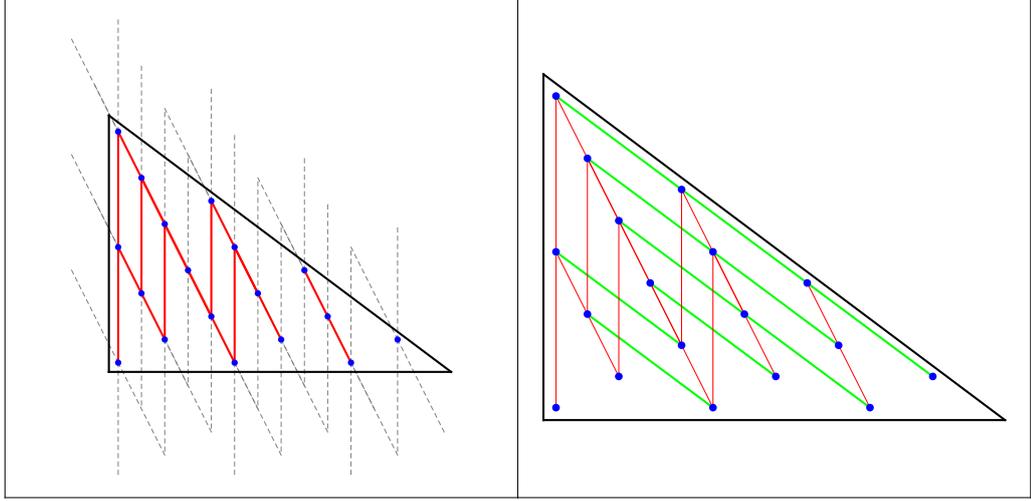


Figure 4.2: $\mathcal{A} = \{3, 4, 5\}$ and $S = \{(2, -4, 2), (0, -5, 4)\}$. Left: Projection on the x_1x_2 plane of $G := (\mathcal{F}_{\mathbf{n}}, \mathcal{E}_{\mathbf{n}})$ for $\mathbf{n} = (6, 1, 2)$. The red lines are the edges of G , while the dashed lines represent edges that are not in G because one endpoint does not belong to \mathbb{N}^3 . The graph $(\mathcal{F}_{\mathbf{n}}, \mathcal{E}_{\mathbf{n}})$ has three connected components. Right: Adding the vector $(4, -3, 0)$ to S turns G into a connected graph.

$$\mathbf{x}^{\mathbf{v}} = \prod_{l=1}^A x_l^{v_l}$$

If $F \subseteq \mathbb{K}[x_1, \dots, x_A]$, we denote by $\langle F \rangle \subseteq \mathbb{K}[x_1, \dots, x_A]$ the *ideal* generated by F :

$$\langle F \rangle := \left\{ f = \sum_{i=1}^k h_i f_i, \quad h_i \in \mathbb{K}[x_1, \dots, x_A], f_i \in F, k \in \mathbb{N} \right\}$$

Consider $S \subseteq \mathcal{L} \subseteq \mathbb{Z}^A$. The *lattice ideal* associated with it is the following:

$$I(S) = \langle \{\mathbf{x}^{\mathbf{u}^+} - \mathbf{x}^{\mathbf{u}^-}, \mathbf{u} \in S\} \rangle$$

where, as usual, for $u \in \mathbb{Z}^A$, $u_l^+ = \max\{u_l, 0\}$ and $u_l^- = -\min\{u_l, 0\}$. The proof uses two main ingredients. The first is a nice connection between generating sets and lattice ideals. A proof can be found in [28]. It is lemma 11.3.

Lemma 4.4.1. *Let \mathcal{L} be a lattice of \mathbb{Z}^A . A set $S \subseteq \mathcal{L}$ is a generating set for \mathcal{L} in the sense of Definition 4.4.3 if and only if $I(S) = I(\mathcal{L})$.*

The second ingredient is Hilbert's basis theorem.(see e.g. Corollary 10.4.20 in [28]).

Theorem 4.4.2. *Every ideal is finitely generated. If I is any ideal, then there exist a finite set M such that $I = \langle M \rangle$*

With these two tools at hand, the proof of Theorem 4.4.1 is relatively simple.

Proof. Let \mathcal{L} be a lattice and $I(\mathcal{L})$ be the associated lattice ideal. Then by Theorem 4.4.2 $I(\mathcal{L})$ is finitely generated. We call M a finite set such that $\langle M \rangle = I(\mathcal{L})$. Since $I(\mathcal{L})$ is generated by the set:

$$\{\mathbf{x}^{u^+} - \mathbf{x}^{u^-}, u \in \mathcal{L}\}$$

we can without loss of generality assume that M is contained in that set. That is, there exist a finite set S such that

$$M = \{\mathbf{x}^{u^+} - \mathbf{x}^{u^-}, u \in S\}$$

This implies that $\langle M \rangle = I(S) = I(\mathcal{L})$. But then by Lemma 4.4.1 the conclusion follows. \square

Generating sets of lattices are hard to compute explicitly. However, it may happen that a lattice basis is a generating set even though this is false in general. Below we give some sufficient condition for this to happen. These conditions are fulfilled by many interesting examples. In that case, the results of the section 4.5 take a very pleasant form, as they allow to work with lattice basis rather than generating sets. To the best of our knowledge, the next proposition is original.

Proposition 4.4.1. *Let \mathcal{B} be a basis of \mathcal{L} . Suppose that one of the following conditions holds:*

i) The basis \mathcal{B} contains an element $\bar{\mathbf{c}}$ such that each coordinate $\bar{c}^l, l = 1, \dots, A$ is positive.

ii) Each vector of the basis \mathcal{B} has non negative coordinates.

Then, the basis \mathcal{B} is a generating set.

Proof. i) Let $\mathbf{n} \in \mathbb{N}^A, \mathbf{m} \in \mathcal{F}_{\mathbf{n}}$. Since \mathcal{B} is a lattice basis there exists $\mathbf{c}_1, \dots, \mathbf{c}_K \subseteq (\mathcal{B} \cup -\mathcal{B})^K$ such that, if we define recursively

$$w_0 = \mathbf{n}, \quad w_k = w_{k-1} + \mathbf{c}_k$$

then we have that $w_K = \mathbf{m}$. Let us consider l large enough such that

$$l \min_{l=1, \dots, A} \bar{c} \geq \left| \min_{\substack{l=1, \dots, A \\ k=1, \dots, K}} w_k^l \right|. \quad (4.30)$$

We then consider the sequence $w'_k, k = 0, \dots, K+2l$ defined as follows:

$$w'_k = \begin{cases} w'_{k-1} + \bar{\mathbf{c}}, & \text{if } 1 \leq k \leq l \\ w'_{k-1} + \mathbf{c}_{k-l}, & \text{if } l+1 \leq k \leq K+l \\ w'_{k-1} - \bar{\mathbf{c}} & \text{if } K+l+1 \leq k \leq K+2l. \end{cases}$$

It is now easy to check, thanks to condition (4.30), that

$$w'_k \in \mathcal{F}_{\mathbf{n}} \quad \forall k \leq K+2l.$$

Since all the shifts involved in the definition of w'_k are associated to vectors in $\mathcal{B} \cup -\mathcal{B}$ we also have that (w'_{k-1}, w'_k) is an edge of $\mathcal{G}(\mathcal{F}_{\mathbf{n}}, \mathcal{B}), k \leq K+2l$.

Moreover we can check that

$$w'_{K+2l} = \mathbf{n} + l\bar{\mathbf{c}} + \sum_{k \leq K} \mathbf{c}_k - l\bar{\mathbf{c}} = \mathbf{m}$$

Therefore \mathbf{n} and \mathbf{m} are connected in $\mathcal{G}(\mathcal{F}_{\mathbf{n}}, \mathcal{B})$ and the conclusion follows since the choice of \mathbf{m} is arbitrary in $\mathcal{F}_{\mathbf{n}}$ and \mathbf{n} any point in \mathbb{N}^A .

- ii) Let $\mathbf{n} \in \mathbb{N}^A, \mathbf{m} \in \mathcal{F}$. Since \mathcal{B} is a lattice basis there exists $K < \infty$ and $\mathbf{c}_1, \dots, \mathbf{c}_K \subseteq (\mathcal{B} \cup -\mathcal{B})^K$ such that if we define recursively:

$$w_0 = \mathbf{n}, \quad w_k = w_{k-1} + \mathbf{c}_k \quad (4.31)$$

then we have that $w_K = \mathbf{m}$

Observe that w.l.o.g there exists K^+ s.t. $\mathbf{c}_k \in \mathcal{B}$ for all $k \leq K^+$ and $\mathbf{c}_k \in -\mathcal{B}, k \in \{K^+ + 1, \dots, K\}$. Applying the hypothesis one can check directly that $\{w_k\}_{0 \leq k \leq K}$ is a path which connects \mathbf{n} to \mathbf{m} in $\mathcal{G}(\mathcal{F}_{\mathbf{n}}, \mathcal{B})$. \square

4.5 Sharp characterizations of conditional laws

In this section we are given a probability law μ supported on \mathbb{N}^A , not necessarily a multivariate Poisson law.

In parallel with section 4.2.2 we study the set of probability measures which are equal to μ when conditioned to the sigma algebra $\sigma(\mathbf{A})$ generated by the random variable $\mathbf{A}\mathbf{n}$. It is the set $\mathcal{R}_{\mathbf{A}}(\mu)$ defined at (4.11).

We recall that, associated to any $v \in \mathbb{Z}^A$ there is a shift transformation:

$$\theta_{\mathbf{v}} : \mathbb{Z}^A \rightarrow \mathbb{Z}^A, \quad \mathbf{z} \mapsto \mathbf{z} + \mathbf{v} \quad (4.32)$$

Consider now the image of μ under $\theta_{\mathbf{v}}$ and its Radon Nykodim decomposition w.r.t. to μ :

$$\mu \circ \theta_{\mathbf{v}}^{-1} = \mu_{\mathbf{v}}^{ac} + \mu_{\mathbf{v}}^{sing}$$

and define $G_{\mathbf{v}}$ as the density of the absolutely continuous part:

$$\frac{d\mu_{\mathbf{v}}^{ac}}{\mu} := G_{\mathbf{v}}(\mathbf{n})$$

Then a generating set for the lattice $\ker_{\mathbb{Z}}(\mathbf{A})$ offers a very efficient characterization of $\mathcal{R}_{\mathbf{A}}(\mu)$.

Proposition 4.5.1. *Let $\mathbf{A} \in \mathbb{R}^{d \times A}$ be any matrix and the lattice $\ker_{\mathbb{Z}}(\mathbf{A})$ be defined as before by $\ker_{\mathbb{Z}}(\mathbf{A}) := \ker(\mathbf{A}) \cap \mathbb{Z}^A$. Assume that S is a generating set of $\ker_{\mathbb{Z}}(\mathbf{A})$ and let μ, ρ be two probability measures on \mathbb{N}^A . Suppose moreover that $\text{supp}\mu(\mathbf{n}) = \mathbb{N}^A$. Then $\rho \in \mathcal{R}_{\mathbf{A}}(\mu)$ if and only if:*

$$\forall \mathbf{v} \in S, \quad \rho(f \circ \theta_{\mathbf{v}}) = \rho(f G_{\mathbf{v}}) \quad \forall f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A), \quad (4.33)$$

where $G_{\mathbf{v}}$ is defined by (4.9).

Proof. (\Rightarrow) goes along the same lines of Proposition 4.2.4, since $S \subseteq \ker_{\mathbb{Z}}(\mathbf{A})$.

(\Leftarrow) Let $\mathbf{n}, \mathbf{m} \in \mathbb{N}^A$ be such that $\mathbf{A}\mathbf{n} = \mathbf{A}\mathbf{m}$ and assume that $\rho(\mathbf{n}) > 0$. Then $\mathbf{m} \in \mathcal{F}_{\mathbf{n}}$ (see (4.29)). Since S is a generating set for $\ker_{\mathbb{Z}}(\mathbf{A})$ there exists a path from \mathbf{m} to \mathbf{n} included in $\mathcal{G}(\mathcal{F}, S)$ i.e. there exists $\mathbf{v}_1, \dots, \mathbf{v}_K \in S$ such that, if we define recursively:

$$w_0 = \mathbf{m}, \quad w_k = \theta_{\mathbf{v}_k} w_{k-1},$$

then $w_k \in \mathbb{N}^A$ for all k and $w_K = \mathbf{n}$. We can choose $f^k = \mathbf{1}_{\{\mathbf{z}=w_k\}}$ and apply (4.33) for $\mathbf{v} = \mathbf{v}_k$:

$$\rho(w_{k-1}) = \frac{\mu(w_{k-1})}{\mu(w_k)} \rho(w_k)$$

which, since μ is a positive probability on \mathbb{N}^A , offers an inductive proof that $\rho(w_k) > 0$. Therefore one obtains

$$\frac{\rho(\mathbf{m})}{\rho(\mathbf{n})} = \prod_{k=1}^K \frac{\rho(w_{k-1})}{\rho(w_k)} = \prod_{k=1}^K \frac{\mu(w_{k-1})}{\mu(w_k)} = \frac{\mu(\mathbf{m})}{\mu(\mathbf{n})}$$

which is equivalent to $d\rho/d\mu(\mathbf{n}) = d\rho/d\mu(\mathbf{m})$, which completes the proof, as we have shown that we can choose an $\mathbf{A}\mathbf{n}$ measurable version of the density. \square

4.5.1 Refining the main theorem

As consequence of Proposition 4.5.1, by choosing μ to be a Poisson multivariate law, we obtain the following Proposition, which improves Proposition 4.2.4.

Proposition 4.5.2. *Let $\rho \in \mathcal{P}(\mathbb{N}^A)$ and S be a generating set of $\ker_{\mathbb{Z}}(\mathbf{A})$ defined by (4.12). Then $\rho \in \mathcal{R}_{\mathbf{A}}(\mathfrak{p}_{\lambda})$ if and only if*

$$\forall \mathbf{v} \in S, \quad \rho(f \circ \theta_{\mathbf{v}}) = \frac{1}{\lambda^{\mathbf{v}}} \rho(f G_{\mathbf{v}}), \quad \forall f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A), \quad (4.34)$$

where $G_{\mathbf{v}}$ is defined in (4.9).

This brings immediately to a refinement of the main result.

Theorem 4.5.1. *Let j be defined by (4.2). $P \in \mathcal{P}(\Omega)$ belongs to the reciprocal class $\mathfrak{R}(R)$ if and only if*

i) *For all $u \in \mathcal{U}$ and all $F \in \mathcal{B}(\Omega)$,*

$$P(F \circ \pi_u) = P\left(F \exp\left(\sum_{l=1}^A \int_0^1 \log \Xi_j(l, t, u(l, t)) \dot{u}(l, t) dN_t^l\right)\right). \quad (4.35)$$

ii) *There exists a generating set $S \subseteq \ker_{\mathbb{Z}}(\mathbf{A})$ such that for every $\mathbf{c} \in S$ and every $f \in \mathcal{B}^{\sharp}(\mathbb{Z}^A)$, the following identity holds:*

$$\rho(f \circ \theta_{\mathbf{c}}) = \Phi_j^{\mathbf{c}} \rho(f G_{\mathbf{c}}), \quad (4.36)$$

where $\rho := P \circ \mathbf{N}_1^{-1} \in \mathcal{P}(\mathbb{N}^A)$ is the law of \mathbf{N}_1 under P .

4.6 Short-time asymptotics of cycles

In this section we interpret the cycle characteristics via short time expansions of the bridges of the reference walk. To keep the notation simple, consider jump intensities j which are time-homogeneous. However, all

the results carry over with minor changes to the time-inhomogeneous case. We also assume condition *ii*) of Proposition 4.4.1 holds:

$$\ker_{\mathbb{Z}}(\mathbf{A}) \text{ admits a lattice basis } \ker_{\mathbb{Z}}^*(\mathbf{A}) \text{ included in } \mathbb{N}^A. \quad (4.37)$$

A lattice basis satisfying (4.37) is a generating set for $\ker_{\mathbb{Z}}(\mathbf{A})$, thanks to Proposition 4.4.1. Therefore it is sufficient to interpret the characteristics Φ_j^c for $c \in \ker_{\mathbb{Z}}^*(\mathbf{A})$.

Assumption (4.37) is not only natural in view of the interpretation we will give in Proposition 4.6.1 but it is satisfied in many interesting situations. One can prove that this is the case when $\mathbb{A} \subseteq \mathbb{Z}$ and \mathbb{A} contains at least one negative and one positive jump.

Assumption (4.37) also holds in several situations when $d > 1$, e.g. in the setting of Example 4.3.3.

In the context of diffusions, various physical interpretation of the reciprocal characteristics have been given, mainly based on analogies with Stochastic Mechanics, see [24], [51], [77] and [78].

We propose here a different interpretation as infinitesimal characteristics, based on the short-time expansions for the probability that the process makes a cycle around its current state.

We recall that, as it was defined in Chapter 2, a *cycle* of $(\mathcal{X}, \rightarrow)$ is a sequence $(x_k)_{k=0}^{|\gamma|} := \gamma$ such that $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_{|\gamma|} = x_0$. In this section (and only here) we make the additional assumption that $x_0 = 0$.

To any cycle γ we can associate an element $\mathbf{N}(\gamma) \in \ker_{\mathbb{Z}}(\mathbf{A}) \cap \mathbb{N}^A$ by counting how many times each jump occurred in it, thus neglecting their ordering:

$$\forall 1 \leq l \leq A \quad \mathbf{N}(\gamma)^l := \#\{k \leq |\gamma| : x_k - x_{k-1} = a^l\}. \quad (4.38)$$

N may be seen as the *skeleton* of the cycle. We will develop this viewpoint in Chapter 5. Note that, for a given $c \in \ker_{\mathbb{Z}}(\mathbf{A})$, we can construct a cycle γ such that $\mathbf{N}(\gamma) = c$ if and only if $c \in \mathbb{N}^A$. Therefore, under assumption (4.37), $\mathbf{N}^{-1}(c)$ is non empty for any $c \in \ker_{\mathbb{Z}}^*(\mathbf{A})$.

Definition 4.6.1. We define the *trace* $\gamma_\varepsilon(\omega)$ of a path $\omega \in \Omega$ as the ordered sequence formed by the displacements from the initial position up to time ε :

$$\Upsilon_\varepsilon(\omega) = (0, X_{T_1} - X_0, \dots, X_{T_{|\mathbf{N}|\varepsilon}} - X_0).$$

The subset of paths whose trace coincides with a given cycle γ over a small time interval $[0, \varepsilon]$ is denoted by

$$L_\varepsilon^\gamma := \{\omega : \Upsilon_\varepsilon(\omega) = \gamma\}.$$

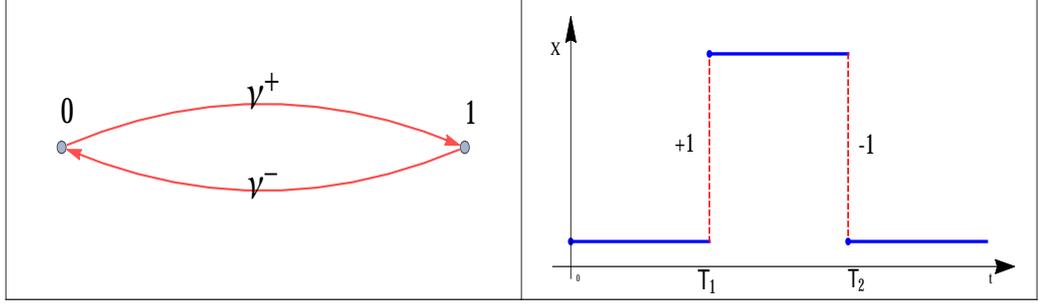


Figure 4.3: Here $\mathcal{A} = \{-1, 1\}$ and $\ker_{\mathbb{Z}}(\mathbf{A}) = (1, 1)\mathbb{Z}$. Left: A representation of the cycle $\gamma = \{0, 1, 0\}$ satisfying $\mathbf{N}(\gamma) = (1, 1)$. Right: A typical path in L_{ε}^{γ} . The probability of L_{ε}^{γ} is equivalent to $\frac{1}{2}(\nu^+ \nu^-)\varepsilon^2$ over the whole reciprocal class, as $\varepsilon \rightarrow 0$.

Finally, we introduce the usual time-shift operator on the canonical space:

$$\tau_t : \mathbb{D}([0, 1], \mathbb{R}^d) \longrightarrow \mathbb{D}([0, 1 - t], \mathbb{R}^d), \quad \tau_t(\omega)_s = \omega_{t+s}, \forall 0 \leq s \leq 1 - t.$$

The following short-time expansion holds under the reference walk.

Proposition 4.6.1. *Let j be a time-homogeneous intensity, $x, x_0 \in \mathcal{X}$. Then for any time $t \geq 0$, $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$ and any cycle γ with $\mathbf{N}(\gamma) = \mathbf{c}$, we have:*

$$R^{x_0}(\tau_t(X) \in L_{\varepsilon}^{\gamma} | X_t = x) = \frac{1}{\Phi_j^{\mathbf{c}} |\mathbf{c}|!} \varepsilon^{|\mathbf{c}|} + o(\varepsilon^{|\mathbf{c}|}) \quad \text{as } \varepsilon \rightarrow 0$$

where $|\mathbf{c}| = \sum_{l=1}^A c^l (= |\gamma|)$.

Proof. First observe that w.l.o.g. we can assume $t = 0$, the general result following from the Markov property of R^{x_0} . Recall that \bar{j} is the total jump rate $\sum_{l=1}^A j^l$. Moreover, we denote by $l(k)$ the unique element of $\{1, \dots, A\}$ such that $X_{T_k} - X_{T_{k-1}} = a^{l(k)}$. This variable is used to identify the arc that the walk uses to jump at T_k . With an elementary computation based on

the explicit distribution of R^{x_0} :

$$\begin{aligned}
R^x(L_\varepsilon^\gamma) &= R^x\left(\{|N|_\varepsilon = |\mathbf{c}|\} \cap \bigcap_{k=1}^{|\mathbf{c}|} \{X_{T_k} - X_{T_{k-1}} = a^{l(k)}\}\right) \\
&= \exp(-\varepsilon \bar{j}) \frac{(\varepsilon \bar{j})^{|\mathbf{c}|}}{|\mathbf{c}|!} \prod_{k=1}^{|\mathbf{c}|} \frac{j^{l(k)}}{\bar{j}} = \exp(-\varepsilon \bar{j}) \varepsilon^{|\mathbf{c}|} \prod_{l=1}^A (j^l)^{\#\{k:l(k)=l\}} \\
&= \exp(-\varepsilon \bar{j}) \frac{\varepsilon^{|\mathbf{c}|}}{|\mathbf{c}|!} \prod_{l=1}^A (j^l)^{n^l} = \exp(-\varepsilon \bar{j}) \frac{1}{\Phi_j^{\mathbf{c}} |\mathbf{c}|!} \varepsilon^{|\mathbf{c}|}
\end{aligned}$$

from which the conclusion follows. \square

Even more interesting, the same time-asymptotics holds under any $P \in \mathfrak{R}(R)$ and in particular under any bridge R^{xy} .

Theorem 4.6.1. *Let j be a time-homogeneous intensity and $P \in \mathfrak{R}(R)$. Then for any time $t \geq 0$, $\mathbf{c} \in \ker_{\mathbb{Z}}^*(\mathbf{A})$ and any cycle γ with $\mathbf{N}(\gamma) = \mathbf{c}$, we have:*

$$P - a.s. \quad P\left(\tau_t(X) \in L_\varepsilon^\gamma \middle| X_t\right) = \frac{1}{\Phi_j^{\mathbf{c}} |\mathbf{c}|!} \varepsilon^{|\mathbf{c}|} + o(\varepsilon^{|\mathbf{c}|}) \quad a.s. \varepsilon \rightarrow 0$$

That is, in a very short time interval, P goes around the cycle γ with the same probability as R .

Proof. Assume that $P \in \mathfrak{R}(R)$. Observe that w.l.o.g we can assume that $P_0 = \delta_{x_0}$ for some $x_0 \in \mathbb{R}^d$, the general result following by mixing over the initial condition. Then by Proposition 2.2.2, $dP/dR^{x_0} = h(X_1)$. We first show the identity:

$$R^{x_0}\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} h(X_1) \middle| X_t\right) = P\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} \middle| X_t\right) R^{X_t}\left(h(X_{1-t})\right). \quad (4.39)$$

Indeed, let us take any test function of the form $\mathbf{1}_{\{X_t \in A\}}$. We have:

$$\begin{aligned}
R^{x_0}\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} h(X_1) \mathbf{1}_{\{X_t \in A\}}\right) &= P\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} \mathbf{1}_{\{X_t \in A\}}\right) \\
&= P\left(P\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} \middle| X_t\right) \mathbf{1}_{\{X_t \in A\}}\right) \\
&= R^{x_0}\left(P\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} \middle| X_t\right) h(X_1) \mathbf{1}_{\{X_t \in A\}}\right) \\
&= R^{x_0}\left(P\left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} \middle| X_t\right) R^{x_0}(h(X_1) \middle| X_t) \mathbf{1}_{\{X_t \in A\}}\right)
\end{aligned}$$

from which (4.39) follows. Consider now the left hand side of (4.39). We have, by applying the Markov property and the fact that γ is a cycle:

$$\begin{aligned}
R^{x_0} \left(h(X_1) \mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} | X_t \right) &= R^{x_0} \left(R^{x_0} (h(X_1) | \mathcal{F}_{[t, t+\varepsilon]}) \mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} | X_t \right) \\
&= R^{x_0} \left(R^{X_{t+\varepsilon}} (h(X_{1-(t+\varepsilon)})) \mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} | X_t \right) \\
&= R^{x_0} \left(R^{X_t} (h(X_{1-(t+\varepsilon)})) \mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} | X_t \right) \\
&= R^{x_0} \left(\mathbf{1}_{\{\tau_t(X) \in L_\varepsilon^\gamma\}} | X_t \right) R^{X_t} \left(h(X_{1-(t+\varepsilon)}) \right).
\end{aligned}$$

Applying (4.39) and Proposition 4.6.1 and the continuity of

$$(\omega, t, \cdot) \mapsto R^{X_t}(h(X_{1-\cdot}))$$

we obtain:

$$\frac{1}{\Phi_j^c |\mathbf{c}|!} R^{X_t}(h(X_{1-t})) = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-|\mathbf{c}|} P(\mathbf{1}_{\{\tau_t \in L_\varepsilon^\gamma\}} | X_t) R^{X_t}(h(X_{1-t})) \quad (4.40)$$

We observe that $R^{X_t}(h(X_{1-t})) = dP_t/d(R^{x_0})_t$ and therefore it is strictly positive $P - a.s.$. This allows us to divide on both sides by $R^{X_t}(h(X_{1-t}))$ and the conclusion follows. \square

We have thus shown that each element of the reciprocal class has the same probability to *spin around* its current state in a very short time interval.

Remark 4.6.1. *In the statement of Proposition 4.6.1 we could have replaced X_t with $X_{[0,t]}$, i.e. the following asymptotics holds true:*

$$P(\tau_t(X) \in L_\varepsilon^\gamma | X_{[0,t]}) = \frac{1}{\Phi_j^c |\mathbf{c}|!} \varepsilon^{|\mathbf{c}|} + o(\varepsilon^{|\mathbf{c}|}) \quad \text{as } \varepsilon \rightarrow 0.$$

Wrapping up the content of this last section, we have seen that the cycle characteristic have a deep probabilistic interpretation. In Chapter 6 we largely extend the results of this section by showing that i) they hold on a general graph ii) the arc characteristics also have a deep interpretation and iii) they provide a characterization of $\mathfrak{R}(R)$.

4.7 Characteristics and concentration of measure

In this section we look at a very specific class of models and derive concentration of measure inequalities based on the reciprocal characteristics for some quantities of interest. The results we obtain point towards a more general relation between conditioning and concentration of measure which would be interesting to study in more generality. Some natural possible generalizations are briefly discussed in the last chapter of the thesis. The concentration rates we obtain are situated in between Poissonian and Gaussian concentration and are sharp at the leading order, and show the right dependence on the reciprocal characteristics in the exponential correction terms. At the moment of writing this thesis, it seems that there is no functional inequality designed to produce systematically such rates, and therefore we need to argue differently. In doing this, we show some improvements on the concentration bounds of the Poisson distributions, by carefully repeating Herbst's argument on an inequality introduced by Dai Pra Paganoni and Posta in [26].

4.7.1 A simple question

In this section we take $\mathbb{A} := \{-1, k\}$, where k is a positive integer. The lattice $\ker_{\mathbb{Z}}(\mathbf{A})$ is then spanned by the vector $(k, 1)$.

We consider time homogeneous intensities for R :³

$$j(t, z \rightarrow z + k) = j_k, \quad j(t, z \rightarrow z - 1) = j_{-1} \quad \forall z \in \mathbb{Z}, t \in [0, 1]$$

In words, we are looking at random walks on \mathbb{Z} that can only jump 1 down or k up at constant intensity. The associated reciprocal characteristic is $j_{-1}^k j_k$. For simplicity, we omit the superscript c and denote it simply Φ_j .

Let us consider the bridge between 0 and 0 of the reference walks, R^{00} . Then we ask a rather simple and natural question:

How many jumps of size k does the bridge R^{00} make?

Under the non-pinned reference walk R , the distribution of the jumps of size k simply follows a Poisson distribution of mean j_k . Under the bridge R^{00} we have no longer a Poisson distribution. In particular, explicit computations for the Laplace transform are not available anymore. However, using the results of the previous sections we have the following characterization of this law.

³We label the intensities j_{-1} and j_k to make our presentation more clear. In the notation of the former sections we should have used j_1 for j_{-1} and j_2 for j_k

Corollary 4.7.1. *Let*

$$\rho(n) := R^{00}(\#\{t : X_t - X_{t-} = k\}) \quad (4.41)$$

Then ρ is characterized by the following equation:

$$\Phi_j \rho(f(n+1)) = \rho\left(f(n)n \prod_{i=0}^{k-1} (n-i)\right) \quad \forall f \in \mathcal{B}(\mathbb{N}) \quad (4.42)$$

Remark 4.7.1. *Another definition of ρ is the following one:*

$$\rho(n) = \frac{1}{Z} \mathbf{p}_{\Phi_j}(kn, n) \quad (4.43)$$

where \mathbf{p}_{Φ_j} is a two dimensional Poisson distribution whose components are independent and have mean $\Phi_j^{1/(k+1)}$ and Z a normalization constant. This can be seen by either checking directly that the right hand side of (4.43) satisfies the equation (4.41) or using Theorem (4.5.1). Note that \mathbf{p}_{Φ_j} is not the joint law of the variables that count the number of jumps of each size under the reference dynamics (such variable is \mathbb{N}_1 in the notation above). Indeed, under R the number of jumps of size k has a Poisson law with mean j_k and under \mathbf{p}_{Φ_j} it has mean $\Phi_j^{1/(k+1)}$. But the reciprocal characteristics associated to these two different laws coincide, and therefore we can use \mathbf{p}_{Φ_j} to describe the bridge R^{00} .

Proof. With a slight abuse of notation (see footnote in the previous page) we call N_{-1} the total number of jumps of size -1 over the time interval $[0, 1]$ and by N_k the total number of jumps of size k , denote by $\tilde{\rho}$ the distribution $R \circ (N_{-1}, N_k)^{-1} \in \mathcal{P}(\mathbb{N}^2)$. With an application of Theorem 4.5.1 we find immediately that that $\tilde{\rho}$ satisfies:

$$\Phi_j \tilde{\rho}(\tilde{f}(n_{-1} + k, n_k + 1)) = \tilde{\rho}\left(\tilde{f}(n_{-1}, n_k) n_k \prod_{i=0}^{k-1} (k n_{-1} - i)\right) \quad \forall \tilde{f} \in \mathcal{B}(\mathbb{N}^2) \quad (4.44)$$

But since we are under the law of the bridge:

$$\tilde{\rho}(k n_k - n_{-1} = 0) = 1$$

Substituting $n_{-1} = k n_k$ in (4.50) and considering \tilde{f} to depend only on n_k we get the conclusion. \square

Note that the formula (4.42) generalizes Chen's characterization of Poisson distribution. We recover it as a special case when the term multiplying

f on the right hand side of (4.42) is linear. But here, such term is a polynomial that can be of arbitrary degree.

Concentration inequalities provide very strong information on the laws of random variables. What is known is that to a linear coefficient corresponds a Poisson distribution, for which a concentration of measure inequality is known. If any concentration of measure corresponds to a non-linear coefficient, and the rate of concentration in terms of it, are the object of study this section. Let us just mention few reason why this question is interesting: one classical way of proving concentration of measure inequalities is to pass through a logarithmic Sobolev inequality of some kind. For discrete probabilities several inequalities exist, (see e.g. [5, 26]). However, they are designed to recover Poissonian concentration, while from solutions of (4.42) we expect a stronger concentration, due to the non-linearity, which is not implied by those inequalities. This is a significant difference with the Gaussian case. Pinning a two dimensional Gaussian vector to a linear subspace gives back another Gaussian vector. Therefore standard techniques to prove concentration of measure can be used. Here, we are pinning a two dimensional Poisson vector to the linear subspace $\{kn_k - n_{-1} = 0\}$ and the situation looks rather different. Another interesting aspect which explains why the standard techniques break down is the following: while changing the constant in the classical logarithmic Sobolev inequality on \mathbb{R}^d influences the *leading order* terms of the associated Gaussian concentration inequality, for both modified logarithmic Sobolev inequalities on a discrete space studied in [5, 26], changing the constant for which the inequality hold only influences the *exponential correction terms*. Therefore if one wants to get a stronger concentration at the leading order, something different has to be found. In the rest of the section, a function g is said to be $o(R)$ if:

$$\lim_{R \rightarrow +\infty} g(R)/R = 0$$

We derived the following result:

Theorem 4.7.1. *Let ρ be the unique solution of (4.50). Then there exist $C_0 > 0$ such that for all f which are 1-Lipschitz and for all $R > C_0$:*

$$\rho(f \geq \rho(f) + R) \leq \exp(-(k+1)R \log R + (\log(\Phi_j) + C_1)R + o(R))$$

The constant C_1 does not depend on Φ_j . C_0 might depend on it.

Remark 4.7.2. (i) *The size of the large jump drives the leading order in the concentration rate, while the reciprocal characteristic is responsible for the exponential correction term.*

- (ii) *The larger k , the more concentrated is the random variable. This is because to compensate a large jump a bridge has to make many small jumps, and this reduces the probability of many large jumps.*
- (iii) *It can be seen, that the leading order term is optimal and that the dependence on the reciprocal characteristic in the exponential terms seems also to be optimal. This is done in [19]*

The proof of this theorem is contained in the next two subsections.

4.7.2 Poissonian concentration revisited

A concentration result is obtained by [5](and proved in [63] with a different technique). They first prove a Modified Logarithmic Sobolev Inequality (MLSI for short) for the Poisson distribution and then use the Herbst's argument (outlined e.g. in chapter 5 of [43]). However, as pointed out by the same authors, the inequality is not optimal in the sense that it predicts heavier tails for the distribution of Lipschitz functionals than what is expected. Indeed, in their Proposition 10 they establish that if ρ is the Poisson distribution of parameter j , and $f : \mathbb{N} \rightarrow \mathbb{R}$ is a 1-Lipschitz function in the sense that $|f(n+1) - f(n)| \leq 1$, then

$$\rho(f \geq \rho(f) + R) \leq \exp\left(-\frac{R}{4} \log\left(1 + \frac{R}{2\lambda}\right)\right) \quad (4.45)$$

However, using the explicit form of the Laplace transform $\rho(\exp(\tau n))$ one can show that (see e.g. Example 7.3 in [70]):

$$\rho(n \geq \rho(n) + R) \leq \exp\left(-R\left(\log\left(1 + \frac{R}{\lambda}\right) - 1\right) - \lambda \log\left(1 + \frac{R}{\lambda}\right)\right) \quad (4.46)$$

Note that the exponent on the right hand side of (4.45) can be written as $-\frac{R}{4} \log R + \log\left(\frac{\lambda}{2}\right)R + o(R)$, while the right hand side of (4.46) can be written as $-R \log R + (\log(\lambda) + 1)R + o(R)$. Hence, in this sense, (4.46) is sharper than (4.45), but it holds only for a specific functional, which is the identity. A first result we obtain is a concentration of measure inequality of the form (4.46) which holds uniformly on the 1-Lipschitz functions. We apply the Herbst's argument to another type of Modified Log Sobolev Inequality, introduced by Dai Pra, Paganononi, and Posta in [26]. In their Proposition 3.1 they show that the Poisson distribution of mean λ satisfies the following inequality:

$$\forall f > 0, \quad \rho(f \log f) - \rho(f) \log(\rho(f)) \leq \lambda \rho(\nabla f \nabla \log f) \quad (4.47)$$

where $\nabla f(n)$ is the discrete gradient $f(n+1) - f(n)$.

Proposition 4.7.1. *Let ρ satisfy (6.40). Then for all $f : \mathbb{N} \rightarrow \mathbb{R}$ 1-Lipschitz:*

$$\rho(f \geq \rho(f) \geq R) \leq \exp\left(- (R + 2\lambda) \log\left(1 + \frac{R}{2\lambda}\right) + R\right) \quad (4.48)$$

Remark 4.7.3. (i) *Note that the right hand side of (4.48) is of the form $\exp(-R \log R + (\log(\lambda) + 1 + \log(2))R + o(R))$. Therefore the concentration estimate is sharp concerning the leading order term, it shows the right dependence on λ in the exponential correction term. Concerning the constants appearing in the exponential terms, we have $1 + \log(2)$. We do not know whether this is sharp or not. However, nothing better than 1 is reasonable to expect because of (4.46)*

(ii) *At the moment, we do not know whether repeating Herbst's argument based on the inequality in [5] could also yield better concentration bounds than those already contained in the paper. One reason to check this is in relation with Ollivier's theory of discrete Ricci curvature [62]. He shows that an inequality similar to that one holds under a discrete Ricci curvature lower bound condition. The concentration results he obtains out of it of the following type: for moderate values of R , i.e. when R is within the "Gaussian window", he obtains Gaussian concentration and for large values of R he obtains exponential concentration. But if one shows that the inequality in [5] yields Poissonian concentration for all values R (i.e. $-R \log R$), it is another interesting notion to attach to a lower bound on the discrete Ricci curvature.*

Proof. Let f be 1-Lipschitz. Then we already know, using e.g. the estimate 4.45 that f has exponential moments of all order. As a consequence, all the expectations we are going to consider in the next lines are finite. Let us define:

$$\varphi_\tau := \rho(\exp(\tau f)), \quad \psi_\tau := \log \rho(\exp(\tau f))$$

We apply the inequality (4.47) to $\exp(\tau f)$. Note that the left hand side reads as $\tau \partial_\tau \varphi_\tau - \varphi_\tau \psi_\tau$. The right hand side can be written as

$$\lambda \tau \rho(\exp(\tau f)) [\exp(\tau \nabla f) - 1] \nabla f$$

which, since f is Lipschitz can be bounded by

$$\lambda \tau (\exp(\tau) - 1) \rho(\exp(\tau f)) = \lambda \tau [\exp(\tau) - 1] \varphi_\tau$$

We thus get the following differential inequality:

$$\tau \partial_\tau \varphi_\tau - \varphi_\tau \psi_\tau \leq \lambda \tau \varphi_\tau (\exp(\tau) - 1) \quad (4.49)$$

Dividing on both sides by φ_τ , and using the chain rule, it can be rewritten as a differential inequality for ψ :

$$\tau \partial_\tau \psi_\tau - \psi_\tau \leq \lambda \tau (\exp(\tau) - 1), \quad \partial_\tau \psi_0 = \rho(f), \psi_0 = 0 \quad (4.50)$$

The ODE corresponding to this inequality is

$$\tau \partial_\tau h_\tau - h_\tau = \lambda \tau (\exp(\tau) - 1), \quad \partial_\tau h_0 = \rho(f), h_0 = 0 \quad (4.51)$$

Note that the condition $h_0 = 0$ is implied by the form of the equation. (4.51) admits a unique solution, given by:

$$h_\tau = \tau \rho(f) + \lambda \tau \gamma(\tau) \quad (4.52)$$

where

$$\gamma(\tau) = \sum_{k=1}^{+\infty} \frac{1}{k} \frac{\tau^k}{k!} \quad (4.53)$$

The fact that (4.52) is the solution to (4.51) can be checked directly by differentiating term by term the series defining γ in (4.53). We claim that

$$\forall \tau \geq 0 \quad \psi_\tau \leq h_\tau \quad (4.54)$$

The proof of this claim, is given separately in Proposition 4.7.2

Given (4.54), a standard argument with Markov inequality yields:

$$\rho(f \geq \rho(f) + R) \leq \exp \left(\inf_{\tau \geq 0} \psi_\tau - \tau \rho(f) - \tau R \right) = \exp \left(\inf_{\tau > 0} \lambda \tau \gamma(\tau) - \tau R \right)$$

We can bound γ in an elementary way:

$$\gamma(\tau) = \sum_{k=1}^{+\infty} \frac{1}{k} \frac{\tau^k}{k!} \leq \frac{2}{\tau} \sum_{k=1}^{+\infty} \frac{\tau^{k+1}}{(k+1)!} = 2 \frac{\exp(\tau) - \tau - 1}{\tau}$$

and therefore:

$$\rho(f \geq \rho(f) + R) \leq \exp \left(\inf_{\tau > 0} 2\lambda \exp(\tau) - (2\lambda + R)\tau - 2\lambda \right)$$

Solving the optimization problem yields the conclusion. \square

Remark 4.7.4. We obtain a slightly sharper result by keeping γ in its implicit form:

$$\rho(f \geq \rho(f) + R) \leq \exp(\inf_{\tau > 0} \lambda \tau \gamma(\tau) - \tau R)$$

Here, we prove (4.54).

Proposition 4.7.2. Let h be defined by (4.51) and ψ be as in (4.50) Then

$$\forall \tau > 0, \quad \psi_\tau \leq h_\tau$$

Proof. Consider $\varepsilon > 0$ and define h_τ^ε as the unique solution of

$$\tau \partial_\tau h_\tau^\varepsilon - h_\tau^\varepsilon = \tau(\exp(\tau) - 1), \quad \partial_\tau h_0^\varepsilon = \rho(f) + \varepsilon \quad (4.55)$$

Then $\eta_0^\varepsilon := \psi_0 - h_0^\varepsilon = 0$ satisfies:

$$\tau \eta_\tau^\varepsilon - \eta_\tau^\varepsilon \leq 0, \quad \partial_\tau \eta_0^\varepsilon = -\varepsilon$$

Since η^ε is continuously differentiable, we have that $T > 0$, where T is defined as

$$T := \inf\{\tau > 0 : \partial_\tau \eta_\tau^\varepsilon = 0\} \quad (4.56)$$

Assume that $T < +\infty$. Then, at T , we have:

$$T \underbrace{\partial_\tau \eta_T^\varepsilon}_{=0} - \eta_T^\varepsilon \leq 0 \Rightarrow \eta_T^\varepsilon \geq 0 \quad (4.57)$$

But this is impossible since $\eta_0^\varepsilon = 0, \eta_\tau^\varepsilon < 0$ for all $\tau < T$. Therefore $\partial_\tau \eta_\tau^\varepsilon < 0$ for all $\tau > 0$. Since $\eta_0^\varepsilon = 0$, we also have that $\psi_\tau^\varepsilon < 0$ for all $\tau > 0$. therefore:

$$\forall \tau > 0, \quad \psi_\tau \leq \inf_{\varepsilon > 0} h_\tau^\varepsilon = h_\tau$$

□

4.7.3 A useful staircase

The idea behind the proof of theorem 4.7.1 is to construct a measure π (see Definition 4.7.1) which “interpolates” ρ and for which the Modified Logarithmic Sobolev Inequality (4.47) gives sharp concentration bounds. We constructed ρ by pinning the two dimensional Poisson law \mathfrak{p}_{Φ_j} to belong to the line $\{kn_k = n_{-1}\}$, see Remark 4.7.1. The interpolation is constructed by interpolating the line with a “staircase” and pinning the Poisson vector to belong to it. Intuitively, one should think it that the MLSI is sharp for this staircase, because, unlike lines, one can travel along the staircase by using the edges of the lattice \mathbb{Z}^2 .

Definition 4.7.1. Let $(\gamma_1, \gamma_2) =: \gamma : \mathbb{N} \rightarrow \mathbb{N}^2$ be defined as follows:

$$\gamma(m) := \begin{cases} (\frac{m}{k+1}, k\frac{m}{k+1}) & \text{if } m \in (k+1)\mathbb{N} \\ (\lfloor \frac{m}{k+1} \rfloor + 1, m - \lfloor \frac{m}{k+1} \rfloor - 1) & \text{otherwise} \end{cases} \quad (4.58)$$

and $\pi \in \mathcal{P}(\mathbb{N})$ be defined as follows:

$$\pi(m) = \frac{1}{Z'} \mathbf{p}_{\Phi_j}(\gamma(m)) \quad (4.59)$$

where \mathbf{p}_{Φ_j} is a two dimensional Poisson law with independent coordinates, each one having mean $\Phi_j^{1/k+1}$ as in Remark 4.7.1, and Z' a normalization constant.

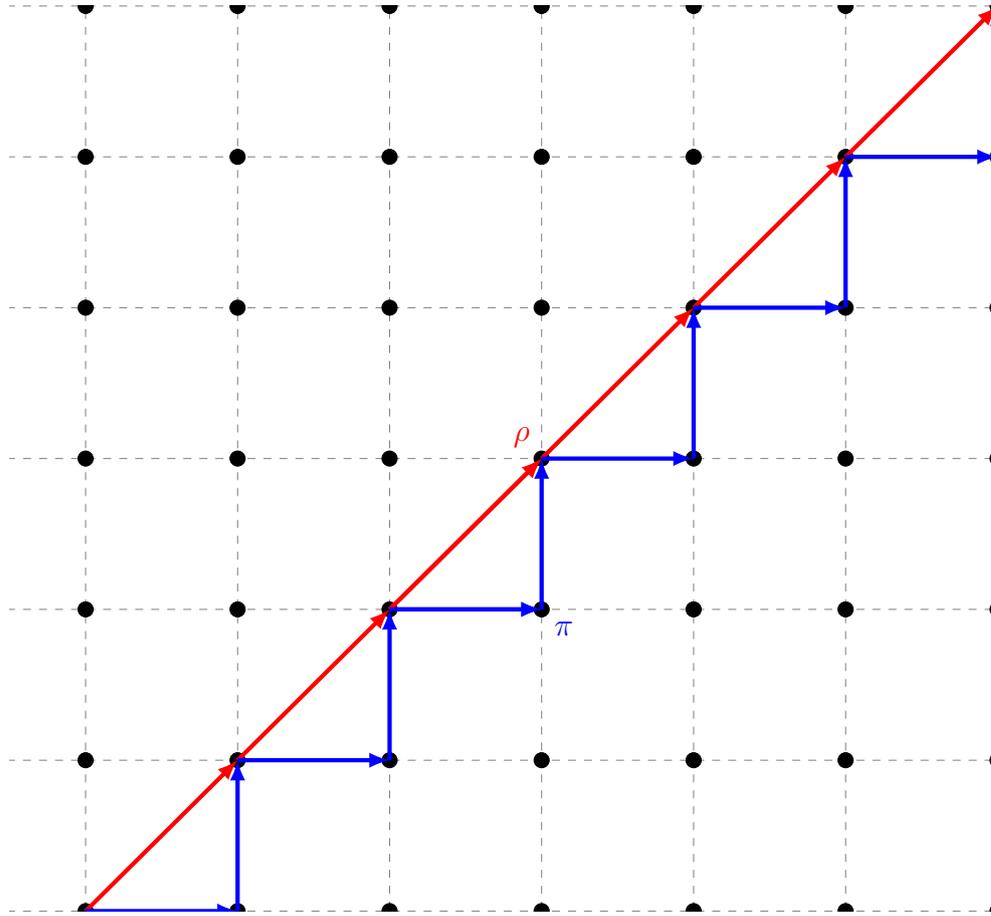


Figure 4.4: An illustration in the case $k = 1$ of the interpolation argument of Proposition 4.7.3. We want to study the measure ρ , obtained by pinning a two dimensional Poisson vector to belong to the subspace $n_1 = n_{-1}$, which is the red line in the picture. For such measure the MLSI does not give a precise concentration bound. But if we consider π , obtained by pinning the same two dimensional vector to belong to the blue "staircase" this bound becomes sharp. The set of blue arcs is precisely the image of the curve γ from Definition 4.7.1

Another ingredient we shall use in the proof is the following criterion for MLSI, due to Caputo and Posta. What we make here is a summary of some of their results in Section 2 of the paper [9], adapted to our scopes. To keep track of the constants, we also use Lemma 1.2 of [43] We do not reprove these results here.

Lemma 4.7.1 (Caputo and Posta criterion for MLSI,[9]). *Let $\pi \in \mathcal{P}(\mathbb{N})$ be such that*

$$c(m) := \frac{\pi(m-1)}{\pi(m)}$$

has the property that for some $v \in \mathbb{N}, c > 0$:

$$\inf_{m \geq 0} c(m+v) - c(m) \geq c > 0 \quad (4.60)$$

and that $\sup_{m \geq 0} c(m+v) - c(m) < +\infty$. Then the function \tilde{c} defined by

$$\tilde{c}(m) := c(m) + \frac{1}{v} \sum_{i=0}^{v-1} \frac{v-i}{v} [c(m+i) + c(m-i) - 2c(m)] \quad (4.61)$$

is uniformly increasing, that is

$$\inf_{m \geq 0} \tilde{c}(m+1) - \tilde{c}(m) \geq \delta \quad (4.62)$$

for some $\delta > 0$. Moreover, if we define $\tilde{\pi} \in \mathcal{P}(\mathbb{N})$ by:

$$\tilde{\pi}(0) = \frac{1}{Z}, \quad \tilde{\pi}(m) = \frac{1}{Z} \prod_{i=1}^m \frac{1}{\tilde{c}(i)} \quad (4.63)$$

then $\tilde{\pi}$ is equivalent to π in the sense that there exist \tilde{C} such that:

$$1/\tilde{C} \leq \frac{\pi(m)}{\tilde{\pi}(m)} \leq \tilde{C} \quad (4.64)$$

Finally, π satisfies the MLSI (4.47) with $\delta^{-1} \exp(4\tilde{C})$ instead of λ .

Using this criterion, we derive MLSI for π .

Proposition 4.7.3. *The measure π satisfies the modified logarithmic Sobolev inequality (4.47) with a constant of the form $\Phi_j^{-1/(k+1)} C$, where C is a constant independent from Φ_j .*

Proof. Using the elementary observation that (see Proposition 4.2.2):

$$\frac{\mathfrak{p}_{\Phi_j}(n_1-1, n_2)}{\mathfrak{p}_{\Phi_j}(n_1, n_2)} = \Phi_j^{-1/(k+1)} n_1, \quad \frac{\mathfrak{p}_{\Phi_j}(n_1, n_2-1)}{\mathfrak{p}_{\Phi_j}(n_1, n_2)} = \Phi_j^{-1/(k+1)} n_2$$

we have the following:

$$c(m) := \frac{\pi(m-1)}{\pi(m)} = \begin{cases} \Phi_j^{-1/(k+1)} \gamma_1(m), & \text{if } m \in (k+1)\mathbb{N} + 1 \\ \Phi_j^{-1/(k+1)} \gamma_2(m), & \text{otherwise} \end{cases} \quad (4.65)$$

Now consider any m . We have that:

$$c(m+k+1) - c(m) \geq \Phi_j^{-1/(k+1)}$$

By Lemma 4.7.1, π satisfies the (4.47) with a constant $\exp(4\tilde{C})\delta^{-1}$, where \tilde{C} is given by (4.64) and δ by (4.62). Using the explicit form of \tilde{c} given in (4.63), we see that \tilde{c} is affine in $\Phi_j^{-1/(k+1)}$, because so is c given by (4.65). This implies that $\pi(m)/\tilde{\pi}(m)$ does not depend on Φ_j , and therefore the same holds for \tilde{C} . Moreover, because of the fact that \tilde{c} is affine in $\Phi_j^{-1/(k+1)}$, then δ is of the form $\Phi_j^{-1/(k+1)}\tilde{\delta}$ for some $\tilde{\delta}$ independent from Φ_j . But then, using $\exp(4\tilde{C})\delta^{-1} = \exp(4\tilde{C})\tilde{\delta}^{-1}\Phi_j^{1/(k+1)} = C\Phi_j^{1/(k+1)}$ for some C independent from Φ_j . \square

We can finally prove Theorem 4.7.1.

Proof. We observe that that, by definition of ρ (see 4.7.1) and π (see Definition 4.7.1) we have that:

$$\rho(n) = \frac{Z'}{Z}\pi((k+1)n), \quad \forall n \in \mathbb{N}$$

where $Z'/Z := Z'' = \mathfrak{p}_{\Phi_j}(n \in \{\gamma(m) : m \in \mathbb{N}\})/\mathfrak{p}_{\Phi_j}(kn_{-1} = n_k)$ is a numerical constant independent from n .

Consider now $f : \mathbb{N} \rightarrow \mathbb{R}$ which is 1-Lipschitz. Then define $g : \mathbb{N} \rightarrow \mathbb{R}$ by putting:

$$g(m) := \left(1 - \frac{m}{k+1} + \lfloor \frac{m}{k+1} \rfloor\right) f(\lfloor \frac{m}{k+1} \rfloor) + \left(\frac{m}{k+1} - \lfloor \frac{m}{k+1} \rfloor\right) f(\lfloor \frac{m}{k+1} \rfloor + 1) \quad (4.66)$$

Then it is immediate to verify that g is $1/(k+1)$ -Lipschitz.

Because of Proposition 4.7.3 there exists C independent from Φ_j such that π satisfies MLSI (4.47) with constant $C\Phi_j^{1/(k+1)}$. Using the concentration bound on from Proposition 4.7.1 on $(k+1)g$ we get that for some C_1 independent from Φ_j , and for any $R' > 0$:

$$\sum_{m: g(m) \geq \pi(g(m)) + R'} \pi(m) \leq \exp(-(k+1)R' \log(R')) + (\log(\Phi_j) + C_2)R' + o(R')$$

Assume that one could prove that there exist a constant C_0 , which may depend on Φ_j such that:

$$\sup_{f: f \text{ 1-Lipschitz}} \pi(g) - \rho(f) \leq C_0$$

We shall prove the existence of C_0 in Lemma 4.7.2. Then we have, for any $R' > 0$:

$$\begin{aligned} \sum_{m:g(m)\geq\pi(g)+R'} \pi(m) &\geq \sum_{m:g(m)\geq\rho(f)+C_1+R'} \pi(m) \\ &\geq \sum_{n:g((k+1)n)\geq\rho(f)+C_1+R'} \pi((k+1)n) \\ &= \frac{1}{Z''} \sum_{n:f(n)\geq\rho(f)+C_1+R'} \rho(n) \end{aligned}$$

which gives for all $R' > 0$, f 1-Lipschitz:

$$\rho(f \geq \rho(f) + C_1 + R') \leq Z'' \exp(-(k+1)R' \log R' + (\log(\Phi_j) + C_2)R' + o(R'))$$

Therefore for any $R \geq C_0$, after putting $R' := R - C_0$ and absorbing Z'' in the $o(R)$ term, we have that:

$$\rho(f \geq \rho(f) + R) \leq \frac{1}{C_1} \exp(-(k+1)(R-C_0) \log(R-C_0) + (\log(\Phi_j) + \log(C_1))(R-C_0) + o(R))$$

Using elementary calculus we can rewrite the right hand side of this last inequality as

$$-(k+1)R \log(R) + (\log(\Phi_j) + \log(C_1))R + o(R) \quad (4.67)$$

Indeed, the difference between $(R - C_0) \log(R - C_0)$ and $R \log(R)$ is of order $\log R$ and therefore substituting $R - C_0$ with R in (4.67) cannot affect the exponential terms, but only lower order terms. This concludes the proof. \square

Lemma 4.7.2. *There exist $C_0 < +\infty$ such that*

$$\sup_{f:f1Lip} \pi(g) - \rho(f) \leq C_0$$

where, for a given f , g is given by (4.66).

Proof. Since, by construction of g , we have that $\rho(f + \alpha) - \pi(g + \alpha) = \rho(f) - \pi(g)$, we can w.l.o.g assume that $f(0) = 0$, which also implies $g(0) = 0$. But then, since π admits a first moment and by construction g is $k+1$ Lipschitz:

$$\sup_{f:f1Lip} \pi(g) - \rho(f) \leq \sup_{g:g \frac{1}{(k+1)}-Lip, g(0)=0} \pi(g) + \sup_{f:f1Lip, f(0)=0} \rho(f) < +\infty$$

which gives the conclusion. \square

Chapter 5

Random walks on Abelian groups

Outline of the chapter In this chapter, which is based on the submitted work [23], we study the reciprocal class of a random walk on Abelian group. Our main result is a duality formula characterizing $\mathfrak{R}(R)$. The difference operator has the clear interpretation of the addition of a loop to the paths of the canonical process. Several examples are discussed.

In this chapter we consider as a reference process a random walk on an Abelian group. The graph where the walk takes place is then the Cayley graph associated to the support of the jump intensity. This class of graphs extends lattices, retaining a translation invariance property. Therefore they are still suitable to derive a characterization of the reciprocal based on duality formulae. The main result of this chapter is indeed a duality formula characterizing $\mathfrak{R}(R)$, where the derivative operator has the clear interpretation of the addition of a random *loop* to the trajectory, and the stochastic integral is a multiple stochastic integral, which takes into account the geometrical complexity of loops. The meaning of the formula is very intuitive: it simply says that *the ratio between the "probability" of two paths which differ by a loop is proportional to the reciprocal characteristic associated with that loop*. This simple property characterizes $\mathfrak{R}(R)$. It is important to highlight the differences between cycles and loop. While a cycle is defined on the graph $(\mathcal{X}, \rightarrow)$, a loop is a path in Ω whose trace is a cycle started in the identity element of the group. Our main technical tool is an iterated version of Mecke formula, which we establish as a new result, and is interesting in its own right. Mecke-Slijvniak formula, which was originally developed to study random measures [75, 54], is useful to characterize stochastic pro-

cesses, as shown by Murr [57] and Privault [63]. It concerns the addition of a jump to the paths of the canonical process. An iteration of this formula concerns the addition of several jumps. When the jumps combine in such a way that the resulting transformation is the addition of a loop, it is valid over the whole $\mathfrak{R}(R)$.

This formula allows to lift at the level of paths all the geometrical considerations on the structure of cycles which we made in Chapter 4, and merge in a single equation the two equations used in Theorem 4.3.1. The fact that some of the shifts considered in Equation (4.15) are associated with vectors having some negative coordinates prevents this unified approach from being possible in general, and required to use generating sets of lattices to achieve a sharp characterization.

Here, we work under a geometrical assumption (see Assumption 5.3.1), which allows to overcome these difficulties. In the language of the previous chapter, it amounts to say that we can find lattice basis which are also generating sets. This assumption makes possible to have a rigorous version of the intuitive statement that, by adding and subtracting a series of well chosen loops, we can connect any two paths having the same initial and final state. The sharpness of our characterization is given by the fact that we work with lattice basis. The loops we add or we cancel always have a very simple geometrical structure (which we will call "skeleton") which is encoded in one of the elements of the basis. All other complex geometrical structures are obtained by superposition of these simple ones.

Organization of the chapter In section 5.1 we present a parallelism between random walks and random measures, which we are going to exploit in our main result. In section 5.2 we prove the Iterated Mecke formula, while in section 5.3 we present our geometrical assumption and its relevance to have a nice decomposition of the set of loops, and state our main result which is proven, with the help of several lemmas in section 5.4. A gallery of examples in section 5.5 concludes the chapter. The following chart helps in connecting the notation of this chapter with the general framework we set up in Chapter 2.

	General def	Local def
State space	\mathcal{X}	Abelian group $(G,+)$
Arcs	\mathcal{A}	G^2
Active arcs	$\mathcal{A}_{\rightarrow}(j)$	$(g \rightarrow g + g'), g \in G, g' \in G_{\nu}$
Reference intensity	$j(t, g \rightarrow g + g')$	$\nu_{g'}$

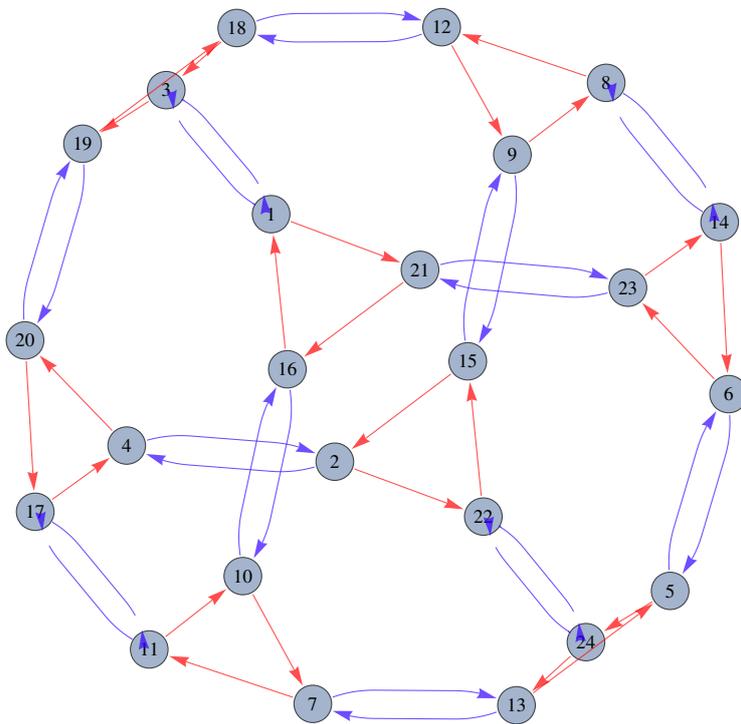


Figure 5.1: A Cayley graph over the symmetric group with 4 elements. In this case the lattice \mathcal{L}^+ (see 5.3.1) has dimension two: all blue cycles correspond to one element of its basis, and the red cycles to the other one.

5.1 Random walks and random measures

In this section, the state space is a countable Abelian group $(G, +)$ with identity element e . The arc set is the whole \mathcal{X}^2 . The reference walk has a space time homogeneous intensity. This means that we fix a non negative finite measure ν on G , and set:

$$j(t, g \rightarrow g + g') = \nu_{g'} \quad \forall g, g' \in G, t \in [0, 1] \quad (5.1)$$

In the rest of the chapter $G_\nu \subseteq G$ denotes the support of ν .

For any other non negative measure $\tilde{\nu}$ different from the reference measure, we call a $\tilde{\nu}$ -random walk any Markov walk whose intensity is defined by (5.1), where $\tilde{\nu}$ replaces ν .

The path space Ω introduced in section 2.3.1 is naturally embedded $\mathbb{D}([0, 1], G)$ the space of càdlàg paths for the topology induced by the discrete metric in G . More precisely, $\omega \in \mathbb{D}([0, 1], G)$ belongs to Ω if and only if $\omega_t - \omega_{t-} \in G_\nu$. Note that, because of the existence of left and right limits, paths in $\mathbb{D}([0, 1], G)$ have finitely many jumps.

For a measurable space $(\Gamma, \mathfrak{B}_\Gamma)$ we denote \mathcal{S}_Γ the set of finite point measures, that is:

$$\mathcal{S}_\Gamma := \left\{ \sum_{i=1}^N \delta_{x_i} : x_i \in \Gamma, N \in \mathbb{N} \right\}. \quad (5.2)$$

This space is also called the configuration space, (see Section 6.1 in [63]) The set of atoms of $\mu \in \mathcal{S}_\Gamma$ is denoted $\mathcal{E}(\mu)$.

We will often choose, as Γ the following product space of time-space elements:

$$\Gamma := [0, 1] \times G \quad \ni \gamma = (t, g).$$

We identify trajectories in $\mathbb{D}([0, 1], G)$ and point measures in \mathcal{S}_Γ via the following canonical bijective map M :

$$X \mapsto M_X := \sum_{0 \leq t \leq 1} \sum_{g \in G} \delta_{(t, g)} \mathbf{1}_{\{\Delta X_t = g\}} \quad (5.3)$$

Mecke's original idea was to characterize any Poisson random measure (see Proposition 6.1.3 of [63] for a definition) by mean of an integral formula (see Satz 3.1 in [54]), via the change of measures which consists to add one (random) atom to the initial point measure, as in the right-hand side of equation (5.4).

Proposition 5.1.1. *For $\tilde{P} \in \mathcal{P}(\mathcal{S}_\Gamma)$ the following assertions are equivalent:*

- i) \tilde{P} is the Poisson random measure with intensity measure $\rho = m \otimes \nu$ on Γ .
ii) For all $\Phi \in \mathcal{B}(\mathcal{M}(\Gamma) \times \Gamma)$,

$$\int \int_{\Gamma} \Phi(\mu, \gamma) \mu(d\gamma) \tilde{P}(d\mu) = \int \int_{\Gamma} \Phi(\mu + \delta_{\gamma}, \gamma) \rho(d\gamma) \tilde{P}(d\mu). \quad (5.4)$$

Proof. We first prove ii) \Rightarrow i). Consider $f \in \mathcal{B}(\Gamma)$. We define

$$\psi^f(\lambda) := \tilde{P}\left(\exp\left(-\lambda \int_{\Gamma} f d\mu\right)\right)$$

Clearly,

$$\partial_{\lambda} \psi^f(\lambda) := \tilde{P}\left(-\underbrace{\int_{\Gamma} \exp\left(-\lambda \int_{\Gamma} f d\mu\right) f(\gamma) \mu(d\gamma)}_{\Phi(\mu, \gamma)}\right)$$

This expression is suitable to apply (5.4) to $\Phi(\mu, \gamma) = \exp\left(-\lambda \int_{\Gamma} f d\mu\right) f(\gamma)$. We obtain:

$$\begin{aligned} \tilde{P}\left(-\int_{\Gamma} \exp\left(-\lambda \int_{\Gamma} f d\mu\right) f(\gamma) \mu(d\gamma)\right) &= \tilde{P}\left(-\int_{\Gamma} \exp\left(-\lambda \int_{\Gamma} f d(\mu + \delta_{\gamma})\right) f(\gamma) d\rho(\gamma)\right) \\ &= \left(-\int_{\Gamma} \exp\left(-\lambda f\right) f d\rho\right) \psi^f(\lambda) \end{aligned}$$

We proved that $\psi^f(\lambda)$ satisfies the ODE:

$$\partial_{\lambda} \psi^f(\lambda) = \left(\int_{\Gamma} \exp\left(-\lambda f\right) f d\rho\right) \psi^f(\lambda), \quad \psi^f(0) = 1 \quad (5.5)$$

Solving it we obtain:

$$\psi^f(\lambda) = \exp\left(-\int_{\Gamma} 1 - \exp(-\lambda f) d\rho\right)$$

Setting $\lambda = 1$:

$$\tilde{P}\left(\exp\left(-\int_{\Gamma} f d\mu\right)\right) = \exp\left(-\int_{\Gamma} 1 - \exp(-f) d\rho\right) \quad (5.6)$$

Now, let $f = \sum_{i=1}^k \eta_i \mathbf{1}_{A_i}$, for some disjoint sets $A_1, \dots, A_k \subseteq \mathcal{B}_{\Gamma}$, $\eta_1, \dots, \eta_k \geq 0$. Moreover, let us denote by $\pi \in \mathcal{P}(\mathbb{N}^A)$ the image measure of the vector $(\mu(A_1), \dots, \mu(A_k))$ under \tilde{P} . With these choices, the equation (5.6) reads as:

$$\pi\left(\exp\left(-\sum_{i=1}^k \eta_i n_i\right)\right) = \exp\left(-\sum_{i=1}^k (1 - \exp(-\eta_i))\rho(A_i)\right)$$

in which we recognize the form of the Laplace transform of a Poisson random vector, that is the random variables $\mu(A_1), \dots, \mu(A_n)$ are independent and $\mu(A_i)$ is distributed according to a Poisson distribution of parameter $\rho(A_i)$. The conclusion then follows. Let us prove the converse implication $i) \Rightarrow ii)$. Take disjoint sets A_1, \dots, A_k and consider the random variables $(\mu(A_1), \dots, \mu(A_k))$, and non negative measurable functions f_1, \dots, f_k . Since $\mu(A_i)$ is a Poisson random variable for all $1 \leq i \leq k$ we can use Chen's characterization (see Chen ??) to deduce that:

$$\tilde{P}\left(f_i(\mu(A_i)) \mu(A_i)\right) = \rho(A_i) \tilde{P}\left(f_i(\mu(A_i) + 1)\right)$$

By using independence and linearity of the expectation:

$$\tilde{P}\left(\prod_{i=1}^k f_i(\mu(A_i)) \sum_{j=1}^k \mu(A_j)\right) = \tilde{P}\left(\sum_{j=1}^k \rho(A_j) \prod_{i=1, i \neq j}^k f_i(\mu(A_i))\right)$$

which is precisely Mecke formula (5.4) for:

$$\Phi(\mu, \gamma) = \prod_{i=1}^k f_i(\mu(A_i)) \mathbf{1}_{\{\gamma \in \cup_{j=1}^k A_j\}}$$

A standard approximation argument yields the conclusion. \square

Adapted to our context it reads as follows. Remark that the left-hand side of (5.4) also reads $\int \sum_{\gamma \in \mu} \Phi(\mu, \gamma) \tilde{P}(d\mu)$ where the notation $\gamma \in \mu$ means that the points $\gamma \in \Gamma$ build the support of the point measure μ : one integrates the function Φ under the Campbell measure associated with \tilde{P} . Thus (5.4) determines the Campbell measure of a Poisson random measure as the shifted product measure of itself with its intensity.

Let us adapt this tool to $\mathbb{D}([0, 1], G)$. First, for $\gamma = (t, g) \in \Gamma$, let us denote by χ^γ the corresponding simple step function $\chi^\gamma := g \mathbf{1}_{[t, 1]} \in \mathbb{D}([0, 1], G)$. Then define the transformation Ψ_γ on the path space which consists in adding one jump g at time t .

Definition 5.1.1. For $\gamma = (t, g) \in \Gamma$, $\Psi_\gamma X := X + \chi^\gamma$, $X \in \mathbb{D}([0, 1], G)$.

Notice that, under any probability $P \in \mathcal{P}(\mathbb{D}([0, 1], G))$ satisfying $P(X_t = X_{t-}) = 1$ for all $t \in [0, 1]$, one has:

$$M_{\Psi_\gamma X} = M_X + \delta_\gamma \quad P - a.s. \quad (5.7)$$

The following diagram summarizes these considerations:

$$\begin{array}{ccc}
\mathbb{D}([0, 1], G) & \xrightarrow{\Psi_\gamma} & \mathbb{D}([0, 1], G) \\
M_X \downarrow & & \uparrow (M_X)^{-1} \\
\mathcal{S}_\Gamma & \xrightarrow{+\delta_\gamma} & \mathcal{S}_\Gamma
\end{array}$$

Let us rewrite Proposition 5.1.1 in the language of random walks.

Proposition 5.1.2. *For $P \in \mathcal{P}(\mathbb{D}([0, 1], G))$ the following assertions are equivalent:*

- i) P is a ν -random walk on G .
- ii) For all $F \in \mathcal{B}(\mathbb{D}([0, 1], G) \times \Gamma)$,

$$P\left(\int_\Gamma F(X, \gamma) M_X(d\gamma)\right) = P\left(\int_\Gamma F(\Psi_\gamma X, \gamma) \rho(d\gamma)\right), \quad (5.8)$$

Proof. i) \Rightarrow ii).

Since R is ν -random walk, M_X is a Poisson random measure with intensity $dt \otimes \nu$. Then Mecke's formula holds for $\tilde{P} := P \circ M^{-1}$. Since M is invertible and its inverse is measurable we can plug into (5.4) test functions Φ of the form $F(X, \gamma)$ and the conclusion follows.

ii) \Rightarrow i).

Let $P \in \mathcal{P}(\mathbb{D}([0, 1], G))$ satisfying (5.8). We define $\tilde{P} := P \circ M^{-1} \in \mathcal{P}(\mathcal{S}_\Gamma)$. Then, by considering test functions of the form $\Phi = F(M_X, \gamma)$ and using the fact that $M_{\Psi_\gamma X} = M_X + \delta_\gamma \mathbb{P} \otimes \rho - a.s.$, we deduce that \tilde{P} is a Poisson random measure with intensity $\rho = dt \otimes \nu$ by Proposition 5.1.1. Observing that

$$X_t = \sum_{g \in G} g M_X([0, t] \times \{g\}) \quad (5.9)$$

the conclusion follows using (5.7). \square

5.2 An iterated formula

To prepare the characterization of bridges which we will present in the next section, we now consider a generalization of the formula (5.8) obtained by iteration. For this purpose, we define n -dimensional analogous of the objects appearing in (5.8), $n \geq 1$ fixed.

$$\text{For } \bar{\gamma} = (\gamma_1, \dots, \gamma_n) \in \Gamma^n, \Psi_{\bar{\gamma}} := \Psi_{\gamma_n} \circ \dots \circ \Psi_{\gamma_1}, \quad \chi^{\bar{\gamma}} = \sum_{i=1}^n \chi^{\gamma_i}. \quad (5.10)$$

Then $\Psi_{\bar{\gamma}}$ is the transformation that adds to X the n -step function $\chi^{\bar{\gamma}}$, which corresponds in the framework of point measures to add n new atoms to the original measure. It remains to specify the set over which we will integrate the point measure $\mu^{\otimes n}$. Indeed, we remove from Γ^n the set Δ_n , union of all diagonals:

$$\Delta_n := \{\bar{\gamma} \in \Gamma^n : \exists i \neq j, \gamma_i = \gamma_j\}. \quad (5.11)$$

With these definitions, we can state the iterated formula satisfied under reference random walk R .

Proposition 5.2.1. *For any test function $F \in \mathcal{B}(\mathbb{D}([0, 1], G) \times \Gamma^n)$,*

$$R\left(\int_{\Gamma^n \setminus \Delta_n} F(X, \bar{\gamma}) \mu^{\otimes n}(d\bar{\gamma})\right) = R\left(\int_{\Gamma^n} F(\Psi_{\bar{\gamma}}X, \bar{\gamma}) \rho^{\otimes n}(d\bar{\gamma})\right). \quad (5.12)$$

Remark 5.2.1. *In general it is not true that $\int_{\Delta_n} F(X, \bar{\gamma}) M_X^{\otimes n}(d\bar{\gamma}) = 0$. Indeed, if γ is an atom of μ then $\underbrace{(\gamma, \dots, \gamma)}_{n \text{ times}}$ is an atom of $M_X^{\otimes n}$ which belongs to Δ_n .*

Proof. We first prove a preliminary result.

Lemma 5.2.1. *Define for $\gamma \in \Gamma$, $A_\gamma := \{\bar{\gamma} \in \Gamma^n : \gamma_i = \gamma \text{ for some } 1 \leq i \leq n\}$. Then*

$$M_{\Psi_\gamma X}^{\otimes n}(\cdot) = M_X^{\otimes n}(\cdot) + M_{\Psi_\gamma X}^{\otimes n}(A_\gamma \cap \cdot), \quad R \otimes \rho \text{ a.e.} \quad (5.13)$$

Proof. We denote by $\mathcal{E}^n(X)$ the set of atoms of $M_X^{\otimes n}$. Clearly $\mathcal{E}^n(X) = (\mathcal{E}(X))^n \subset \Gamma^n$. From (5.7),

$$\mathcal{E}(\Psi_\gamma X) = \mathcal{E}(X) \cup \{\gamma\} \quad R \otimes \rho \text{ a.e.} \quad (5.14)$$

so that the atoms of $M_{\Psi_\gamma X}^{\otimes n}$ are $\mathcal{E}^n(\Psi_\gamma X) = \left(\mathcal{E}(X) \cup \{\gamma\}\right)^n$ $R \otimes \rho$ a.e.. In this case $\mathcal{E}^n(\Psi_\gamma X) \setminus \mathcal{E}^n(X) = \mathcal{E}^n(\Psi_\gamma X) \cap A_\gamma$, which leads to the conclusion. \square

To prove the proposition by induction, we adopt the following notation: we decompose any element $\bar{\gamma} = (\gamma_1, \dots, \gamma_{n+1})$ of Γ^{n+1} into $(\tilde{\gamma}, \gamma)$ where $\tilde{\gamma}$ is its projection on Γ^n , and $\gamma = \gamma_{n+1}$ is its last coordinate.

Proof of the statement for $n = 1$: It is exactly Proposition 5.1.2.

Let us now assume that the statement is true for n , that is (5.12) holds true for all test functions $f \in \mathcal{B}(\mathbb{D}([0, 1], G) \times \Gamma^n)$. Let us now consider $v \in \mathcal{B}(\Gamma)$ and prove that (5.12) holds for any function of the form $F(X, \bar{\gamma}) =$

$f(X, \tilde{\gamma})v(\gamma)$. The extension to a general $F \in \mathcal{P}(\mathbb{D}([0, 1], G) \times \Gamma^{n+1})$ will be then standard. We have:

$$R\left(\int_{\Gamma^{n+1}} F(\Psi_{\tilde{\gamma}}X, \tilde{\gamma})\rho^{\otimes n+1}(d\tilde{\gamma})\right) = R\left(\int_{\Gamma} \int_{\Gamma^n} f(\Psi_{\gamma}(\Psi_{\tilde{\gamma}}X), \tilde{\gamma})\rho^{\otimes n}(d\tilde{\gamma})v(\gamma)\rho(d\gamma)\right) \quad (5.15)$$

Exchanging the order of integration, and applying the inductive hypothesis to $f(\Psi_{\gamma}X, \tilde{\gamma})$ allows to rewrite the right hand side of the last identity as:

$$\int_{\Gamma} R\left(\int_{\Gamma^n \setminus \Delta_n} f(\Psi_{\gamma}X, \tilde{\gamma})M_X^{\otimes n}(d\tilde{\gamma})\right)v(\gamma)\rho(d\gamma) \quad (5.16)$$

We can apply Lemma 5.2.1 to rewrite the former integral as:

$$\int_{\Gamma} R\left(\int_{(\Gamma^n \setminus \Delta_n) \setminus A_{\gamma}} f(\Psi_{\gamma}X, \tilde{\gamma})M_{\Psi_{\gamma}X}^{\otimes n}(d\tilde{\gamma})\right)v(\gamma)\rho(d\gamma). \quad (5.17)$$

We apply Proposition 5.1.2 to $(X, \gamma) \mapsto \int_{(\Gamma^n \setminus \Delta_n) \setminus A_{\gamma}} f(X, \tilde{\gamma})M_X^n(d\tilde{\gamma})v(\gamma)$ and we obtain

$$\begin{aligned} & R\left(\int_{\Gamma} \left(\int_{(\Gamma^n \setminus \Delta_n) \setminus A_{\gamma}} f(X, \tilde{\gamma})M_X^{\otimes n}(d\tilde{\gamma})\right)v(\gamma)M_X(d\gamma)\right) \\ &= R\left(\int_{\{\tilde{\gamma}:\gamma \in \Gamma, \tilde{\gamma} \in (\Gamma^n \setminus \Delta_n) \setminus A_{\gamma}\}} F(X, \tilde{\gamma})M_X^{\otimes n+1}(d\tilde{\gamma})\right) \end{aligned}$$

It is easy to see that $\{\tilde{\gamma} = (\tilde{\gamma}, \gamma) : \gamma \in \Gamma, \tilde{\gamma} \in (\Gamma^n \setminus \Delta_n) \setminus A_{\gamma}\} = \Gamma^{n+1} \setminus \Delta_{n+1}$ and the conclusion follows. \square

5.3 Loops and their skeletons

We call *loop* a path in $\mathbb{D}([0, 1], G)$ that starts and ends at the identity element e . It shall not be confused with a cycle as defined in Definition 3.2.1. A cycle is a sequence of vertices of $\mathcal{X} = G$, while a loop is a path in Ω . If X is a loop, and n is the total number of jumps of the loop, then the sequence $(X_0, X_{T_1}, \dots, X_{T_n})$ is a cycle such that $X_0 = X_{T_n} = e$. For each path $X \in \Omega$ we define its *skeleton* as the application $\varphi_X : G_{\nu} \rightarrow \mathbb{N}$ defined by:

$$\varphi_X(g) := \mu([0, 1] \times \{g\}). \quad (5.18)$$

Thus $\varphi_X(g)$ counts how many times the jump g occurs along the path X . If X is a loop, we observe that

$$\sum_{g \in G_{\nu}} \varphi_X(g) g = e. \quad (5.19)$$

Therefore, as X varies in the set of all possible loops of Ω , φ_X varies in the set

$$\mathcal{L}^+ := \{\varphi \in \mathbb{N}^{G_\nu} : \sum_{g \in G_\nu} \varphi(g) g = e, \ell(\varphi) < +\infty\} \quad (5.20)$$

where $\ell(\varphi) := \sum_{g \in G_\nu} |\varphi(g)|$ is the *length* of φ . Enlarging this set to the maps φ with negatives values by considering

$$\mathcal{L} := \{\varphi \in \mathbb{Z}^{G_\nu} : \sum_{g \in G_\nu} \varphi(g) g = e, \ell(\varphi) < +\infty\},$$

one recovers for \mathcal{L} a lattice structure, which will be very useful. To any $\varphi^* \in \mathcal{L}$ we can associate the - non empty - set of loops whose skeleton is φ^* :

$$\Omega_{e, \varphi^*} := \{X \in \Omega : X_0 = X_1 = e \text{ and } \varphi_X = \varphi^*\}. \quad (5.21)$$

These paths have exactly $\varphi^*(g)$ jumps of type g , for all $g \in G_\nu$.

We make the following assumption on \mathcal{L} :

Assumption 5.3.1. *The lattice \mathcal{L} is such that:*

i) *The skeleton of loops generate \mathcal{L} :*

$$\text{Span}(\mathcal{L}^+) = \mathcal{L} \quad (\text{H1})$$

where $\text{Span}(\mathcal{L}^+)$ is, as usual, the set of all integer combinations of elements of \mathcal{L}^+ .

ii) *Each jump in G_ν belongs to (at least) the skeleton of one loop, that is, the following assumption holds:*

$$\forall g \in G_\nu, \text{ there exists } \varphi \in \mathcal{L} \text{ such that } \varphi(g) > 0. \quad (\text{H2})$$

Let us note that (H1) ensures that \mathcal{L} admits a *basis* $\mathcal{B} \subset \mathcal{L}^+$. From now on, we fix such a basis \mathcal{B} . W.l.o.g. we can assume that (H2) is satisfied replacing \mathcal{L} with \mathcal{B} . As we shall see in Section 5.4, assumptions (H1) and (H2) allow a fruitful decomposition of the path space Ω . Heuristically, one can transform one path into any other one having the same initial and final values, by subsequently adding and removing loops whose skeleton belongs to \mathcal{B} . However, let us first state our main result.

5.3.1 Main result: an integral characterization of the reciprocal class

In the next theorem we state that the identity (5.12) appeared in Proposition 5.2.1 is not only valid over the whole reciprocal class $\mathfrak{R}(R)$ but indeed characterizes it, if one restricts the set of test functions F to some well chosen subset.

For each skeleton φ^* in the basis \mathcal{B} , consider the following set of test functions:

$$\mathcal{H}_{\varphi^*} := \left\{ F \in \mathcal{P}(\mathbb{D}([0, 1], G) \times \Gamma^{\ell(\varphi^*)}) : F(X, \bar{\gamma}) \equiv \mathbf{1}_{\{\chi^{\bar{\gamma}} \in \Omega_{e, \varphi^*}\}} F(X, \bar{\gamma}) \right\}. \quad (5.22)$$

Therefore, we will restrict our attention to perturbations of the sample paths consisting in adding a loop $\chi^{\bar{\gamma}}$ whose skeleton is equal to φ^* . Now we are ready for stating and proving the main result.

Theorem 5.3.1. *The probability measure $P \in \mathcal{P}(\Omega)$ belongs to the reciprocal class $\mathfrak{R}(R)$ if and only if for any skeleton φ^* in the basis \mathcal{B} and for all test functions $F \in \mathcal{H}_{\varphi^*}$, we have:*

$$P\left(\int_{\Gamma^n \setminus \Delta_n} F(X, \bar{\gamma}) M_X^{\otimes n}(d\bar{\gamma})\right) = \Phi_{\varphi^*}^{\nu} P\left(\int_{\Gamma^n} F(\Psi_{\bar{\gamma}} X, \bar{\gamma}) (dt \otimes \lambda)^{\otimes n}(d\bar{\gamma})\right), \quad (5.23)$$

where $n = \ell(\varphi^*)$, $\lambda := \sum_{g \in G} \delta_g$ is the counting measure on G and

$$\Phi_{\varphi^*}^{\nu} := \prod_{g \in G_{\nu}} \nu(g)^{\varphi^*(g)} \in \mathbb{R}^+. \quad (5.24)$$

In particular, if (5.23) holds true under P satisfying $P(X_0 = x, X_1 = y) = 1$ for some (x, y) , then P is nothing else but the bridge R^{xy} .

The positive number $\Phi_{\varphi^*}^{\nu}$ is the **reciprocal characteristics** associated to the jump measure ν and the skeleton φ^* .

Corollary 5.3.1. *The reciprocal characteristics are invariants of the reciprocal class in the following sense. Let ν and μ two non negative finite measures on G with the same support. The reciprocal classes $\mathfrak{R}(R)$ and $\mathcal{R}ec(\tilde{R})$ coincide if and only if their family of reciprocal characteristics coincide:*

$$\Phi_{\varphi^*}^{\mu} = \Phi_{\varphi^*}^{\nu}, \quad \forall \varphi^* \in \mathcal{B}. \quad (5.25)$$

In that case the bridges of both ν - and μ -random walk on G coincide too.

Remark 5.3.1. *There is a remarkable probabilistic interpretation of the reciprocal characteristics $\Phi_{\varphi^*}^{\nu}$ as the leading factor, in the short-time expansion, of the probability that the ν -random walk follows a loop with skeleton φ^* . This result will be proven in chapter 6.*

5.4 Proof of the main theorem

Proof. (\Rightarrow)

We use, as main argument, the specific form of the density with respect to R of any probability measure in $\mathfrak{A}(R)$ as it is given by Proposition 2.2.2:

$$P \in \mathfrak{A}(R) \Rightarrow P \ll R, \quad \text{and} \quad \frac{dP}{dR} = h(X_0, X_1) \text{ for some } h : G \times G \rightarrow \mathbb{R}^+. \quad (5.26)$$

Take now any $F \in \mathcal{H}_{\varphi^*}$. Then, using Identity (5.12), the definition of \mathcal{H}_{φ^*} and the fact that $(\Psi_{\bar{\gamma}}X)_0 = X_0$, $(\Psi_{\bar{\gamma}}X)_1 = X_1 + e = X_1$, one gets

$$\begin{aligned} P\left(\int_{\Gamma^n} F(\Psi_{\bar{\gamma}}X, \bar{\gamma})(dt \otimes \nu)^{\otimes n}(d\bar{\gamma})\right) &= P\left(\int_{\Gamma^n} F(\Psi_{\bar{\gamma}}X, \bar{\gamma}) \prod_{g \in G_\nu} \nu(g)^{\varphi_{X\bar{\gamma}}(g)}(dt \otimes \lambda)^{\otimes n}(d\bar{\gamma})\right) \\ &\stackrel{F \in \mathcal{H}_{\varphi^*}}{=} \Phi_{\varphi^*}^\nu P\left(\int_{\Gamma^n} F(\Psi_{\bar{\gamma}}X, \bar{\gamma})(dt \otimes \lambda)^{\otimes n}(d\bar{\gamma})\right) \\ &= \Phi_{\varphi^*}^\nu R\left(h(X_0, X_1) \int_{\Gamma^n} F(\Psi_{\bar{\gamma}}X, \bar{\gamma})(dt \otimes \lambda)^{\otimes n}(d\bar{\gamma})\right) \\ &= \Phi_{\varphi^*}^\nu R\left(\int_{\Gamma^n} h((\Psi_{\bar{\gamma}}X)_0, (\Psi_{\bar{\gamma}}X)_1) F(\Psi_{\bar{\gamma}}X, \bar{\gamma})(dt \otimes \lambda)^{\otimes n}(d\bar{\gamma})\right) \\ &= R\left(\int_{\Gamma^n \setminus \Delta_n} h(X_0, X_1) F(X, \bar{\gamma}) M_X^{\otimes n}(d\bar{\gamma})\right) \\ &= P\left(\int_{\Gamma^n \setminus \Delta_n} F(X, \bar{\gamma}) M_X^{\otimes n}(d\bar{\gamma})\right) \end{aligned}$$

which completes the proof of the first implication.

(\Leftarrow)

The converse implication is more sophisticated and needs several steps.

Let us introduce the set of paths which correspond to the support of $e y$ -bridges, $y \in G$:

$$\Omega_y := \{X \in \Omega : X_0 = e, X_1 = y\}. \quad (5.27)$$

Now we partition Ω_y according to the skeleton of its elements:

$$\Omega_y = \bigcup_{\varphi \in \mathcal{L}_y^+} \Omega_{y,\varphi}, \quad \Omega_{y,\varphi} := \Omega_y \cap \{X \in \Omega : \varphi_X = \varphi\} \quad (5.28)$$

$$\text{where } \mathcal{L}_y^+ = \{\varphi \in \mathbb{N}^{G_\nu} : \sum_{g \in G_\nu} g\varphi(g) = y, \ell(\varphi) < +\infty\}. \quad (5.29)$$

In order to discretize the time, we introduce a mesh $h \in \mathbb{N}^*$ and partition $\Omega_{y,\varphi}$ by specifying the number of different jumps occurred in each h -dyadic interval. That is, we consider functions $\theta : \{0, \dots, 2^h - 1\} \times G_\nu \longrightarrow \mathbb{N}$ and we look for paths which have $\theta(k, g)$ jumps of type g during the time interval $I_k^h := (2^{-h}k, 2^{-h}(k+1)]$, for each k and each $g \in G_\nu$. For each skeleton φ we define the set

$$\Theta_\varphi^h := \left\{ \theta : \{0, \dots, 2^h - 1\} \times G_\nu \longrightarrow \mathbb{N}, \sum_{0 \leq k \leq 2^h - 1} \theta(k, g) = \varphi(g), \quad \forall g \in G_\nu \right\} \quad (5.30)$$

of all possible h -dyadic time repartition of the jumps, compatible with the skeleton φ . We thus obtain $\Omega_{y,\varphi} = \bigcup_{\theta \in \Theta_\varphi^h} \Omega_{y,\varphi}^{h,\theta}$ where

$$\Omega_{y,\varphi}^{h,\theta} := \{X \in \Omega_y : \mu(I_k^h \times \{g\}) = \theta(k, g), \quad 0 \leq k < 2^h, \quad g \in G_\nu\}. \quad (5.31)$$

Consider the set

$$V := \left\{ v = (\varphi, \theta) \text{ with } \varphi \in \mathcal{L}_y^+, \theta \in \Theta_\varphi^h \right\} \quad (5.32)$$

of pairs of skeletons connecting e to y and h -dyadic time repartition of their jumps. Elements of this set are discrete versions of paths of Ω : the spatial structure of the path is given by the skeleton φ , and the time structure is approximated by θ . One equips V with the following l^1 -metric:

$$d(v, \tilde{v}) := \sum_{(k,g) \in \{0, \dots, 2^h - 1\} \times G_\nu} |\theta - \tilde{\theta}|(k, g) \in \mathbb{N}, \quad v = (\varphi, \theta), \tilde{v} = (\tilde{\varphi}, \tilde{\theta}) \in V. \quad (5.33)$$

Take now two paths $X, X' \in \Omega_y$ and their trace v, v' on V . Our aim is to find a way to transform X into X' (resp. v into v') by adding or removing a finite number of loops whose skeletons belong to the basis \mathcal{B} . Let us introduce the following relation:

$$v_1 = (\varphi_1, \theta_1) \hookrightarrow v_2 = (\varphi_2, \theta_2) \text{ if } \varphi_2 \in \varphi_1 + \mathcal{B} \text{ and } \theta_2 - \theta_1 \in \Theta_{\varphi_2 - \varphi_1}^h. \quad (5.34)$$

We shall now use assumptions (H1) and (H2).

Lemma 5.4.1. *For each v and $\tilde{v} \neq v \in V$ on can construct a connecting finite sequence $v_1, \dots, v_N = \tilde{v}_{\tilde{N}}, \tilde{v}_{\tilde{N}-1}, \dots, \tilde{v}_1$ in V such that*

$$v \hookrightarrow v_1 \hookrightarrow \dots \hookrightarrow v_N = \tilde{v}_{\tilde{N}} \leftrightarrow \tilde{v}_{\tilde{N}-1} \dots \leftrightarrow \tilde{v}_1 \leftrightarrow \tilde{v}.$$

Proof. We distinguish two cases:

Case i) The skeletons φ and $\tilde{\varphi}$ coincide.

In this case, it is sufficient to show that we can construct v_1 and \tilde{v}_1 in V such that $v \hookrightarrow v_1, \tilde{v} \hookrightarrow \tilde{v}_1, \varphi_1 = \tilde{\varphi}_1$ and $d(v_1, \tilde{v}_1) \leq d(v, \tilde{v}) - 1$. The conclusion would then follow by iterating this procedure until $d(v_K, \tilde{v}_K) = 0$, i.e. $v_K = \tilde{v}_K$. At this point, we have constructed a chain from v to v_K , and another one from \tilde{v} to \tilde{v}_K . Joining them, we obtain a chain from v to \tilde{v} and the conclusion follows.

Therefore, let us indicate how to construct v_1 and \tilde{v}_1 . Since $\theta \neq \tilde{\theta}$ but $\varphi = \tilde{\varphi}$ there exists a jump $g \in G_\nu$ and two time intervals I_k^h and I_l^h such that $\theta(k, g) \geq \tilde{\theta}(k, g) + 1$ and $\theta(l, g) \leq \tilde{\theta}(l, g) - 1$. Moreover, thanks to (H2) there exists at least one skeleton φ^* in the basis \mathcal{B} containing the jump g : $\varphi^*(g) > 0$. Consider now any time repartition $\theta_1 \in \Theta_{\varphi^*}^h$ such that $\theta_1(l, g) \geq 1$. We then construct $\tilde{\theta}_1$ as follows:

$$\tilde{\theta}_1 = \theta_1 + \mathbf{1}_{\{(k,g)\}} - \mathbf{1}_{\{(l,g)\}}.$$

It is simple to check that $v_1 := (\varphi + \varphi^*, \theta + \theta_1), \tilde{v}_1 := (\varphi + \varphi^*, \tilde{\theta} + \tilde{\theta}_1)$ fulfill the desired requirements. By construction, $v \hookrightarrow v_1, \tilde{v} \hookrightarrow \tilde{v}_1$ and v_1, \tilde{v}_1 have the same skeleton. Moreover

$$|\theta + \theta_1 - (\tilde{\theta} + \tilde{\theta}_1)| = |\theta - \tilde{\theta}| - \mathbf{1}_{\{(k,g),(l,g)\}} \quad (5.35)$$

so that $d(v_1, \tilde{v}_1) = d(v, \tilde{v}) - 2$.

Case ii) The skeletons φ and $\tilde{\varphi}$ differ.

We first observe that, if $\varphi, \tilde{\varphi} \in \mathcal{L}_y^+$ thus $\varphi - \tilde{\varphi} \in \mathcal{L}$. Since \mathcal{B} is a basis of the lattice \mathcal{L} (see (H1)), there exist $(\varphi_j^*)_{j=1}^K, (\tilde{\varphi}_i^*)_{i=1}^{\tilde{K}} \subseteq \mathcal{B}$ such that

$$\varphi + \sum_{j=1}^K \varphi_j^* = \tilde{\varphi} + \sum_{i=1}^{\tilde{K}} \tilde{\varphi}_i^*. \quad (5.36)$$

Let us now choose for all j and i a time repartition $\theta_j \in \Theta_{\varphi_j^*}^h$ and $\tilde{\theta}_i \in \Theta_{\tilde{\varphi}_i^*}^h$. It is straightforward to verify that, if we define

$$v_0 = v, v_j := (\varphi + \sum_{j'=1}^j \varphi_{j'}^*, \theta + \sum_{j'=1}^j \theta_{j'}), \tilde{v}_0 = \tilde{v}, \tilde{v}_i := (\tilde{\varphi} + \sum_{i'=1}^i \tilde{\varphi}_{i'}^*, \tilde{\theta} + \sum_{i'=1}^i \tilde{\theta}_{i'}) \quad (5.37)$$

then $(v_j)_{j=0}^K, (\tilde{v}_i)_{i=0}^{\tilde{K}}$ are two sequences connecting v to v_K and \tilde{v} to $\tilde{v}_{\tilde{K}}$. By construction $v_K, \tilde{v}_{\tilde{K}}$ have the same skeleton and one can use case i) again. □

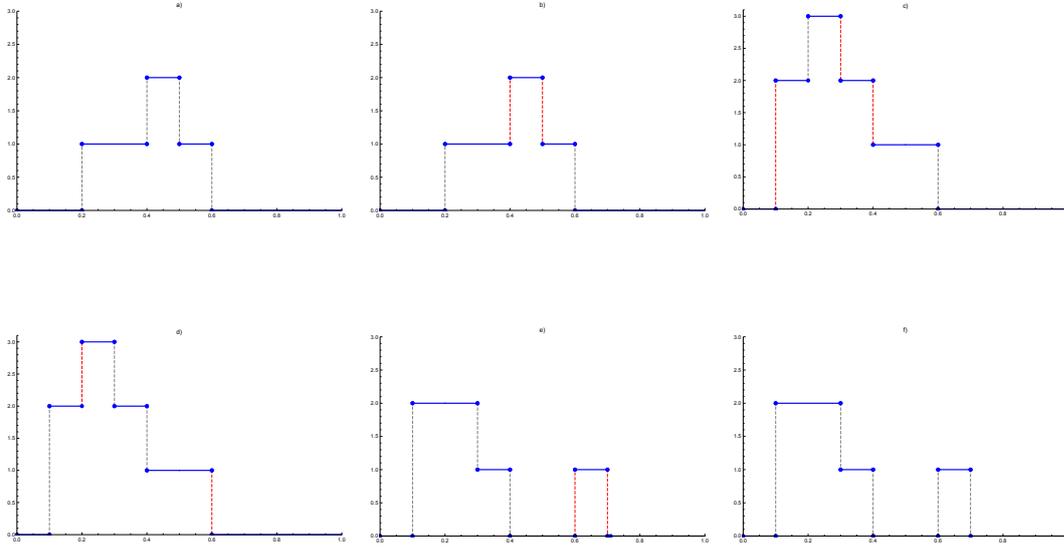


Figure 5.2: In this picture we illustrate by an example the proof of Lemma 20. Take $G = (\mathbb{Z}, +)$, and $G_\nu = \{-1, 1, 2\}$, situation which is treated in Section 3.1.1. $\mathcal{B} = \{\varphi_1, \varphi_2\}$, where $\varphi_1 := \mathbf{1}_1 + \mathbf{1}_{-1}$ and $\varphi_2 := \mathbf{1}_2 + 2\mathbf{1}_{-1}$, is a basis fulfilling H1) and H2). The picture shows how to transform the path a) in the path f) by mean of addition and cancellation of loops whose skeleton belongs to \mathcal{B} . All loops that are either added or removed are denoted by red dashed lines, which correspond to their jumps. At first, following case ii), we have to modify the loop a) to match its skeleton $(2,2,0)$ with that of f), $(3,1,1)$. Therefore in b) we remove a loop with skeleton φ_1 , then in c) add back a loop with skeleton φ_2 . The skeleton is now the desired one. Now we follow case i): we shift one jump of height -1 and one of height 1 further right. Since those two jumps form a loop with skeleton φ_1 we simply delete them in d) and add a new loop with the same skeleton, but now with the desired jump times in e).

In the next lemma we compare the probability of the paths in $\Omega_{y,\varphi}^{h,\theta}$ and those obtained by adding a loop with skeleton $\varphi^* \in \mathcal{B}$, under P and under R .

Lemma 5.4.2. *Let $y \in G, h \in \mathbb{N}^*, \varphi \in \mathcal{L}_y^+, \theta \in \Theta_\varphi^h$ be fixed. Suppose (5.23) holds under P . Then, for any $\varphi^* \in \mathcal{B}$ and $\theta^* \in \Theta_{\varphi^*}^h$,*

$$\frac{P(\Omega_{y,\varphi+\varphi^*}^{h,\theta+\theta^*})}{R(\Omega_{y,\varphi+\varphi^*}^{h,\theta+\theta^*})} = \frac{P(\Omega_{y,\varphi}^{h,\theta})}{R(\Omega_{y,\varphi}^{h,\theta})}. \quad (5.38)$$

Proof. Take an arbitrary ordering of the support of θ^* : $(k_1, g_1), \dots, (k_N, g_N)$. To simplify the notation, we write θ_j (resp. θ_j^*) for $\theta(k_j, g_j)$ (resp $\theta^*(k_j, g_j)$). Consider the test function $F(X, \bar{\gamma}) = f(X) v(\bar{\gamma})$, where

$$f = \mathbf{1}_{\Omega_{y, \varphi + \varphi^*}^{h, \theta + \theta^*}}, \text{ and } v(\bar{\gamma}) = \mathbf{1}_{\Omega_{e, \varphi^*}^{h, \theta^*}}(\chi^{\bar{\gamma}}). \quad (5.39)$$

It is straightforward that

$$f \circ \Psi_{\bar{\gamma}}(X) v(\bar{\gamma}) = \mathbf{1}_{\Omega_{y, \varphi}^{h, \theta}}(X) v(\bar{\gamma}) \quad P \otimes \rho^n \text{ a.e.} \quad (5.40)$$

Therefore, since $F \in \mathcal{H}_{\varphi^*}$, (5.23) holds and its right hand side rewrites as

$$\Phi_{\varphi^*}^\nu \left(\int_{\Gamma^n} v(\bar{\gamma}) (dt \otimes \lambda)^{\otimes n} (d\bar{\gamma}) \right) P(\Omega_{y, \varphi}^{h, \theta}). \quad (5.41)$$

Concerning the left hand side, let us first rewrite it as

$$P \left(f(X) \int_{\Gamma^n \setminus \Delta_n} v(\bar{\gamma}) M_X^{\otimes n} (d\bar{\gamma}) \right). \quad (5.42)$$

Our aim is to show by a direct computation that the (discrete) stochastic integral $\int_{\Gamma^n \setminus \Delta_n} v(\bar{\gamma}) dM_X^{\otimes n} (d\bar{\gamma})$ is actually constant for that choice of v if $X \in \Omega_{y, \varphi + \varphi^*}^{h, \theta + \theta^*}$.

First, we observe that an atom $\bar{\gamma} \in \Gamma^n \setminus \Delta_n$ of $M_X^{\otimes n}$ contributes (with the value 1) to the integral if and only if $\chi^{\bar{\gamma}} \in \Omega_{e, \varphi^*}^{h, \theta^*}$, that is if

$$\#\{i : \gamma_i \in I_{k_j}^h \times \{g_j\}\} = \theta_j^*, \quad 1 \leq j \leq N. \quad (5.43)$$

We then need to count the atoms of $M_X^{\otimes n}$ satisfying (5.43). This is equivalent to count all ordered lists of $n = \ell(\varphi^*)$ atoms of M_X verifying that

- 1) the list contains no repetitions
- 2) for all $1 \leq j \leq N$, the number of elements in the list which belong to $I_{k_j}^h \times \{g_j\}$ is θ_j^* .

Therefore, for each j , we first choose a subset of cardinality θ_j^* among $\theta_j + \theta_j^*$ elements (recall that $X \in \Omega_{y, \varphi + \varphi^*}^{h, \theta + \theta^*}$). To do that we have $\binom{\theta_j + \theta_j^*}{\theta_j^*}$ choices. Then we decide how to sort the list, and for this, there are $n!$ possibilities.

Therefore

$$\mathbf{1}_{\Omega_{y, \varphi + \varphi^*}^{h, \theta + \theta^*}}(X) \int_{\Gamma^n \setminus \Delta_n} v(\bar{\gamma}) M_X^{\otimes n} (d\bar{\gamma}) = \mathbf{1}_{\Omega_{y, \varphi + \varphi^*}^{h, \theta + \theta^*}}(X) n! \prod_{j=1}^N \binom{\theta_j + \theta_j^*}{\theta_j^*} \quad (5.44)$$

and (5.23) rewrites as

$$\Phi_{\varphi^*}^\nu \int_{\Gamma^n} v(\bar{\gamma})(dt \otimes \lambda)^{\otimes n}(d\bar{\gamma}) P(\Omega_{y,\varphi}) = n! \prod_{j=1}^N \begin{pmatrix} \theta_j + \theta_j^* \\ \theta_j^* \end{pmatrix} P(\Omega_{y,\varphi+\varphi^*}^{h,\theta+\theta^*}). \quad (5.45)$$

Since equation (5.23) holds under R , equation (5.45) holds under R as well. Since R gives positive probability to both events $\Omega_{y,\varphi}$ and $\Omega_{y,\varphi+\varphi^*}^{h,\theta+\theta^*}$, the identity (5.38) follows. \square

Remark that, with the notation of the above lemma, if we define $v := (\varphi, \theta)$ and $w := (\varphi + \varphi^*, \theta + \theta^*)$, then $v \leftrightarrow w$.

Lemma 5.4.1 allows us to extend the conclusion of Lemma 5.4.2 to the whole set of skeletons, as we will prove now.

Lemma 5.4.3. *Let $y \in G$, $h \in \mathbb{N}^*$, $\varphi, \tilde{\varphi} \in \mathcal{L}_y^+$, $\theta \in \Theta_\varphi^h$, $\tilde{\theta} \in \Theta_{\tilde{\varphi}}^h$ be fixed. Suppose (5.23) holds under P . Then,*

$$\frac{P(\Omega_{y,\tilde{\varphi}}^{h,\tilde{\theta}})}{R(\Omega_{y,\tilde{\varphi}}^{h,\tilde{\theta}})} = \frac{P(\Omega_{y,\varphi}^{h,\theta})}{R(\Omega_{y,\varphi}^{h,\theta})}. \quad (5.46)$$

Proof. We observe that $v = (\varphi, \theta)$ and $\tilde{v} = (\tilde{\varphi}, \tilde{\theta})$ are elements of V . As proved above, there exists a connecting sequence $(v_i)_{i=0}^K := (\varphi_i, \theta_i)_{i=0}^K$, with $v_0 = v, v_K = \tilde{v}$, linking v to \tilde{v} , and such that either $v_i \leftrightarrow v_{i+1}$ or $v_i \hookrightarrow v_{i+1}$. This entitles us to apply recursively Lemma 5.4.2 to any pair v_i, v_{i+1} and obtain

$$\frac{P(\Omega_{y,\varphi_{i+1}}^{h,\theta_{i+1}})}{R(\Omega_{y,\varphi_{i+1}}^{h,\theta_{i+1}})} = \frac{P(\Omega_{y,\varphi_i}^{h,\theta_i})}{R(\Omega_{y,\varphi_i}^{h,\theta_i})} = \dots = \frac{P(\Omega_{y,\varphi}^{h,\theta})}{R(\Omega_{y,\varphi}^{h,\theta})}. \quad (5.47)$$

The conclusion follows with $i = N - 1$. \square

We can now complete the proof of the converse implication of the main theorem.

Fix $x, y \in G$ with $P(X_0 = x, X_1 = y) > 0$. W.l.o.g. we assume that $x = e$. Thanks to Lemma 5.4.3 we know that for any mesh h , there exists a positive constant c_h such that

$$P(\Omega_{y,\varphi}^{h,\theta}) = c_h R(\Omega_{y,\varphi}^{h,\theta}), \quad \forall \varphi \in \mathcal{L}_y^+, \theta \in \Theta_\varphi^h. \quad (5.48)$$

Now we show that the proportionality constant does not depend on the scale of the time discretisation: $c_h = c_{h+1}$. To this aim, let us observe that

$$P(\Omega_y) = \sum_{(\varphi,\theta) \in V} P(\Omega_{y,\varphi}^{h,\theta}) = \sum_{(\varphi,\theta) \in V} c_h R(\Omega_{y,\varphi}^{h,\theta}) = c_h R(\Omega_y). \quad (5.49)$$

In the same way one gets $P(\Omega_y) = c_{h+1}R(\Omega_y)$ which implies that $c_h = c_{h+1}$. Therefore, there exists a constant $c > 0$ such that

$$P(\Omega_{y,\varphi}^{h,\theta}) = c R(\Omega_{y,\varphi}^{h,\theta}), \quad \forall h \in \mathbb{N}^*, \varphi \in \mathcal{L}_y^+, \theta \in \Theta_\varphi^h. \quad (5.50)$$

By standard approximation arguments this implies the equality between P and cR on $\Omega_y \cap \mathcal{F}$ which then implies $P^{ey} = R^{ey}$. The conclusion follows. \square

Remark 5.4.1. Consider the identities (5.23) for $G = \mathbb{R}^d$ and compute them for particular test functions F which only depend on the skeleton of the paths. These equations, indexed by the skeletons in \mathcal{B} , then characterize the (marginal) distribution of the random vector defined as the number of jumps of any type occurred during the time interval $[0, 1]$, as it was done in [20]. Note that for the unconstrained random walk the distribution of this random vector is a multivariate Poisson law, see e.g. [20] Section 2.2.1.

5.5 Examples

In this section, we present several examples of random walks defined on finite or infinite Abelian groups G .

For each example, we verify if assumptions (H1) and (H2) are satisfied by computing a basis \mathcal{B} of skeleton of loops. We give explicitly the associated reciprocal characteristics (5.24). In some cases, we also write down the integral formula (5.23), highlighting how it is influenced by the geometrical properties of the underlying group G .

Finally, using Corollary 5.3.1 we answer the question of whether a μ -random walk belongs or not to the reciprocal class of the reference walk R . (Recall for definition of μ random walk as given by (5.1)). To denote a generic μ -random walk we adopt the convenient notation P_μ . But this does not apply to the reference walk, which is a ν -random walk, and we keep calling it R rather than P_ν . That is, we solve equation (5.25) and identify the set of probability measures:

$$\mathfrak{R}(R) \cap \{P_\mu : \mu \text{ finite measure on } G_\nu\}. \quad (5.51)$$

We will see that, in some cases, this set reduces to the singleton R and in other cases, this set is non trivial.

5.5.1 The group is infinite

The finite support G_ν of the jump measure ν contains $\{-1, 1\}$.

For any $z \in G_\nu \setminus \{1\}$ we define on G_ν the non negative map φ_z as follows:

$$\varphi_z = \mathbf{1}_z + |z| \mathbf{1}_{-sgn(z)}.$$

It corresponds to the skeleton of paths with one jump of type z and $|z|$ jumps of type $-sgn(z)$. As candidate for the lattice basis of \mathcal{L} , we propose

$$\mathcal{B} := \{\varphi_z\}_{z \in G_\nu \setminus \{1\}}. \quad (5.52)$$

Assumption (H2) is trivially satisfied and it is clear that the elements of \mathcal{B} are linearly independent. Therefore we only need to check if \mathcal{B} spans \mathcal{L} , that is, if for each $\phi \in \mathcal{L}$, there exist integer coefficients $\alpha_z \in \mathbb{Z}$ such that

$$\forall \bar{z} \in G_\nu, \quad \phi(\bar{z}) = \sum_{\substack{z \in G_\nu \\ z \neq 1}} \alpha_z \varphi_z(\bar{z}). \quad (5.53)$$

We now verify that the following choice is the right one:

$$\text{For } z \in G_\nu \setminus \{-1, +1\} \alpha_z = \phi(z) \text{ and } \alpha_{-1} = \phi(-1) - \sum_{\substack{z \in G_\nu \\ z > 1}} z \phi(z). \quad (5.54)$$

- $\bar{z} \notin \{-1, +1\}$. Since $\varphi_{\bar{z}}$ is the only element of \mathcal{B} whose support contains \bar{z} , we have

$$\phi(\bar{z}) = \alpha_{\bar{z}} \varphi_{\bar{z}}(\bar{z}) = \sum_{\substack{z \in G_\nu \\ z \neq 1}} \alpha_z \varphi_z(\bar{z}) \quad (5.55)$$

- $\bar{z} = -1$. Notice that -1 belongs to the support of any φ_z , as soon as $z > 1$. Therefore

$$\phi(-1) = \sum_{\substack{z \in G_\nu \\ z > 1}} \phi(z)z + \alpha_{-1} = \sum_{\substack{z \in G_\nu \\ z > 1}} \alpha_z \varphi_z(-1) + \alpha_{-1} \varphi_{-1}(-1) = \left(\sum_{\substack{z \in G_\nu \\ z \neq 1}} \alpha_z \varphi_z \right)(-1). \quad (5.56)$$

- $\bar{z} = 1$. Notice that $+1$ belongs to the support of any φ_z , as soon as $z \leq -1$. Recall that $\phi \in \mathcal{L}$. Therefore

$$\begin{aligned} \phi(1) &= - \sum_{\substack{z \in G_\nu \\ z \neq 1}} \phi(z)z = \sum_{\substack{z \in G_\nu \\ z < 1}} -\phi(z)z + \phi(-1) \\ &= \sum_{\substack{z \in G_\nu \\ z \leq -1}} \alpha_z \varphi_z(1) = \left(\sum_{\substack{z \in G_\nu \\ z \neq 1}} \alpha_z \varphi_z \right)(1). \end{aligned}$$

Let us now compute the reciprocal characteristics associated to each skeleton in \mathcal{B} :

$$\Phi_{\varphi_z}^\nu = \nu(-\text{sgn}(z))^{|z|} \nu(z), \quad z \in G_\nu \setminus \{1\}. \quad (5.57)$$

Finally, thanks to Corollary 5.3.1, we obtain

$$\begin{aligned} P_\mu \in \mathfrak{R}(R) &\Leftrightarrow \forall z \in G_\nu \setminus \{1\}, \quad \mu(-\text{sgn}(z))^{|z|} \mu(z) = \nu(-\text{sgn}(z))^{|z|} \nu(z) \\ &\Leftrightarrow \exists \alpha > 0 \text{ such that } \frac{d\mu}{d\nu}(z) = \alpha^z. \end{aligned}$$

Example 5.5.1. Simple random walks: $G_\nu = \{-1, 1\}$.

Due to the above computations, the basis \mathcal{B} of the lattice reduces to the singleton $\{\varphi_{-1}\}$ and the unique reciprocal characteristics is given by

$$\Phi_{\varphi_{-1}}^\nu = \nu(-1)\nu(1).$$

Therefore the only loops which appear in the integral characterization (5.23) have length $n = \ell(\varphi_{-1}) = 2$.

Test functions of the form

$$F(X, (\gamma_1, \gamma_2)) = f(X) \mathbf{1}_{\{g_1=1, g_2=-1\}} h(t_1, t_2) \quad (5.58)$$

belong to $\mathcal{H}_{\varphi_{-1}}$. Such functions are supported by pairs (γ_1, γ_2) such that the transformation $\Psi_{\gamma_1, \gamma_2}$ adds to any path a jump +1 at time t_1 and a jump -1 at time t_2 . The identity (5.23) now reads as:

$$P\left(f(X) \sum_{\substack{(t_1, t_2): \Delta X_{t_1}=1 \\ \Delta X_{t_2}=-1}} h(t_1, t_2)\right) = \nu(-1)\nu(1) \int_{[0,1]^2} P(f(\Psi_{\gamma_1, \gamma_2} X)) h(t_1, t_2) dt_1 dt_2. \quad (5.59)$$

As in Remark 5.4.1, if we consider test functions f which only depend on the skeletons of the paths, $f(X) = v(\varphi_X)$, we obtain that the distribution $\chi(dn_{-1}, dn_1) \in \mathcal{P}(\mathbb{N}^2)$ of the number n_{-1} (resp. n_1) of negative (resp. positive) jumps is characterized by the system of equations: for all $v \in \mathcal{P}(\mathbb{N}^2)$,

$$\begin{aligned} \int v(n_{-1}, n_1) n_{-1} n_1 \chi(dn_{-1}, dn_1) &= \nu(-1)\nu(1) \int v(n_{-1} + 1, n_1 + 1) \chi(dn_{-1}, dn_1), \\ \chi(n_1 = n_{-1}) &= 1. \end{aligned}$$

This result coincides with [20], Example 2.18.

$$G_\nu = \{1, 2\}.$$

In that case, since -1 does not belong to the support of the jump measure, it leads to a case where (H2) is not satisfied. It is straightforward to prove that the lattice \mathcal{L} is one-dimensional and is equal to $\{z\varphi^*, z \in \mathbb{Z}\}$ where

$$\varphi^*(1) = 2, \quad \varphi^*(2) = -1. \quad (5.60)$$

Clearly \mathcal{L} does not admit a non negative basis.

5.5.2 G is the cyclic group $\mathbb{Z}/N\mathbb{Z}$

We now consider the finite cyclic group $G := \mathbb{Z}/N\mathbb{Z} =: \{0, 1, 2, \dots, N-1\}$.

The support G_ν of the jump measure reduces to $\{-1, 1\}$.

This case corresponds to nearest neighbour random walks. The non negative basis $\mathcal{B} := \{\varphi_{N-1}, \varphi^*\}$ where

$$\varphi_{N-1} = \mathbf{1}_1 + \mathbf{1}_{N-1} = \mathbf{1}_1 + \mathbf{1}_{-1} \text{ and } \varphi^* = N\mathbf{1}_1 \quad (5.61)$$

is suitable. The associated reciprocal characteristics are

$$\Phi_{\varphi_{N-1}}^\nu = \nu(\mathbf{1})\nu(-1) \text{ and } \Phi_{\varphi^*}^\nu = \nu(\mathbf{1})^N. \quad (5.62)$$

The existence of the second invariant $\Phi_{\varphi^*}^\nu$ corresponding to the loop around the cycle $\{0, 1, 2, \dots, N-1\}$ implies that R is the unique nearest neighbor random walk of the reciprocal class $\mathfrak{R}(R)$. This differs from the nearest neighbor random walk on \mathbb{Z} , treated in Example 5.5.1. We proved there that any random walk P_μ , with μ satisfying $\mu(-1)\mu(1) = \nu(-1)\nu(1)$, induces the same reciprocal class.

The distribution χ of the random vector (n_{-1}, n_1) under the $\mathbf{00}$ -bridge is given by the following system of integral equations, satisfied for any test function v on \mathbb{N}^2 :

$$\begin{aligned} & \int v(n_{-1}, n_1) n_{-1} n_1 \chi(dn_{-1}, dn_1) \\ &= \nu(-1)\nu(1) \int v(n_{-1} + 1, n_1 + 1) \chi(dn_{-1}, dn_1), \\ & \int v(n_{-1}, n_1) n_1 \cdots (n_1 - (N-1)) \chi(dn_{-1}, dn_1) \\ &= \nu(\mathbf{1})^N \int v(n_{-1}, n_1 + N) \chi(dn_{-1}, dn_1), \\ & \chi(n_1 - n_{-1} \in N\mathbb{Z}) = 1. \end{aligned}$$

The support G_ν of the jump measure covers $\mathbb{Z}/N\mathbb{Z} \setminus \{0\}$.

We now consider a random walk on $\mathbb{Z}/N\mathbb{Z}$ which can jump anywhere: $G_\nu = \mathbb{Z}/N\mathbb{Z} \setminus \{0\}$. Here, we focus for simplicity on the case $N = 4$, which is the first non trivial example, and disintegrate the jump measure ν as follows:

$$\nu = \nu(\mathbf{1}) \delta_1 + \nu(\mathbf{2}) \delta_2 + \nu(\mathbf{3}) \delta_3, \quad \nu(\mathbf{1})\nu(\mathbf{2})\nu(\mathbf{3}) > 0. \quad (5.63)$$

It can be proven along the same lines as in the previous examples, that a suitable non negative basis for the lattice \mathcal{L} is given by $\mathcal{B} = \{\varphi^*, \eta^*, \xi^*\}$ where

$$\varphi^* = \mathbf{1}_1 + \mathbf{1}_3, \quad \eta^* = 4\mathbf{1}_1, \quad \xi^* = 2\mathbf{1}_1 + \mathbf{1}_2. \quad (5.64)$$

Hence the associated reciprocal characteristics are:

$$\Phi_{\varphi^*}^\nu = \nu(\mathbf{1})\nu(\mathbf{3}), \quad \Phi_{\eta^*}^\nu = \nu(\mathbf{1})^4, \quad \Phi_{\xi^*}^\nu = \nu(\mathbf{1})^2\nu(\mathbf{2}). \quad (5.65)$$

We now turn our attention to the integral formula (5.23). Simple functions $F \in \mathcal{H}_{\xi^*}$ are of the form:

$$F(X, (\gamma_1, \gamma_2, \gamma_3)) = f(X)\mathbf{1}_{\{g_1=g_2=1, g_3=2\}}h(t_1, t_2, t_3). \quad (5.66)$$

$(\gamma_1, \gamma_2, \gamma_3)$ is in the support of F if two jumps of value $\mathbf{1}$ happen at times t_1, t_2 and one jump of value $\mathbf{2}$ at time t_3 , leading to a global null displacement since $4 = \mathbf{0}$. The formula (5.23) reads:

$$P\left(f(X) \sum_{\substack{(t_1, t_2, t_3): t_1 \neq t_2, \\ \Delta X_{t_1} = \Delta X_{t_2} = \mathbf{1} \\ \Delta X_{t_3} = \mathbf{2}}} h(t_1, t_2, t_3)\right) = \nu(\mathbf{1})^2\nu(\mathbf{2})P\left(\int_{[0,1]^3} f(\Psi_{\gamma_1, \gamma_2, \gamma_3} X)h(t_1, t_2, t_3)\right) dt_1 dt_2 dt_3. \quad (5.67)$$

The distribution of the random vector (n_1, n_2, n_3) under the $\mathbf{00}$ -bridge is given by the following identities, valid for any $v : \mathbb{N}^3 \rightarrow \mathbb{R}$:

$$\begin{aligned} \int v(n_1, n_2, n_3)n_1n_3 \chi(dn_1, dn_2, dn_3) &= \nu(\mathbf{1})\nu(\mathbf{3}) \int (v(n_1 + 1, n_2, n_3 + 1)\chi(dn_1, dn_2, dn_3) \\ \int v n_1(n_1 - 1)(n_1 - 2)(n_1 - 3)\chi(dn_1, dn_2, dn_3) &= \nu(\mathbf{1})^4 \int v(n_1 + 4, n_2, n_3)\chi(dn_1, dn_2, dn_3) \\ \int v(n_1, n_2, n_3)n_1(n_1 - 1)n_2\chi(dn_1, dn_2, dn_3) &= \nu(\mathbf{1})^2\nu(\mathbf{2}) \int v(n_1 + 2, n_2 + 1, n_3)\chi(dn_1, dn_2, dn_3) \\ \chi(n_1\mathbf{1} + n_2\mathbf{2} + n_3\mathbf{3} = \mathbf{0}) &= 1. \end{aligned}$$

In this situation, again R is the unique random walk of its reciprocal class.

5.5.3 The state space is a product group

Consider the product of two groups, say G and G' , and two non negative finite measures on them, say ν and ν' , such that in both cases (H1) and (H2) are satisfied. Then, the product group $G \times G'$ equipped with the product measure $\nu \otimes \nu'$ fulfills (H1) and (H2) too. The key idea is as follows: if \mathcal{B} and \mathcal{B}' are suitable basis of G and G' then we can define for all $\eta \in \mathcal{B}$,

$$\varphi_\eta : G_\nu \times G_{\nu'} \rightarrow \mathbb{N}, \quad \varphi_\eta(g, g') = \eta(g) \quad (5.68)$$

and for all $\eta' \in \mathcal{B}'$,

$$\varphi_{\eta'} : G_\nu \times G_{\nu'} \rightarrow \mathbb{N}, \quad \varphi_{\eta'}(g, g') = \eta'(g'). \quad (5.69)$$

The set $\mathcal{B}_\otimes = \{\varphi_\eta\}_{\eta \in \mathcal{B}} \cup \{\varphi_{\eta'}\}_{\eta' \in \mathcal{B}'}$ is an appropriate basis for the lattice of skeletons defined on the product group.

Example 5.5.2. Random walk on the d -dimensional discrete hypercube $(\mathbb{Z}/2\mathbb{Z})^d$.

The d -dimensional discrete hypercube is the d -product of the cyclic group with two elements. We denote by (e_1, \dots, e_d) its canonical basis.

A random walk on the hypercube is defined uniquely through its jump measure $\nu = \sum_{i=1}^d \nu(i) \delta_{e_i}$. Since it can be realized as the product of d random walks on $\mathbb{Z}/2\mathbb{Z}$, the basis $\mathcal{B} := \{\varphi_i^*\}_{1 \leq i \leq d}$, $\varphi_i^* = 2\mathbf{1}_{e_i}$, is a suitable choice.

For the integral characterization it is enough to consider loops of length $\ell = 2$. However, we have here d different skeletons to consider. Test functions of the form

$$F(X, \gamma) = f(X) \mathbf{1}_{\{g_1=g_2=e_i\}} h(t_1, t_2), \quad 1 \leq i \leq d, \quad (5.70)$$

belong to $\mathcal{H}_{\varphi_i^*}$. For any $i \in \{1, \dots, d\}$ fixed, (5.23) reads as:

$$P\left(f(X) \sum_{\substack{(t_1, t_2): t_1 \neq t_2, \\ \Delta X_{t_1} = \Delta X_{t_2} = e_i}} h(t_1, t_2)\right) = \nu(i)^2 \int_{[0,1]^2} P\left(f(\Psi_{\gamma_1, \gamma_2} X)\right) h(t_1, t_2) dt_1 dt_2. \quad (5.71)$$

Concerning the distribution of the random vector $(n_{e_1}, \dots, n_{e_d})$, it has independent marginals χ_i , $1 \leq i \leq d$, which are characterized through the system of equations: for all $v \in \mathcal{P}(\mathbb{N})$,

$$\begin{aligned} \int v(n) n(n-1) \chi_i(dn) &= \nu(i)^2 \int v(n+1) \chi_i(dn), \\ \chi_i(n \in 2\mathbb{N}) &= 1. \end{aligned}$$

Chapter 6

Random walks on a general graph

Outline of the chapter The reciprocal class of a Markov random walk on a general graph is characterized. Duality formulae are no longer a useful tool, because of the lack of symmetry in the graph. We propose two other methods: we use short-time asymptotic probabilities and stochastic calculus. Several examples are discussed. This chapter is based on the submitted paper [21].

In this chapter we study the reciprocal class of a random walk on a general graph. In this broad framework, there is no possibility to use duality formulae as in the previous chapters. This is simply because the graph does not possess the necessary translation-invariant structure, and perturbations of the sample paths cannot be defined in a canonical way. Therefore we rely mainly on Girsanov Theorem, h -transform techniques and short time asymptotic to characterize the reciprocal class. The characteristics are computed in full generality, recovering all the previous computations as subcases. A complete probabilistic interpretation of them is given. Recall that one of the many possible descriptions of a Markov random walk with intensity j (see e.g. Th 3.2 in [60]) and initial distribution μ is by saying that it is the only Markov process whose law at time 0 is given by μ and such that at any time t , if it sits in X_t , it will jump to z in a small time interval of length h with probability $j(X_{t-} \rightarrow z)h + o(h)$.

That is, we prescribe the initial distribution, Markov property and some short time expansions along the arcs.

Here we generalize this to reciprocal walks. We prove that a reciprocal walk in $\mathfrak{R}(R)$ is completely identified by the law of its endpoints, the

reciprocal property and two families of short-time expansions: the first family concerns the distribution of jump times and identifies the *arc characteristics*. The second one concerns the probability of going around cycles and identifies the *cycle characteristics*.

As a by product, we obtain an alternative dynamic characterization of the bridge of the reference walk which, instead of looking at conditioning, tells what properties of the reference walk are carried over to its bridges.

The results of this section, even when restricted to the graphs we already encountered in the previous chapters, bring new results: indeed we do not assume that the reference intensity is space homogeneous. Therefore we also revisit these situations. Concerning sharpness of the results, an effort is also put in trying to optimize over the number of cycles which we use in our characterization. This leads us to introduce what we call a generating set for the set of cycles. In some interesting cases we can compute such sets explicitly: for example, in the case of planar graphs, we shall see that a generating set is given by the faces of the graph. They play a role analogous to that of lattice basis which we discussed in Chapters 4 and 5. Another relevant example where such a generating set can be computed is the discrete hypercube.

We are also interested in the form of the (predictable) intensity associated with a reciprocal process, and we prove some representation results, which extend and clarify some results about h -transforms, which are better known for diffusion processes.

A list of examples is discussed in quite some detail. We consider, in particular: directed trees, birth and death processes, some planar graphs, the hypercube, the complete graph and some Cayley graphs. We calculate their reciprocal characteristics and sometimes solve the associated characteristic equation.

Organization of the material

Next Section 6.1 is devoted to some preliminaries about directed graphs, Markov walks and their intensities. Our main results are stated at Section 6.2. They are Theorems 6.2.1 and 6.2.2, together with their Corollary 6.2.1. Their proofs are given in section 6.3. The key preliminary result is Lemma 6.3.2. Some more results about the elements of the reciprocal class are proved at Section 6.4. In particular, we give at Proposition 6.4.1 another characterization of $\mathfrak{R}(R)$ in terms of the shape of the jump intensity of any element of the class. The characteristic equation (6.25) seen as an equation of the unknown P^x where j is given is also solved at Theorem

6.4.1. Several examples are treated at Section 6.5.

6.1 Preliminaries

We are in the general setting described in section 2.3.1, $(\mathcal{X}, \rightarrow)$ is a directed graph. On it is defined a jump intensity $j : [0, 1] \times \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$. The graph and the intensity together satisfy Assumption 2.3.1. A reference random walk of intensity j and initial measure of full support is given. We want to study its reciprocal class $\mathfrak{R}(R)$. We now introduce some extra notation which complements the one introduced in chapter 2.

Remark 6.1.1. *Since any element of $\mathfrak{R}(R)$ is absolutely continuous with respect to R , by Girsanov's theorem it admits a predictable intensity of jumps, see [36, Thm. 5.1] and the discussion at Section 2.3.3. This will be used constantly in the rest of the chapter.*

6.1.1 Directed subgraphs associated with an intensity

Here we generalize in a coherent way the definition of $\mathcal{A}_{\rightarrow}(j)$ given at point iv) of Assumption 2.3.1. There are two relevant graph structures that are associated with the intensity k . The subgraph of k -active arcs at time t is the subset

$$\mathcal{A}_{\rightarrow}(t, k) := \{(z \rightarrow z') \in \mathcal{A} : k(t, z \rightarrow z') > 0\}$$

and its symmetric extension is denoted by

$$\mathcal{A}_{\leftrightarrow}(t, k) := \{(z \rightarrow z'), (z' \rightarrow z); (z \rightarrow z') \in \mathcal{A}_{\rightarrow}(t, k)\}.$$

Only Markov intensities such that these structures do not depend on time will be encountered, see (6.1) below.

Definition 6.1.1 (The directed subgraphs associated to k). *In the situation where $\mathcal{A}_{\rightarrow}(t, k)$ does not depend on t , i.e.*

$$\mathcal{A}_{\rightarrow}(t, k) = \mathcal{A}_{\rightarrow}(k), \quad \forall t \in [0, 1), \quad (6.1)$$

$\mathcal{A}_{\rightarrow}(k)$ is called the subgraph of k -active arcs and $\mathcal{A}_{\leftrightarrow}(k)$ denotes its symmetric extension.

The symmetrized subgraph will be necessary for considering cycles.

6.1.2 Directed subgraphs associated with a random walk

If $P \in \mathcal{P}(\Omega)$ is a Markov random walk with an intensity k that satisfies (6.1). Since $(\mathcal{X}, \rightarrow)$ of bounded degree, and P -almost every sample path performs finitely many jumps, the assumption (6.1) implies that the support of $P_t \in \mathcal{P}(\mathcal{X})$ remains constant for each time $0 < t < 1$. We denote this set by

$$\mathcal{X}(P) := \text{supp}(P_{1/2}) \subset \mathcal{X}.$$

It is the set of all vertices that are visited by the random walk P . Note that the initial and final times $t = 0$ and $t = 1$ are excluded to allow for P to be a bridge. We always have $\text{supp}P_0 \subseteq \mathcal{X}(P)$ and $\text{supp}P_1 \subseteq \mathcal{X}(P)$; these inclusions may be strict.

The directed subgraph $(\mathcal{X}(P), \mathcal{A}_{\rightarrow}(P))$ of all the P -active arcs is the subgraph of $(\mathcal{X}, \mathcal{A}_{\rightarrow}(k))$ which is defined by

$$\mathcal{A}_{\rightarrow}(P) := \{(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(k) : z, z' \in \mathcal{X}(P)\}.$$

Its symmetric extension is

$$\mathcal{A}_{\leftrightarrow}(P) := \{(z \rightarrow z'), (z' \rightarrow z); (z \rightarrow z') \in \mathcal{A}_{\rightarrow}(k), z, z' \in \mathcal{X}(P)\}.$$

Let us provide some comment to make the relation between $\mathcal{A}_{\rightarrow}(k)$ and $\mathcal{A}_{\rightarrow}(P)$ clearer. If the initial marginal P_0 is supported by a proper subset of \mathcal{X} , it might happen that $\mathcal{A}_{\rightarrow}(P)$ is a proper subset of $\mathcal{A}_{\rightarrow}(k)$ and also that $\mathcal{A}_{\leftrightarrow}(P)$ is a proper subset of $\mathcal{A}_{\leftrightarrow}(k)$. For instance, let k be the intensity of the Poisson process on $\mathcal{X} = \mathbb{Z}$ given by $k(t, n \rightarrow n+1) = \lambda > 0$ for all $0 \leq t \leq 1$ and $n \in \mathbb{Z}$. Let P be the Poisson random walk with intensity λ and initial state $n_0 \in \mathbb{Z}$. Then, $\mathcal{A}_{\rightarrow}(k) = \{(n \rightarrow n+1); n \in \mathbb{Z}\}$ and $\mathcal{A}_{\rightarrow}(P) = \{(n \rightarrow n+1); n \in \mathbb{Z}, n \geq n_0\}$.

Definition 6.1.2 (About the j -active arcs from x to \mathcal{Y}). *Let $x \in \mathcal{X}$ be any vertex and \mathcal{Y} be any nonempty subset of $\text{supp}R_1^x$.*

(i) *We define the subgraph*

$$\mathcal{A}_{\rightarrow}^R(x, \mathcal{Y}) := \cup_{y \in \mathcal{Y}} \mathcal{A}_{\rightarrow}(R^{xy})$$

of all the arcs that constitute the $\mathcal{A}_{\rightarrow}(j)$ -walks from x to \mathcal{Y} .

(ii) *We denote $\mathcal{A}_{\leftrightarrow}^R(x, \mathcal{Y})$ its symmetric extension.*

(iii) *We define the set*

$$\mathcal{X}^R(x, \mathcal{Y}) := \text{proj}_{\mathcal{X}} \mathcal{A}_{\rightarrow}^R(x, \mathcal{Y})$$

of all vertices visited by the $\mathcal{A}_{\rightarrow}$ -walks starting at x and ending in \mathcal{Y} .

The $\mathcal{A}_{\rightarrow}^R(x, \mathcal{Y})$ is constructed by considering all arcs of $(\mathcal{X}, \rightarrow)$ that can be traversed by a bridge R^{xy} whose final state belongs to \mathcal{Y} . $\mathcal{X}^R(x, \mathcal{Y})$ is then the set of all vertices touched by arcs in $\mathcal{A}_{\rightarrow}^R(x, \mathcal{Y})$.

Remark 6.1.2. Remark that $\{x\}$ and \mathcal{Y} may be proper subsets of $\mathcal{X}^R(x, \mathcal{Y})$ as the example of a bridge R^{xy} suggests in many situations. We also have for any $P \in \mathcal{P}(\Omega)$,

$$(\{x\} \cup \text{supp}P_1^x) \subset \mathcal{X}(P^x) = \text{proj}_{\mathcal{X}} \mathcal{A}_{\rightarrow}(P^x)$$

where the inclusion may be strict.

6.1.3 Gradients and generating sets of cycles

We provide some basics about gradients on graphs. Most of the notions are close to what is known for gradients of functions in \mathbb{R}^d .

Definition 6.1.3 (gradient, cycles of length two). (i) Let $\mathcal{A}_{\leftrightarrow}$ be a symmetric set of arcs. The set of cycles of length two is denoted \mathcal{E}

$$\mathcal{E} = \{(z \rightarrow z' \rightarrow z); (z \rightarrow z') \in \mathcal{A}_{\leftrightarrow}\} \quad (6.2)$$

(ii) Let $\mathcal{A} \subset \mathcal{X}$ specify a directed graph $(\mathcal{X}, \rightarrow)$ on \mathcal{X} with. An arc function $\ell : \mathcal{A} \rightarrow \mathbb{R}$ is the gradient of the vertex function $\psi : \mathcal{X} \rightarrow \mathbb{R}$ if

$$\ell(z \rightarrow z') = \psi(z') - \psi(z), \quad \forall (z \rightarrow z') \in \mathcal{A}. \quad (6.3)$$

(iii) For any arc function ℓ , and any walk $\mathbf{w} = (x_0 \rightarrow \dots \rightarrow x_n)$ we denote $\ell(\mathbf{w}) := \ell(x_0 \rightarrow x_1) + \dots + \ell(x_{n-1} \rightarrow x_n)$.

The following is a well known result.

Lemma 6.1.1. Let $\mathcal{A}_{\leftrightarrow}$ be a symmetric set of arcs. The function $\ell : \mathcal{A}_{\leftrightarrow} \rightarrow \mathbb{R}$ is a gradient if and only if $\ell(\mathbf{c}) = 0$, for any closed $\mathcal{A}_{\leftrightarrow}$ -walk \mathbf{c} .

Proof. At first, let us observe that we can assume w.l.o.g. that $(\mathcal{X}, \mathcal{A})$ is connected. Otherwise, we can simply repeat the proof for each connected component. Since $\mathcal{A}_{\leftrightarrow}$ is a symmetric set, connected components are defined in a natural way, and there is no ambiguity. If ℓ is the gradient of ψ , then $\ell(\mathbf{w}) = \psi(x_n) - \psi(x_0)$, which vanishes when $\mathbf{w} = (x_0 \rightarrow \dots \rightarrow x_n)$ is a cycle.

Conversely, let ℓ be such that $\ell(\mathbf{c}) = 0$, for all $\mathcal{A}_{\leftrightarrow}$ -cycle \mathbf{c} . As $(z \rightarrow z' \rightarrow z)$ is a cycle, we have

$$\ell(z \rightarrow z') + \ell(z' \rightarrow z) = 0, \quad \forall z \leftrightarrow z' \in \mathcal{X}. \quad (6.4)$$

Choose a tagged vertex $* \in \mathcal{X}$, set $\psi(*) = 0$ and for any $x \neq *$, define

$$\psi(x) := \ell(\mathbf{w}), \quad \text{for any } \mathbf{w} \in \{(* \rightarrow x_1 \rightarrow \cdots \rightarrow x_n = x), \text{ for some } n \geq 1\}.$$

To see that this is a meaningful definition, take two paths $\mathbf{w} = (* \rightarrow x_1 \cdots \rightarrow x_n)$ and $\mathbf{w}' = (* \rightarrow y_1 \cdots \rightarrow y_m)$ such that $x_n = y_m = x$. As $(* \rightarrow x_1 \cdots \rightarrow x_n = x = y_m \rightarrow y_{m-1} \rightarrow \cdots \rightarrow *)$ is a cycle, we have $0 = \ell(* \rightarrow x_1 \cdots \rightarrow x) + \ell(x \rightarrow y_{m-1} \rightarrow \cdots \rightarrow *) = \ell(\mathbf{w}) - \ell(\mathbf{w}')$, where the last equality is obtained with (6.4). Therefore, ψ is well defined. Finally, it follows immediately from our definition of ψ that $\ell(z \rightarrow z') = \psi(z') - \psi(z)$, for all $(z \rightarrow z') \in \mathcal{A}_{\leftrightarrow}$. \square

We introduce the notion of generating set of cycles which turns out to be useful when deriving sharp characterizations of reciprocal classes. It generalizes the notion of lattice set and generating set for a lattice, which we widely use in Chapter 4 and 5.

Definition 6.1.4 (Generating set of \mathcal{C}). *We say that a subset \mathcal{C}_o of the set of cycles \mathcal{C} generates \mathcal{C} if for any arc function $\ell : \mathcal{A}_{\leftrightarrow} \rightarrow \mathbb{R}$, we have:*

$$\ell(\mathbf{c}) = 0, \quad \forall \mathbf{c} \in \mathcal{C}_o \Rightarrow \ell(\mathbf{c}) = 0, \quad \forall \mathbf{c} \in \mathcal{C}.$$

Let us point out that we do not ask \mathcal{C}_o to be minimal with respect to the inclusion. The whole set of cycles \mathcal{C} is always a generating set. Of course, the smaller \mathcal{C}_o is, the sharper is the characterization of the reciprocal class. In this definition it is not assumed that \mathcal{A} is symmetric.

6.2 Main results

Before stating the main results of the chapter, we still need to define the reciprocal characteristics. The definition we are going to give in a few lines generalizes all the previous ones.

6.2.1 Reciprocal characteristics

Definition 6.2.1 (Reciprocal characteristics of a Markov random walk). *Let $P \in \mathcal{P}(\Omega)$ be a Markov random walk with its jump intensity k which is assumed*

to satisfy (6.1) and to be continuously t -differentiable, i.e. for any $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(P)$ the function $t \mapsto k(t, z \rightarrow z')$ is continuously differentiable on the semi-open time interval $[0, 1)$.

(i) We define for all $t \in [0, 1)$ and all $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(P)$,

$$\chi_a[P](t, z \rightarrow z') := \partial_t \log k(t, z \rightarrow z') + \bar{k}(t, z') - \bar{k}(t, z)$$

where $\bar{k}(t, z)$ is the total intensity of jump, see (2.11).

(ii) We define for all $t \in [0, 1)$ and any cycle $\mathbf{c} = (x_0 \rightarrow \cdots \rightarrow x_{|\mathbf{c}|} = x_0)$ on the directed graph $(\mathcal{X}, \mathcal{A}_{\leftrightarrow}(P))$ associated with the symmetric extension $\mathcal{A}_{\leftrightarrow}(P)$ of $\mathcal{A}_{\rightarrow}(P)$,

$$\chi_c[P](t, \mathbf{c}) := \prod_{(x_i \rightarrow x_{i+1}) \in \mathcal{A}_{\rightarrow}(P)} k(t, x_i \rightarrow x_{i+1}) / \prod_{(x_i \rightarrow x_{i+1}) \in \mathcal{A}_0(P)} k(t, x_{i+1} \rightarrow x_i)$$

where

$$\mathcal{A}_0(P) := \mathcal{A}_{\leftrightarrow}(P) \setminus \mathcal{A}_{\rightarrow}(P) = \{(z \rightarrow z') \in \mathcal{A}_{\leftrightarrow}(P) : k(t, z \rightarrow z') = 0, \forall t \in [0, 1)\}$$

is the set of all k -inactive arcs. No graph structure is associated to $\mathcal{A}_0(P)$.

(iii) We call $\chi[P] = (\chi_a[P], \chi_c[P])$ the reciprocal characteristic of P . The term $\chi_a[P]$ is the arc component and $\chi_c[P]$ is the cycle component of $\chi[P]$.

(iv) We often write

$$\chi[R] =: \chi[j]$$

to emphasize the role of the reference intensity j .

(v) A cycle \mathbf{c} as in item (b) above is shortly called a $\mathcal{A}_{\leftrightarrow}(P)$ -cycle.

Note that no division by zero occurs and that under our regularity assumption on k , ∂_t acts on a differentiable function: $\chi[P]$ is well defined.

Consistency with all the previous definitions We shall recover the definitions we gave in Chapter 3,4,5 as special cases of Definition 6.2.1. Concerning Chapter 3, let us observe that the graph is $(\mathbb{Z}, \rightarrow)$, with $z \rightarrow z' \Leftrightarrow z' = z + 1$. The cycle characteristic is trivially 1 on this graph, and since the only arc outgoing from z goes towards $z + 1$ we have that $\bar{j}(t, z) = j(t, z)$ for all t, z , where $j(t, z)$ is the convention for the intensity we adopted in

Chapter 3. Therefore the arc characteristic at point (i) of Definition (6.2.1) coincides with that of Definition 3.2.1.

In chapter 4, $(\mathcal{X}, \rightarrow)$ was a lattice. The arc characteristic introduced in Definition 4.2.2 along an arc of the type $(x \rightarrow x + a^l)$ is worth

$$j^l(t)/j^l(s).$$

Here, according to our definition, since in the hypothesis of Chapter 4 the total intensity is constant among all vertices at any time, we have:

$$\chi_a[R](t, x \rightarrow x + a^l) = \partial_t \log j(t, x \rightarrow x + a^l) = \partial_t \log j^l(t)$$

This definition is not numerically equal to $j^l(t)/j^l(s)$ but, reasoning as in Remark 4.2.3 we get their equivalence, in the sense that they encode the same information on j .

Concerning Definition 4.3.1, using the notation of Chapter 4, for any \mathbf{c} in the lattice $\ker_{\mathbb{Z}}(\mathbf{A})$, we can choose a function $\eta : \{1, \dots, |\mathbf{c}|\} \rightarrow \{1, \dots, A\}$ with the property that:

$$\#\{m : \eta(m) = l\} = c^l \quad \forall 1 \leq l \leq A$$

and define, for an arbitrary $x_0 \in \mathcal{X}$, the cycle $\tilde{\mathbf{c}}$ of $(\mathcal{X}, \rightarrow)$

$$\tilde{\mathbf{c}} := (x_0 \rightarrow x_0 + a^{\eta_1} \rightarrow x_0 + a^{\eta_1} + a^{\eta_2} \rightarrow \dots \rightarrow x_0)$$

Then we have:

$$\Phi_j^{\tilde{\mathbf{c}}} = \chi_{\tilde{\mathbf{c}}}[R](t, \tilde{\mathbf{c}})$$

An argument completely identical to the last one yields that the definitions of Chapter 5 can be inscribed in the current framework.

6.2.2 The main results

They are stated at Theorems 6.2.1, 6.2.2 and Corollary 6.2.1. Theorem 6.2.1 gives a characterization of the reciprocal class of j in terms of the reciprocal characteristics. Theorem 6.2.2 provides an interpretation of the reciprocal characteristics of a reciprocal walk by means of short-time asymptotic expansions of some conditional probabilities. Putting together these theorems leads us to Corollary 6.2.1 which states a characterization of the reciprocal class in terms of these short-time asymptotic expansions.

Theorem 6.2.1 (Characterization of $\mathfrak{R}(R)$). *Let $(\mathcal{X}, \rightarrow)$ and j satisfy Assumption 2.3.1.*

A random walk $P \in \mathcal{P}(\Omega)$ belongs to $\mathfrak{R}(R)$ if and only if the following assertions hold

- (i) It is a reciprocal walk and $P \ll R$
- (ii) For all $x \in \text{supp}P_0$ the conditioned random walk P^x admits an intensity $k^x(t, z \rightarrow z')$ which is t -differentiable on $[0, 1)$.
- (iii) The subgraph of all P^x -active arcs doesn't depend on t and is given by

$$\mathcal{X}(P^x) = \mathcal{X}^R(x, \text{supp}P_1^x), \quad \mathcal{A}_{\rightarrow}(P^x) = \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x).$$

- (iv) For any $t \in [0, 1)$ and any $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x)$, we have

$$\chi_a[P^x](t, z \rightarrow z') = \chi_a[j](t, z \rightarrow z'). \quad (6.5)$$

- (v) For any $t \in [0, 1)$ and any $\mathcal{A}_{\leftrightarrow}^R(x, \text{supp}P_1^x)$ -cycle \mathbf{c} , we have

$$\chi_{\mathbf{c}}[P^x](t, \mathbf{c}) = \chi_{\mathbf{c}}[j](t, \mathbf{c}). \quad (6.6)$$

Remark 6.2.1. In point (ii) we tacitly assumed that the intensity is of the form $k^x(t, z \rightarrow z')$, i.e. it is the intensity of a Markov walk. This is not a contradiction. Indeed we have shown at Proposition 2.1.2 that pinning a reciprocal probability gives a Markov probability. Therefore, since P is reciprocal, P^x is Markov.

In some cases where the graph enjoys regularity, the property (iv) above can be weakened by only considering the identity (6.6) on a generating set of the $\mathcal{A}_{\leftrightarrow}^R(x, \text{supp}P_1^x)$. This is made precise below at Proposition 6.2.1.

The reciprocal characteristics come with a natural probabilistic interpretation which is expressed in terms of short-time asymptotic of the distribution of bridges. We shall show that they can be recovered as quantities related to Taylor expansions as $h > 0$ tends to zero of conditional probabilities of the form $P(X_{[t, t+h]} \in \cdot \mid X_t, X_{t+h})$. This is the content of Theorem 6.2.2 below.

Let us introduce the notation needed for its statement. For any integer $k \geq 1$ and any $0 \leq t < 1$, we denote by T_k^t the k -th instant of jump after time t . It is defined for $k = 1$ by $T_1^t := \inf \{s \in (t, 1] : X_{s-} \neq X_s\}$ and for any $k \geq 2$ by $T_k^t := \inf \{s \in (T_{k-1}^t, 1] : X_{s-} \neq X_s\}$ with the convention $\inf \emptyset = +\infty$.

Theorem 6.2.2 (Interpretation of the characteristics). *Let $(\mathcal{X}, \rightarrow)$ and j satisfy Assumption 2.3.1. Let P be any random walk in $\mathfrak{R}(R)$.*

(a) For any $t \in [0, 1)$, any $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(P^x)$ and any measurable subset $I \subset [0, 1]$, we have

$$\begin{aligned} P(T_1^t \in t + hI \mid X_t = z, X_{t+h} = z', T_2^t > t + h) \\ = \int_I dr + h\chi_a[j](t, z \rightarrow z') \int_I (r - 1/2) dr + o_{h \rightarrow 0^+}(h). \end{aligned} \quad (6.7)$$

(b) For any $t \in [0, 1)$ and any closed $\mathcal{A}_{\rightarrow}(P^x)$ -walk \mathbf{c} , we have

$$\begin{aligned} P\left((X_t \rightarrow X_{T_1^t} \rightarrow \cdots \rightarrow X_{T_{|\mathbf{c}|}^t} = X_t) = \mathbf{c}, T_{|\mathbf{c}|}^t < t + h < T_{|\mathbf{c}|+1}^t \mid X_t = X_{t+h}\right) \\ = \chi_c[j](t, \mathbf{c})h^{|\mathbf{c}|}/|\mathbf{c}|! + o_{h \rightarrow 0^+}(h^{|\mathbf{c}|}). \end{aligned} \quad (6.8)$$

Note that in statement (b), only cycles with respect to $\mathcal{A}_{\rightarrow}(P^x)$ and not its symmetrized version $\mathcal{A}_{\leftrightarrow}(P^x)$ must be taken into account.

This theorem extends by a large amount the early findings of section 4.6.

In the same spirit that a Markov walk is specified by the Markov property and its jump intensity which can be obtained as the limit in small time of a conditional expectation, we obtain the following characterization of $\mathfrak{R}(R)$. The same observations made in Remark 6.2.1 hold true for the following results:

Corollary 6.2.1 (Short-time expansions characterize $\mathfrak{R}(R)$). *Let $(\mathcal{X}, \rightarrow)$ and j satisfy Assumption 2.3.1. In addition, we also assume that the directed graph is symmetric, i.e.*

$$\mathcal{A}_{\rightarrow}(j) = \mathcal{A}_{\leftrightarrow}(j).$$

Then a random walk $P \in \mathbb{P}(\Omega)$ belongs to $\mathfrak{R}(R)$ if and only if the following assertions hold.

(i) P is reciprocal and $P \ll R$.

(ii) For all $x \in \text{supp}P_0$ the conditioned random walk P^x admits an intensity k^x which is t -differentiable on $[0, 1)$.

(iii) The subgraph of all P^x -active arcs doesn't depend on t and is given by

$$\mathcal{X}(P^x) = \mathcal{X}^R(x, \text{supp}(P_1^x)), \quad \mathcal{A}_{\rightarrow}(P^x) = \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x).$$

(iv) For any $t \in (0, 1)$, any $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(P^x)$ and any measurable subset $I \subset [0, 1]$, the identity (6.7) is satisfied with $P = P^x$.

(v) For any $t \in (0, 1)$ and any $\mathcal{A}_{\leftrightarrow}(P^x)$ -cycle \mathbf{c} the identity (6.8) is satisfied with $P = P^x$.

The assumption $\mathcal{A}_{\rightarrow}(j) = \mathcal{A}_{\leftrightarrow}(j)$ is needed for Corollary 6.2.1 to hold. Without any restriction on the structure of the graph, this result is false in general. However, it is possible to relax this restriction in some specific situations. For instance, it is the case of the non-oriented triangle at page 163.

It is possible to improve the statements of Theorem 6.2.1 and Corollary 6.2.1 as follows.

Proposition 6.2.1. *The conclusions of Theorem 6.2.1 and Corollary 6.2.1 remain unchanged when weakening the properties (iv) by only considering cycles \mathbf{c} in any generating subset of the closed $\mathcal{A}_{\leftrightarrow}(P^x)$ -walks, see Definition 6.1.4.*

6.3 Proofs of the main results

Let P be any element of $\mathfrak{R}(R)$, $x \in \text{supp}P_0$. We know with Proposition 2.2.2 and 2.1.2 that $P^x \ll R^x$ and that P^x is Markov. Therefore P admits an intensity of the form $k(t, X_{t-} \rightarrow X_t)$ and the related Girsanov formula (see section 2.3.3) is for each $x \in \text{supp}P_0$,

$$\begin{aligned} \frac{dP^x}{dR^x} = \mathbf{1}_{\{\tau=\infty\}} \exp \left(- \int_0^1 (\bar{k} - \bar{j})(t, X_{t-}) dt \right. \\ \left. + \sum_{0 < t \leq 1: X_{t-} \neq X_t} \log \frac{k}{j}(t, X_{t-} \rightarrow X_t) \right) \end{aligned} \quad (6.9)$$

where the stopping time τ is given by

$$\begin{aligned} \tau := \inf \left\{ t \in [0, 1); k(t, X_{t-} \rightarrow X_t) = 0 \right. \\ \left. \text{or } \int_0^t \bar{k}(s, X_s) ds = \infty \right\} \in [0, 1] \cup \{\infty\} \end{aligned}$$

with the convention $\inf \emptyset = \infty$.

Lemma 6.3.1 (HJB equation). *For any $x \in \mathcal{X}$ and any nonnegative function $h_1 : \mathcal{X} \rightarrow [0, \infty)$ such that $E_{R^x} h_1(X_1) = 1$, the function ψ^x defined by*

$$\begin{cases} \psi^x(t, z) := \log E_{R^x}[h_1(X_1) | X_t = z] \in \mathbb{R}, & t \in (0, 1), z \in \mathcal{X}^R(x, \text{supp}h_1) \\ \psi^x(0, x) := 0, & t = 0, z = x, \end{cases}$$

is a well-defined real function which satisfies the following regularity properties:

(i) for all $z \in \mathcal{X}^R(x, \text{supph}_1)$, $t \mapsto \psi_t^x(z)$ is continuously differentiable on $(0, 1)$,

(ii) $\lim_{t \rightarrow 0^+} \psi_t^x(x) =: \psi_0^x(x) = 0$,

(iii) $\lim_{t \rightarrow 1^-} \psi_t^x(z) =: \psi_1^x(z) \in \mathbb{R}$ exists for all $z \in \text{supph}_1$,

and for each $0 \leq T \leq 1$, the Itô formula

$$\begin{aligned} \psi_T^x(X_T) := \log h_T(X_T) &= \sum_{0 < t \leq T; X_{t^-} \neq X_t} [\psi_t^x(X_t) - \psi_t^x(X_{t^-})] \\ &\quad + \int_{[0, T]} \partial_t \psi_t^x(X_t) dt \in \mathbb{R}, \quad h_1(X_1)R^x - a.s. \end{aligned} \quad (6.10)$$

is meaningful.

Furthermore, ψ^x is a classical solution of the Hamilton-Jacobi-Bellman equation

$$\begin{cases} \partial_t \psi_t(z) + \sum_{z': (z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \text{supph}_1)} j(t, z \rightarrow z') [e^{\psi_t(z') - \psi_t(z)} - 1] = 0, & t \in (0, 1), z \in \mathcal{X}^R(x, \text{supph}_1) \\ \lim_{s \rightarrow 1^-} \psi_s(y) = \log h_1(y), & t = 1, y \in \text{supph}_1. \end{cases} \quad (6.11)$$

Remark 6.3.1. (a) It is important to see that the identity (6.10) is only valid almost surely with respect to $\mathbf{1}_{\text{supph}_1}(X_1)R$, but not with respect to R .

(b) Note that if $z \rightarrow z' \in \mathcal{A}_{\rightarrow}^R(x, \text{supph}_1)$, then both z, z' are in $\mathcal{X}^R(x, \text{supph}_1)$. Therefore both $\psi_t^x(z)$ and $\psi_t^x(z')$ are well defined, and then so is the sum appearing in the HJB equation (6.11).

Proof. The function $h(t, z) := E_{R^x}[h_1(X_1) \mid X_t = z]$, $0 < t \leq 1$, $z \in \mathcal{X}(R^x)$ is space-time harmonic. We are going to show that it satisfies the Kolmogorov equation

$$(\partial_t + \mathcal{G}_t)h(t, z) = 0, \quad 0 < t < 1, z \in \mathcal{X}(R^x).$$

in the classical sense. \mathcal{G}_t is the generator of the reference walk R , see Definition 2.3.2 Remark that it is needed that $z \in \text{supp}R_t^x$ for the conditional expectation to be well defined. But the assumption 2.3.1 implies that

$\text{supp}R_t^x = \mathcal{X}(R^x)$ for all $0 < t \leq 1$. We obtain $h(t, \cdot) = \overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)(h_1)$ where the ordered exponential is defined by:

$$\overleftarrow{\text{exp}}\left(\int_t^1 \mathcal{G}_s ds\right) := \text{Id} + \sum_{n \geq 1} \int_{t \leq s_1 \leq \dots \leq s_n \leq 1} \mathcal{G}_{s_1} \cdots \mathcal{G}_{s_n} ds_1 \cdots ds_n.$$

The continuity of \mathcal{G}_t (recall that j is t -continuous) ensures that its formal left t -derivative is $-\mathcal{G}_t \overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)$. Furthermore, $\overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)$ and $\mathcal{G}_t \overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)$ are absolutely summable series. More precisely, for any nonnegative h_1 in $L^1(R_1^x)$, we know by a martingale argument that $h_t = \overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)h_1$ is in $L^1(R_t^x)$. Hence, $h(\cdot, z)$ is continuous at $t = 1$ for all $z \in \mathcal{X}(R^x)$ and $h(\cdot, x)$ is continuous at $t = 0$. In addition, with the assumed uniform boundedness of \mathcal{G}_t (see iii) of Assumption 2.3.1), we see that $\mathcal{G}_t \overleftarrow{\text{exp}}(\int_t^1 \mathcal{G}_s ds)h_1 = \mathcal{G}_t h_t$ is also in $L^1(R_t^x)$. Therefore, h is continuously left t -differentiable on $[0, 1)$. But the continuity of the left derivative implies both the existence and the continuity of the derivative. Consequently, the backward differential system

$$\begin{cases} (\partial_t + \mathcal{G}_t)h(t, z) = 0, & 0 < t < 1, z \in \mathcal{X}(R^x), \\ \lim_{t \rightarrow 1^-} h(t, y) =: h(1, y) = h_1(y), & t = 1, y \in \mathcal{X}(R^x), \end{cases} \quad (6.12)$$

can be considered in the classical sense and

$$\lim_{t \rightarrow 0^+} h(t, x) =: h(0, x) = 1$$

since $h(0, x) = 1$ is fixed by hypothesis.

On the other hand, point iv) of Assumption 2.3.1 implies that for all $0 < t < 1$ and $z \in \mathcal{X}^R(x, \text{supp}h_1) \subset \mathcal{X}(R^x)$, $h(t, z)$ is positive. It follows that we are allowed to define $\psi_t^x(z) := \log h(t, z)$ as a real number for any $0 < t < 1$ and any $z \in \mathcal{X}^R(x, \text{supp}h_1)$. Of course, for $t = 0$ one must only consider $z = x$ and $\lim_{t \rightarrow 0^+} \psi_t^x(x) = \psi_0^x(x) = 0$.

We have shown that the regularity properties (i,ii,iii) are satisfied.

The Itô formula

$$\begin{aligned} \psi_T^x(X_T) = \psi_S^x(X_S) + & \sum_{S < t \leq T; X_{t-} \neq X_t} [\psi_t^x(X_t) - \psi_t^x(X_{t-})] \\ & + \int_{[S, T]} \partial_t \psi_t^x(X_t) dt \in (-\infty, \infty), \quad h_1(X_1) R^x - a.s. \end{aligned} \quad (6.13)$$

is meaningful for all $0 < S \leq T < 1$. Indeed, under Assumption 2.3.1 there are finitely many jumps $R^x - a.s.$ and we have already seen that $\psi_t^x(X_t)$

is finite for every $0 \leq t < 1$, $h_1(X_1) R^x - a.s.$ Therefore, $\psi_T^x(X_T)$, $\psi_S^x(X_S)$ and $\sum_{S < t \leq T; X_{t^-} \neq X_t} [\psi_t^x(X_t) - \psi_t^x(X_{t^-})]$ are finite. It follows that the integral $\int_{(S, T]} \partial_t \psi_t^x(X_t) dt$ is also well-defined $h_1(X_1) R^x - a.s.$

Letting S tend to 0 and T to 1 in (6.13), with the limits (ii) and (iii) we obtain (6.10) where the integral $\int_{[0, T)} \partial_t \psi_t^x(X_t) dt$ is well defined $h_1(X_1) R^x - a.s.$

Finally, considering $h = e^{\psi^x}$ in (6.12) gives the HJB equation (6.11) and completes the proof of the lemma. \square

The following result is the key lemma of the proof of Theorem 6.2.1.

Lemma 6.3.2. *If the random walk $P \in \mathcal{P}(\Omega)$ belongs to $\mathfrak{R}(R)$ then $P \ll R$ and the following assertions are satisfied.*

- (i) *It is a reciprocal walk and $P \ll R$*
- (ii) *For all $x \in \text{supp}P_0$ the conditioned random walk P^x admits an intensity $k^x(t, z \rightarrow z')$ which is t -differentiable on $[0, 1)$.*
- (iii) *The subgraph of all P^x -active arcs doesn't depend on t and is given by*

$$\mathcal{X}(P^x) = \mathcal{X}^R(x, \text{supp}P_1^x), \quad \mathcal{A}_{\rightarrow}(P^x) = \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x).$$

- (iv) *The function $\psi^x : (0, 1) \times \mathcal{X}^R(x, \text{supp}P_1^x) \rightarrow \mathbb{R}$ defined by :*

$$\psi^x(t, z) := \log E_{R^x}(h(X_1) | X_t = z), \quad h(z) := \frac{dP_1^x}{dR_1^x}(z)$$

is such that for all $z \in \mathcal{X}^R(x, \text{supp}P_1^x)$, $t \mapsto \psi_t^x(z)$ is continuously differentiable on $(0, 1)$ and k^x and ψ^x are linked by the relations

$$\log \frac{k^x}{j}(t, z \rightarrow z') = \psi_t^x(z') - \psi_t^x(z), \quad (6.14)$$

$$\partial_t \psi_t^x(z) + (\bar{k}^x - \bar{j})(t, z) = 0, \quad (6.15)$$

for all $t \in (0, 1)$, $z \in \mathcal{X}^R(x, \text{supp}P_1^x)$ and $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x)$.

In (6.15), the average frequency of jumps $\bar{k}^x(t, z) := \sum_{z': z \rightarrow z'} k^x(t, z \rightarrow z')$ is finite everywhere on $[0, 1) \times \mathcal{X}^R(x, \text{supp}P_1^x)$.

Recall that $P^x \ll R^x$ implies that P^x admits an intensity of jumps k^x .

Proof. Let us take some $P \in \mathfrak{R}(R)$ and show that it satisfies the announced properties. Point (i) is a direct consequence of the definition of reciprocal class and Proposition 2.2.1. As $\text{supp}P_{01} \subset \text{supp}R_{01}$, we can apply Proposition 2.2.2 which states that $P = h(X_0, X_1)R$. This implies that for every $x \in \text{supp}P_0$, we have $P^x \ll R^x$, and there exist $h^x : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ such that

$$P^x = h^x(X_1) R^x \quad (6.16)$$

By taking conditional expectations we get $h^x = dP_1^x/dR_1^x R_1^x - a.s.$. Comparing this expression of the density with the one given by Girsanov theorem (6.9), we see that the events $\{\tau = \infty\}$ and $\{X_1 \in \text{supp}h^x = \text{supp}P_1^x\}$ match, up to an R^x -negligible set. This proves that $\mathcal{A}_{\rightarrow}(t, P^x)$ doesn't depend on t and that it is equal to $\mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x)$, which is (iii). Let us define

$$\psi_t^x(z) := \log E_{R^x}(h^x(X_1) \mid X_t = z).$$

We know that ψ^x shares the regularity properties (i), (ii) & (iii) of Lemma 6.3.1. Applying (6.10) with $T = 1$ to $P^x = h^x(X_1) R^x$ leads us to rewrite (6.16) in the following form:

$$\frac{dP^x}{dR^x} = \mathbf{1}_{\{X_1 \in \text{supp}P_1^x\}} \exp \left(\sum_{0 < t \leq 1; X_{t-} \neq X_t} [\psi_t^x(X_t) - \psi_t^x(X_{t-})] + \int_{(0,1]} \partial_t \psi_t^x(X_t) dt \right).$$

Comparing with (6.9), we obtain

$$\begin{aligned} & \sum_{0 < t \leq 1; X_{t-} \neq X_t} \log \frac{k^x}{j}(t, X_{t-} \rightarrow X_t) - \int_{[0,1]} (\bar{k}^x - \bar{j})(t, X_t) dt \\ &= \sum_{0 < t \leq 1; X_{t-} \neq X_t} [\psi_t^x(X_t) - \psi_t^x(X_{t-})] + \int_{(0,1]} \partial_t \psi_t^x(X_t) dt, \quad \mathbf{1}_{\{X_1 \in \text{supp}P_1^x\}} R^x - a.s. \end{aligned}$$

Identifying the jumps, we see that

$$\mathbf{1}_{\{X_{t-} \neq X_t\}} \log \frac{k^x}{j}(t, X_{t-} \rightarrow X_t) = \psi_t^x(X_t) - \psi_t^x(X_{t-}), \quad \mathbf{1}_{\{X_1 \in \text{supp}P_1^x\}} R^x - a.s. \quad (6.17)$$

More precisely, (6.17) gives us (6.14). By Lemma 6.3.1 we know that ψ^x satisfies the HJB equation (6.11). Replacing the expression $\psi_t^x(z') - \psi_t^x(z)$ in that equation with the one given by (6.14), we obtain (6.15).

Remark that (6.14) also implies that k^x is t -continuously differentiable on $[0, 1)$, which is (ii). This completes the proof of the lemma. \square

6.3.1 Proof of Theorem 6.2.1

In order to simplify notation, for a given $x \in \text{supp}P_0$, we write $\mathcal{Z}_{\rightarrow} = \mathcal{A}_{\rightarrow}^R(x, \text{supp}P_1^x)$ and $\mathcal{Z}_{\leftrightarrow} = \mathcal{A}_{\leftrightarrow}^R(x, \text{supp}P_1^x)$ during this proof.

• **Proof of the necessary condition.** Let us show that $P \in \mathfrak{R}(R)$ satisfies the announced properties. The first items (i) and (ii) are already proved at Lemma 6.3.2.

Now, we rely on Lemma 6.3.2(iv). Differentiating (6.14) and plugging (6.15) into the resulting identity gives us $\chi_a[k^x] = \chi_a[j]$ on $\mathcal{Z}_{\rightarrow}$ which is (iii).

Let us prove (iv). For any $t \in (0, 1)$ and any $(z \rightarrow z') \in \mathcal{Z}_{\rightarrow}$, we denote $\ell(t, z \rightarrow z') = \log \frac{k^x}{j}(t, z \rightarrow z')$. If the reversed arc $(z' \rightarrow z)$ is also in $\mathcal{Z}_{\rightarrow}$, we see with (6.14) that $\ell(t, z' \rightarrow z) = -\ell(t, z \rightarrow z')$. Otherwise, we extend $\ell(t, \cdot)$ from $\mathcal{Z}_{\rightarrow}$ to $\mathcal{Z}_{\leftrightarrow}$ by means of this identity. Therefore,

$$\ell(t, z \rightarrow z') = \psi_t^x(z') - \psi_t^x(z), \quad \forall (z \rightarrow z') \in \mathcal{Z}_{\leftrightarrow}$$

and we are allowed to apply Lemma 6.1.1 to obtain $\chi_c[k^x](t, \mathbf{c}) = \chi_c[j](t, \mathbf{c})$ for all the $\mathcal{Z}_{\leftrightarrow}$ -cycles, which is the desired result.

• **Proof of the sufficient condition.** Take $P \in \mathcal{P}(\Omega)$ such that for every $x \in \text{supp}P_0$, P^x is Markov and its intensity k^x satisfies the properties (i-iv) of Theorem 6.2.1. Fix $x \in \text{supp}P_0$.

We start exploiting the property (v). Because $\chi_c[k^x](t, \mathbf{c}) = \chi_c[j](t, \mathbf{c})$ for any $t \in (0, 1)$ and any closed $\mathcal{Z}_{\leftrightarrow}$ -walk \mathbf{c} , by Lemma 6.1.1 there exists a function φ^x which satisfies

$$\varphi^x(t, z') - \varphi^x(t, z) = \log \frac{k^x}{j}(t, z \rightarrow z'), \quad \forall t \in (0, 1), (z \rightarrow z') \in \mathcal{Z}_{\rightarrow}. \quad (6.18)$$

On the other hand, the property (iv) implies that $\partial_t \varphi^x(t, z) + (\bar{k}^x - \bar{j})(t, z) = \partial_t \varphi^x(t, z') + (\bar{k}^x - \bar{j})(t, z')$ for all $t \in (0, 1)$ and $(z \rightarrow z') \in \mathcal{Z}_{\rightarrow}$. Since φ^x is defined up to some time-dependent additive function, we can w.l.o.g. assume that: $\partial_t \varphi^x(t, x) + (\bar{k}^x - \bar{j})(t, x) = 0$, for all $t \in (0, 1)$. Therefore, we obtain with the property (iii) and our assumption (6.5) that

$$\partial_t \varphi^x(t, z) + (\bar{k}^x - \bar{j})(t, z) = 0, \quad \forall t \in (0, 1), z \in \mathcal{X}^R(x, \text{supp}P_1^x). \quad (6.19)$$

We know with the property (ii) that $P^x \ll R^x$. Restricting the path measures to the sub- σ -field $\sigma(X_{[0,t]})$ for any $0 \leq t < 1$, and plugging (6.18) and (6.19) into Girsanov's formula (6.9), we obtain

$$\frac{dP_{[0,t]}^x}{dR_{[0,t]}^x} = \mathbf{1}_{\{\tau_t = \infty\}} \exp \left(\sum_{0 < s \leq t: X_s \neq X_{s-}} [\varphi^x(s, X_s) - \varphi^x(s, X_{s-})] + \int_{(0,t)} \partial_s \varphi^x(s, X_s) ds \right)$$

where

$$\tau_t := \inf \left\{ r \in [0, t]; k^x(r, X_{r-} \rightarrow X_l) = 0 \text{ or } \int_{(0,r]} \partial_s \varphi^x(s, X_s) ds = -\infty \right\} \in [0, t] \cup \{\infty\}.$$

Thanks to property (iii), k^x never vanishes on $[0, 1) \times \mathcal{Z}_\rightarrow$. Hence, τ_t is finite if and only if $\int_{[0,r]} \partial_s \varphi^x(s, X_s) ds = -\infty$ for some $0 < r \leq t$. But for any $0 \leq t < 1$, we have

$$\sum_{0 < s \leq t; X_s \neq X_{s-}} [\varphi^x(s, X_s) - \varphi^x(s, X_{s-})] + \int_{(0,t]} \partial_s \varphi^x(s, X_s) ds = \varphi^x(t, X_t) - \varphi^x(0, X_0).$$

This implies that $\int_{(0,t]} \partial_s \varphi^x(s, X_s) ds$ is finite for every $0 \leq t < 1$ and it follows that τ_t is infinite $R^x - a.s.$ for all $0 \leq t < 1$. Consequently,

$$\frac{dP_{[0,t]}^x}{dR_{[0,t]}^x} = \exp(\varphi^x(t, X_t) - \varphi^x(0, x)) \quad (6.20)$$

since the prefactor $1_{\{\tau_t = \infty\}}$ does not vanish. Let us denote $Z := \frac{dP^x}{dR^x}$ and $Z_t := E_{R^x}(Z \mid X_{[0,t]}) = \frac{dP_{[0,t]}^x}{dR_{[0,t]}^x}$ for all $0 \leq t < 1$. We see with (6.20) that Z_t is X_t -measurable. This implies that Z is $X_{[t,1]}$ -measurable for all $0 \leq t < 1$ and consequently that Z is X_1 -measurable. Mixing over $x \in \text{supp}P_0$ and using Proposition 2.2.2 we conclude that P belongs to $\mathfrak{R}(R)$. \square

6.3.2 Proof of Theorem 6.2.2

Let us fix $x \in \text{supp}P_0$ and $t \in (0, 1)$. Note that for $h > 0$ such that $t + h < 1$ and $(z \rightarrow z') \in \mathcal{A}_\rightarrow^R(x, \text{supp}P_1^x)$, the conditional distribution $P(\cdot \mid X_t = z, X_{t+h} = z')$ is well defined. Because of (??) and the reciprocal property of P , we have

$$\begin{aligned} P(T_1^t \in t + hI \mid X_t = z, X_{t+h} = z', T_2^t > t + h) \\ = R(T_1^t \in t + hI \mid X_t = z, X_{t+h} = z', T_2^t > t + h). \end{aligned} \quad (6.21)$$

Therefore it suffices to do the proof with R instead of P .

• Proof of (a). Recall that for a Poisson process with intensity $\lambda(t)$ the density of the law of the first instant of jump is $t \mapsto \lambda(t) \exp(-\int_0^t \lambda(s) ds)$, $t \geq 0$. Therefore,

$$\begin{aligned} & R(T_1^t \in t + hI, X_{t+h} = z', T_2^t > t + h \mid X_t = z) \\ &= \int_{hI} \bar{j}(t+r, z) \exp\left(-\int_0^r \bar{j}(t+s, z) ds\right) \frac{j(t+r, z \rightarrow z')}{\bar{j}(t+r, z)} \exp\left(-\int_r^h \bar{j}(t+s, z') ds\right) dr \\ &= h \int_I \exp\left(-\int_0^{hr} \bar{j}(t+s, z) ds\right) j(t+hr, z \rightarrow z') \exp\left(-\int_{hr}^h \bar{j}(t+s, z') ds\right) dr. \end{aligned}$$

Using the following expansions as h tends to zero:

$$\begin{aligned}\exp\left(-\int_0^{hr}\bar{j}(t+s,z)ds\right) &= 1 - \bar{j}(t,z)hr + o(h), \\ \exp\left(-\int_{hr}^h\bar{j}(t+s,z')ds\right) &= 1 - \bar{j}(t,z')(1-r)h + o(h), \\ j(t+hr, z \rightarrow z') &= j(t, z \rightarrow z') + \partial_t j(t, z \rightarrow z')hr + o(h),\end{aligned}$$

we obtain

$$\begin{aligned}R(T_1^t \in t+hI, X_{t+h} = z', T_2^t > t+h \mid X_t = z) \\ &= hj(t, z \rightarrow z') \int_I \left(1 + h \left[\frac{\partial_t j(t, z \rightarrow z')}{j(t, z \rightarrow z')}r - \bar{j}(t, z)r - \bar{j}(t, z')(1-r)\right]\right) dr + o(h^2) \\ &= hj(t, z \rightarrow z') \int_I \left(1 + h\{\chi_a[j](t, z \rightarrow z')r - \bar{j}(t, z')\}\right) dr + o(h^2).\end{aligned}$$

In particular, with $I = [0, 1]$ this implies that

$$\begin{aligned}R(X_{t+h} = z', T_2^t > t+h \mid X_t = z) \\ &= hj(t, z \rightarrow z')(1 + h\{\chi_a[j](t, z \rightarrow z')/2 - \bar{j}(t, z')\}) + o(h^2).\end{aligned}$$

Taking the ratio of these probabilities leads us to

$$\begin{aligned}R(T_1^t \in t+hI \mid X_t = z, X_{t+h} = z', T_2^t > t+h) \\ &= \int_I \frac{1 + h\{\chi_a[j](t, z \rightarrow z')r - \bar{j}(t, z')\} + o(h)}{1 + h\{\chi_a[j](t, z \rightarrow z')/2 - \bar{j}(t, z')\} + o(h)} dr \\ &= \int_I \left(1 + h\{\chi_a[j](t, z \rightarrow z')(r - 1/2)\}\right) dr + o(h).\end{aligned}$$

With (6.21) this gives (6.7).

- **Proof of (b).** Since $R(X_t = X_{t+h} = z) = R(X_t = z)(1 + o(1))$ as $h \rightarrow 0^+$, we can write the proof with $R(\cdot \mid X_t = z)$ instead of $R(\cdot \mid X_t = X_{t+h} = z)$.

Therefore, if $\mathbf{c} = (z = x_0 \rightarrow x_1 \cdots \rightarrow x_{|\mathbf{c}|} = z)$,

$$\begin{aligned}
R\left((X_t \rightarrow X_{T_1^t} \rightarrow \cdots \rightarrow X_{T_{|\mathbf{c}|}^t}) = \mathbf{c}, T_{|\mathbf{c}|}^t < t+h < T_{|\mathbf{c}|+1}^t \mid X_t = z\right) \\
&= \int_{\{t < t_1 < \cdots < t_{|\mathbf{c}|} < t+h\}} \prod_{i=1}^{|\mathbf{c}|} \exp\left[-\int_{t_{i-1}}^{t_i} \bar{j}(s, x_i) ds\right] j(t_i, x_i \rightarrow x_{i+1}) \\
&\quad \times \exp\left[-\int_{t_{|\mathbf{c}|}}^{t+h} \bar{j}(s, z) ds\right] dt_1 \cdots dt_{|\mathbf{c}|} \\
&= \chi_c[j](t, \mathbf{c})(1 + o(1)) \int_{\{t < t_1 < \cdots < t_{|\mathbf{c}|} < t+h\}} \exp\left[-\sum_{i=1}^{\mathbf{c}} \int_{t_i}^{t_{i+1}} \bar{j}(s, x_i) ds\right. \\
&\quad \left.- \int_{t_{|\mathbf{c}|}}^{t+h} \bar{j}(s, z) ds\right] dt_1 \cdots dt_{|\mathbf{c}|} \\
&= \chi_c[j](t, \mathbf{c})h^{|\mathbf{c}|}/|\mathbf{c}|! + o(h^{|\mathbf{c}|})
\end{aligned}$$

where we used the convention that $t_0 := t$. This completes the proof of the theorem. \square

6.3.3 Proof of Corollary 6.2.1

Proof. The necessary condition is a direct consequence of Theorems 6.2.1 and 6.2.2. For the sufficient condition, all we have to show is that the properties (a) and (b) of Theorem 6.2.2 respectively imply the properties (iii) and (iv) of Theorem 6.2.1.

Based on identity (6.21), we see that the same calculations as in Theorem 6.2.2's proof at page 149 shows that replacing R^x by P^x and j by k^x lead to the same conclusions with k^x instead of j . It remains to compare the resulting expansions to conclude that (6.5) and (6.6) are satisfied. \square

6.3.4 Proof of Proposition 6.2.1

Proof. In the case of Theorem 6.2.1 observe that (iv) is equivalent by lemma 6.1.1 to the fact that for all t the function ℓ defined on $\mathcal{A}_{\leftrightarrow}(j)$ by

$$\ell(z \rightarrow z') = \begin{cases} \log\left(\frac{k^x}{j}(t, z \rightarrow z')\right), & \text{if } (z \rightarrow z') \in \mathcal{A}_{\rightarrow}(j) \\ -\log\left(\frac{k^x}{j}(t, z' \rightarrow z)\right), & \text{otherwise} \end{cases}$$

is such $\ell(\mathbf{c}) = 0$ for any cycle. But, by definition, this is equivalent to impose $\ell(\mathbf{c}) = 0$ on a generating set of cycles. In the same way, repeating

the proof of Corollary 6.2.1 we see that restricting (iv) to a generating set of cycles implies that $\ell(\mathbf{c}) = 0$ on that set, where ℓ has been defined just above. But then, again by definition this implies $\ell(\mathbf{c}) = 0$ for any cycle, which is precisely point (iv) of Theorem 6.2.1. The same Theorem gives the conclusion. \square

6.4 More results

We give some additional results on the characterization of the reciprocal class $\mathfrak{R}(R)$.

6.4.1 Jump intensity

Proposition 6.4.1 below expresses characterizations of the reciprocal class in terms of the intensities of jumps by exploiting the fact that for any $P \in \mathfrak{R}(R)$ and each x , P^x is an h -transform of R^x .

Proposition 6.4.1 (Representation of the intensity of an element of $\mathfrak{R}(R)$).
Let $P \in \mathcal{P}(\Omega)$ be a random walk. The following assertions are equivalent.

- (a) $P \in \mathfrak{R}(R)$.
(b) There exists $h : \mathcal{X} \rightarrow [0, \infty)$ such that

$$P^x = h(x, X_1) R^x$$

with $\sum_{y \in \mathcal{X}} R_1^x(y) h(x, y) = 1$, for all $x \in \text{supp} P_0$.

- (c) There exists $g : \mathcal{X} \rightarrow [0, \infty)$ such that

- (i) $\text{supp} P_0 = \{x : g(x, y_0) > 0 \text{ for some } y_0\}$,
(ii) for all $x \in \text{supp} P_0$, $\sum_{y \in \mathcal{X}} R_1^x(y) g(x, y) < \infty$ and
(iii) P is a random walk with intensity

$$k(t, X_{[0,t)}, X_{t^-} \rightarrow z) = \mathbf{1}_{\{X_{t^-} \in \mathcal{X}^{R(x, \text{supp} g(x, \cdot))}\}} \frac{g_t^{X_0}(z)}{g_t^{X_0}(X_{t^-})} j(t, X_{t^-} \rightarrow z), \quad R\text{-a.s.} \quad (6.22)$$

where for any $0 \leq t \leq 1$ and $z \in \mathcal{X}(R^x)$,

$$g_t^x(z) := E_R[g(x, X_1) \mid X_t = z] = \sum_{y \in \mathcal{X}} r(t, z; 1, y) g(x, y),$$

with $r(t, z; 1, y) := R(X_1 = y \mid X_t = z)$.

In addition, for each $x \in \text{supp}P_0$, g^x solves the heat equation

$$\begin{cases} (\partial_t + \mathcal{G}_t)g^x = 0, & 0 \leq t < 1, \\ g_1^x = g(x, \cdot), & t = 1. \end{cases} \quad (6.23)$$

where \mathcal{G}_t is the generator of R .

Moreover, the link between h and g is $h(x, y) = g(x, y)/g_0^x(x)$ where $g_0^x(x) = \sum_{y \in \mathcal{X}} R_1^x(y)g(x, y)$ is finite.

Proof. The equivalence of (a) and (b) is proved at Proposition 2.2.2. Let us prove the equivalence of (b) and (c).

Statement (b) tells us that P^x is an h -transform of R^x . It is a general result of [44] that the extended generator of this h -transform is given for any function u with a finite support by

$$A_t^x u(X_{t-}) = \mathcal{G}_t u(X_{t-}) + \Gamma_t(g_t^x, u)(X_{t-})/g_t^x(X_{t-}), \quad P^x - a.s.$$

where $\Gamma_t(g_t, u)(z) := \sum_{z': z \rightarrow z'} j(t; z \rightarrow z') [g_t^x(z') - g_t^x(z)] [u(z') - u(z)]$ is the carré du champ operator. This identity characterizes the h -transformation. Note that it is only valid P^x -almost surely and not R^x -almost surely.

As for any $t \in [0, 1)$ and $z \in \mathcal{X}(R^x)$, $g_t^x(z) > 0 \Leftrightarrow z \in \mathcal{X}^R(x, \text{supp}g(x, \cdot))$, we see that $A_t^x u(z) = \mathbf{1}_{\{z \in \mathcal{X}^R(x, \text{supp}g(x, \cdot))\}} \sum_{z': z \rightarrow z'} j(t, z \rightarrow z') g_t^x(z')/g_t^x(z) [u(z') - u(z)]$ which gives (6.22). This completes the proof of Proposition 6.4.1.

A direct proof of the equivalence of (b) and (c), which does not rely on a general result about the extended generator of an h -transform, consists of identifying $dP^x/dR^x = h(x, X_1)$ by means of Girsanov's formula (6.9) and to apply the representation result (6.23) under its HJB form (6.11), via the transformation $g = e^\psi$, as in the proof of Lemma 6.3.2. \square

As a special case of Proposition 6.4.1, we recover the known fact that for each $(x, y) \in \text{supp}R_{01}$, the jump intensity k^{xy} of the bridge R^{xy} is

$$k^{xy}(t, z \rightarrow z') = \frac{r(t, z'; 1, y)}{r(t, z; 1, y)} j(t, z \rightarrow z'), \quad 0 \leq t < 1, (z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \{y\}). \quad (6.24)$$

6.4.2 Characteristic equation

We see with Theorem 6.2.1 that for any given Markov intensity j , the description of the reciprocal class $\mathfrak{R}(R)$ is linked to the solution of some equation of the form

$$\begin{cases} \mathcal{A}_{\rightarrow}(k^x) = \mathcal{A}_{\rightarrow}^R(x, \mathcal{Y}^x), \\ \chi[k^x] = \chi[j], \end{cases} \quad x \in S \quad (6.25)$$

where we use notation

$$\chi[P^x] =: \chi[k^x]$$

to emphasize the role of the intensity. In (6.25), the given subsets $S \subset \mathcal{X}$ and

$$\mathcal{Y}^x \subset \mathcal{X}(R^x), \quad x \in S$$

are non-empty and the unknown is the collection of Markov intensities $(k^x; x \in S)$. More precisely, (6.25) is a shorthand for the following list of properties that must hold for all $x \in S$.

- (i) The intensity k^x is t -differentiable on $[0, 1)$.
- (ii) The subgraph of all k^x -active arcs doesn't depend on t and is

$$\mathcal{A}_{\rightarrow}(k^x) := \{(z \rightarrow z') : k^x(t, z \rightarrow z') > 0\} = \mathcal{A}_{\rightarrow}^R(x, \mathcal{Y}^x).$$

- (iii) For any $t \in [0, 1)$ and any $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \mathcal{Y}^x)$, we have

$$\chi_a[k^x](t, z \rightarrow z') = \chi_a[j](t, z \rightarrow z').$$

- (iv) For any $t \in [0, 1)$ and any closed $\mathcal{A}_{\leftrightarrow}^R(x, \mathcal{Y}^x)$ -walk \mathbf{c} , we have

$$\chi_c[k^x](t, \mathbf{c}) = \chi_c[j](t, \mathbf{c}).$$

Because of Theorem 6.2.1, we say that (6.25) is a characteristic equation. It is natural to ask for the solutions $(k^x; x \in S)$ of (6.25) where j , S and $(\mathcal{Y}^x, x \in S)$ are given.

Theorem 6.4.1 (Solving the characteristic equation (6.25)).

- (a) Take any nonnegative function $g : \mathcal{X} \rightarrow [0, \infty)$ such that $\text{supp}g \subset \text{supp}R_{01}$ and $\sum_{y \in \mathcal{X}} R_1^x(y)g(x, y) < \infty$ for all $x \in \mathcal{X}$. Let us denote $g_t^x(z) := E_R[g(x, X_1) \mid X_t = z] > 0$, for any $t \in [0, 1)$ and $z \in \mathcal{X}^R(x, \text{supp}g(x, \cdot))$. Then,

$$k^x(t, z \rightarrow z') := \frac{g_t^x(z')}{g_t^x(z)} j(t, z \rightarrow z'), \quad t \in [0, 1), (z \rightarrow z') \in \mathcal{A}_{\rightarrow}^R(x, \text{supp}g(x, \cdot)) \quad (6.26)$$

solves (6.25) with $S := \{x \in \mathcal{X}; g(x, y) > 0 \text{ for some } y \in \mathcal{X}\}$ and $\mathcal{Y}^x = \text{supp}g(x, \cdot)$.

(b) Conversely, any solution $(k^x; x \in S)$ of (6.25) which verifies the additional requirement

$$\forall x \in S, \forall 0 \leq t < 1, \quad \sup_{y \in \mathcal{Y}^x} \int_0^t \bar{k}^x(s, y) ds < \infty, \quad (6.27)$$

has the above form (6.26) for some function g and for any $x \in S$,

$$P^x := \frac{g(x, X_1)}{g_0^x(x)} R^x \in \mathcal{P}(\Omega),$$

defines a Markov probability measure on Ω with intensity k^x given by (6.26).

The main point about this result is that unlike Theorem 6.2.1 and Proposition 6.4.1, it is not assumed that k^x given at (6.26) is the intensity of a random walk. It might happen a priori that such a random process “explodes” in finite time with a positive probability. The assumption (6.27) rules this bad behavior out.

Proof. The proof mainly consists is essentially contained in that of Theorem 6.2.1.

The first statement (a) follows from a straightforward computation. The regularity issues are direct consequences of Lemma 6.3.1.

Statement (b) is proved by considering the proof of the sufficient condition of Theorem 6.2.1 at page 148. Let k^x be a solution of (6.25). Mimicking the Girsanov formula (6.9), let us define

$$Q^x := Z_1^x R^x$$

with

$$Z_t^x := \mathbf{1}_{\{\tau > t\}} \exp \left(\sum_{0 < s < t: X_{s-} \neq X_s} \log \frac{k^x}{j}(s, X_{s-} \rightarrow X_s) - \int_0^t (\bar{k}^x - \bar{j})(s, X_s; y) dt \right), \quad 0 \leq t \leq 1$$

and the stopping time τ defined by

$$\tau := \inf \left\{ s \leq 1; k^x(s, X_{s-} \rightarrow X_s) = 0 \text{ or } \int_0^s \bar{k}^x(r, X_r) dr = \infty \right\} \in [0, 1] \cup \{\infty\}.$$

Let us show that under the assumption (6.27), Q^x is a probability measure. The process Z^x is a nonnegative local R^x -martingale. As such it is also an R^x -supermartingale. In particular, $Q^x(\Omega) = E_{R^x} Z_1^x \leq 1$, but it might happen that $Q^x(\Omega) < 1$, in which case Q^x is not a probability measure. However, the property (ii) implies that $k^x(t, X_{t-} \rightarrow X_t) > 0, \forall 0 \leq t < 1$,

$R^x - a.s.$ and the assumption (6.27) implies that for all $0 \leq t < 1$, $\tau > t$, $R^x - a.s.$ and $(Z_s^x)_{0 \leq s \leq t}$ is an R^x -martingale. In particular, $E_{R^x} Z_t^x = 1$ and $Q_{[0,t]}^x$ is probability measure for all $0 \leq t < 1$, which in turns implies that Q^x is a probability measure, since without loss of generality, one can modify the path space Ω by throwing away the R -negligible event of all paths that jump at $t = 1$.

Following almost verbatim the proof of the sufficient condition of Theorem 6.2.1, we show that there exists a function φ^x such that, as in (6.20),

$$Z_t^x = \exp(\varphi^x(t, X_t) - \varphi^x(0, x)), \quad 0 \leq t < 1.$$

It follows from the above modification of Ω at time 1 that $t \mapsto Z_t^x$ admits a version which is left-continuous at 1 and

$$Q^x = \exp(\varphi_1^x(X_1) - \varphi^x(0, x)) R^x,$$

where the limit $\varphi_1^x(X_1) := \lim_{t \rightarrow 1^-} \varphi^x(t, X_t) \in [-\infty, \infty)$ exists R^x -almost surely. Now, we are back to Proposition 6.4.1 with the functions $h(x, y) = \exp(\varphi_1^x(y) - \varphi^x(0, x))$ and $g(x, y) = \exp(\varphi_1^x(y))$ where as a convention $\exp(-\infty) = 0$. \square

6.4.3 Homogeneous walks

In this subsection only time homogeneous intensities are considered. We have seen in the previous chapters that it is often the case that two random walks of different intensities belong to the same reciprocal class, even if we restrict to the time homogeneous case. This is done for instance in Theorem 4.3.2 and Corollary 5.3.1. Here we show that this fact can only happen if the state space is infinite.

Proposition 6.4.2. *Let \mathcal{X} be finite, and j be time homogeneous, positive on \mathcal{A} , and $(\mathcal{X}, \rightarrow)$ be a connected graph. Then, if P is a time homogeneous Markov walk in $\mathfrak{R}(R)$, then the intensity of P is j .*

Proof. Assume that k is another intensity such that there exist a time homogeneous Markov walk $P \in \mathfrak{R}(R)$ of intensity k . By condition (iii) and the connectedness of $(\mathcal{X}, \rightarrow)$ it follows that k is also everywhere positive. From point (v) and Lemma 6.1.1 we have that there exist a positive function $h : \mathcal{X} \rightarrow \mathbb{R}_+$ s.t.

$$k(z \rightarrow z') = \frac{h(z')}{h(z)} j(z \rightarrow z') \quad \forall z \rightarrow z' \in \mathcal{A} \quad (6.28)$$

Imposing the equality of the arc characteristics one finds that

$$\bar{j}(z') - \bar{j}(z) = \bar{k}(z') - \bar{k}(z), \quad \forall z \rightarrow z' \in A \quad (6.29)$$

Consider now an arbitrary pair of vertices $z'' \in \mathcal{X}$. Since $(\mathcal{X}, \rightarrow)$ is connected there exist a path $\mathbf{w} := (z = x_0 \rightarrow x_1 \dots \rightarrow x_{|\mathbf{w}|} = z'')$. Summing the relation (6.29) along the arcs of \mathbf{w} , yields the same relation for z and z'' . Therefore:

$$\bar{j}(z') - \bar{j}(z) = \bar{k}(z') - \bar{k}(z), \quad \forall z, z' \in \mathcal{X}^2 \quad (6.30)$$

This gives that there exist a constant α such that $\bar{j}(z) = \alpha + \bar{k}(z)$ everywhere on \mathcal{X} . By inverting the roles of k and j we can w.l.o.g. assume that $\alpha \leq 0$. But then we have, using (6.28) and the non negativity of α :

$$\frac{1}{\bar{j}(z)} \sum_{z \rightarrow z' \in A} j(z \rightarrow z') h(z') \geq h(z) \quad \forall z \in \mathcal{X} \quad (6.31)$$

Consider now any z realizing $\max\{h(z'), z' \in \mathcal{X}\}$. Since the weights $\frac{j(z \rightarrow z')}{\bar{j}(z)}$ are positive and sum up to one, (6.31) tells that $h(z')$ is constantly equal to $h(z)$. But then by (6.28) $k(z \rightarrow z') = j(z \rightarrow z')$ for all $z \rightarrow z'$, from which the conclusion follows. \square

Remark 6.4.1. *On the other hand, if we allow for time dependent intensities, we can always find one intensity with the property that the associated walks are in $\mathfrak{R}(R)$. It suffices to take the bridges, or any fg transform.*

6.5 Examples

In this series of examples, we illustrate Theorem 6.2.1 improved by Proposition 6.2.1. We compute the reciprocal characteristic $\chi[j]$ and sometimes we consider the characteristic equation (6.25).

Directed tree

Let R be the simple random walk on a directed tree $(\mathcal{X}, \mathcal{A}_{\rightarrow})$. By “directed tree”, it is meant that $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(j)$ implies that $(z' \rightarrow z) \notin \mathcal{A}_{\rightarrow}(j)$, while “simple” means that $j(t, z \rightarrow z') = 1$ for all $(z \rightarrow z') \in \mathcal{A}_{\rightarrow}(j)$, $0 \leq t < 1$. The cycles of the corresponding undirected tree $\mathcal{A}_{\leftrightarrow}(j)$ are clearly generated by the set \mathcal{E} of all cycles of length two, see Definition 6.1.3 which matches with $\mathcal{A}_{\leftrightarrow}(j)$. Therefore, the cycle characteristic is trivial:

$\chi_c[k](t, z \leftrightarrow z') = 1$ for all $0 \leq t < 1$, $(z \leftrightarrow z') \in \mathcal{A}_{\leftrightarrow}$ and any intensity k such that $\mathcal{A}_{\rightarrow}(k) = \mathcal{A}_{\rightarrow}$. In this situation, only the arc component

$$\chi_a[j](t, z \rightarrow z') = \deg(z') - \deg(z), \quad 0 \leq t < 1, (z \rightarrow z') \in \mathcal{A}_{\rightarrow},$$

is relevant, where $\deg(z) := \#\{z' \in \mathcal{X} : (z \rightarrow z') \in \mathcal{A}_{\rightarrow}\}$ is the outer degree, i.e. the number of offsprings of z . The characteristic equation is

$$\begin{aligned} \partial_t \log k^x(t, z \rightarrow z') + \bar{k}^x(t, z') - \bar{k}^x(t, z) &= \deg(z') - \deg(z), \\ 0 \leq t < 1, (z \rightarrow z') &\in \mathcal{A}_{\rightarrow}^R(x, \mathcal{Y}), \end{aligned}$$

for some $\mathcal{Y} \subset \mathcal{X}(R^x)$.

Intensity of a bridge

In particular, the intensity j^{xy} of any bridge R^{xy} satisfies $\mathcal{A}_{\rightarrow}[j^{xy}] = \mathbf{w}^{xy}$: the only walk leading from x to y , and

$$\begin{aligned} \partial_t \log j^{xy}(t, z \rightarrow z') + \mathbf{1}_{\{z' \neq y\}} j^{xy}(t, z' \rightarrow z'') - j^{xy}(t, z \rightarrow z') \\ = \mathbf{1}_{\{z' \neq y\}} \deg(z') - \deg(z), \quad 0 \leq t < 1, (z \rightarrow z') \in \mathbf{w}^{xy}, \end{aligned}$$

where $z \rightarrow z' \rightarrow z''$ are consecutive vertices.

Birth and death process

The vertex set is $\mathcal{X} = \mathbb{N}$ with the usual graph structure. The reference walk R is governed by the time-homogeneous Markov intensity $j(z \rightarrow (z+1)) = \lambda > 0$, $z \geq 0$ and $j(z \rightarrow (z-1)) = \mu > 0$, $z \geq 1$. Clearly, the of edges $\mathcal{E} = \{(z \leftrightarrow z+1), z \in \mathbb{N}\}$ generates \mathcal{C} and the characteristics of the reference intensity are:

$$\begin{cases} \chi_a[j](z \rightarrow z+1) = \chi_a[j](z+1 \rightarrow z) = 0, & z \geq 1, \\ \chi_a[j](0 \rightarrow 1) = -\chi_a[j](1 \rightarrow 0) = \mu, \\ \chi_c[j](z \leftrightarrow z+1) = \lambda\mu, & z \geq 0. \end{cases}$$

Time-homogeneous Markov walks in $\mathfrak{R}(R)$.

Let us search for such a random walk $P \in \mathcal{P}(\Omega)$. We denote $\tilde{\lambda}(z)$ the intensity of $(z \rightarrow z+1)$ and $\tilde{\mu}(z+1)$ the intensity $(z+1 \rightarrow z)$ of the Markov walk P . By Theorem 6.2.1, $P \in \mathcal{R}[j]$ if and only if

$$\begin{cases} \tilde{\lambda}(z+1) + \tilde{\mu}(z+1) - \tilde{\lambda}(z) - \tilde{\mu}(z) = 0, & z \geq 1 \\ \tilde{\lambda}(1) + \tilde{\mu}(1) - \tilde{\lambda}(0) = \mu, \\ \tilde{\lambda}(z)\tilde{\mu}(z+1) = \lambda\mu, & z \geq 0. \end{cases}$$

The solutions to the the above set of equations can be parametrized by choosing $\tilde{\lambda}(0)$ arbitrarily and finding $\tilde{\lambda}(z+1), \tilde{\mu}(z+1)$ recursively as follows

$$\begin{cases} \tilde{\mu}(z+1) = \tilde{\lambda}(z)^{-1}\lambda\mu, & z \geq 0, \\ \tilde{\lambda}(z+1) = \mu + \tilde{\lambda}(0) - \tilde{\mu}(z+1), & z \geq 1. \end{cases}$$

With some simple computations one can see that for any large enough $\tilde{\lambda}(0)$, the above system admits a unique positive and bounded solution. Hence, the corresponding Markov walk has its sample paths in Ω and it is in $\mathfrak{R}(R)$.

Hypercube

Let $\mathcal{X} = \{0, 1\}^d$ be the d -dimensional hypercube with its usual directed graph structure and let $\{g_i\}_{i=1}^d$ be the canonical basis. For $x \in \mathcal{X}$, we set $x^i := x + g_i$ and $x^{ik} = x + g_i + g_k$ where we consider the addition modulo 2.

Proposition 6.5.1. *Let*

$$\mathcal{S} := \{(x \rightarrow x^i \rightarrow x^{ik} \rightarrow x^k \rightarrow x), x \in \mathcal{X}, 1 \leq i, k \leq d\}$$

be the set of all directed squares. The subset

$$\mathcal{S} \cup \mathcal{E} \tag{6.32}$$

is a generating set for the cycles as in Definition 6.1.4.

Proof. Fix a function ℓ such that $\ell(\mathbf{c}) = 0$ for all $\mathbf{c} \in \mathcal{S} \cup \mathcal{E}$. We first observe that a walk $\mathbf{w} = (x_0 \rightarrow \dots \rightarrow x_n)$ is completely described by the pair (x_0, I) , where $I = (i_1, \dots, i_n)$ is defined through

$$g_{i_r} = x_r - x_{r-1}, \quad \forall 1 \leq r \leq n$$

In the rest of the proof we will label walks equivalently by \mathbf{w} or (x_0, I) . We also define for all I :

$$\mathbb{N}^d \ni N(I), \quad N(I)_j = \#\{r : i_r = j\}, \quad \forall 1 \leq j \leq d$$

Note that \mathbf{w} is a cycle if and only if $N(I) \in (2\mathbb{N})^d$. We also define some other notation. We denote by \oplus the concatenation of two walks. If $\mathbf{w} = (x_0 \rightarrow \dots \rightarrow x_n = y_0)$ and $\mathbf{w}' = (y_0 \rightarrow \dots \rightarrow y_m)$, $\mathbf{w} \oplus \mathbf{w}'$ is $(x_0 \rightarrow \dots \rightarrow x_{n-1} \rightarrow y_0 \rightarrow \dots \rightarrow y_m)$. If $\mathbf{c} = (x_0, I)$ is a cycle, we define for every s the cycle $\mathbf{c}^s = (x_0, I^s)$ as the one identified by $I^s = (i_1, \dots, i_{s+1}, i_s, \dots, i_n)$. The proof can be organized in three main steps:

Step 1 In this first step we show that for every $x_0 \in \mathcal{X}$, $\mathbf{n} \in (2\mathbb{N})^d$ there exist at least a cycle $\mathbf{c} = (x_0, I)$ such that:

(a) $N(I) = \mathbf{n}$

(b) $\ell(\mathbf{c}) = 0$

We construct \mathbf{c} directly. According to our notation, the walk $(x_0, (i, i))$ is $(x_0 \rightarrow x_0^i \rightarrow x_0)$, and such a walk belong to \mathcal{E} . If we take

$$\mathbf{c} = \bigoplus_{i=1}^d \underbrace{(x_0, (i, i)) \oplus \dots \oplus (x_0, (i, i))}_{\frac{n_i}{2} \text{ times}} \quad (6.33)$$

we see immediatly that it satisfies (a). To see that it satisfies also (b) we use the hypothesis on the cycles of length two. We have: $\ell(\mathbf{c}) = \sum_i n_i \ell((x_0, (i, i))) = 0$, since $(x_0, (i, i)) \in \mathcal{E}$.

Step 2 In this step we show that if $\ell(\mathbf{c}) = 0$, then $\ell(\mathbf{c}^s) = 0$ as well. Let $\mathbf{c} = (x_0 \rightarrow \dots \rightarrow x_n)$. Consider the cycles $\mathbf{c}_a := (x_{s-1} \rightarrow x_{s-1} + g_{i_{s+1}} \rightarrow x_{s-1}), \mathbf{c}_b := (x_{s-1} + g_{i_{s+1}} \rightarrow x_{s+1} \rightarrow x_{s-1} + g_{i_{s+1}}) \in \mathcal{E}$ and the cycle $\mathbf{c}_c = (x_{s-1} \rightarrow x_s \rightarrow x_{s+1} \rightarrow x_{s+1} + g_{i_s} \rightarrow x_{s-1}) \in \mathcal{S}$. it can be verified directly that:

$$\ell(\mathbf{c}^s) = \ell(\mathbf{c}) - \ell(\mathbf{c}_c) + \ell(\mathbf{c}_a) + \ell(\mathbf{c}_b)$$

Since by hypothesis, $\ell(\mathbf{c}) = 0$, and by hypothesis $\ell(\mathbf{c}_c) = \ell(\mathbf{c}_a) = \ell(\mathbf{c}_b) = 0$, we obtain $\ell(\mathbf{c}^s) = 0$.

Step 3 In this step, we take an arbitrary cycle $\mathbf{c}' = (x_0, I')$ and consider $\mathbf{c} = (x_0, I)$ defined in (6.33) where $\mathbf{n} = N(I')$. It is easy to see that there exist a sequence s_1, \dots, s_K such that $\mathbf{c}' = \mathbf{c}^{s_1, \dots, s_K}$. Using iteratively Step 2 we then get that $\ell(\mathbf{c}') = 0$. This concludes the proof. \square

Remark 6.5.1. *One can produce a smaller generator for the cycles by imposing $i < k$ in the definition of \mathcal{S} . However, no more than this can be done since \mathcal{E} does not generate \mathcal{C} there are no cycles of length 3.*

The bridge of a simple random walk on the discrete hypercube

Let j be the simple random walk on the hypercube. The intensity $j^{xy}(t, z \rightarrow z')$ of the xy -bridge can be computed explicitly with (6.24) since the transition density of the random walk is known explicitly. We have

$$j^{xy}(t, z \rightarrow z') = \begin{cases} \cosh(1-t)/\sinh(1-t), & \text{if } z_i \neq y_i, \\ \sinh(1-t)/\cosh(1-t), & \text{if } z_i = y_i, \end{cases} \quad (6.34)$$

where z_i and $y_i \in \{0, 1\}$ are the i -th coordinates of z and $y \in \mathcal{X}$.

We provide an alternate proof based on the characteristic equation (6.25). First, it is immediate to see that under any bridge, all arcs of the hypercube

are active at any time. From $\chi_c[j^{xy}] = \chi_c[j]$, we deduce that the arc function $\log(j/j^{xy})(t, \cdot)$ is the gradient of some potential ψ_t , see Lemma 6.1.1. The equality of the arc characteristics implies that for all $t \in (0, 1)$ and $z \in \mathcal{X}$

$$\partial_t \psi(t, z) + \sum_{i=1}^d [\exp(\psi_t(z^i) - \psi_t(z)) - 1] = \partial_t \psi(t, x) + \sum_{i=1}^d [\exp(\psi_t(x^i) - \psi_t(x)) - 1].$$

Since ψ is defined up to the addition of a function of time, we can assume without loss of generality that for all $0 < t < 1$, $\partial_t \psi_t(x) + \sum_{i=1}^d [\exp(\psi_t(x^i) - \psi_t(x)) - 1] = 0$. Hence ψ solves the HJB equation:

$$\partial_t \psi(t, z) + \sum_{i=1}^d [\exp(\psi_t(z^i) - \psi_t(z)) - 1] = 0, \quad t \in [0, 1], z \in \mathcal{X}. \quad (6.35)$$

Going along the lines of the proof of Theorem 6.2.1, in particular equation (6.20), allows to deduce that the boundary data for ψ are

$$\lim_{t \rightarrow 1} \psi_t(z) = \begin{cases} -\infty, & \text{if } z \neq y, \\ 0, & \text{if } z = y. \end{cases} \quad (6.36)$$

One can check with a direct computation that the solution (6.35) & (6.36) is

$$\psi(t, z) = \sum_{i=1}^d \log[1 + (-1)^{(z_i - y_i)} e^{2(1-t)}] \quad (6.37)$$

where the subtraction is considered modulo two. By the definition of ψ , we have

$$j^{xy}(t, z \rightarrow z^i) = j(t, z \rightarrow z^i) \exp(\psi_t(z^i) - \psi_t(z)) = \exp(\psi_t(z^i) - \psi_t(z))$$

and (6.34) follows with a simple computation.

Two triangles

We look at two simple directed trees based on triangles.

Oriented triangle

Let $\mathcal{X} = \{A, B, C\}$ and $\mathcal{A} = \{(A \rightarrow B), (B \rightarrow C), (C \rightarrow A)\}$. The reference intensity is j_Δ on each arc and we want to find the intensity of the AB -bridge: $j^{AB}(t, \cdot)$, using the characteristic equation. Imposing the equality



Figure 6.1: Left: an oriented triangle. Right: a non-oriented triangle

of the cycle characteristics implies that $\log(j^{AB}/j)(t, \cdot)$ is the gradient of some potential $\psi_t : \mathcal{X} \rightarrow \mathbb{R}$. The equality of the arc characteristics implies that ψ solves the HJB equation

$$\begin{cases} \partial_t \psi_t(A) + j_\Delta [\exp(\psi_t(B) - \psi_t(A)) - 1] = 0, \\ \partial_t \psi_t(B) + j_\Delta [\exp(\psi_t(C) - \psi_t(B)) - 1] = 0, \\ \partial_t \psi_t(C) + j_\Delta [\exp(\psi_t(A) - \psi_t(C)) - 1] = 0, \\ \lim_{t \rightarrow 1} \psi_t(A) = \lim_{t \rightarrow 1} \psi_t(C) = -\infty, \\ \lim_{t \rightarrow 1} \psi_t(B) = 0, \end{cases} \quad (6.38)$$

where the boundary conditions for ψ follow from (6.20). Since the HJB equation is the logarithm of the Kolmogorov backward equation, we obtain the following solutions

$$\begin{cases} \psi_t(A) = \log\left\{\frac{1}{3} + \frac{2}{3} \exp\left(-\frac{3}{2}j_\Delta(1-t)\right) \sin\left[\frac{\sqrt{3}}{2}j_\Delta(1-t) - \frac{\pi}{6}\right]\right\}, \\ \psi_t(B) = \log\left\{\frac{1}{3} + \frac{2}{3} \exp\left(-\frac{3}{2}j_\Delta(1-t)\right) \cos\left[\frac{\sqrt{3}}{2}j_\Delta(1-t)\right]\right\}, \\ \psi_t(C) = \log\left\{\frac{1}{3} - \frac{2}{3} \exp\left(-\frac{3}{2}j_\Delta(1-t)\right) \sin\left[\frac{\sqrt{3}}{2}j_\Delta(1-t) + \frac{\pi}{6}\right]\right\}. \end{cases} \quad (6.39)$$

We deduce the following identities

$$\begin{cases} j_T^{AB}(t, A \rightarrow B) = j_\Delta \frac{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) + 2 \cos\left(\frac{\sqrt{3}}{2}j_\Delta(1-t)\right)}{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) + 2 \sin\left(\frac{\sqrt{3}}{2}j_\Delta(1-t) - \frac{\pi}{6}\right)}, \\ j_T^{AB}(t, B \rightarrow C) = j_\Delta \frac{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) - 2 \sin\left(\frac{\sqrt{3}}{2}j_\Delta(1-t) + \frac{\pi}{6}\right)}{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) + 2 \cos\left(\frac{\sqrt{3}}{2}j_\Delta(1-t)\right)}, \\ j_T^{AB}(t, C \rightarrow A) = j_\Delta \frac{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) + 2 \sin\left(\frac{\sqrt{3}}{2}j_\Delta(1-t) - \frac{\pi}{6}\right)}{\exp\left(\frac{3}{2}j_\Delta(1-t)\right) - 2 \sin\left(\frac{\sqrt{3}}{2}j_\Delta(1-t) + \frac{\pi}{6}\right)}. \end{cases} \quad (6.40)$$

Non-oriented triangle

The triangle $\mathcal{X} = \{A, B, C\}$ is now equipped with the directed graph structure $\mathcal{A} = \{(A \rightarrow B), (B \rightarrow C), (A \rightarrow C)\}$ where we reverted the direction of the arc on the edge AC , as shown at Figure 6.1. The characteristic associated to the cycle $(A \rightarrow B \rightarrow C \rightarrow A)$ is

$$j(A \rightarrow B)j(B \rightarrow C)/j(C \rightarrow A).$$

Since it is not a cycle of the graph $(\mathcal{X}, \mathcal{A})$, the interpretation given at Theorem 6.2.2 is not available for this characteristic. However, we can reason in a similar way to obtain a probabilistic interpretation of this characteristic as well.

Let R^A be the reference walk conditioned to start from A . If the walk reaches C at time h it is easy to see that as $h \rightarrow 0$, this has happened essentially only by using directly the arc $(A \rightarrow C)$. Therefore, we obtain

$$R^A(X_h = C) = j(A \rightarrow C)h + o(h).$$

Similarly, the probability of going from A to C using the path $(A \rightarrow B \rightarrow C)$ is

$$\begin{aligned} R^A\left((X_0 \rightarrow X_{T_1} \rightarrow X_{T_2}) = (A \rightarrow B \rightarrow C), T_2 \leq h, T_3 > h\right) \\ = j(A \rightarrow B)j(B \rightarrow C)h^2/2 + o(h^2). \end{aligned}$$

Consequently,

$$\begin{aligned} R^A\left((X_0 \rightarrow X_{T_1} \rightarrow X_{T_2}) = (A \rightarrow B \rightarrow C), T_2 \leq h, T_3 > h \mid X_h = C\right) \\ = \frac{j(A \rightarrow B)j(B \rightarrow C)}{2j(A \rightarrow C)}h + o(h). \end{aligned}$$

We see that the characteristic is twice the driving factor of the expansion of this probability as h tends to zero.

Note that while the characteristic in the oriented triangle is associated to a probability of order h^3 , in the present case it is associated to a probability of order h .

Planar graphs

Let $(\mathcal{X}, \leftrightarrow)$ be an undirected symmetric planar graph. We fix a planar representation and consider the set \mathcal{F} of all the counter-clockwise cycles along the faces. We denote by \mathcal{E} the set of all the edges seen as closed 2-walks. Then,

Proposition 6.5.2. *The set*

$$\mathcal{F} \cup \mathcal{E} \quad (6.41)$$

generates the cycles of the planar graph.

To keep the proof into a reasonable size, we will use some basic vocabulary about planar graphs, which we do not define. We believe that this will not generate any confusion. Figure 6.2 should be self explanatory, and clarify any possible doubt.

Proof. Consider a function $\ell : \mathcal{A} \rightarrow \mathbb{R}^+$ such that $\ell(\mathbf{c}) = 0$ along any cycle $\mathbf{c} \in \mathcal{F} \cup \mathcal{E}$. Consider any other simple cycle \mathbf{c} , counterclockwise oriented. Then we have that

$$\ell(\mathbf{c}) = \sum_{\mathbf{f} \in \mathcal{F}, \mathbf{f} \text{ internal to } \mathbf{c}} \ell(\mathbf{f}) - \sum_{z \rightarrow z' \text{ internal to } \mathbf{c}} \ell((z \rightarrow z' \rightarrow z)) \quad (6.42)$$

Since all the summands are zero by hypothesis, then $\ell(\mathbf{c}) = 0$. Consider now the cycle \mathbf{c} , clockwise oriented. The same reasoning as above, using the formula

$$\begin{aligned} \ell(\mathbf{c}) = - \sum_{\mathbf{f} \in \mathcal{F}, \mathbf{f} \text{ internal to } \mathbf{c}} \ell(\mathbf{f}) + \sum_{z \rightarrow z' \text{ internal to } \mathbf{c}} \ell((z \rightarrow z' \rightarrow z)) \\ + \sum_{(z \rightarrow z') \in \mathbf{c}} \ell((z \rightarrow z' \rightarrow z)) \end{aligned} \quad (6.43)$$

completes the proof. \square

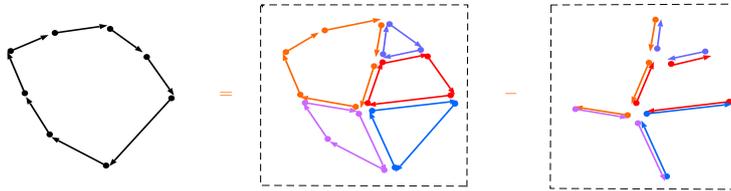


Figure 6.2: An explanation of the formula in (6.42). On the left is depicted a counterclockwise oriented cycle, in the middle its internal faces, and on the right the cycles length two which are internal to the cycle.

Triangular prism

The set $\mathcal{X} = \{A_0, B_0, C_0, A_1, B_1, C_1\}$ is endowed with the directed graph structure as in Figure 6.3 where one should see the left triangle $A_0B_0C_0$ on the picture as the bottom face of the prism and the horizontal arcs of the picture as flowing along the three vertical edges of the prism. The intensity j is time-homogeneous and $j(z \rightarrow z') = j_\Delta$ if $(z \rightarrow z')$ belongs to a triangular face with one given orientation and $j(z \rightarrow z') = j_v$ if $(z \rightarrow z')$ connects the triangular faces. The cycle characteristics of the triangular

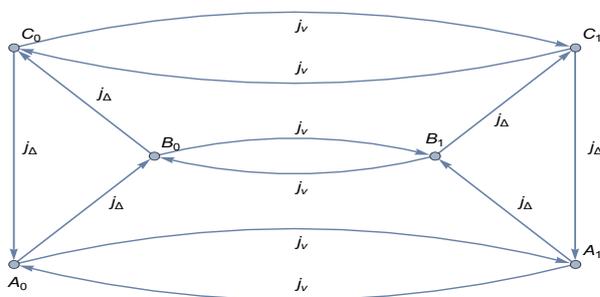


Figure 6.3: Triangular prism.

faces is $\chi_\Delta = j_\Delta^3$ and for the cycles of length two corresponding to the vertical edges we have $\chi_v = j_v^2$.

We are going to derive an explicit expression of the jump intensity $j^{A_0B_1}$ of the bridge from A_0 to B_1 , see (6.44) below.

The nice feature of this example is that it is a non trivial planar graph where the intensity of the bridge can be explicitly computed. This is achieved by putting together some already done calculations about the hypercube and the oriented triangle. Without getting into details, the fact that the prism is the product of the oriented triangle treated at page 161 and the complete graph with two vertices, which is the discrete hypercube of dimension 1, is the key of the following computation.

As in the previous example, imposing the characteristic equation leads to the fact that $j^{A_0, B_1}(t, \cdot)/j(\cdot)$ is the gradient of some potential ψ which solves the following HJB equation:

$$\begin{cases} \partial_t \psi_t(A_0) + j_\Delta [\exp(\psi_t(B_0) - \psi_t(A_0)) - 1] + j_v [\exp(\psi_t(A_1) - \psi_t(A_0)) - 1] = 0 \\ \partial_t \psi_t(B_0) + j_\Delta [\exp(\psi_t(C_0) - \psi_t(B_0)) - 1] + j_v [\exp(\psi_t(B_1) - \psi_t(B_0)) - 1] = 0 \\ \partial_t \psi_t(C_0) + j_\Delta [\exp(\psi_t(A_0) - \psi_t(C_0)) - 1] + j_v [\exp(\psi_t(C_1) - \psi_t(C_0)) - 1] = 0 \\ \partial_t \psi_t(A_1) + j_\Delta [\exp(\psi_t(B_1) - \psi_t(A_1)) - 1] + j_v [\exp(\psi_t(A_0) - \psi_t(A_1)) - 1] = 0 \\ \partial_t \psi_t(B_1) + j_\Delta [\exp(\psi_t(C_1) - \psi_t(B_1)) - 1] + j_v [\exp(\psi_t(B_0) - \psi_t(B_1)) - 1] = 0 \\ \partial_t \psi_t(C_1) + j_\Delta [\exp(\psi_t(A_1) - \psi_t(C_1)) - 1] + j_v [\exp(\psi_t(C_0) - \psi_t(C_1)) - 1] = 0 \end{cases}$$

with the boundary conditions

$$\begin{cases} \lim_{t \rightarrow 1} \psi_t(B_1) = 0 \\ \lim_{t \rightarrow 1} \psi_t(B_0) = \lim_{t \rightarrow 1} \psi_t(A_i) = \lim_{t \rightarrow 1} \psi_t(C_i) = -\infty, \quad i \in \{0, 1\} \end{cases}$$

The symmetric structure of the graph allows to guess the solution. It can be verified with a direct computation that

$$\psi_t(A_i) = \psi_t^T(A) + \psi_t^H(i)$$

where $\psi_t^T : (0, 1) \times \{A, B, C\} \rightarrow \mathbb{R}$ is the solution of the HJB equation on the triangle (6.38), which is solved at (6.39) and ψ_t^H is the solution of the following HJB equation on the complete graph with two vertices (which is nothing but the discrete hypercube in dimension 1):

$$\begin{cases} \partial_t \psi_t^H(0) + j_v [\exp(\psi_t^H(1) - \psi_t^H(0)) - 1] = 0 \\ \partial_t \psi_t^H(1) + j_v [\exp(\psi_t^H(0) - \psi_t^H(1)) - 1] = 0 \\ \lim_{t \rightarrow 1} \psi_t^H(1) = 0 \\ \lim_{t \rightarrow 1} \psi_t^H(0) = -\infty \end{cases}$$

which is solved in a more general form at (6.37). The same reasoning is valid for the calculations of $\psi(B_i)$ and $\psi(C_i)$. Using these explicit formulas, we obtain ψ and

$$j^{A_0, B_1}(t, z \rightarrow z') = \exp(\psi_t(z') - \psi_t(z)) j(z \rightarrow z') \quad (6.44)$$

is the explicit expression of the jump intensity $j^{A_0 B_1}$ of the bridge from A_0 to B_1 . Therefore we have, for instance:

$$\begin{aligned} j^{A_0, B_1}(t, A_0 \rightarrow B_0) &= \exp(\psi_t(B_0) - \psi_t(A_0)) j_\Delta \\ &= \exp(\psi_t^T(B) - \psi_t^T(A) + \psi_t^H(0) - \psi_t^H(0)) j_\Delta \\ &= \exp(\psi_t^T(B) - \psi_t^T(A)) j_\Delta \\ &= j_\Delta \frac{\exp\left(\frac{3}{2} j_\Delta (1-t)\right) + 2 \cos\left(\frac{\sqrt{3}}{2} j_\Delta (1-t)\right)}{\exp\left(\frac{3}{2} j_\Delta (1-t)\right) + 2 \sin\left(\frac{\sqrt{3}}{2} j_\Delta (1-t) - \frac{\pi}{6}\right)} \end{aligned}$$

Complete graph

The directed graph structure of the complete graph on a finite set $\mathcal{X} = \{1, \dots, |\mathcal{X}|\}$ consists of all the couples of distinct vertices, the set of arcs is $\mathcal{A}_{\rightarrow} = \mathcal{X} \setminus \{(x, x); x \in \mathcal{X}\}$. Pick an arbitrary vertex $* \in \mathcal{X}$ and consider the set

$$\mathcal{T}_* := \{(* \rightarrow z \rightarrow z' \rightarrow *); z, z', * \text{ distinct}\}$$

of all directed triangles containing $*$.

Proposition 6.5.3. *The set*

$$\mathcal{E} \cup \mathcal{T}_*$$

generates the set of cycles of the complete graph.

Proof. Let ℓ be a function such that $\ell(\mathbf{c}) = 0$ on all cycles in $\mathcal{E} \cup \mathcal{T}_*$. Consider any simple cycle $\mathbf{c} = (x_0 \rightarrow \dots \rightarrow x_n = x_0)$. Two cases are possible:

- (a) $*$ is not touched by \mathbf{c} . In this case for all $0 \leq i \leq n - 1$ we consider the triangle. $\mathbf{t}_i = (* \rightarrow x_i \rightarrow x_{i+1} \rightarrow *)$. The formula

$$\ell(\mathbf{c}) = \sum_{i=0}^{n-1} \ell(\mathbf{t}_i) \tag{6.45}$$

whose validity can be checked with a direct computation, yields the conclusion together with the assumption.

- (b) If $*$ is touched by \mathbf{c} , then w.l.o.g. we can assume that $* = x_1$. We define the triangle $\mathbf{t} = (* \rightarrow x_2 \rightarrow x_0 \rightarrow *)$, the cycle of length two $\mathbf{e} = (x_0 \rightarrow x_2 \rightarrow x_0)$. We observe that the cycle $\mathbf{c}' = (x_0 \rightarrow x_2 \rightarrow \dots \rightarrow x_n = x_0)$ does not contain $*$ and therefore by point (a) $\ell(\mathbf{c}') = 0$. The formula

$$\ell(\mathbf{c}) = \ell(\mathbf{c}') + \ell(\mathbf{t}) - \ell(\mathbf{e})$$

gives the desired result.

□

Some sampler

Let us analyze in a bit more detail one example of a walk on the complete graph. Take $\pi \in \mathcal{P}(\mathcal{X}^2)$ a positive probability distribution on the finite set

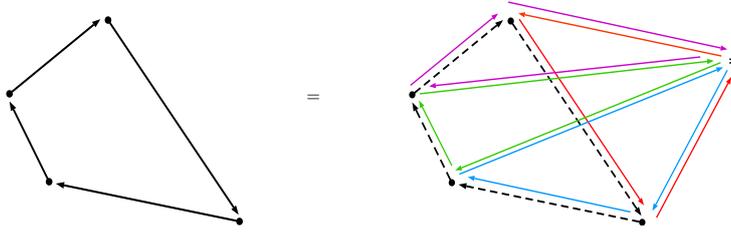


Figure 6.4: Decomposition of a cycle into 3-cycles and 2-cycles

\mathcal{X} . The detailed balance conditions: $\pi(z)j(z \rightarrow z') = \pi(z')j(z' \rightarrow z), \forall z, z'$, tell us that the intensity

$$j(z \rightarrow z') = \sqrt{\frac{\pi(z')}{\pi(z)}} \quad z, z' \in \mathcal{X}$$

admits π as its reversing measure. The characteristics associated with j are

$$\begin{aligned} \chi_a[j](t, (z \rightarrow z')) &= \left[\sum_{x \in \mathcal{X}} \pi(x)^{1/2} \right] (\pi(z')^{-1/2} - \pi(z)^{-1/2}) \\ \chi_c[j](t, \mathbf{c}) &= 1 \end{aligned}$$

for any $0 \leq t \leq 1$, any arc $(z \rightarrow z')$ and any cycle \mathbf{c} .

Cayley graphs: a second look

In this section we revisit some examples we have already made in Chapter 5, in view of our new results. Some new graphs are studied as well, and the non Abelian case is treated.

Let $(\mathcal{X}, *)$ be a group and $\mathcal{G} = \{g_i; i \in I\}$ be a finite subset generating \mathcal{X} . The directed graph structure associated with \mathcal{G} is defined for any $z, z' \in \mathcal{X}$ by $z \rightarrow z'$ if $z' = zg$ for some $g \in \mathcal{G}$. We introduce the time independent reference intensity j given by

$$j(z \rightarrow zg_i) := j_i, \quad \forall z \in \mathcal{X}, g_i \in \mathcal{G},$$

where $j_i > 0$ only depends of the direction g_i . The dynamics of the random walk R is Markov and both time-homogeneous and invariant with respect to left translations, i.e. for all $z_o, z, z' \in \mathcal{X}$, $j(z_o z \rightarrow z_o z') = j(z \rightarrow z')$. For all arc $(z \rightarrow z')$, we have

$$\chi_a(z \rightarrow z') = 0$$

and the cycle characteristic χ_c is translation invariant.

Proposition 6.5.4. *Let j and k be two positive Markov intensities on this Cayley graph which are time-homogeneous and invariant with respect to the left translations. Then, they share the same bridges if and only if for any $n \geq 1$ and $(i_1, \dots, i_n) \in I^n$ with $g_{i_1} \cdots g_{i_n} = e$, we have $j_{i_1} \cdots j_{i_n} = k_{i_1} \cdots k_{i_n}$.*

As usual, we have denoted e the neutral element.

Proof. We have already seen that $\chi_a[j] = \chi_a[k] = 0$. On the other hand, the relation $g_{i_1} \cdots g_{i_n} = e$ means that $\mathbf{c} := (e \rightarrow g_{i_1} \rightarrow g_{i_1}g_{i_2} \rightarrow \cdots \rightarrow g_{i_1}g_{i_2} \cdots g_{i_{n-1}} \rightarrow e)$ is a cycle and the identity $j_{i_1} \cdots j_{i_n} = k_{i_1} \cdots k_{i_n}$ means that $\chi_c[j](\mathbf{c}) = \chi_c[k](\mathbf{c})$. We conclude with Theorem 6.2.1, Proposition 6.2.1 and the invariance with respect to left translations. \square

Remark 6.5.2. *If the group \mathcal{X} is Abelian, Proposition 6.5.4 is covered by corollary 5.3.1.*

Triangular lattice

The triangular lattice is the Cayley graph generated by $g_i = (\cos(\frac{2\pi}{3}(i - 1)), \sin(\frac{2\pi}{3}(i - 1)))$, $i = 1, 2, 3$, and we consider a time-homogeneous and translation invariant Markov intensity j .

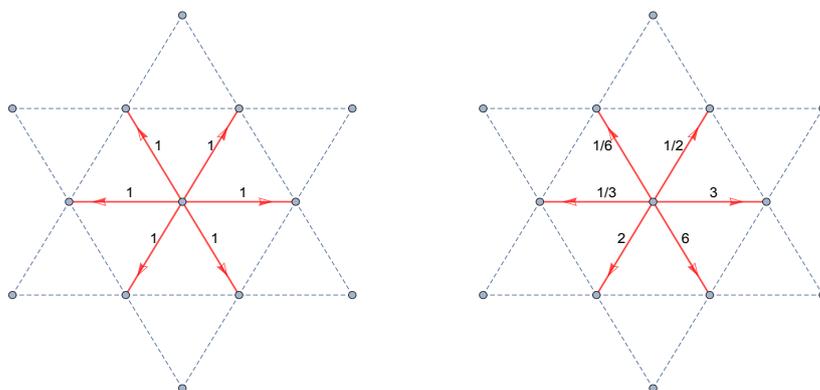


Figure 6.5: Two different space-time homogeneous random walks on the triangular lattice which belong to the same reciprocal class

For any cycle $(z \leftrightarrow z + g_i)$ associated with an edge, we have

$$\chi_c[j](t, z \leftrightarrow z + g_i) = j_i j_{-i}.$$

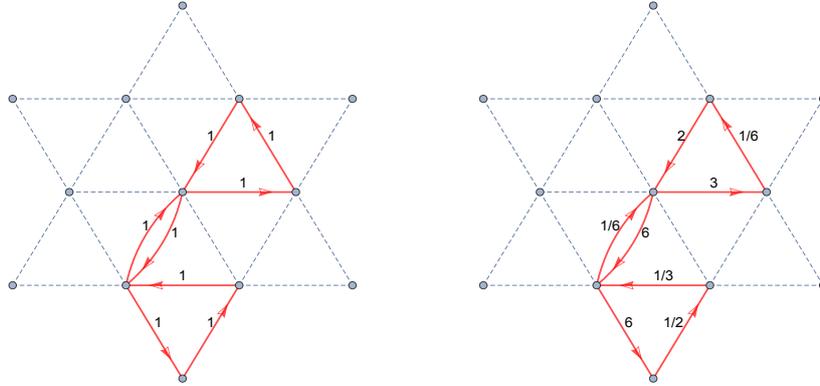


Figure 6.6: The cycle characteristics coincide

If we take any counterclockwise oriented face, i.e. a cycle of the form $\Delta_z := (z \rightarrow z + g_1 \rightarrow z + g_1 + g_2 \rightarrow z)$ for $z \in X$ we have

$$\chi_c[j](t, \Delta_z) = j_1 j_2 j_3.$$

We address the question of finding another space-time homogeneous assignment $\{k_{\pm i}\}_{1 \leq i \leq 3}$ such that the corresponding walk belongs to $\mathfrak{R}(R)$. Applying Theorem 6.2.1 and (6.41), or equivalently invoking Proposition 6.5.4, we can parametrize the solutions k as follows

$$\begin{cases} k_1 = \alpha j_1, & k_{-1} = \alpha^{-1} j_{-1} \\ k_2 = \beta j_2, & k_{-2} = \beta^{-1} j_{-2} \\ k_3 = (\alpha\beta)^{-1} j_3, & k_{-3} = \alpha\beta j_{-3} \end{cases}$$

where $\alpha, \beta > 0$. Corollary 6.2.1 gives some details about the dynamics of the bridge R^{xy} as the unique Markov walk (modulo technical conditions) that starts in x , ends in y and such that, if $h > 0$ is a very small duration:

1. At any time t and independently from the current state, it goes back and forth along the direction i during $[t, t+h]$ with probability $j_i j_{-i} h^2 / 2 + o(h^2)$.
2. At any time t and independently from the current state, it goes around the perimeter of a triangular cell of the lattice in the counterclockwise sense during $[t, t+h]$ with probability $j_1 j_2 j_3 h^3 / 6 + o(h^3)$.
3. If exactly one jump occurs during $[t, t+h]$, then the density of the instant of jump is constant up to a correction factor of order $o(h)$. This follows from $\chi_a[j](t, z \rightarrow z') = 0$ for all t and $(z \rightarrow z')$.

Rooted regular directed tree

It is an infinite directed tree such that each vertex admits exactly $m \geq 1$ offsprings. Except for the root, all the vertices have the same index $m+1$. It is the Cayley tree rooted at $* = e$ and generated by $\mathcal{G} = \{g_1, \dots, g_m\}$ where these m branches are free from each other: they do not satisfy any relation (of the type $g_{i_1} \cdots g_{i_n} = e$). This freedom is equivalent to the nonexistence of cycles which is the defining property of a tree.

As a direct consequence of Proposition 6.5.4 we obtain the following

Corollary 6.5.1. *Two positive, time-homogeneous and translation invariant Markov intensities j and k on a rooted regular directed tree generate the same bridges: $\mathfrak{R}(j) = \mathfrak{R}(k)$.*

In particular, this implies that these bridges are insensitive to time scaling: $\mathfrak{R}(\alpha j) = \mathfrak{R}(j)$, for all $\alpha > 0$.

The lattice \mathbb{Z}^d

The usual directed graph structure on the vertex set $\mathcal{X} = \mathbb{Z}^d$ is the Cayley graph structure generated by $\mathcal{G} = \{g_i, g_{-i}; 1 \leq i \leq d\}$ with $g_i = (0, \dots, 0, 1, 0, \dots, 0)$ where 1 is the i -th entry and we denote $g_{-i} = -g_i$. As another consequence of Proposition 6.5.4 we obtain the following

Corollary 6.5.2. *Two time-homogeneous and translation invariant positive Markov intensities j and k on \mathbb{Z}^d generate the same bridges if and only if for all $1 \leq i \leq d$, they satisfy*

$$j_i j_{-i} = k_i k_{-i}, \quad \forall 1 \leq i \leq d$$

where j_{-i} and k_{-i} are the intensities of jump in the direction $g_{-i} = -g_i$.

Proof. This set of equalities corresponds to the identification of the cycle characteristic along the edges. It is then easy to verify, using Proposition 6.5.4 and Remark 6.5.2 that it suffices to check only the combinations of the form $g_i g_{-i}$, which gives the conclusion. \square

Hypercube, again

Let us visit once more the hypercube $\mathcal{X} = (\mathbb{Z}/2\mathbb{Z})^d$ which is seen now as the Cayley graph generated by the canonical basis $g_i = (0, \dots, 0, 1, 0, \dots, 0)$, $1 \leq i \leq d$, where 1 is the i -th entry. As another consequence of Proposition 6.5.4 we obtain the following

Corollary 6.5.3. *Two time-homogeneous and translation invariant positive Markov intensities on the hypercube generate the same bridges if and only if they coincide.*

Proof. The proof is the same as Corollary 6.5.2's one. But this time $g_{-i} = g_i$, so that $j_i j_{-i} = k_i k_{-i}$ is equivalent to $j_i^2 = k_i^2$. \square

Chapter 7

Perspectives

We conclude by listing some problems that may be interesting to study, and are related to the content of this thesis.

Quantitative estimates based on reciprocal characteristics In this thesis we have studied reciprocal classes of random walks on a general graphs. Several natural questions arise concerning the possibility of deriving quantitative estimates based on the reciprocal characteristics. This seems to be a difficult task, and we partially addressed in the Sections 3.3. and 4.7. These results are very likely to hold in a general framework, and we simply spotted some particular cases. However, how reciprocal characteristics affect the long time behavior of bridges, and a quantitative estimate in terms of the characteristics on how fast the solution to the Schrödinger problem converges to the target final distribution are not known. They are natural questions. Curiously, even the simple question if there is a unique invariant measure for a reciprocal class which describes the distribution at time 0 of a long bridge between $[-t, t]$ has not been addressed so far. To mention some of the possible difficulties that arise when dealing with these problems let us just mention that we noticed how the problem of studying the long time behavior of a bridge is somehow equivalent to study the value of spin in 0 in a spin system with boundary conditions at $-n, n$, and how this value is affected by them. As far as we know, there is not a natural functional inequality attached to this problem, whereas this is the case when studying convergence to equilibrium for a time homogeneous Markov process. The use of *entropic interpolations*, based on solutions the Schrödinger Problem, to derive functional inequalities, is a theory which is being developed. At the moment, it is not based on the reciprocal characteristics, even though we know that the entropic interpolation is in the

reciprocal class of the reference dynamics. Understanding this connection, will probably shed more light on all the above mentioned problems.

Lévy processes and Lévy driven diffusions Reciprocal characteristics have not been computed for such processes. However, we did some preliminary work and it seems very likely that the methods we employed in this work, including duality formulae, carry over to these situations. This would allow to study the reciprocal class of a Lévy process or a Lévy driven diffusion.

A general notion of conditional characteristic Reciprocal characteristics express some quantities which are invariant over the set of probabilities that, when conditioned to the initial and final endpoints, are equal to a reference probability. The same can be done for a general conditioning, different from the bridges, and not necessarily on a path space. One takes a reference distribution, chooses an observation function, and a class of models, classified through a set of parameters. Then one asks what models are *conditionally equal* to the reference distribution, and tries to answer this question in terms of the parameter set. A characteristic should then simply be a particular functional of the parameters that is equal over all the conditionally equivalent models. As a first step, in [18] we have obtained preliminary results on this, generalizing the Schrödinger problem to the case when one is allowed to observe the full trajectory of the sample paths of some coordinates of a multidimensional diffusion. We compute the characteristics and show that the duality formulae approach can be transferred to this case.

The last question



Figure 7.1: A very nice planar graph: a mosaic from Parco Güell in Barcelona, realized with a technique called Trencadis. What are its reciprocal characteristics?

Bibliography

- [1] P.K. Andersen, O. Borgan, R.Gill, and N. Keiding. *Statistical models based on counting processes*. Springer Science and Business Media, 2012.
- [2] I. Benjamini and S. Lee. Conditioned diffusions which are Brownian bridges. *Journal of Theoretical Probability*, 10(3):733–736, 1997.
- [3] S. Bernstein. Sur les liaisons entre les grandeurs aléatoires. In *Vehr. des intern. Mathematikerkongress Zürich*, volume I. 1932.
- [4] A. Beurling. An automorphism of product measures. *Annals of Mathematics*, 72:189–200, 1960.
- [5] S. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *Journal of functional analysis*, 156(2):347–365, 1998.
- [6] P. Brémaud. The martingale theory of point processes over the real half line admitting an intensity. In A. Bensoussan and J.L. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, volume 107 of *Lecture Notes in Economics and Mathematical Systems*, pages 519–542. Springer, 1975.
- [7] P. Brémaud. *Point Processes and Queues: Martingale Dynamics: Martingale Dynamics*. Springer, 1981.
- [8] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media, 2013.
- [9] P. Caputo and G. Posta. Entropy dissipation estimates in a zero-range dynamics. *Probability theory and related fields*, 139(1-2):65–87, 2007.
- [10] E.A. Carlen and E. Pardoux. Differential calculus and integration by parts on Poisson space. In S. Albeverio, Ph. Blanchard, and D. Testard, editors, *Stochastics, Algebra and Analysis in Classical and Quantum*

- Dynamics*, volume 59 of *Mathematics and Its Applications*, pages 63–73. Springer, 1990.
- [11] J.P. Carmichael, J.C. Masse, and R. Theodorescu. Processus gaussiens stationnaires réciproques sur un intervalle. *C. R. Acad. Sci., Paris, Sér. I*, 295:291–293, 1982.
- [12] S.C. Chay. On quasi-Markov random fields. *Journal of Multivariate Analysis*, 2(1):14–76, 1972.
- [13] L.H.Y. Chen. Poisson approximation for dependent trials. *The Annals of Probability*, 3(3):534–545, 1975.
- [14] Y. Chen, T. Georgiou, and M. Pavon. On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *preprint arXiv:1412.4430*, 2014.
- [15] Y. Chen, T. Georgiou, and M. Pavon. Optimal steering of a linear stochastic system to a final probability distribution. *preprint arXiv:1408.2222*, 2014.
- [16] Y. Chen, T. Georgiou, and PavonM. Optimal mass transport over bridges. *arXiv preprint arXiv:1503.00215*, 2015.
- [17] J.M.C. Clark. A local characterization of reciprocal diffusions. *Applied Stochastic Analysis*, 5:45–59, 1991.
- [18] G. Conforti. On some generalizations of the Schrödinger Problem, in preparation.
- [19] G. Conforti. Reciprocal characteristics and concentration of measure, in preparation.
- [20] G. Conforti, P. Dai Pra, and S. Roelly. Reciprocal classes of jump processes. *Preprint arxiv:1406.7495*, 2014.
- [21] G. Conforti and C. Léonard. Reciprocal classes of a random walk on graphs. *Preprint arxiv:1505.01323*.
- [22] G. Conforti, C. Léonard, R. Murr, and S. Roelly. Bridges of Markov counting processes. Reciprocal classes and duality formulas. *Electron. Commun. Probab.*, 20:1–12, 2015.
- [23] G. Conforti and S. Roelly. Reciprocal class of random walks on an Abelian group. *Preprint*, <http://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/7019>.

- [24] A.B. Cruzeiro and J.C. Zambrini. Malliavin calculus and Euclidean quantum mechanics. I. Functional calculus. *Journal of Functional Analysis*, 96(1):62–95, 1991.
- [25] P. Dai Pra. A stochastic control approach to reciprocal diffusion processes. *Applied Mathematics and Optimization*, 23(1):313–329, 1991.
- [26] P. Dai Pra, A.M. Paganoni, and G. Posta. Entropy inequalities for unbounded spin systems. *Annals of probability*, 30:1959–1976, 2002.
- [27] P. Dai Pra and M. Pavon. On the Markov processes of Schrödinger, the Feynman-Kac formula and stochastic control. In *Realization and Modelling in System Theory*, volume 3, pages 497–504. Springer, 1990.
- [28] J.A. De Loera, R. Hemmecke, and M. Köppe. *Algebraic and Geometric Ideas in the Theory of Discrete Optimization*. MOS-SIAM Series on Optimization. SIAM, 2013.
- [29] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009.
- [30] J.L. Doob. Conditional Brownian motion and the boundary limits of harmonic functions. *Bulletin de la Société Mathématique de France*, 85:431–458, 1957.
- [31] R.J. Elliott and A.H. Tsoi. Integration by parts for Poisson processes. *Journal of Multivariate Analysis*, 44(2):179 – 190, 1993.
- [32] P.J. Fitzsimmons. Markov processes with identical bridges. *Electron. J. Probab*, 3, 1998.
- [33] T.R. Fleming and D.P. Harrington. *Counting Processes and Survival Analysis*. Wiley Series in Probability and Statistics. Wiley, 2011.
- [34] H. Föllmer. Random fields and diffusion processes. In *École d’Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pages 101–203. Springer, 1988.
- [35] R. Fortet. Résolution d’un système d’équations de m. schrodinger. *J. Math. Pures Appl.* 9, 83:1040, 1940.
- [36] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(3):235–253, 1975.

- [37] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*. Grundlehren der mathematischen Wissenschaften. Springer, 2003.
- [38] B. Jamison. Reciprocal processes: The stationary Gaussian case. *The Annals of Mathematical Statistics*, 41(5):1624–1630, 1970.
- [39] B. Jamison. Reciprocal processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 30(1):65–86, 1974.
- [40] B. Jamison. The Markov processes of Schrödinger. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 32(4):323–331, 1975.
- [41] A. J. Krener. Reciprocal diffusions and stochastic differential equations of second order. *Stochastics*, 24(4):393–422, 1988.
- [42] A.J. Krener. Reciprocal diffusions in flat space. *Probability Theory and Related Fields*, 107(2):243–281, 1997.
- [43] M. Ledoux. *The Concentration of Measure Phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2001.
- [44] C. Léonard. Stochastic derivatives and generalized h-transforms of markov processes. *Preprint arxiv:1102.3172v1*.
- [45] C. Léonard. Convex conjugates of integral functionals. *Acta Mathematica Hungarica*, 93(4):253–280, 2001.
- [46] C. Léonard. From the Schrödinger problem to the Monge–Kantorovich problem. *Journal of Functional Analysis*, 262(4):1879–1920, 2012.
- [47] C. Léonard. Girsanov theory under a finite entropy condition. In C. Donati-Martin, A. Lejay, and A. Rouault, editors, *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Mathematics*, pages 429–465. Springer, 2012.
- [48] C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014.
- [49] C. Léonard, S. Roelly, and J.C. Zambrini. Reciprocal processes. A measure-theoretical point of view. *Probability Surveys*, 11:237–269, 2014.

- [50] B.C. Levy. Characterization of multivariate stationary Gaussian reciprocal diffusions. *Journal of Multivariate Analysis*, 62(1):74 – 99, 1997.
- [51] B.C. Levy and A.J. Krener. Dynamics and kinematics of reciprocal diffusions. *Journal of Mathematical Physics*, 34(5):1846–1875, 1993.
- [52] B.C. Levy and A.J. Krener. Stochastic mechanics of reciprocal diffusions. *Journal of Mathematical Physics*, 37(2):769–802, 1996.
- [53] T.M. Liggett. *Interacting Particle Systems*. Classics in Mathematics. Springer, 2006.
- [54] J. Mecke. Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 9(1):36–58, 1967.
- [55] T. Mikami. Monge’s problem with a quadratic cost by the zero-noise limit of h-path processes. *Probability Theory and Related Fields*, 129(2):245–260, 2004.
- [56] R. Murr. *Reciprocal classes of Markov processes. An approach with duality formulae*. PhD thesis, Universität Potsdam, 2012. Available at opus.kobv.de/ubp/volltexte/2012/6301/pdf/premath26.pdf.
- [57] R. Murr. Characterization of infinite divisibility by duality formulas. Application to Lévy processes and random measures. *Stochastic Processes and their Applications*, 123(5):1729 – 1749, 2013.
- [58] E. Nelson. *Dynamical theories of Brownian motion*, volume 2. Princeton university press Princeton, 1967.
- [59] J. Neukirch. *Algebraic Number Theory*. Grundlehren der mathematischen Wissenschaften : a series of comprehensive studies in mathematics. Springer, 1999.
- [60] J.R. Norris. *Markov Chains*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, 1999.
- [61] D. Nualart. *The Malliavin calculus and related topics*. Probability and its applications. Springer, 2006.
- [62] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *Journal of Functional Analysis*, 256(3):810–864, 2009.

- [63] N. Privault. *Stochastic Analysis in Discrete and Continuous Settings: With Normal Martingales*. Lecture Notes in Mathematics. Springer, 2009.
- [64] N. Privault and J.C. Zambrini. Markovian bridges and reversible diffusion processes with jumps. *Annales de l'Institut Henri Poincaré (B), Probabilités et Statistiques*, 40(5):599–633, 2004.
- [65] E.R. Rodrigues and J.A. Achcar. *Applications of Discrete-time Markov Chains and Poisson Processes to Air Pollution Modeling and Studies*. Springer Briefs in Mathematics. Springer, 2012.
- [66] S. Røelly. Reciprocal processes. A stochastic analysis approach. In V. Korolyuk, N. Limnios, Y. Mishura, L. Sakhno, and G. Shevchenko, editors, *Modern Stochastics and Applications*, volume 90 of *Optimization and Its Applications*, pages 53–67. Springer, 2014.
- [67] S. Røelly and M. Thieullen. A characterization of reciprocal processes via an integration by parts formula on the path space. *Probability Theory and Related Fields*, 123(1):97–120, 2002.
- [68] S. Røelly and M. Thieullen. Duality formula for the bridges of a brownian diffusion: Application to gradient drifts. *Stochastic Processes and their Applications*, 115(10):1677–1700, 2005.
- [69] S. Røelly and H. Zessin. Une caractérisation des diffusions par le calcul des variations stochastiques. *CR Acad. Sci. Paris Sér. I Math*, 313(5):309–312, 1991.
- [70] N. Ross. Fundamentals of Stein's method. *Probab. Surv*, 8:210–293, 2011.
- [71] L. Rüschemdorf and W. Thomsen. Note on the Schrödinger equation and I-projections. *Statistics & probability letters*, 17(5):369–375, 1993.
- [72] K.I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.
- [73] E. Schrödinger. Über die Umkehrung der Naturgesetze. *Sitzungsberichte Preuss. Akad. Wiss. Berlin. Phys. Math.*, 144:144–153, 1931.
- [74] E. Schrödinger. La théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Ann. Inst Henri Poincaré*, (2):269 – 310, 1932.

- [75] I. M. Slivnjak. Some properties of stationary streams of homogeneous random events. *Teor. Verojatnost. i Primenen.*, 7:347–352, 1962. In Russian.
- [76] C. Stein. *Approximate Computation of Expectations*. IMS Lecture Notes. Institute of Mathematical Statistics, 1986.
- [77] M. Thieullen. Second order stochastic differential equations and non-Gaussian reciprocal diffusions. *Probability Theory and Related Fields*, 97(1-2):231–257, 1993.
- [78] M. Thieullen and J. C. Zambrini. Symmetries in the stochastic calculus of variations. *Probability Theory and Related Fields*, 107(3):401–427, 1997.
- [79] A. Wakolbinger. A simplified variational characterization of Schrödinger processes. *Journal of Mathematical Physics*, 30(12):2943–2946, 1989.
- [80] S. Watanabe. On discontinuous additive functionals and Lévy measures of a Markov process. *Japanese journal of Mathematics*, 1964.
- [81] J.C. Zambrini. Variational processes and stochastic versions of mechanics. *Journal of Mathematical Physics*, 27(9):2307–2330, 1986.