université
PARIS-SACLAY

From rarity to typicality: the improbable journey of a large deviation

De la rareté à la typicité : le parcours improbable d'une grande déviation

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 564 Physique en Île de France (EDPIF)
Spécialité de doctorat: Physique
Unité de recherche: Université Paris-Saclay, CNRS, IJCLab, 91405, Orsay, France.
Référent: Faculté des sciences d'Orsay

Thèse présentée et soutenue à Paris-Saclay, le 26 novembre 2021, par

Lydia CHABANE

Composition du jury:

Cécile MONTHUS Directrice de Recherche, CEA Paris-Saclay	Présidente du jury
Eric BERTIN Directeur de Recherche, Université Grenoble Alpes	Rapporteur et examinateur
Hugo TOUCHETTE Professeur, Stellenbosch University	Rapporteur et examinateur
Francesco TURCI Associé de recherche postdoctoral sénior, University of Bristol	Examinateur

Direction de la thèse:

Gatien VERLEY Maître de Conférence, Université Paris-Saclay	Directeur
Raphaël CHETRITE Chargé de Recherche, Université de Nice Sophia-Antipolis	Codirecteur

Thèse de doctorat
NNT: 2021UPASP118

Abstract

The problem of conditioning time-homogeneous Markov processes on a rare fluctuation has been studied within the framework of large deviation theory. On this basis, a new process equivalent to the conditioned process has been introduced using the generalized Doob transform: it is the "driven process". In this thesis, we aim to generalize these results to a larger class of Markov processes. In the first part of this manuscript, we consider periodically driven Markov processes, characterized by their time-periodic generators. We are interested in conditioning these processes on observables defined through time-periodic functions. Adapting the results of the time-homogeneous case, we derive the driven process for which the typical values of our observables after a large number of periods correspond to the values used for the conditioning. In the periodic case, time-independent generators become time-periodic, matrix exponentials become time-ordered exponentials and spectral problems become first order differential equations. The driven process can be derived either using path ensemble equivalence, or from an optimization problem on large deviation functions. In the second part of this manuscript, we extend these results to the general case of nonlinear Markov processes described by time-independent Lagrangians and Hamiltonians. In this new formalism, the generalized Doob transform leading to the driven process translates into a canonical transformation on Hamiltonians. This transformation — that we call "rectification" — requires to investigate the nonlinear counterpart of the Perron-Frobenius theorem. This investigation led us to conjecture a classification of the solutions of a Hamilton-Jacobi equation. We conclude this part by an opening on the problem of conditioning periodically driven nonlinear processes.

Résumé

Le problème du conditionnement de processus de Markov homogènes en temps sur une fluctuation rare a été étudié dans le cadre de la théorie des grandes déviations. Sur cette base, un nouveau processus équivalent au processus conditionné a été introduit en utilisant la transformée de Doob généralisée : il s'agit du « processus drivé ». Dans cette thèse, on ambitionne de généraliser ces résultats à une classe plus large de processus de Markov. Dans la première partie de ce manuscrit, on considère des processus de Markov conduits périodiquement, caractérisés par des générateurs périodiques. On veut conditionner ces processus sur des observables définies via des fonctions périodiques en temps. En adaptant les résultats du cas homogène en temps, on construit le processus drivé pour lequel les valeurs typiques de nos observables après un grand nombre de périodes correspondent aux valeurs utilisées pour le conditionnement. Dans le cas périodique, les générateurs indépendants du temps deviennent périodiques, les exponentielles de matrices deviennent des exponentielles ordonnées en temps et les problèmes spectraux deviennent des équations différentielles du premier ordre. Le processus drivé s'obtient soit en utilisant l'équivalence de probabilités de chemin, soit à partir d'un problème d'optimisation de fonctions de grandes déviations. Dans la deuxième partie de ce manuscrit, nous étendons ces résultats au cas général des processus de Markov non linéaires décrits par des lagrangiens et des hamiltoniens indépendants du temps. Dans ce nouveau formalisme, la transformée de Doob généralisée menant vers le processus drivé se traduit par une transformation canonique sur les hamiltoniens. Cette transformation — que l'on appellera « rectification » — nécessite d'étudier l'analogue non linéaire du théorème de Perron-Frobenius. Cette étude nous a conduits à conjecturer une classification des solutions d'une équation de Hamilton-Jacobi. Nous concluons cette partie par une ouverture sur le problème du conditionnement des processus non linéaires conduits périodiquement.

Remerciements

"C'est dans les rêves que la magie existe, mais c'est dans la vie que les rêves se réalisent."

Lydia Chabane

Tout d'abord, merci à toi Gatien d'avoir été un super directeur de thèse. Merci pour ta présence, ton humour et tes petits cours improvisés. Merci d'avoir répondu 10 fois à la même question parce que je n'ai pas noté la réponse la première fois. Merci pour ta patience inépuisable, je sais que je peux me montrer très (trop) têtue parfois. Merci pour ton écoute, tes conseils, ta confiance et la liberté que tu m'as laissée. Merci d'avoir toujours été disponible pour moi, même quand tu avais un cours de mécanique analytique à préparer pour le lendemain ;) J'ai été heureuse de t'avoir eu comme prof, puis maître de stage (deux fois), puis directeur de thèse : je crois qu'il est temps de laisser ma place à quelqu'un d'autre !

Merci Raphaël pour ta gentillesse et ton soutien. Merci d'avoir toujours fait passer mon intérêt scientifique et personnel avant tout le reste. Chaque passage à Nice a été une bouffée d'oxygène, et c'est grâce à toi. Je suis sincèrement heureuse d'avoir fait ta connaissance, aussi bien scientifiquement qu'humainement, et je te souhaite plein de bonnes choses pour la suite, tu le mérites !

Merci à tous les membres du jury pour l'intérêt scientifique apporté à ma thèse et ma soutenance. Merci Cécile, Eric, Hugo et Francesco pour votre bienveillance, vos commentaires positifs et vos encouragements. Eric et Hugo, merci beaucoup pour votre lecture méticuleuse du manuscrit et vos remarques constructives. Cécile, merci d'avoir été en plus ma tutrice scientifique.

Un grand merci à Alexandre pour ta patience et ta gentillesse. Merci pour ton aide scientifique, ton temps et toutes les réponses apportées à mes questions. Merci de t'être adapté à moi et de m'avoir laissé une place entière dans notre collaboration. J'ai adoré travailler avec toi.

Je tiens à remercier toute l'équipe de physique statistique. Nous sommes une petite équipe mais je suis heureuse et fière d'y avoir fait partie. Un grand merci à toi Henk pour ta compagnie et tes discussions intéressantes à l'heure du déjeuner. Merci pour la passion que tu mets dans ton travail, tu es une inspiration pour tout jeune chercheur. Merci énormément Cécile pour ta gentillesse et ton aide précieuse jusqu'au jour de la soutenance. Tu as grandement contribué à son bon déroulé. Merci également pour tout ce que tu fais au niveau scientifique, que ce soit ton implication dans la vie du labo ou l'effort que tu fais pour rendre les sciences accessibles (clin d'œil à Ecoclim et Pint of Science). Merci à vous

deux, vous allez me manquer ! Merci Hadrien pour ton aide au début de ma thèse, tu m'as passé le flambeau (ou plutôt ta chaise rouge) et aujourd'hui je le passe à Paul. Merci Paul pour les essais zoom pré-soutenance, bon courage pour la suite :)

Merci à tous mes amis du LPT. Merci Elie pour ton soutien réciproque pendant nos galères administratives. Merci pour toutes les fois où tu m'as écoutée me plaindre après une énième irruption dans ton bureau. J'admire ton calme et ta patience face à toutes ces difficultés, et bizarrement ça me rassurait un peu de te voir rassuré malgré la situation. Amaury, merci pour ton écoute et toutes nos discussions (pas que déprimantes je t'assure :p). J'ai toujours apprécié nos échanges, et j'ai toujours appris quelque chose avec toi. Continue d'être aussi impliqué dans les causes que tu défends, c'est inspirant. Merci Florian d'avoir toujours demandé après moi depuis notre rencontre en L3, et merci d'avoir facilité mon intégration dans le labo ! Merci Martín pour ta bonne humeur, tes intrusions impromptues dans mon bureau et pour toutes les fois où tu m'as empêchée de travailler. Tes cœurs sur mon tableau me manqueront énormément. Enfin, merci à Giulia, Florentin, Tim, Simone, Nicolas et aux autres doctorants et stagiaires de passage pour les rires partagés, les discussions légères et les pauses café sympathiques.

Merci à tous les membres du laboratoire qui ont croisé ma route. Marie, tu es l'âme du LPT. Il n'y a pas assez de mots pour te remercier pour ta gentillesse et pour tout ce que tu fais pour nous, merci pour tout. Asmaa, merci d'avoir été ma marraine et merci pour ton écoute et ta gentillesse à chaque comité de suivi. Merci Olivier pour toute l'aide informatique que tu m'as apportée, on sait tous les deux que ce n'est pas mon point fort ! Merci à Sébastien et Samuel d'avoir toujours encouragé les déplacements et échanges scientifiques pour nous, doctorants. Merci Philippe et Jocelyne pour votre réactivité et votre aide administrative.

Merci à toute l'équipe MISS de m'avoir permis de vivre cette expérience incroyable. J'ai toujours dit que la MISS était le rayon de soleil de ma thèse et j'en garderai un souvenir merveilleux. Merci à Valérie, Elisabeth et Sylvie d'avoir créé cette bulle de bonheur pour les enfants et pour nous, les doctorants ! Merci à tous les animateurs MISS avec qui j'ai partagé de super moments, que ce soit sur place ou lors de la conception des vidéos. Merci aux équipes Musique, Nano, Insecte et Histoire (avec une petite préférence pour ce dernier atelier, je dois l'avouer). Julien, Marie et Mendes, merci énormément pour votre vidéo, c'était adorable et vous allez me manquer !

Un grand merci à toute l'équipe Quercy pour ce magnifique séjour scientifique (et pas que !). Merci pour nos discussions scientifiques, pour nos rires et nos merveilleux repas qui n'en finissaient plus. J'étais super contente de passer ce moment avec chacun d'entre vous, on était une équipe d'enfer ! Merci Andrés d'avoir permis tout cela. Merci pour ta gentillesse, ta gourmandise et ta générosité. Tu es une personne en or et tu es l'une des plus belles rencontres que j'ai faites cette année. À bientôt pour Quercy 2026, et comme dirait quelqu'un de très sage : « c'est la vie qu'on mérite ».

Je tenais à remercier tous les professeurs qui m'ont encouragée et soutenue tout au long de mes études. La liste est trop longue pour tous les citer, mais chacun a joué un rôle dans ma construction scientifique et personnelle. Je pense à Madame Toumi qui a

été la première à avoir cru en moi. J'avais 8 ans, mais je ne l'oublie pas. Un grand merci à mon professeur de mathématiques de lycée, M. Benabidallah. Si vous passez par là, sachez que vous avez grandement contribué à entretenir ma passion pour les maths et à rendre mes deux dernières années de lycée génialissimes. Mes pensées vont ensuite à mes professeurs de prépa. Un énorme merci à M. Arnoult, M. Berger, M. Hervé et M. Dumont. Merci pour votre patience, votre soutien inconditionnel, vos encouragements pendant les moments difficiles. Merci de transmettre vos connaissances avec passion et de penser au bien-être des élèves (surtout en prépa). Je suis très reconnaissante de vous avoir eu comme professeurs et je ne garde de cette période que les bons moments ! Je tiens à remercier tous mes professeurs du magistère pour leur effort de pédagogie et l'ambiance bienveillante des cours. Je pense en particulier à Robin Zegers, Bartjan van Tent, François Naulin, Giuseppe Foffi, Mathieu Langer, Frédéric Moisy, Christophe Texier et Nikki Franses-Jones. Je tiens à faire une mention spéciale à toi Claire. Merci pour ton soutien, ton écoute, ta gentillesse et tout ce que tu fais pour les étudiants du magistère. Merci également à toi Frédéric pour les super moments pendant l'option vulgarisation et Pint of Science. Merci à vous deux et à toute l'équipe FPT pour cette aventure extraordinaire (et non sans émotion), j'adore repenser à ces moments ! Enfin, je te remercie toi, Monsieur Garrido, pour ta patience, ta pédagogie et tes blagues de mauvais goût (en présentiel et par procuration). Merci d'avoir rendu les TPs supportables — j'étais tellement heureuse quand j'ai su que c'était toi qui allais nous encadrer. Mon dernier cours au magistère était avec toi, merci d'avoir clos cette période sur une belle note :)

Merci à tous mes amis qui ont assisté à ma soutenance, et à ceux qui ont vécu cette aventure à mes côtés. Merci ma camarade de galère de M2 Sarah, ce fût un plaisir de souffrir avec toi et de dévaliser les banquets de l'ENS (et à peu près n'importe quel autre banquet avec des macarons). Je pense à vous Noë, Eugénie et Valentin. Merci pour nos soirées jeux, nos soirées films et nos soirées Mama Kitchen. Merci de me réexpliquer les règles des jeux pour la 3ème fois parce que je suis trop lente à la détente. Merci pour tous ces petits moments qui ont rendu la thèse plus légère. Noë, tu mérites un passage pour toi tout seul. Tu te souviens, il y a 3 ans et 3 mois dans la crêperie, quand je te disais : « ça y est, c'est la fin, la thèse est bientôt finie » ? Eh bien, voilà, j'avais raison ! (En même temps, j'ai toujours raison, et toi tu as toujours tort, voilà, c'est dit). Je te remercie pour notre aventure Pint of Science, nos débats d'anglais et nos goûters tiramisu. Je ne te remercie pas pour ta mauvaise foi pendant nos débats stupides et inutiles ! Plus sérieusement, merci pour tous ces souvenirs, que ce soit avec le canard du lac d'Orsay, ou devant le Grinch ou Happy birthdead (oui, oui, j'affiche notre mauvais goût cinématographique devant tout le monde). Bref, merci pour notre amitié, tout simplement.

Merci à toute ma famille et mes proches. Merci à mes parents qui m'ont permis d'être ici aujourd'hui et de vivre une vie meilleure, je vous en serai toujours reconnaissante. Merci à ma maman sans qui je ne serai pas arrivée où j'en suis. Merci de m'avoir toujours soutenue et encouragée, peu importe la direction que je voulais prendre. Tu es mon rayon de soleil et je te dois toute ma vie. Merci à Sousou, ma deuxième maman, pour tout ce que tu fais pour nous, et pour tout ce que tu as fait pour moi depuis toute petite. C'est à vous deux que je dédie ce travail, je vous aime fort. Merci Tonton pour ton aide sur tous les aspects de la vie. Tu m'as plusieurs fois sorti du pétrin, même à distance ! Merci à mes petits frères et sœur Anis, Aline et Willem d'exister tout simplement. Ma vie ne serait pas

la même sans vous. Aline, bon courage pour la suite, j'attends tes remerciements dans 10 ans ;) Merci Radou et Tata Fatiha pour votre aide le jour de la soutenance. Enfin, merci à mes oncles et tantes, cousins et cousines, et amis de la famille qui m'ont aidée de près ou de loin dans cette aventure !

Pour finir, merci à toi Anis. Pour tout. Sans ton soutien, je n'y serais jamais arrivée. Merci pour toutes ces belles années passées à tes côtés, pour ta présence dans les moments difficiles et pour ta patience sans limites. Merci d'avoir toujours fait passer mon bien-être avant le tien. Merci pour notre complicité, nos rires et nos soirées Koh-Lanta & burgers. Tu as rendu ces trois dernières années plus douces, tu es mon pilier...

A ma maman…

…à Sousou

♡

Contents

Introduction

"Si tu peux pas faire de grandes choses, fais de petites choses avec grandeur."

Youssoupha

Nonequilibrium statistical physics is at the core of current researches. As almost all systems found in nature are not at thermodynamic equilibrium, this field finds applications in all branches of physics, but also in other domains such as biology, chemistry and environmental science [1, 2]. It aims to generalize the theory of equilibrium statistical mechanics, whose primary goal was to formalize the results of classical thermodynamics. The latter studies the thermal behavior of bodies and the transformation of matter. It is an axiomatic theory based on empirical laws that deals with exchanges of energy, heat and work in macroscopic systems. However, it does not provide a microscopic interpretation of what is happening. For that reason, statistical mechanics — also called statistical thermodynamics — emerged, providing a rigorous probabilistic formalism that explains the thermodynamic phenomena occurring within equilibrium macroscopic systems from the microscopic properties of their constituents.

Brief history of thermodynamics and statistical mechanics [3]

Classical thermodynamics was initially interested in understanding the functioning of steam engines and thermal concepts like heat and temperature. In 1783, Antoine Lavoisier introduced the caloric theory which states that heat is an invisible fluid called caloric that flows from hot bodies to cold bodies. The first to seriously criticize the caloric theory was Graf von Rumford (Benjamin Thompson) who supported the idea that heat is related to motion after observing heat production in boring cannons, in 1789.

The origins of thermodynamics as a modern science stem from Sadi Carnot's work about a generalized theory of heat engines *Reflections on the Motive Power of Fire* [4], published in 1824, in which a first version of the second law of thermodynamics is reported. In 1843, Julius Robert Mayer stated one of the first versions of the first law of thermodynamics. The full statement of the law was established by Rudolf Clausius in 1850. In 1865, Clausius introduced an extensive state function named *entropy* that formalizes the notion of irreversibility. However, the work of Clausius did not provide a molecular interpretation for the entropy. It was not until 1877 that the Austrian physicist Ludwig Boltzmann related the entropy to the number of microstates of a thermodynamic system. This result has significantly contributed to the development of statistical mechanics, first initiated by

the Scottish physicist James Clerk Maxwell who established the Maxwell distribution of molecular velocities in 1860.

In 1902, the American mathematical physicist J. Willard Gibbs published *Elementary Principles in Statistical Mechanics* [5] in which he formalized statistical mechanics by introducing the theory of ensembles. This work is considered as the foundation of modern statistical mechanics.

Evolution of statistical physics

Since then, statistical mechanics has known a great evolution, leading to the exploration of new subjects such as quantum fluids, nonlinear chemical physics, critical phenomena, transport theory, or biophysics [6]. One of the first extensions of statistical physics was to take into account the indistinguishability of quantum particles. This work was initiated by Satyendranath Bose in 1924 who treated black-body radiation as a gas of photons. Soon after, Albert Einstein completed Bose's results in two papers published in 1924 and 1925, leading to the so-called *Bose–Einstein statistics* [7]. In 1926, Enrico Fermi showed that the statistics obeyed by fermions is different from the Bose–Einstein statistics, giving rise to the so-called *Fermi-Dirac statistics* [7]. Thereafter, Lev Davidovitch Landau and John von Neumann introduced the density matrix [7], laying the foundations of *quantum statistical mechanics*.

Later on, significant advances led to the emergence of the subdomain of *nonequilibrium statistical physics*. The first investigations of this field were directed toward the near-equilibrium regime, giving rise to major results like Onsager's reciprocal relations which relate fluxes to their corresponding thermodynamic affinities [8, 9], or the fluctuation–dissipation theorem [10, 11] which relates the response of a system to a small perturbation and its fluctuations when it is at equilibrium. Subsequently, these investigations extended to the arbitrarily far from equilibrium regime by the appearance of important results among which the fluctuations theorems [12–16], the Jarzynski relation [17, 18], the Crooks fluctuation theorem [19, 20], the Hummer-Szabo relation [21] and the Hatano-Sasa relation [22]. This led to the emergence of *stochastic thermodynamics* [23–25].

Stochastic thermodynamics uses the theory of stochastic processes to study *small* systems in which fluctuations are no longer negligible with respect to their mean values. Examples of such systems are molecular motors, biochemical networks, colloids, microorganisms, etc. To take into account fluctuations, a new approach to thermodynamics has been developed. Heat and work have been defined at the level of a single trajectory [26, 27], engendering the framework of *stochastic energetics* [28]. A similar definition has been provided for the entropy production and the stochastic entropy of the system [29–31]. Beyond these theoretical developments, experimental achievements have been done, see Ref. [24, 32, 33] for a review. An example consists in measuring the work distribution for a colloidal particle pushed periodically by a laser toward a repulsive substrate [34]. The experimental results were in good agreement with theoretical predictions based on solving the Fokker–Planck equation. One of the motivations of stochastic thermodynamics is work extraction at micro and nanoscales in the same way traditional thermodynamics was motivated by the appearance of heat engines. Microscopic heat engines using trapped and time-periodically driven colloidal particles have been realized experimentally [35, 36], providing a better understanding of the physics of energy conversion at all scales.

In the same vein, nonequilibrium statistical physics aims to generalize the concepts

of equilibrium statistical mechanics to nonequilibrium systems. Lately, R. M. L. Evans proposed a nonequilibrium version of statistical ensembles for sheared fluids [37–39]. Such ensembles have been used for the study of glassy phases and dynamical phase transitions [40–51]. The equivalence between these ensembles has been studied in Ref. [52] using the theory of large deviations which describes the asymptotic behavior of likely and rare events [53–55]. This theory has been applied to equilibrium statistical mechanics, providing a general language for equilibrium notions such as *entropy*, *free energy*, *maximum entropy principle* or *minimum free energy principle* [56–70]. On this basis, large deviation theory is a promising theoretical foundation to nonequilibrium statistical physics [71–79]. Contrary to equilibrium, time now needs to be taken into account by looking at the dynamical fluctuations of time-integrated observables and their *typical* values in the long-time limit, or both in the long-time and large-size limit when considering many-body systems. In keeping with these concepts, R. Chétrite and H. Touchette formalized the *problem of conditioning a Markov process on a large deviation* [52,80,81] which aims to find an equivalent process in the long-time limit to a process for which observables are conditioned on rare fluctuations. This problem was initially considered by J. L Doob [82,83] who studied a Markov process conditioned on leaving the state-space at a certain position. After that, other conditionings were considered [84–98]. The conditioned process generalizes the equilibrium microcanonical ensemble in which energy is fixed. Instead of fixing the value of observables, one can rather fix the value of their conjugate variables, in the same way we fix temperature instead of energy in the equilibrium canonical ensemble: this defines the *biased process* [44,99,100]. In Ref. [80], this process has been shown to be related to the nonequilibrium version of the canonical ensemble and to reproduce the conditioned process in the long-time limit [52,101,102]. However, the biased process does not satisfy all the requirements of a proper process. Thereafter, R. Chétrite and H. Touchette proposed a procedure allowing one to build a proper Markov process from the biased process that is equivalent in the long-time limit to the conditioned process.

Contribution of this manuscript

The guiding principle of this manuscript is the problem of conditioning of Refs. [52,80,81]. This problem was dealt with in the case of time-homogeneous Markov processes. The objective of this thesis is to generalize this work to a larger class of Markov processes: first to periodically driven Markov processes, then to nonlinear Markov processes.

The former extension is motivated by the fact that many thermodynamic machines, including engines, operate via cycles or under periodic control. Such machines are experimentally studied nowadays at the fluctuating level [103–106]. Fluctuations in periodically driven systems modeled by Markov processes with time-periodic transition rates have also attracted interest at the theoretical level [107, 108].

In the latter extension, we mainly focus on time-homogeneous nonlinear Markov processes, the case of periodically driven nonlinear Markov processes being only approached as an opening. In addition to time, this class of processes involves a large size-type parameter such as volume or number of particles, hence allowing one to model systems with many degrees of freedom. According to statistical physics, we aim to provide a reduced description of those systems using only a small set of variables. This is done using large deviation theory with respect to the size-type parameter, leading to a *Lagrangian/Hamiltonian formalism* [109–123], formalized by A. Lazarescu et al. in Ref. [124] for chemical reaction

networks. The results of Ref. [80] are then translated within this formalism, generalizing the problem of conditioning to nonlinear Markov processes.

Outline

The outline of this manuscript is as follows:

- Chapters 1 and 2 lay the foundations of the mathematical formalism used in this manuscript. Chapter 1 is an introduction to the theory of stochastic processes and large deviations. Chapter 2 is a brief review of statistical mechanics and stochastic thermodynamics, and defines the mathematical models on which this thesis is based, namely Markov processes.

- Chapter 3 addresses the problem of conditioning periodically driven Markov processes on a large deviation. The first part of this chapter is about Markov jump processes and the second part is about Markov diffusion processes. This problem is approached from two perspectives: path ensemble equivalence and variational problem.

- Chapter 4 proposes a theory for the problem of conditioning nonlinear Markov processes. This theory focuses on time-homogeneous nonlinear processes, but we hope that it paves the way for a generalization to periodically driven nonlinear processes. The first part of this chapter applies the results of Ref. [80] to a class of nonlinear jump processes to derive their counterpart in the Lagrangian/Hamiltonian formalism of Ref. [124]. On this basis, we propose in the second part of this chapter a general theory for this problem that is model-free. We conclude this chapter by applying this theory to specific nonlinear models.

N.B.: A glossary and a table of notations are made available to the reader at the end of the manuscript.

Chapter 1

Mathematical introduction

"Shoot for the moon. Even if you miss, you'll land among the stars."

Norman Vincent Peale

This chapter introduces the mathematical framework used in this manuscript. For pedagogical reasons, we may avoid some technical details that would be required for perfect rigour.

1.1 Stochastic processes: preliminary definitions

In this section, we remind some definitions that will be useful throughout the manuscript. They come mainly from Refs. [125–127].

1.1.1 Random variables

A random variable X is a function defined on a set of outcomes of a random phenomenon or experiment and assigning a value to each outcome. For example, if we toss a coin, we can get a head or a tail: these are the outcomes of our random experiment. We choose to assign the value 0 for heads and the value 1 for tails. If we call $\Omega \equiv \{\text{Head, Tail}\}$ the set of outcomes and $\mathsf{V} \equiv \{0, 1\}$ the set of values, then the function $X : \Omega \longrightarrow \mathsf{V}$ is a random variable. When X takes its values in a countable (resp. continuous) set, we say that it is a discrete (resp. continuous) random variable.
Contrary to deterministic functions, the value of the random variable $X(t)$ at a given time t is unknown. Thereby, such objects are studied within probability theory and one focuses mainly on the probability of having a given value of the random variable. For instance, if we consider a non-biased coin, there is as much chance of getting tails as heads at each time t: $P(X(t) = 0) = P(X(t) = 1) = \frac{1}{2}$.

1.1.2 Stochastic processes

A stochastic process or random process is informally defined as a family of random variables $\{X_t \mid t \in \boldsymbol{E}\}$, where $\boldsymbol{E}$ is a subset of the set of real numbers. Usually, t is the time and the stochastic process contains the information on the time evolution of the random variable.

By abuse of notation, we usually refer to the stochastic process by X_t or $X(t)$. When $\boldsymbol{E}$ is finite or countable, X_t is a discrete-time stochastic process. If not, X_t is a continuous-time stochastic process. Note the difference between the nature of the stochastic process X_t and the nature of the set $\boldsymbol{E}$. For instance, the continuous-time stochastic process X_t can take discrete values. We denote by $P(x_n, t_n; x_{n-1}, t_{n-1}; \dots x_0, t_0)$ the joint probability that $X(t)$ takes the values $x_n, x_{n-1}, \dots, x_0$ at times $t_n, t_{n-1}, \dots, t_0$ respectively, with the implicit ordering $t_n > t_{n-1} > \dots t_0$. The knowledge of all possible joint probabilities describes completely the system. Using the chain rule, the joint probability can be written in terms of conditional probabilities:

$$\begin{aligned} P(x_n, t_n; x_{n-1}, t_{n-1}; \dots x_0, t_0) = P(x_n, t_n \mid x_{n-1}, t_{n-1}; \dots x_0, t_0) \\ P(x_{n-1}, t_{n-1} \mid x_{n-2}, t_{n-2}; \dots x_0, t_0) \dots P(x_1, t_1 \mid x_0, t_0) P(x_0, t_0) \end{aligned} \tag{1.1}$$

Note that P can be a probability or a probability density according to the nature of X.

1.1.3 Markov processes

Markov processes are stochastic processes with a Markov property, meaning that the future evolution of the stochastic process depends only on its present and not on its past. Mathematically, this means that the conditional probability of the process future state given its present and past values depends only on the present state:

$$P\left(x_{n+1}, t_{n+1} \mid x_n, t_n; x_{n-1}, t_{n-1}, \dots, x_0, t_0\right) = P\left(x_{n+1}, t_{n+1} \mid x_n, t_n\right). \tag{1.2}$$

Eq. (1.2) is known as the *Markov property*. In nature, true Markov processes rarely exist. Systems display a memory time during which memory effects are important [128]. However, if the memory time is much smaller than the other time scales of the system, it is reasonable to approximate the system by a Markov process [125]. When the Markov property (1.2) applies, any joint propability $P(x_n, t_n; x_{n-1}, t_{n-1}; \cdots x_0, t_0)$ expresses in terms of simple conditional probabilities $P(x_i, t_i \mid x_{i-1}, t_{i-1})$ — called *transition probabilities* — and the chain rule (1.1) simplifies to:

$$\begin{aligned} P(x_n, t_n; x_{n-1}, t_{n-1}; \cdots x_0, t_0) = P(x_n, t_n \mid x_{n-1}, t_{n-1}) P(x_{n-1}, t_{n-1} \mid x_{n-2}, t_{n-2}) \\ \dots P(x_1, t_1 \mid x_0, t_0) P(x_0, t_0). \end{aligned} \tag{1.3}$$

Markov processes are used to model many random systems [129, 130]. Some of the various applications of Markov processes are from meteorology [131–135], biology [136–139], chemistry [140–142], bioinformatics [143–145], music [146–150] and finance [151–153]. They also constitute the basis of the well-known Markov chain Monte Carlo algorithms [154] which counts many applications among which computational biology [155–158], physics [159–162] and linguistics [163]. Beyond these large range of applications, the Markov property brings a huge simplification in mathematical calculations as we will see in the following.

Chapman-Kolmogorov Equation

We have for any stochastic process [125]:

$$P(x_3,t_3 \mid x_1,t_1) = \int \mathrm{d}x_2 P(x_3,t_3;x_2,t_2 \mid x_1,t_1)$$
$$= \int \mathrm{d}x_2 P(x_3,t_3 \mid x_2,t_2;x_1,t_1) P(x_2,t_2 \mid x_1,t_1). \tag{1.4}$$

For Markov processes, Eq. (1.4) becomes

$$P(x_3,t_3 \mid x_1,t_1) = \int \mathrm{d}x_2 P(x_3,t_3 \mid x_2,t_2) P(x_2,t_2 \mid x_1,t_1). \tag{1.5}$$

Eq. (1.5) is known as the *Chapman-Kolmogorov equation*. Note that according to the nature of X, the sum $\int$ may be an integral or a discrete sum $\sum_x$. In the following, we will use the integral notation $\int$ for both discrete and continuous random variables. In his paper *Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung* (1931) [164], Andreï Nikolaïevitch Kolmogorov has studied the solutions of the Chapman-Kolmogorov equation (1.5). He showed that, under certain conditions, the transition probabilities evolve according to a differential equation involving two processes [125,165]: a *jump process* and a *diffusion process*. These two processes will be defined in Sections 2.3 and 2.4.

1.2 Generating functions

Let A_n be a random variable indexed by some parameter n and $P(A_n = a)$ the probability that A_n takes the value a. We recall that P may be a probability density if A_n takes its values in a continuous set.

1.2.1 Moment generating function

The *moment generating function*, often abbreviated as *generating function*, is an alternative way to study the statistical properties of a random variable. The generating function G of A_n is defined by

$$G(\gamma) \equiv \mathbb{E}[\mathrm{e}^{\gamma A_n}], \tag{1.6}$$

where the mean $\mathbb{E}$ is defined by

$$E[f(A_n)] \equiv \int \mathrm{d}a\, P(A_n = a) f(a), \tag{1.7}$$

for any function f. When they exist, G generates the moments $\mathbb{E}[(A_n)^\ell]$ of A_n ($\ell \in \mathbb{N}$) by differentiation since

$$G(\gamma) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} \gamma^\ell \, \mathbb{E}[(A_n)^\ell], \tag{1.8}$$

implying

$$\mathbb{E}[(A_n)^\ell] = G^{(\ell)}(0), \tag{1.9}$$

with $G^{(\ell)}$ the ℓ^{th} derivative of G. The generating function characterizes completely the probability $P(A_n = a)$ since both are related by a Fourier transform[1] and P can be

[1] The Fourier transform g of the function f is defined by $g(y) \equiv \int_{-\infty}^{+\infty} \mathrm{d}x \mathrm{e}^{ixy} f(x)$.

obtained from G by inverting the Fourier transform.

1.2.2 Cumulant generating function

Two important statistical quantities of the random variable A_n are its mean value $\mathbb{E}[A_n]$ and its variance $\mathbb{E}[(A_n)^2] - \mathbb{E}(A_n)^2]$ characterizing the fluctuations of A_n around its mean value. The mean and the variance are called respectively the first and second second cumulant of A_n. Other cumulants are obtained from the cumulant generating function (CGF) $\tilde{\Gamma}$ defined by the logarithm of the generating function:

$$\tilde{\Gamma}(\gamma) \equiv \ln G(\gamma). \tag{1.10}$$

As for the moments, the ℓ^{th} cumulant $\mathfrak{c}_l$ of A_n is obtained by deriving the CGF ℓ times:

$$\mathfrak{c}_l = \tilde{\Gamma}^{(\ell)}(0). \tag{1.11}$$

1.2.3 Scaled cumulant generating function

If we are now interested in the statistics of A_n for large n, it is appropriate to introduce the so-called scaled cumulant generating function (SCGF) defined by

$$\Gamma(\gamma) \equiv \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[\mathrm{e}^{n\gamma A_n}]. \tag{1.12}$$

The importance of the SCGF will appear in the next section when dealing with the theory of large deviations.

1.3 Large deviation theory

This section is at the core of the results that will be presented in this manuscript. The following content comes mainly from Refs. [53–55, 60, 70].

In broad terms, large deviation theory deals with the exponential decay of probabilities of rare events with a large parameter (e.g. time or system size). The first large deviation results are due to Cramer [166] and Chernoff [167], but the formalization of this theory was developed from 1975 by Donsker and Varadhan [168–171] and Freidlin and Wentzell [172].

Even if their probability to occur is small, it is interesting to look at rare fluctuations. On the one hand, the rare fluctuation of an observable for one process is the typical fluctuation for another process [80]. Hence, one motivation for studying these fluctuations [124] is to find a way to enhance them [108, 173] or suppress them [174]. On the other hand, rare fluctuations may have a huge impact on systems. One example concerns climate change in which rare events such as extreme droughts, heat waves, rain-fall, and storms have non-negligible consequences on eco and socioeconomic systems [174–177]. Amir Dembo and Ofer Zeitouni wrote in their book [53]: "*To understand why rare events are important at all, one only has to think of a lottery to be convinced that rare events (such as hitting the jackpot) can have an enormous impact.*"

1.3.1 Convergence of random variables

Since large deviation theory deals with the asymptotic behavior of random variables, we need to define properly the convergence of sequences of random variables A_n to a limit random variable A:

$$\lim_{n\to\infty} A_n = A. \tag{1.13}$$

There is not a unique way to define such a limit. In the following, we define the main types of convergence.

Convergence in law

A_n converges to A in law if for any continuous bounded function f,

$$\lim_{n\to\infty} \mathbb{E}\left[f(A_n)\right] = \mathbb{E}\left[f(A)\right]. \tag{1.14}$$

Convergence in probability

A_n converges to A in probability if for all $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(|A_n - A| > \epsilon\right) = 0. \tag{1.15}$$

Convergence in probability implies convergence in law.

Convergence in r^{th} order mean

A_n converges to A in r^{th} order mean ($r \geq 1$) if

$$\lim_{n\to\infty} \mathbb{E}\left[|A_n - A|^r\right] = 0. \tag{1.16}$$

We call *convergence in mean* the case $r = 1$ and *convergence in mean square* the case $r = 2$. Convergence in r^{th} order mean implies convergence in probability. Moreover, for $r > s \geq 1$, convergence in r^{th} order mean implies convergence in s^{th} order mean.

Almost sure convergence

A_n converges to A almost surely if

$$P\left(\lim_{n\to\infty} A_n = A\right) = 1 \tag{1.17}$$

except for a set of events of null probability. We say that A_n converges to A surely if (1.17) holds for all events. Almost sure convergence implies convergence in probability.

1.3.2 Large deviation principle

Large deviation theory is built on a mathematical principle describing the asymptotic behavior of probabilities that we call the *large deviation principle* (LDP). We say that A_n satisfies a LDP with rate $I(a)$ if the limit

$$\lim_{n\to\infty} -\frac{1}{n} \ln P(A_n = a) = I(a) \tag{1.18}$$

exists and we write

$$P(A_n = a) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nI(a)}, \tag{1.19}$$

where $A_n \underset{n\to\infty}{\asymp} B_n$ signifies $\lim_{n\to\infty} \frac{1}{n} \ln \frac{A_n}{B_n} = 0$. The function I is called the *rate function* or the *large deviation function* (LDF). Eq. (1.19) means that we neglect the sub-exponential terms in the asymptotic expansion of the probability as n becomes large:

$$\ln P(A_n = a) \underset{n\to\infty}{=} -nI(a) + o(n). \tag{1.20}$$

Thus, the probability decays exponentially for any value a except for the zeros $a_\star$ of the LDF: $I(a_\star) = 0$. The random variable A_n converges in probability to these zeros that we call *typical values* of A_n. Conversely, any value $a \neq a_\star$ is referred to as *a rare fluctuation* since the probability of finding such values tends to 0 as n becomes very large. Note that large deviation theory is based on the fact that the limit in Eq. (1.18) exists and that I is finite and non-zero everywhere. If the rate of decay of $P(A_n = a)$ is sub-exponential then the LDF is zero and one needs to consider the next orders in Eq. (1.20). On the contrary, if the rate of decay of $P(A_n = a)$ is super-exponential, the LDF is infinite. In these two cases, a LDP with a speed of decay that is different from n may exists. It is important to note that LDFs have at least one zero to ensures the normalization of the probability and are always positive. Indeed, negative values of I yield probabilities greater than 1.

1.3.3 Illustrative example

Let's come back to our example of coin toss. We flip the coin n times and we label X_i the outcome of the i^{th} toss. The random variable X_i takes the value 0 (head) or 1 (tail) with equal probability. We are interested in the asymptotic behavior of the frequency of tails during the n tosses:

$$A_n \equiv \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{1.21}$$

Since the tosses are independent, the probability that A_n takes the value a satisfies

$$P(A_n = a) = \frac{1}{2^n} \binom{n}{na} \equiv \frac{1}{2^n} \frac{n!}{(na)!(n-na)!}. \tag{1.22}$$

Using Stirling formula $\ln(n!) = n \ln n - n + o(n)$, we obtain

$$\ln\left(P(A_n = a)\right) = -n\left[\ln 2 + a \ln a + (1-a)\ln(1-a)\right] + o(n), \tag{1.23}$$

implying

$$P(A_n = a) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nI(a)}, \tag{1.24}$$

with

$$I(a) \equiv \ln 2 + a \ln a + (1-a)\ln(1-a). \tag{1.25}$$

Hence, A_n satisfies a LDP with rate function $I(a)$. Note that in the limit $n \to \infty$, A_n can be considered as a continuous random variable taking its values in $[0, 1]$ and the probability $P(A_n = a)$ becomes a probability density. As shown in Fig. 1.1, I has a unique zero $a_\star = \frac{1}{2}$ implying that A_n converges in probability toward $\frac{1}{2}$, which is expected since the coin is non-biased.

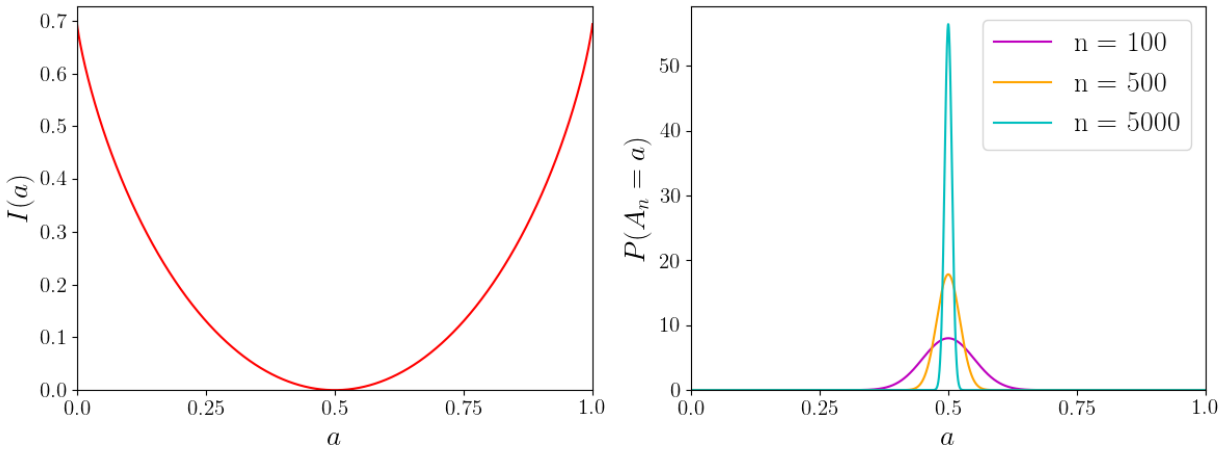

Figure 1.1 – (Left) LDF of the model of coin toss. Unsurprisingly, the LDF vanishes at $a = \frac{1}{2}$, corresponding to the case where we obtain as many heads as tails when the number of tosses goes to infinity. (Right) Probability density $P(A_n = a)$ for growing values of n. The probability density becomes narrower around the typical value $a_\star = \frac{1}{2}$ as n increases, and converges to a delta function centered in $a_\star$ when $n \to \infty$.

1.3.4 Gärtner-Ellis theorem

Large deviation theory aims to prove the existence of the limit in Eq. (1.18) and to compute LDFs. In physics, we usually focus on deriving the expression of the LDF since managing to do so implies the existence of the limit. In simple examples, one can do so by computing directly the probability distribution of the random variable and deriving the LDF from its asymptotic limit as done for the coin toss of Section 1.3.3. Yet, it is in general difficult, even impossible, to do such a direct calculation. This problem is remedied by the *Gärtner-Ellis theorem* [178, 179]. This theorem states that if the SCGF

$$\Gamma(\gamma) = \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[\mathrm{e}^{n\gamma A_n}] \tag{1.26}$$

exists and is differentiable everywhere, then A_n satisfies a LDP and the LDF $I(a)$ is related to the SCGF $\Gamma(\gamma)$ by a Legendre-Fenchel (LF) transform (cf. Appendix 1.A):

$$I(a) = \sup_\gamma \{\gamma a - \Gamma(\gamma)\}\,. \tag{1.27}$$

To give some intuition to Eq. (1.27), suppose that A_n satisfies a LDP (1.19). It follows from Eq. (1.7) and Eq. (1.19) that:

$$\mathbb{E}[\mathrm{e}^{n\gamma A_n}] \underset{n\to\infty}{\asymp} \int \mathrm{d}a\ \mathrm{e}^{n[\gamma a - I(a)]}. \tag{1.28}$$

For large n, we can approximate the integral by its dominant integrand[2]:

$$\mathbb{E}[\mathrm{e}^{n\gamma A_n}] \underset{n\to\infty}{\asymp} \mathrm{e}^{n \sup_\gamma \{\gamma a - I(a)\}}. \tag{1.29}$$

From the definition of the SCGF (1.26), if follows

$$\Gamma(\gamma) = \sup_\gamma \{\gamma a - I(a)\}\,. \tag{1.30}$$

[2]This is known as the Laplace's method (real integrals) or the saddle-point method (extension of the Laplace's method to integrals in the complex plane).

If now Γ is differentiable everywhere, then the LF transform is involutive (cf. Appendix 1.A) and we recover the result of Eq. (1.27). Note that this reasoning is not a proof of the Gärtner-Ellis but rather an heuristic argument. Even so, it shows that if a random variable satisfies a LDP, then the SCGF is the LF transform of the LDF. More generally, if a random variable satisfies a LDP then for any function f, the functional defined by

$$\Gamma(f) \equiv \mathbb{E}[\mathrm{e}^{f(A_n)}] \tag{1.31}$$

satisfies

$$\Gamma(f) = \sup_{\gamma} \{f(a) - \Gamma(\gamma)\}. \tag{1.32}$$

This is *Varadhan's Theorem* [180]. For further details, two derivations of the Gärtner-Ellis have been suggested in Ref. [70]. In equilibrium statistical mechanics, the entropy plays the role of a LDF and the free energy the role of a SCGF. Both are related by a LF transform [70].

The Gärtner-Ellis theorem is of practical interest since it provides a mathematical procedure to compute LDFs. Moreover, LDFs coming from this theorem are all **strictly** convex [70, 181]. In this case, I has a unique zero $a_\star$ and the typical value of A_n coincides with its mean value in the limit $n \to \infty$:

$$\mathbb{E}[A_n] = \int \mathrm{d}a\, P(A_n = a) a \underset{n\to\infty}{\to} \int \mathrm{d}a\, \delta(a - a_\star) a = a_\star, \tag{1.33}$$

where we used the fact that $\lim_{n\to\infty} P(A_n = a) = \delta(a - a_\star)$ when I has a unique zero. Moreover, expanding I around $a_\star$ up to second order yields

$$P(A_n = a) \underset{n\to\infty}{\sim} \mathrm{e}^{\frac{1}{2} I''(a_\star)(a-a_\star)^2}, \tag{1.34}$$

where we used the fact that $I(a_\star) = I'(a_\star) = 0$. Values around $a_\star$, i.e. *small deviations* from the typical value, are normally-distributed. Hence, large deviation theory is a generalization of the Central Limit theorem[3]: not only does it contain information about normally-distributed small deviations, but it also provides information about *large* fluctuations of random variables.

1.3.5 Contraction principle

It sometimes happens that we know about the LDF of a random variable X_n but we are interested in the large deviations of the random variable $A_n = f(X_n)$, with f a continuous function. The *contraction principle* [171] states that the LDF $I(a)$ of A_n is obtained from the LDF $J(x)$ of X_n by

$$I(a) = \inf_{x|f(x)=a} J(x), \tag{1.35}$$

where the argument of the infimum must satisfy $f(x) = a$. Eq. (1.35) results from the Laplace's method. Indeed, expressing the probability $P(A_n = a)$ in terms of the probability $P(X_n = x)$ yields

$$P(A_n = a) = \int \mathrm{d}x \delta(a - f(x)) P(X_n = x) \underset{n\to\infty}{\asymp} \int \mathrm{d}x \delta(a - f(x)) \mathrm{e}^{-nJ(x)} \underset{n\to\infty}{\asymp} \mathrm{e}^{-n \inf_{x|f(x)=a} J(x)}, \tag{1.36}$$

[3] The central limit theorem establishes the convergence of the sum of random variables toward a normal distribution.

where we used the LDP on X_n in the second equality and Laplace's method in the last equality.

Appendices

"Quand le sage désigne la Lune, l'idiot regarde le doigt."

Proverbe chinois

1.A Convexity and Legendre-Fenchel transform

The content of this section comes mainly from Ref. [182]. For further reading on convex analysis, see Refs. [181, 183, 184].

The Legendre-Fenchel (LF) transform $f^*(y)$ of a function $f(x)$ is defined by

$$f^*(y) \equiv \sup_x \{yx - f(x)\} \tag{1.37}$$

Note that if x and y are vectors, the product xy is replaced by the scalar product $x \cdot y$ in Eq. (1.37). When f is differentiable and strictly convex[4], the LF transform reduces to the Legendre transform

$$f^*(y) = yx_\star - f(x_\star), \tag{1.38}$$

where $x_\star$ satisfies $y \equiv f'(x_\star)$.

Important results

- The LF transform f^* of any function f is always convex, even if f is not convex [183]. Hence, the LF transform is not necessarily involutive, i.e. the LF transform f^{**} of f^* is not necessarily equal to f.
- f^{**} is the convex envelope (or convex hull) of f, i.e. f^{**} is the largest convex function such that $f^{**} \leq f$.
- By the Fenchel–Moreau theorem, $f^{**} = f$ if and only if f is convex (f is continuous for simplicity).
- If f^* is differentiable at y, then $f^{**}(x) = f(x)$ at $x = f^{*\prime}(y)$.

[4] We remind that a function f is convex on an interval $\mathcal{I}$ if for any x_1, x_2 of $\mathcal{I}$, and for any $s \in [0,1]$, we have $f(sx_1 + (1-s)x_2) \leq sf(x_1) + (1-s)f(x_2)$. If the inequality is strict, we say that f is *strictly* convex. Geometrically, it means that the curve of the function $f(x)$ on the interval $[x_1, x_2]$ is under the segment $[A_1, A_2]$ formed by the point A_1 (resp. A_2) of coordinates $(x_1, f(x_1))$ (resp. $(x_2, f(x_2))$).

Chapter 2

Stochastic thermodynamics

"Vite, Patrick, sans réfléchir : si tu pouvais avoir quelque chose maintenant, qu'est-ce que ce serait?
– Um... plus de temps pour réfléchir."

Bob l'éponge

In this chapter, we introduce the general field of stochastic thermodynamics and the mathematical models used to describe nonequilibrium systems. Before proceeding with stochastic thermodynamics, I found it pedagogical to go through a brief review of some elements of equilibrium statistical mechanics that will be useful for the rest of this manuscript. The reader can refer to Refs. [185–193] for a more detailed description of equilibrium statistical mechanics and thermodynamics.

2.1 Equilibrium statistical mechanics

Statistical mechanics makes use of probability theory to deduce the behavior and physical properties of macroscopic systems, i.e. with a large number of particles, from the laws governing its microscopic constituents. For instance, the pressure of a gas arises from the collisions of the molecules, and the temperature is a measure of the mean kinetic energy of the particles. Examples of macroscopic systems are gases, liquids, solids, liquid crystals, plasma, biological matter, stellar matter, etc. In these systems, it is usually impossible to follow the evolution of each individual particle. Statistical mechanics aims to define the relevant macroscopic physical quantities that characterize the system (volume, temperature, pressure...) and to relate them to the microscopic quantities describing the particles. To do so, the many-body system is described within statistical–mechanical frameworks in line with the external conditions applied to the system. These frameworks are referred to by *equilibrium ensembles* and the simplest ones are the microcanonical and canonical ensembles.

2.1.1 Microcanonical ensemble

The microcanonical ensemble deals with isolated systems, i.e. systems that cannot exchange energy or particles with their environment. In this ensemble, the energy of the

system is fixed $E = constant$ and the accessible microstates, i.e. those of energy E, are equiprobable. All other microstates have a null probability.

2.1.2 Canonical ensemble

The canonical ensemble, introduced by Josiah Willard Gibbs, describes a system of fixed number of particles in equilibrium with a heat reservoir of fixed temperature β^{-1}. The system exchanges energy with the heat reservoir and the states of the system have different value of energy, contrary to the case of the microcanonical ensemble. The canonical probability of the microstate ℓ of energy E_ℓ is given by

$$P_l = \frac{\mathrm{e}^{-\beta E_\ell}}{Z}, \tag{2.1}$$

where $Z \equiv \sum_l \mathrm{e}^{-\beta E_\ell}$ is the partition function.

Remark: If we also allow particle exchange with the reservoir, the system is studied within the grand canonical ensemble.

2.1.3 Ensemble equivalence

In the canonical ensemble, the energy is allowed to fluctuate. However, in the limit of very large number of particles, known as the *thermodynamic limit*, the fluctuations of the energy around its mean value becomes negligible and the system can be considered to have one definite value of energy, which recovers the framework of the microcanonical ensemble. Thereby, in the thermodynamic limit, canonical and microcanonical ensembles are equivalent and working in either ensembles is just a matter of mathematical convenience (providing some concavity conditions on the microcanonical entropy [194, 195]).

2.1.4 Laws of thermodynamics

Equilibrium thermodynamics is ruled by a set of empirical principles, later formalized by equilibrium statistical mechanics. We remind here the statement of the first and second laws.

First law of thermodynamics: the first law is a principle of energy conservation. It states that the variation of the internal energy U of a closed thermodynamic system involves two types of energy transfer between the system and its surrounding: *work* (W) and *heat* (Q). Work is a transfer of energy due to a macroscopic force and which can be used to move an object [193, 196, 197]. For instance, the pressure of a piston on a gas will cause the displacement of the piston in one direction, and this is associated with a pressure-volume work [196]. Heat is a transfer of energy due to the random agitation of microscopic degrees of freedom because of a difference of temperature between the system and its environment. The first principle reads

$$\Delta U = W + Q. \tag{2.2}$$

The energy U is a state function, meaning that its variation depends only on the initial and final states while W and Q depend on the whole trajectory between the initial and final states.

Second law of thermodynamics: the second law introduces a new thermodynamic quantity called *entropy* which formalizes the notion of *irreversibility*, and states that for any transformation of a thermodynamic system, the variation of the global entropy (system + environment) cannot decrease:

$$\Sigma \geq 0, \tag{2.3}$$

defining this way the arrow of time of the transformation. Σ is called *entropy production* and for a reversible transformation, $\Sigma = 0$. If we consider a system in contact with an environment of temperature T, the entropy production reads

$$\Sigma = \Delta S + S_{\text{env}}, \tag{2.4}$$

where ΔS is the variation of entropy of the system and where the variation of entropy of the environment $S_{\text{env}} = -\frac{Q}{T}$ is related to the heat dissipation between the system and the environment. Notice that contrary to the entropy of the environment and the entropy production, the entropy of the system is a state function. If the system is isolated, $Q = 0$ and $\Delta S = \Sigma \geq 0$. Hence, the entropy of isolated systems can only increase. Developments in statistical mechanics have given a microscopic description of entropy and established that the thermodynamic entropy of a system coincides with the Shannon entropy defined by $S \equiv -\sum_\ell p_\ell \ln p_\ell$, where the sum runs over the states of the system and where p_ℓ is the probability of state ℓ. This entropy contains all the statistical information of the system and determines the equilibrium state of the system given its external conditions as the one maximizing the entropy.

2.2 Stochastic thermodynamics

Equilibrium statistical mechanics is very useful to study the equilibrium properties of large systems. Yet, it is not suitable for the description of the dynamics of nonequilibrium systems, or systems with few degrees of freedom for which the fluctuations of physical observables are no longer negligible. These fluctuations are studied within the framework of stochastic thermodynamics which relies on the theory of stochastic processes to define physical observables and study their statistics.

There are many ways to drive a system out of equilibrium. For example, a system can be out of equilibrium because it is subjected to an external force, or is in contact with several reservoirs leading to a gradient of temperature or chemical potential for instance, or because it is relaxing toward an equilibrium state. Stochastic thermodynamics provides tools to define for these systems thermodynamic observables such as heat, work and entropy production, generalizing the concepts of equilibrium statistical thermodynamics. This theory allows studying real systems such as colloidal particles, biopolymers (DNA, RNA, proteins...), enzymes, molecular motors, etc., and many experimental achievements have supported the theoretical developments [34–36, 198–207].

Theoretically, stochastic thermodynamics uses Markov processes to build models and define fluctuating observables. In this manuscript, we will only consider systems modeled by continuous-time Markov processes and we use the term *Markov processes* for *continuous-time Markov processes* by abuse of language. We will focus on the two bricks of a general Markov process: Markov jump processes and diffusion processes [125, 126]. For a general review on stochastic thermodynamics, see Refs. [23–25].

2.3 Markov jump processes

2.3.1 Markov generator and master equation

We consider a continuous-time Markov jump process $z(t)$ defined on a discrete set of states that we label x or y. The state of the system at time t is denoted by $z(t)$. This Markov process is characterized by the transition rates $k_{xy} \equiv k_{xy}(t)$ to jump from a state y to another state x at time t ($x \neq y$). These transition rates are nonnegative and are defined from the transition probability $P(x, t+\delta t \mid y, t)$ by

$$k_{xy}(t) \equiv \lim_{\delta t \to 0} \frac{1}{\delta t} P(x, t+\delta t \mid y, t). \tag{2.5}$$

For $x = y$, we define $-k_{yy} \equiv \sum_{x \neq y} k_{xy}$ as the escape rate from state y at time t. The matrix $\boldsymbol{k}$ of components k_{xy} is known as the *transition rate matrix* or the *generator* of the Markov process. We denote by $\pi_x \equiv \pi_x(t)$ the probability to be in state x at time t. The probability vector $\boldsymbol{\pi}$ of components π_x satisfies the so-called *master equation*

$$\frac{\partial \boldsymbol{\pi}}{\partial t} = \boldsymbol{k}\boldsymbol{\pi}, \tag{2.6}$$

with initial condition $\boldsymbol{\pi}(0)$. The master equation is just a particular case of the differential version of the Chapman-Kolmogorov equation[1] (1.5). The norm of the probability is conserved by the master equation since by construction $\sum_x k_{xy} = 0$, $\forall y$, implying that $\frac{\partial}{\partial t}\left(\sum_x \pi_x\right) = 0$: we say that $\boldsymbol{k}$ generates a norm-conserving Markov process. Thereby, the master equation and the normalized initial probability ensure the normalization of the probability at all times. The master equation is a continuity equation for the probability and can be expressed under the form:

$$\frac{\partial \pi_x}{\partial t} = \sum_y j^{\boldsymbol{\pi}}_{xy}, \tag{2.7}$$

where we introduced the probability current $j^{\boldsymbol{\pi}}_{xy} \equiv j^{\boldsymbol{\pi}}_{xy}(t)$ from state y to state x at time t:

$$j^{\boldsymbol{\pi}}_{xy}(t) \equiv k_{xy}(t)\pi_y(t) - k_{yx}(t)\pi_x(t). \tag{2.8}$$

Another way to characterize a Markov process is by means of path probabilities. A *path* or a *trajectory* is the succession of states visited by the system and the data of the times at which transitions occur, and in some cases the mechanism by which a new state is reached (for instance, which heat reservoir the system exchanges energy with when a transition occurs). In the following, we denote a path by $[z]$. We label $\{z_i\}_{i=0}^M$ the visited states and $\{t_i\}_{i=0}^M$ the times at which the system jumps so that

$$z(t) = z_i \qquad \text{for} \qquad t_i \leq t < t_{i+1}. \tag{2.9}$$

[1] To be more precise, the Chapman-Kolmogorov equation is about the transition probabilities $P(x, t \mid y, t')$ as seen in Sec. 1.1.3. The more intuitive form (2.6) is just a matter of specifying the initial condition: $\pi_x(t) = \sum_y P(x, t \mid y, 0)\, \pi_y(0)$.

Time $t_0 = 0$ is the initial time and t_M is the last jump time before the final time t. The path $[z]$ is characterized by its probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ obtained from the transition probabilities $P(z_{i+1}, t_{i+1} \mid z_i, t_i)$ and the Markov property by

$$\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] = \pi_{z_0}(0) \prod_{i=0}^{M-1} \left\{ P(z_{i+1}, t_{i+1} \mid z_i, t_i) \prod_{\tau \in [t_i, t_{i+1}[} P(z_i, \tau + \delta t \mid z_i, \tau) \right\}, \tag{2.10}$$

where δt is an infinitesimal time interval and where the first term in the brackets refers to the probability to jump and the second term refers to the probability to stay in the same state when there is no jump. Using the fact that

$$P(z_{i+1}, t_{i+1} \mid z_i, t_i) = k_{z_{i+1},z_i}(t_{i+1}) \delta t_{i+1}, \tag{2.11}$$

$$P(z_i, \tau + \delta t \mid z_i, \tau) = 1 - \sum_{x \neq z(\tau)} k_{x,z_i}(\tau) \delta \tau, \tag{2.12}$$

we finally obtain the path probability

$$\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] = \pi_{z_0}(0) \exp\left[\sum_{i=0}^{M-1} \ln\left(k_{z_{i+1},z_i}(t_{i+1})\right) - \int_0^t \sum_{x \neq z(\tau)} k_{x,z(\tau)}(\tau) \mathrm{d}\tau \right] \prod_{i=0}^{M-1} \delta t_{i+1}. \tag{2.13}$$

2.3.2 Observables and first law

As mentioned in the introduction, stochastic thermodynamics uses the framework of stochastic processes to define fluctuating physical observables. In this manuscript, we look at general two-component observables of the form

$$\boldsymbol{A}_t[z] \equiv \begin{pmatrix} \frac{1}{t} \sum_{i=0}^{M-1} g_{z_{i+1},z_i}(t_{i+1}) \\ \frac{1}{t} \int_0^t \mathrm{d}\tau h_{z(\tau)}(\tau), \end{pmatrix} \tag{2.14}$$

where $\boldsymbol{g}$ is a function of the jumps and $\boldsymbol{h}$ a function of the states. Note that since $\boldsymbol{A}_t$ depends on the random variable $[z]$, it is also a random variable. This general observable represents many physical quantities including the main observables in thermodynamics. For instance, consider a system in contact with a heat reservoir of inverse temperature β and a particle reservoir of chemical potential μ, and assume that its energy-levels $U_x(t)$ has been made time-dependent by an external driving. We denote by $N_x(t)$ the number of particles in state x at time t. Taking $g_{xy}(t) = \mu\left[N_x(t) - N_y(t)\right]$ or $g_{xy}(t) = \left[U_x(t) - \mu N_x(t)\right] - \left[U_y(t) - \mu N_y(t)\right]$, and taking $h_x(t) = \partial_t U_x(t)$, the first component of $t\boldsymbol{A}_t[z]$ represents the chemical work $W^{\mathrm{chem}}[z]$ or the heat $Q[z]$ along the path $[z]$, and the second component of $t\boldsymbol{A}_t[z]$ represents the work $W[z]$ along the path $[z]$ [208]:

$$W_{\mathrm{chem}}[z] \equiv \sum_{i=0}^{M-1} \mu \left[N_{z_{i+1}}(t_{i+1}) - N_{z_i}(t_{i+1})\right] \tag{2.15}$$

$$Q[z] \equiv \sum_{i=0}^{M-1} \left[U_{z_{i+1}}(t_{i+1}) - U_{z_i}(t_{i+1})\right] - W_{\mathrm{chem}}[z], \tag{2.16}$$

$$W[z] \equiv \int_0^t \mathrm{d}\tau \partial_t U_{z(\tau)}(\tau). \tag{2.17}$$

The first principle of thermodynamics ensuring energy conservation holds at the level of trajectories and the variation of energy between times 0 and t satisfies

$$\Delta U(t) = W[z] + Q[z] + W_{\text{chem}}[z], \tag{2.18}$$

meaning that the variation of the energy comes either from its variation with respect to time (work) or from the variation of the state of the system (heat and chemical work). Note that the variation of energy is a state function and does not depend on the trajectory contrary to work and heat. For other thermodynamic observables involved in the first law (information reservoir, non-conservative forces, multiple reservoirs, etc.), refer to Refs. [209, 210].

2.3.3 Fluctuation theorem and second law

Another interesting observable is the entropy production along the path $[z]$ [31]

$$\Sigma[z] \equiv \ln \frac{\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]}{\mathbb{P}_{\bar{\boldsymbol{k}},\boldsymbol{\pi}(t)}[\bar{z}]}, \tag{2.19}$$

where the $[\bar{z}]$ is the reversed trajectory of $[z]$:

$$\bar{z}(\tau) \equiv z(t-\tau), \tag{2.20}$$

and $\bar{\boldsymbol{k}}$ is the generator of the reversed dynamics:

$$\bar{\boldsymbol{k}}(\tau) \equiv \boldsymbol{k}(t-\tau). \tag{2.21}$$

From (2.13) and the definitions of $[\bar{z}]$ and $\bar{\boldsymbol{k}}$, it follows

$$\Sigma[z] = \Delta S(t) + S_{\text{env}}[z], \tag{2.22}$$

where

$$\Delta S(t) \equiv \left[s^{\text{int}}_{z_N} - \ln \pi_{z_N}(t)\right] - \left[s^{\text{int}}_{z_0} - \ln \pi_{z_0}(0)\right] \tag{2.23}$$

is the change in the entropy of the system defined by

$$S_x(t) \equiv s^{\text{int}}_x - \ln \pi_x(t), \tag{2.24}$$

and involving the stochastic entropy $-\ln \pi_x(t)$[2] and the equilibrium entropy s^{int}_x related to the internal structure of the state x [211]; and where

$$S_{\text{env}}[z] \equiv \sum_{i=0}^{N-1} \ln \frac{k_{z_{i+1},z_i}(t_{i+1})}{k_{z_{N-i-1},z_{N-i}}(t_{i+1})} - \left[s^{\text{int}}_{z_N} - s^{\text{int}}_{z_0}\right] \tag{2.25}$$

is the change in the entropy of the environment along the trajectory $[z]$ due either to the energy exchange between the system and the environment caused by the jumps (first term) or the change in the internal entropy of the system (second term). If we assume that the transition rates satisfy the *instantaneous detailed balance* defined by

$$\frac{k_{xy}(t)}{k_{yx}(t)} = \mathrm{e}^{-\beta(U_x(t)-U_y(t))+\beta\mu(N_x(t)-N_y(t))+(s^{\text{int}}_x - s^{\text{int}}_y)}, \tag{2.26}$$

[2] Note that the mean of the stochastic entropy is the Shannon entropy: $\langle \Delta S \rangle \equiv -\sum_x \pi_x \ln \pi_x$.

then the entropy of the environment satisfies

$$S_{\rm env}[z] = \sum_{i=0}^{N-1} \left[-\beta(U_{z_{i+1}}(t_{i+1}) - U_{z_i}(t_{i+1})) + \beta\mu(N_{z_{i+1}}(t_{i+1}) - N_{z_i}(t_{i+1}))\right] = -\beta Q[z], \tag{2.27}$$

recovering the result of equilibrium thermodynamics. We now consider the entropy production of the reversed dynamics $\bar{\Sigma}[\bar{z}]$ given by

$$\bar{\Sigma}[\bar{z}] \equiv \ln \frac{\mathbb{P}_{\bar{\boldsymbol{k}},\boldsymbol{\pi}(t)}[\bar{z}]}{\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]} = -\Sigma[z]. \tag{2.28}$$

We can relate the probability that $\Sigma[z]$ takes the value s to the probability that $\bar{\Sigma}[\bar{z}]$ takes the opposite value $-s$:

$$P(\Sigma[z] = s) = \sum_{[z]} \delta(\Sigma[z] - s)\, \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] = \mathrm{e}^{s} \sum_{[z]} \delta(\bar{\Sigma}[\bar{z}] + s)\, \mathbb{P}_{\bar{\boldsymbol{k}},\boldsymbol{\pi}(t)}[\bar{z}], \tag{2.29}$$

leading to

$$\frac{P(\Sigma[z] = s)}{\bar{P}(\bar{\Sigma}[z] = -s)} = \mathrm{e}^{s} \tag{2.30}$$

where we introduced $\bar{P}(\bar{\Sigma}[z] = -s)$ the probability in the time-reversed and path reversed dynamics that $\bar{\Sigma}[\bar{z}]$ takes the value $-s$:

$$\bar{P}(\bar{\Sigma}[z] = -s) \equiv \sum_{[z]} \delta(\bar{\Sigma}[\bar{z}] + s) \mathbb{P}_{\bar{\boldsymbol{k}},\boldsymbol{\pi}(t)}[\bar{z}]. \tag{2.31}$$

Eq. (2.30) is known as the *detailed fluctuation theorem* [12–16, 24, 31, 212–217]. Note that the fluctuation theorem compares the probability of the entropy production of a given experiment (direct dynamics) with the entropy production of another experiment (reversed dynamics). We consider the same observable (entropy production) but for two different experiments. If the protocol is time reversal symmetric, i.e. $\bar{\boldsymbol{k}} = \boldsymbol{k}$, one obtains the stronger fluctuation relation:

$$\frac{P(\Sigma[z] = s)}{P(\Sigma[z] = -s)} = \mathrm{e}^{s}, \tag{2.32}$$

where the probabilities appearing in the numerator and denominator are the same. In this case, Eq. (2.32) compares the probability of the entropy production of a given experiment along a trajectory with the probability of the entropy production of the same experiment along the reversed trajectory. Note that if we assume that $\frac{1}{t}\Sigma[z]$ and $\frac{1}{t}\bar{\Sigma}[z]$ satisfy LDPs:

$$P(\Sigma[z] = ts) \underset{t\to\infty}{\asymp} \mathrm{e}^{-tI(s)}, \tag{2.33}$$

$$\bar{P}(\bar{\Sigma}[z] = ts) \underset{t\to\infty}{\asymp} \mathrm{e}^{-t\bar{I}(s)}, \tag{2.34}$$

then from Eq. (2.30), the fluctuation theorem reads for LDFs:

$$I(s) - \bar{I}(-s) = -s. \tag{2.35}$$

Now, averaging the exponential of the entropy production with respect to the path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$

$$\mathbb{E}\left[\mathrm{e}^{-\Sigma[z]}\right] = \int \mathrm{d}s P(\Sigma[z]=s)\mathrm{e}^{-\Sigma[z]} \tag{2.36}$$

yields the *integral fluctuation theorem*:

$$\mathbb{E}\left[\mathrm{e}^{-\Sigma[z]}\right] = 1, \tag{2.37}$$

where we used the detailed fluctuation theorem (2.30). Applying Jensen's inequality

$$\mathrm{e}^{\mathbb{E}[X]} \;\leq\; \mathbb{E}[\mathrm{e}^{X}], \tag{2.38}$$

we obtain

$$\mathbb{E}\left(\Sigma[z]\right) \geq 0. \tag{2.39}$$

Hence, even though the entropy production can take negative values at the level of trajectories, its mean value is always positive: the second law of thermodynamics holds on average. There has been a large number of experimental illustrations of the fluctuation theorem [200–202, 204, 218].

2.3.4 Biased generator and generating functions

Let us come back to the general observable $\boldsymbol{A}_t$ of Eq. (2.14). We are interested in the fluctuations of $\boldsymbol{A}_t$ at time t, or more precisely of the time-extensive observable $t\boldsymbol{A}_t$. As discussed in Sec. 1.2, all the statistics of $\boldsymbol{A}_t$ follows from the generating function

$$G(t,\boldsymbol{\gamma}) \equiv \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right], \tag{2.40}$$

where $\boldsymbol{\gamma} = \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix}$ is the conjugate variable of $\boldsymbol{A}_t$ and where the average value is defined with respect to the path probability (2.13) with initial probability $\boldsymbol{\pi}(0)$:

$$\mathbb{E}_{\boldsymbol{\pi}(0)}[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{f}[z]}] \equiv \int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{f}[z]} \tag{2.41}$$

for any $\boldsymbol{f}$ functional of the path. Note that using the fact that

$$P(\boldsymbol{A}_t[z]=\boldsymbol{a}) = \int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]\delta(\boldsymbol{A}_t[z]-\boldsymbol{a}), \tag{2.42}$$

we recover the definition of the average given in Eq. (1.7). Let us consider the generating function conditioned on the final state

$$\tilde{G}_x(t,\boldsymbol{\gamma}) \equiv \mathbb{E}_{\boldsymbol{\pi}(0)}\left[e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\delta_{x,z(t)}\right], \tag{2.43}$$

so that $G(t,\boldsymbol{\gamma}) = \sum_x \tilde{G}_x(t,\boldsymbol{\gamma})$. Notice that this generating function is related to the joint probability

$$P_x(\boldsymbol{A}_t[z]=\boldsymbol{a},t) \equiv \int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]\delta_{x,z(t)}\delta(\boldsymbol{A}_t[z]=\boldsymbol{a}) \tag{2.44}$$

that the system is in state x and that the observable $\boldsymbol{A}_t[z]$ takes the value $\boldsymbol{a}$ at time t by:

$$\tilde{G}_x(t,\boldsymbol{\gamma}) = \int \mathrm{d}\boldsymbol{a} P_x(\boldsymbol{A}_t[z]=\boldsymbol{a},t)e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{a}}. \tag{2.45}$$

We aim to describe the time evolution of $\tilde{G}$. To do so, we introduce for clarity the observable $\tilde{\boldsymbol{A}}_t[z] \equiv t\boldsymbol{A}_t[z]$. During an infinitesimal time δt, either there is a jump and the first component $\tilde{A}_t^{(1)}$ of $\tilde{\boldsymbol{A}}_t$ varies by a quantity $g_{z(t+\delta t),z(t)}(t)$ and the second component $\tilde{A}_t^{(2)}$ varies by a quantity $h_{z(t)}\delta t$, or there is no jump and only $\tilde{\boldsymbol{A}}_t^{(2)}$ varies by $h_{z(t)}(t)\delta t$. The joint probability at time $t+\delta t$ reads then [210]

$$P_x(\tilde{\boldsymbol{A}}_{t+\delta t}[z]=\tilde{\boldsymbol{a}},t+\delta t) = \sum_{y\neq x} k_{xy}(t)\delta t P_y\left(\tilde{A}_t^{(1)}[z]=\tilde{a}^{(1)}-g_{xy}(t),\tilde{A}_t^{(2)}[z]=\tilde{a}^{(2)}-h_y(t)\delta t,t\right) + \left[1-\sum_{y\neq x}k_{yx}(t)\delta t\right]P_x\left(\tilde{A}_t^{(1)}[z]=\tilde{a}^{(1)},\tilde{A}_t^{(2)}[z]=\tilde{a}^{(2)}-h_x(t)\delta t,t\right) \tag{2.46}$$

with $\tilde{a}^{(1)}$ and $\tilde{a}^{(2)}$ respectively the first and second component of $\tilde{\boldsymbol{a}}$. Expanding up to order δt, we obtain

$$\frac{\partial}{\partial t}P_x(\tilde{\boldsymbol{A}}_t[z]=\tilde{\boldsymbol{a}},t) = \sum_{y\neq x}k_{xy}(t)P_y\left(\tilde{A}_t^{(1)}[z]=\tilde{a}^{(1)}-g_{xy}(t),\tilde{A}_t^{(2)}[z]=\tilde{a}^{(2)},t\right) - \sum_{y\neq x}k_{yx}(t)P_x(\tilde{\boldsymbol{A}}_t[z]=\tilde{\boldsymbol{a}},t) - h_x(t)\frac{\partial}{\partial\tilde{a}^{(2)}}P_x(\tilde{\boldsymbol{A}}_t[z]=\tilde{\boldsymbol{a}},t). \tag{2.47}$$

Using Eqs. (2.45) and (2.47), we finally obtain the differential equation ruling the time evolution of the generating function $\tilde{G}_x$ [16, 219]:

$$\frac{\partial}{\partial t}\tilde{G}_x(t,\boldsymbol{\gamma}) = \sum_y \kappa_{xy}(t,\boldsymbol{\gamma})\tilde{G}_y(t,\boldsymbol{\gamma}), \tag{2.48}$$

where we have introduced the matrix $\boldsymbol{\kappa}$ of components

$$\kappa_{xy}(t,\boldsymbol{\gamma}) \equiv \begin{cases} k_{xy}(t)\mathrm{e}^{\gamma_1 g_{xy}(t)} & \text{if } x\neq y, \\ -\sum_{x'\neq x}k_{x'x}(t)+\gamma_2 h_x(t) & \text{if } x=y. \end{cases} \tag{2.49}$$

The matrix $\boldsymbol{\kappa}$ is called the *tilted generator* or the *biased generator* [44, 99, 100]. The difference with the master equation (2.6) is that contrary to the transition rate matrix $\boldsymbol{k}$, the matrix $\boldsymbol{\kappa}$ does not satisfy the norm-conservation property: $\sum_x \kappa_{xy}(t,\boldsymbol{\gamma})\neq 0$, for any y. The biased matrix is the generator of a Markov dynamics whose "transition probability"

$$\tilde{G}(x,t\mid y,0) = \int \mathfrak{D}[z]\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\mathbb{P}[z\mid y,0]\delta_{z(t),x}, \tag{2.50}$$

solution of Eq. (2.48) with $\tilde{G}(x,0\mid y,0)=\delta_{xy}$ [28], is not normalized. We call $\tilde{G}(x,t\mid y,0)$ the *biased transition probability* by abuse of language. The *biased path probability* $\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]$ corresponding to the non-normalized "path probability" of the dynamics generated by $\boldsymbol{\kappa}$ is obtained by replacing $\boldsymbol{k}$ by $\boldsymbol{\kappa}$ in Eq. (2.13):

$$\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z] \equiv \mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]. \tag{2.51}$$

Notice that the initial generating function coincides with the initial probability:

$$\tilde{G}_x(0) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\delta_{x,z(t)}\right] = \pi_x(0), \tag{2.52}$$

for any state x. In Section 2.5.2, we will see that the canonical path probability is defined from the normalization of the biased path probability of the tilted process.

2.4 Markov diffusion processes

In the following, we express the results of the previous section in the language of diffusion processes [125–127, 220], which constitute the continuous-state counterpart of jump processes.

2.4.1 Langevin equation, Markov generator and Fokker-Planck equation

We consider a continuous-time Markov diffusion process z_t taking its values on the set of real numbers. The state of the system at time t is denoted by z_t. For simplicity, we consider a one-dimensional process, i.e. z_t is a scalar. This Markov process is characterized by two functions: the drift $b(z_t, t)$ and the diffusion coefficient $\sigma(z_t, t)$. We denote by $\varrho(x, t)$ the probability density to occupy the state x at time t. The random variable z_t evolves according a stochastic differential equation called the *overdamped*[3] *Langevin equation*:

$$\dot{z}_t = b(z_t, t) + \sigma(z_t, t)\xi_t, \tag{2.54}$$

where ξ_t is a Gaussian random variable with

$$\langle \xi_t \rangle = 0, \tag{2.55}$$
$$\langle \xi_t \xi_{t'} \rangle = \delta(t - t'), \tag{2.56}$$

meaning that ξ_t and $\xi_{t'}$ are statistically independent for $t \neq t'$. ξ_t is formally defined by

$$\xi_t = \frac{\mathrm{d}W_t}{\mathrm{d}t}, \tag{2.57}$$

where W_t is a Wiener process. Mathematically speaking, the Langevin equation does not exist as the derivative of a Wiener process is not defined. The correct interpretation of the equation is in terms of the following integral equation

$$z_t = z_0 + \int_0^t b(z_\tau, \tau)\mathrm{d}\tau + \int_0^t \sigma(z_\tau, \tau)\,\mathrm{d}W_\tau, \tag{2.58}$$

where it remains to give an appropriate definition of the stochastic integral $\int_0^t \sigma(z_\tau, \tau)\,\mathrm{d}W_\tau$. More generally, let us consider the following stochastic integral:

$$\mathcal{I}(t) = \int_0^t f(\tau)\mathrm{d}W_\tau, \tag{2.59}$$

where f is a stochastic process whose stochasticity depends of W and which is square integrable $\mathbb{E}[\int_0^t f(\tau)^2\mathrm{d}\tau] < \infty$ [127]. To define properly $\mathcal{I}$, we discretize the time interval

[3] In general, the Langevin equation describes the motion of a particle of mass m subjected to a friction ν, an external force F and a stochastic force X:

$$m\ddot{z}_t = -\nu\dot{z}_t + F(z_t, t) + X_t. \tag{2.53}$$

In the limit of large friction, we can neglect the inertial term and consider the overdamped Langevin equation of Eq. (2.54), see Ref. [127] for more details.

$[0,t]$ into M parts by introducing $t_i \equiv i\Delta t$ with $i \in [\![0, M-1]\!]$ and $\Delta t = \frac{t}{M}$. Let $\alpha \in [0,1]$ and $\tau_i \equiv (1-\alpha)t_i + \alpha t_{i+1}$. Then, the stochastic integral $\mathcal{I}$ is defined by

$$\mathcal{I}(t) \equiv \lim_{M\to\infty} \sum_{i=0}^{M-1} f(\tau_i)\left[W(t_{i+1}) - W(t_i)\right]. \tag{2.60}$$

The two main conventions used in stochastic calculus are the *Itô convention* obtained for $\alpha = 0$ and the *Stratonovich convention* obtained for $\alpha = \frac{1}{2}$. The Îto convention is favored by the mathematics community as it has the nice property to be a Martingale. On the other hand, physicists rather use the Stratonovich integral since the chain rule holds in this convention. Moreover, the Stratonovich convention may be more adapted to approach real systems (see the Wong–Zakai theorem [221, 222]). From now on, we will work with the Stratonovich convention and use the notation $\mathcal{I}(t) = \int_0^t f(\tau) \circ \mathrm{d}W_\tau$ to refer to the Stratonovich integral. The data z_τ at any time $\tau \in [0,t]$ constitutes a path denoted by $[z]$. As before, we discretize the interval $[0,t]$ into M segments such that $z_i \equiv z_{t_i}$. The transition probability density reads [222]

$$P(z_i, t_i \mid z_{i-1}, t_{i-1}) \equiv \left\langle \delta\left(z_i - z_{i-1} - \int_{t_{i-1}}^{t_i} \left[b(z_\tau,\tau) + \sigma(z_\tau,\tau)\circ\xi_\tau\right]\mathrm{d}\tau\right)\right\rangle_\xi, \tag{2.61}$$

where the average $\langle\ldots\rangle_\xi$ is defined with respect to the noise ξ. The probability density of the path $[z]$ is obtained from the transition probabilities as follows

$$\mathbb{P}_{b,\sigma,\varrho(0)}[z] = \varrho(z_0,0)\prod_{i=0}^{M-1} P(z_i,t_i \mid z_{i-1},t_{i-1}), \tag{2.62}$$

where $\varrho(0)$ is the initial probability density. Following the derivation of Ref. [222], it leads to

$$\mathbb{P}_{b,\sigma,\varrho(0)}[z] = \varrho(z_0,0)\exp\left\{-\int_0^t \mathrm{d}\tau\left[\frac{1}{2D(z_\tau,\tau)}\left(\dot{z}_\tau - \hat{b}(z_\tau,\tau)\right)^2 + \frac{1}{2}\nabla b(z_\tau,\tau)\right]\right\}, \tag{2.63}$$

where $\nabla \equiv \partial_x$ is the (1D) spatial derivative and where we defined the modified drift $\hat{b}(x,t) \equiv b(x,t) - \frac{1}{2}\sigma(x,t)\nabla\sigma(x,t)$ and the variance $D(x,t) \equiv \sigma(x,t)^2$. We point out the fact that this path probability does not exist strictly speaking since the term $\dot{z}_t^2$ is not defined mathematically. Furthermore, taking another convention in the definition of the stochastic integral leads to another path probability. This problem is avoided by considering in further calculations not the path probability of Eq. (2.63), but the ratio between two path probabilities associated with different diffusion processes so that the diverging term $\dot{z}_t^2$ disappears. This quantity is convention-independent and the physical consistency is preserved.

The Langevin equation describes the evolution of the process z_t. Another level of description is via the probability density $\varrho(x,t)$ to occupy the state x at time t. It satisfies the so-called *Fokker-Planck equation*:

$$\frac{\partial\varrho(x,t)}{\partial t} = -\nabla\left[\hat{b}(x,t)\varrho(x,t) - \frac{1}{2}D(x,t)\nabla\varrho(x,t)\right], \tag{2.64}$$

with initial condition $\varrho(0)$. In Appendix 2.A, we derive the Fokker-Planck equation from the Langevin equation. Similarly to the Master equation, the Fokker-Planck equation is also a particular case of the Chapman-Kolmogorov equation and governs the evolution of the transition probability density $P(x,t \mid y,t')$. The transition probability densities $P(x,t+\mathrm{d}t \mid y,t)$ for an infinitesimal time $\mathrm{d}t$, and hence the path probability density $\mathbb{P}_{b,\sigma,\varrho(0)}[z]$, can also be obtained by solving directly the Fokker-Planck equation [220]. Note that the Fokker-Planck equation is nothing but a continuity equation

$$\frac{\partial \varrho(x,t)}{\partial t} = -\nabla J^{\varrho}(x,t), \tag{2.65}$$

where we introduced the probability current J^{ϱ}:

$$J^{\varrho}(x,t) \equiv \hat{b}(x,t)\varrho(x,t) - \frac{1}{2}D(x,t)\nabla\varrho(x,t). \tag{2.66}$$

The Fokker-Planck equation can also be expressed in the same form as the Master equation by introducing the so-called *Fokker-Planck operator* $\mathfrak{L}$ defined by its action on an arbitrary function f:

$$(\mathfrak{L}f)(x,t) \equiv -\nabla\left[\hat{b}(x,t)f(x,t) - \frac{1}{2}D(x,t)\nabla f\right], \tag{2.67}$$

leading to

$$\frac{\partial \varrho(x,t)}{\partial t} = (\mathfrak{L}\varrho)(x,t). \tag{2.68}$$

The Fokker-Planck operator is the generator of the diffusion process. It conserves the norm of the probability density since $\int_{-\infty}^{+\infty}\mathrm{d}x(\mathfrak{L}\varrho)(x,t) = 0$ for any time t, where we used the fact that ϱ vanishes at infinity.

2.4.2 Observables and laws of thermodynamics

In the framework of diffusion processes, the observable we are interested in reads

$$\boldsymbol{A}_t[z] = \begin{pmatrix} \frac{1}{t}\int_0^t g(z_\tau,\tau)\circ\mathrm{d}z_\tau \\ \frac{1}{t}\int_0^t h(z_\tau,\tau)\mathrm{d}\tau \end{pmatrix}, \tag{2.69}$$

where g and h are functions of position and time. When specifying g and h, $\boldsymbol{A}_t$ may represent many physical quantities. For instance, the first component of $\boldsymbol{A}_t$ is the empirical current at x if $g(y,\tau) = \delta(x-y)$, and the second component of $\boldsymbol{A}_t$ is the fraction of time the system spends in x if $h(y,\tau) = \delta(x-y)$ [80]. Let us now consider the simple example of a system in contact with a heat reservoir of inverse time-dependent temperature $\beta(t)$ and a having space-time dependent energy U. The variation of the internal energy between times 0 and t is given by:

$$\Delta U(t) = \int_0^t \mathrm{d}\tau \frac{\mathrm{d}U}{\mathrm{d}\tau}(z_\tau,\tau) = W[z] + Q[z], \tag{2.70}$$

where the work stems from the variation of the energy with respect to time:

$$W[z] \equiv \int_0^t \partial_t U(z_\tau,\tau)\,\mathrm{d}\tau \tag{2.71}$$

and the heat from the variation of the energy with respect to the position

$$Q[z] \equiv \int_0^t \nabla U(z_\tau, \tau) \circ \mathrm{d}z_\tau = \int_0^t \delta Q(z_\tau, \tau), \tag{2.72}$$

with $\delta Q(z_\tau, \tau) \equiv \nabla U(z_\tau, \tau) \circ \mathrm{d}z_\tau$. Eq. (2.70) is the expression of the first law of thermodynamics and was demonstrated experimentally along a single trajectory for a colloidal particle in a time-dependent non-harmonic potential in Ref. [34]. The entropy production along the path $[z]$ is defined by [31]

$$\Sigma[z] \equiv \ln \frac{\mathbb{P}_{b,\sigma,\varrho(0)}[z]}{\mathbb{P}_{\bar{b},\bar{\sigma},\varrho(t)}[\bar{z}]} \tag{2.73}$$

where the $[\bar{z}]$ is the reversed trajectory of $[z]$:

$$\bar{z}_\tau \equiv z_{t-\tau}, \tag{2.74}$$

and $\bar{b}$ and $\bar{\sigma}$ are respectively the drift and the diffusion coefficient of the reversed dynamics:

$$\bar{b}(x, \tau) \equiv b(x, t-\tau), \tag{2.75}$$
$$\bar{\sigma}(x, \tau) \equiv \sigma(x, t-\tau) \tag{2.76}$$

for all x. From (2.63) and the definitions of $\bar{z}$, $\bar{b}$ and $\bar{\sigma}$, it follows

$$\Sigma[z] = \Delta S(t) + S_{\mathrm{env}}[z], \tag{2.77}$$

where $\Delta S(t)$ is the change in stochastic entropy $S(x, \tau) \equiv -\ln \varrho(x, \tau)$ between times 0 and t:

$$\Delta S(t) \equiv -\ln \varrho(z_t, t) - \big(-\ln \varrho(z_0, 0)\big), \tag{2.78}$$

and $S_{\mathrm{env}}[z]$ is the change in the entropy of the environment along the trajectory $[z]$:

$$S_{\mathrm{env}}[z] \equiv \int_0^t \mathrm{d}\tau \frac{2}{D(z_\tau, \tau)} \hat{b}(z_\tau, \tau) \circ \mathrm{d}z_\tau. \tag{2.79}$$

If we assume that the drift and the diffusion coefficient satisfy the instantaneous detailed balance defined by

$$\frac{1}{2} D(x, t) = -\frac{\hat{b}(x, t)}{\beta(t) \nabla U(x, t)}, \tag{2.80}$$

then the entropy of the environment is related to the heat by $S_{\mathrm{env}}[z] = -\int_0^t \beta(\tau) \delta Q(z_\tau, \tau)$, recovering the result of equilibrium thermodynamics. For a Brownian particle, the variance does not depend on x and the drift $\hat{b}(x, t) = b(x, t)$ is $\frac{F(x,t)}{\nu}$, where $F(x, t) \equiv -\nabla U(x, t)$ is the conservative force associated with U, and ν is the friction coefficient [223]. In this case, the detailed balance condition of Eq. (2.80) becomes Einstein's relation $\frac{1}{2} D = \frac{1}{\beta \nu}$. Finally, following the same procedure as for jump processes, one can derive the fluctuation theorems for the entropy production (2.30, 2.35, 2.37) for diffusion processes and recover the second law on average.

2.4.3 Biased generator and generating functions

Let us come back to the general observable $\boldsymbol{A}_t$ of Eq. (2.69). As in Sec. 2.3.4, we are interested in the fluctuations of $\boldsymbol{A}_t$ at time t. The generating function associated with $\boldsymbol{A}_t$ reads

$$G_{\boldsymbol{\gamma}}(t) \equiv \mathbb{E}_{\varrho(0)}\left[e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right], \tag{2.81}$$

where $\boldsymbol{\gamma} = \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix}$ is the conjugate variable of $\boldsymbol{A}_t$ and where the average is defined with respect to the path probability (2.63) with initial probability density $\varrho(0)$:

$$\mathbb{E}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{f}[z]}\right] \equiv \int \mathfrak{D}[z]\mathbb{P}_{b,\sigma,\varrho(0)}[z]\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{f}[z]}, \tag{2.82}$$

for any $\boldsymbol{f}$ functional of the path. We define the generating function conditioned on the final state by

$$\tilde{G}_{\boldsymbol{\gamma}}(x,t) \equiv \mathbb{E}_{\varrho(0)}\left[e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\delta(z_t - x)\right], \tag{2.83}$$

so that $G_{\boldsymbol{\gamma}}(t) = \int \tilde{G}_{\boldsymbol{\gamma}}(x,t)\mathrm{d}x$. In the following, we lighten the notations by dropping the dependency of the functions b, σ and $\tilde{G}$ on x and t but one keeps in mind their presence. The generating function $\tilde{G}_{\boldsymbol{\gamma}}$ evolves according to

$$\dot{\tilde{G}}_{\boldsymbol{\gamma}} = (\Lambda_{\boldsymbol{\gamma}}\tilde{G}_{\boldsymbol{\gamma}}), \tag{2.84}$$

where we introduced the biased Fokker-Planck operator $\Lambda_{\boldsymbol{\gamma}}$ defined by

$$(\Lambda_{\boldsymbol{\gamma}}f) \equiv (-\nabla + \gamma_1 g)(\hat{b}f) + \frac{1}{2}(-\nabla + \gamma_1 g)\left[D(-\nabla + \gamma_1 g)f\right] + \gamma_2 hf, \tag{2.85}$$

for any space-time function f. Eq. (2.85) is derived in Appendix 2.B. Similarly to the discrete case, the biased Fokker-Planck operator $\Lambda_{\boldsymbol{\gamma}}$ generates a Markov process that is not norm-conserving since

$$\int_{-\infty}^{+\infty} \mathrm{d}x(\Lambda_{\boldsymbol{\gamma}}\tilde{G}_{\boldsymbol{\gamma}})(x,t) \neq 0, \tag{2.86}$$

and whose path probability reads [80]

$$\mathbb{P}_{\Lambda_{\boldsymbol{\gamma}},\varrho(0)}[z] \equiv \mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\mathbb{P}_{b,\sigma,\varrho(0)}[z]. \tag{2.87}$$

In the next section, we will see how this path probability is related to the canonical path ensemble.

2.5 Generalization of microcanonical and canonical ensembles

We saw in Sec. 2.1 that statistical properties of equilibrium macroscopic systems are studied within statistical ensembles. The most usual ones are the microcanonical ensemble in which the energy is fixed, and the canonical ensemble for which the constraint is not on energy anymore but on temperature. Nonequilibrium analogues of these ensembles have been introduced in Refs. [37–39] and formalized by H. Touchette and R. Chétrite in Ref. [52] in which they defined nonequilibrium canonical and microcanonical path ensembles and discussed their equivalence in the long-time limit. In this section, we review the main results of Refs. [52, 195, 224] which will be useful for the rest of this manuscript. In the following, we consider the case of Markov jump processes but similar results are valid for diffusion process by simply adapting the notations in the path probabilities.

2.5.1 Nonequilibrium microcanonical process

By analogy with the equilibrium case, the nonequilibrium microcanonical ensemble corresponds to the process $z(t)$ whose path probability is conditioned on $\boldsymbol{A}_t = \boldsymbol{a}$:

$$\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \equiv \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z \mid \boldsymbol{A}_t[z] = \boldsymbol{a}]. \tag{2.88}$$

In general, there is no Markov generator that can exactly generate this microcanonical path ensemble, which makes this process difficult to study. Yet, we will see in the following that it is asymptotically equivalent (in a way that will be precised) to the canonical process, defined below, and which is associated with a Markov generator (see Ref. [80] and Chapter 3).

2.5.2 Nonequilibrium canonical process

In the equilibrium canonical ensemble, instead of fixing the value of energy, we fix the value of its Legendre conjugate parameter, namely the temperature. By analogy, the nonequilibrium canonical path probability is defined by connecting the original process to an exponential tilting of the path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$:

$$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] \equiv \frac{\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right]}, \tag{2.89}$$

where $\boldsymbol{A}_t[z]$ generalizes the energy E_ℓ in the canonical distribution (2.1), the Legendre conjugate variable $\boldsymbol{\gamma}$ generalizes the inverse temperature $-\beta$, and $G(t,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right]$ generalizes the partition function Z. Note that the canonical path probability is obtained by normalizing the biased path probability (2.51). This definition has already been used in many articles, for instance for the simulation of transition paths associated with glassy systems [40, 42, 43, 47, 49].

2.5.3 Equivalence between microcanonical and canonical path ensembles

Let us first define mathematically the notion of equivalence. Two path probabilities $\mathbb{P}_t$ and $\mathbb{Q}_t$ are said to be logarithmically equivalent if the random variable defined by

$$R_t[z] \equiv \frac{1}{t}\ln\frac{\mathbb{P}_t[z]}{\mathbb{Q}_t[z]} \tag{2.90}$$

converges in probability to 0 with respect to both probabilities $\mathbb{P}_t$ and $\mathbb{Q}_t$:

$$\lim_{t\to\infty}\frac{1}{t}R_t = 0 \tag{2.91}$$

almost everywhere (i.e. except for events of null probability), and we denote it

$$\mathbb{P}_t \underset{t\to\infty}{\asymp} \mathbb{Q}_t. \tag{2.92}$$

In order to discuss the equivalence in the sense of (2.91) between canonical and microcanonical path probabilities, we need to assume that $\boldsymbol{A}_t$ satisfies a LDP:

$$P(\boldsymbol{A}_t = \boldsymbol{a}) \underset{t\to\infty}{\asymp} \mathrm{e}^{-tI(\boldsymbol{a})}, \tag{2.93}$$

where the probability $P(\boldsymbol{A}_t = \boldsymbol{a})$, and hence the LDF I, is defined using the reference path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$:

$$P(\boldsymbol{A}_t = \boldsymbol{a}) = \int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]\delta(\boldsymbol{A}_t[z] - \boldsymbol{a}). \tag{2.94}$$

The equivalence between $\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z]$ and $\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$ is strictly related to the convexity of the LDF I according to [52]:

- If I is strictly convex at $\boldsymbol{a}$ then there exists $\boldsymbol{\gamma}$ such that $\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \underset{t\to\infty}{\asymp} \mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$. If I is differentiable[4], the equivalence is satisfied for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$, which is a nonequilibrium version of the relation between entropy and temperature in equilibrium statistical mechanics: $\beta = \frac{\partial S}{\partial E}$. In case of non-strict convexity, this result holds for the convergence in probability with respect to $\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z]$ but this limit is not defined for the convergence in probability with respect to $\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$.

- If I is not convex at $\boldsymbol{a}$, there is no logarithmic equivalence between $\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z]$ and $\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$.

It has been shown in Ref. [195] that the path probability equivalence in (2.91) is related to the equivalence between typical values of observables with respect to both path probabilities. Sketchily, it means that if an observable $\boldsymbol{B}_t$ satisfies a LDP with respect to $\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$ $\left(\text{resp. } \mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z]\right)$ with LDF $J_{\boldsymbol{\gamma}}$ (resp. $J_{\boldsymbol{a}}$), and in case of logarithmic equivalence between canonical and microcanonical path probabilities, then the LDFs $J_{\boldsymbol{\gamma}}$ and $J_{\boldsymbol{a}}$ have the same zeros. This result holds only if I is strictly convex. In case of non-strict convexity, the set of zeros of $J_{\boldsymbol{a}}$ is contained in the set of zeros of $J_{\boldsymbol{\gamma}}$, and we talk about *partial equivalence*.

[4] If I is not differentiable, the equivalence holds for all $\boldsymbol{\gamma} \in \partial I(\boldsymbol{a})$, where $\partial I(\boldsymbol{a})$ is the *subdifferential* of I at $\boldsymbol{a}$. The notion of subdifferential generalizes the notion of derivative for non-differentiable convex function. The subdifferential of a convex function f at a point $\boldsymbol{x}$ is defined as the set of values $\boldsymbol{s}$ such that

$$\forall \boldsymbol{y},\ f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{s}\cdot(\boldsymbol{y} - \boldsymbol{x}).$$

$\boldsymbol{s}$ is called a *subderivative*. In 1D, the geometrical interpretation of the subdifferential of f at x is given by the ensemble of slopes of the lines passing by the point $(x, f(x))$ and that are below or touch the curve of f everywhere else. For instance, the subdifferential of the function $f : x \mapsto |x|$ at 0 is $\partial f(0) = [-1, 1]$.

Appendices

"Si vous ne pouvez expliquer un concept à un enfant de six ans, c'est que vous ne le comprenez pas complètement."

Albert Einstein

2.A Derivation of the Fokker-Planck equation from the Langevin equation

We derive the Fokker-Planck equation by following the procedure described in Ref. [222]. First, we compute the mean and the variance of $\delta z \equiv z_{t+\delta t} - z_t$. From the Langevin equation (2.54), we have

$$z_{t+\delta t} = z_t + \int_t^{t+\delta t} [b(z_\tau, \tau) + \sigma(z_\tau, \tau)\, \xi_\tau]\, \mathrm{d}\tau, \tag{2.95}$$

where the stochastic integral (2.60) is defined in the α-convention, with $\alpha \in [0, 1]$. Noticing that

$$\alpha z_{t+\delta t} + (1-\alpha) z_t = z_t + \alpha \delta z, \tag{2.96}$$

we have for arbitrarily small δt:

$$\int_t^{t+\delta t} b(z_\tau, \tau) \mathrm{d}\tau \simeq b(z_t + \alpha\delta z, t + \alpha\delta t)\delta t \tag{2.97}$$

$$\int_t^{t+\delta t} [\sigma(z_\tau, \tau)\xi_\tau]\, \mathrm{d}\tau \simeq \sigma(z_t + \alpha\delta z, t + \alpha\delta t) \int_t^{t+\delta t} \mathrm{d}\tau\, \xi_\tau. \tag{2.98}$$

Hence, from Eq. (2.95) and Eqs. (2.97–2.98),

$$\delta z = b(z_t + \alpha\delta z, t + \alpha\delta t)\delta t + \sigma(z_t + \alpha\delta z, t + \alpha\delta t) \int_t^{t+\delta t} \mathrm{d}\tau\, \xi_\tau. \tag{2.99}$$

Since $(\int_t^{t+\delta t} \mathrm{d}t\, \xi_t)^2$ is of order δt, $\int_t^{t+\delta t} \mathrm{d}t\, \xi_t$ is of order $\sqrt{\delta t}$ implying that δz is also of order $\sqrt{\delta t}$ via the Langevin equation. We expand Eq. (2.99) up to order δt:

$$\delta z \simeq b(z_t, t)\delta t + \sigma(z_t, t) \int_t^{t+\delta t} \mathrm{d}\tau\, \xi_\tau + \alpha\sigma(z_t, t)\nabla\sigma(z_t, t) \left(\int_t^{t+\delta t} \mathrm{d}\tau\, \xi_\tau \right)^2, \tag{2.100}$$

$$(\delta z)^2 \simeq \sigma(z_t, t)^2 \left(\int_t^{t+\delta t} \mathrm{d}\tau\, \xi_\tau \right)^2. \tag{2.101}$$

Taking the average with respect to the noise ξ, the mean and variance of δz are given by

$$\langle \delta z \rangle_\xi \simeq b(z_t,t)\delta t + \alpha\sigma(z_t,t)\nabla\sigma(z_t,t)\delta t, \tag{2.102}$$

$$\left\langle (\delta z)^2 \right\rangle_\xi \simeq \sigma(z_t,t)^2\delta t, \tag{2.103}$$

where we used Eqs. (2.55–2.56). We now compute the probability density to be in x at time $t+\delta t$:

$$\varrho(x,t+\delta t) = \int \mathrm{d}y P(x,t+\delta t \mid y,t)\varrho(y,t), \tag{2.104}$$

where the transition probability density is given by

$$P(x,t+\delta t \mid y,t) = \langle \delta(z_{t+\delta t}-x)\rangle_\xi \tag{2.105}$$

$$= \langle \delta(z_t+\delta z-x)\rangle_\xi \tag{2.106}$$

$$\simeq \langle \delta(z_t-x)\rangle_\xi + \langle \delta z\,\nabla\delta(z_t-x)\rangle_\xi + \frac{1}{2}\left\langle (\delta z)^2\nabla^2\delta(z_t-x)\right\rangle_\xi, \tag{2.107}$$

with the condition $z_t = y$ and where we used Taylor's formula up to second order in δz in the last equation. Using Eqs. (2.102–2.103) in Eq. (2.107), we obtain

$$\begin{aligned}\varrho(x,t+\delta t) = &\int \mathrm{d}y\, \varrho(y,t)\delta(x-y)\\ &+ \int \mathrm{d}y\, \varrho(y,t)\left[b(y,t)+\alpha\sigma(y,t)\nabla_y\sigma(y,t)\right]\,\delta t\,\nabla_y\delta(x-y)\\ &+ \frac{1}{2}\int \mathrm{d}y\, \varrho(y,t)\,\sigma(y,t)^2\,\delta t\,(\nabla_y)^2\delta(x-y),\end{aligned} \tag{2.108}$$

where ∇_y is the derivative with respect to y. Finally, using integrations by parts, we obtain

$$\varrho(x,t+\delta t) = \varrho(x,t) - \nabla\left[\left(b(x,t)+\alpha\sigma(x,t)\nabla\sigma(x,t)\right)\varrho(x,t)\right]\delta t + \nabla^2\left[\frac{1}{2}\sigma(x,t)^2\varrho(x,t)\right]\delta t, \tag{2.109}$$

leading to the Fokker-Planck equation in the α-convention:

$$\frac{\partial\varrho(x,t)}{\partial t} = -\nabla\left[\left(b(x,t)+\alpha\sigma(x,t)\nabla\sigma(x,t)\right)\varrho(x,t) - \frac{1}{2}\nabla\left[\sigma(x,t)^2\varrho(x,t)\right]\right]. \tag{2.110}$$

Taking $\alpha = \frac{1}{2}$ and introducing $\hat{b}(x,t) \equiv b(x,t) - \frac{1}{2}\sigma(x,t)\nabla\sigma(x,t)$ and $D(x,t) \equiv \sigma(x,t)^2$, the Fokker-Planck equation (2.64) in the Stratonovich convention follows from Eq. (2.110).

2.B Derivation of the biased Fokker-Planck equation

In this appendix, we derive the biased Fokker-Planck equation (2.85) for an arbitrary $\alpha \in [0,1]$ in the convention of the stochastic integral. The Stratonovich convention will be recovered for $\alpha = \frac{1}{2}$. To do so, we compute the value of the generating function $\tilde{G}_{\boldsymbol{\gamma}}$ (2.83) conditioned on x at time $t+\delta t$:

$$\tilde{G}_{\boldsymbol{\gamma}}(x,t+\delta t) = \int \mathrm{d}y \tilde{G}_{\boldsymbol{\gamma}}(x,t+\delta t \mid y,t)\tilde{G}_{\boldsymbol{\gamma}}(y,t), \tag{2.111}$$

where the biased transition probability density from y at time t to x at time $t+\delta t$ is defined by

$$\tilde{G}_{\boldsymbol{\gamma}}(x,t+\delta t \mid y,t) = \left\langle \delta(z_{t+\delta t}-x) \mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}} \right\rangle_\xi \tag{2.112}$$

$$= \delta(y-x)\left\langle \mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}} \right\rangle_\xi + \nabla_y \delta(y-x)\left\langle \delta z\, \mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}} \right\rangle_\xi$$
$$+ \frac{1}{2}{\nabla_y}^2 \delta(y-x)\left\langle (\delta z)^2 \mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}}) \right\rangle_\xi + o(\delta t), \tag{2.113}$$

where we used Eq. (2.107) in the second line and where the observable $\boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}$ evaluated on the fraction of the path $[z]$ between times t and $t+\delta t$ reads in the α-convention:

$$\boldsymbol{A}_{\delta t}[z]_t^{t+\delta t} = \begin{pmatrix} \frac{1}{\delta t}\int_t^{t+\delta t} g(z_\tau,\tau)\mathrm{d}z_\tau \\ \frac{1}{\delta t}\int_t^{t+\delta t} h(z_\tau,\tau)\mathrm{d}\tau \end{pmatrix} \simeq \begin{pmatrix} \frac{1}{\delta t} g(z_t+\alpha\delta z, t+\alpha\delta t)\delta z \\ \frac{1}{\delta t} h(z_t+\alpha\delta z, t+\alpha\delta t)\delta t \end{pmatrix}. \tag{2.114}$$

Keeping in mind that δz is of order $\sqrt{\delta t}$, the expansion of $\mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}}$ up to order δt gives

$$\mathrm{e}^{\delta t \boldsymbol{\gamma}\cdot \boldsymbol{A}_{\delta t}[z]_t^{t+\delta t}} = 1 + \gamma_1 g(z_t+\alpha\delta z, t+\alpha\delta t)\delta z + \frac{1}{2}\left(\gamma_1 g(z_t+\alpha\delta z, t+\alpha\delta t)\right)^2 (\delta z)^2$$
$$+ \gamma_2 h(z_t+\alpha\delta z, t+\alpha\delta t)\delta t + o(\delta t), \tag{2.115}$$

$$= 1 + \gamma_1 g(z_t,t)\delta z + \left[\alpha\gamma_1 \nabla g(z_t,t) + \frac{1}{2}\left(\gamma_1 g(z_t,t)\right)^2\right](\delta z)^2 + \gamma_2 h(z_t,t)\delta t + o(\delta t). \tag{2.116}$$

Using Eqs. (2.100–2.103) and Eq. (2.116) in Eq. (2.113) and after long but not tricky calculation, it follows

$$\tilde{G}_{\boldsymbol{\gamma}}(x,t+\delta t \mid y,t) = \delta(x-y) + \delta(x-y)\delta t\Big[b(y,t)\gamma_1 g(y,t) + \alpha\sigma(y,t)\nabla\sigma(y,t)\gamma_1 g(y,\tau)$$
$$+ \alpha\sigma(y,t)^2\gamma_1\nabla g(y,t) + \frac{1}{2}\gamma_1^2 g(y,t)^2\sigma(y,t)^2 + \gamma_2 h(y,t)\Big]$$
$$+ \nabla_y\delta(x-y)\delta t\left[b(y,t) + \alpha\sigma(y,t)\nabla\sigma(y,t) + \gamma_1 g(y,t)\sigma(y,t)^2\right]$$
$$+ \left[{\nabla_y}^2\delta(x-y)\right]\frac{1}{2}\delta t\,\sigma(y,t)^2. \tag{2.117}$$

Combined with Eq. (2.111), it yields

$$\frac{\partial}{\partial t}\tilde{G}_{\boldsymbol{\gamma}}(x,t) = (\Lambda^\alpha_{\boldsymbol{\gamma}}\tilde{G}_{\boldsymbol{\gamma}})(x,t), \tag{2.118}$$

where the biased operator in the α-convention $\Lambda^\alpha_{\boldsymbol{\gamma}}$ is defined by its action on a arbitrary function f:

$$\Lambda^\alpha_{\boldsymbol{\gamma}} f \equiv \left[-\nabla + \gamma_1 g\right]\left[(b+\alpha\sigma\nabla\sigma)f\right] + \gamma_2 h f + \left[\alpha\sigma^2\gamma_1\nabla g + \frac{1}{2}{\gamma_1}^2 g^2\sigma^2\right] f$$
$$- \nabla\left[\gamma_1 g\sigma^2 f - \frac{1}{2}\nabla(\sigma^2 f)\right] \tag{2.119}$$

where we dropped the space-time dependency for clarity. In the Îto convention ($\alpha = 0$), the biased operator simplifies to

$$\Lambda^I_\gamma f = [-\nabla + \gamma_1 g]\,(bf) + \frac{1}{2}\,[-\nabla + \gamma_1 g]^2\,\left(\sigma^2 f\right) - \frac{1}{2}\nabla(\gamma_1 g)\sigma^2 f + \gamma_2 h f. \tag{2.120}$$

Taking $\alpha = \frac{1}{2}$, and after some algebraic manipulations, we recover the biased operator in the Stratonovich convention of Eq. (2.85).

Chapter 3

Conditioning periodically driven Markov processes on large deviations

"La vie n'a pas de sens. C'est juste l'occasion de faire des choses que tu trouves intéressantes."

Orochimaru – Naruto

In Refs. [80, 81], R. Chétrite and H. Touchette treated the problem of conditioning time-homogeneous Markov processes on a rare fluctuation of an observable $\boldsymbol{A}$ satisfying a LDP. They showed that this conditioned process is equivalent in the long-time limit to an effective Markov process called *the driven process* which leads $\boldsymbol{A}$ to have the conditioning value as a typical value. This result can be shown using a path probability approach or a variational approach.

The problem of conditioning a Markov process was first raised by Doob [82, 83] who considered a Wiener process[1] conditioned to leave the interval $[0, L]$ via the position L. He solved this problem by introducing a transformation on the Wiener process, later referred to as *Doob's h-transform*. This transformation, and more generally this problem of conditioning, have been generalized in Refs. [80, 81] to general Markov processes conditioned on the large deviations of a general observable. This work involves fundamental questions including the nonequilibrium generalization of equilibrium notions such as the definitions of canonical and microcanonical ensembles at the path level, ensemble equivalence [52] and Jaynes' maximum entropy principle using large deviation theory. From a more practical perspective, this work shows that the large deviation of an observable for a given process corresponds to a typical value of this observable for another process, and provides a way to build the Markov generator of this other process. An application of this result has been provided for conditioned time-homogeneous Langevin processes with weak noise in Refs. [225, 226].

In this section, we extend the work of R. Chétrite and H. Touchette to periodically driven Markov processes and for observables involving time-periodic functions. For the sake of completeness, we address this problem in the mathematical framework of both jump and diffusion processes. Even if the theory is essentially the same, it may be useful for

[1] The Wiener process is a continuous-time Markov process used to model the Brownian motion.

readers to have both formalisms separately. The content of Section 3.1 has been published in Ref. [227].

3.1 Conditioning of periodically driven Markov jump processes

3.1.1 Periodically driven Markov jump processes

We consider a continuous-time Markov jump process defined on a finite state space and whose Markov generator $\boldsymbol{k}$ is time-periodic with period T:

$$\boldsymbol{k}(t+T) = \boldsymbol{k}(t), \ \forall t. \tag{3.1}$$

We set the initial time to $t = 0$ and the final time to $t = nT$, where the integer n refers to the number of periods that have elapsed, and we assume that the solution of the master equation $\frac{\partial \boldsymbol{\pi}}{\partial t} = \boldsymbol{k}\boldsymbol{\pi}$ reaches a periodic solution $\boldsymbol{\pi}^{\mathrm{TiPS}}$ when $n \to \infty$ — where TiPS stands for *Time Periodic State* — i.e. $\boldsymbol{\pi}^{\mathrm{TiPS}}(t+T) = \boldsymbol{\pi}^{\mathrm{TiPS}}(t)$, $\forall t$. All the definitions introduced in Sec. 2.3 apply in this section with the difference that the final time is fixed to $t = nT$. As discussed previously, we are interested in the fluctuations of the observable $\boldsymbol{A}_{nT}$, functional of paths up to the final time nT, defined in Eq. (2.14) by replacing t by nT, and where we consider T-periodic functions $\boldsymbol{g}$ and $\boldsymbol{h}$: $\boldsymbol{g}(t+T) = \boldsymbol{g}(t)$ and $\boldsymbol{h}(t+T) = \boldsymbol{h}(t)$, $\forall t$. We assume that $\boldsymbol{A}_{nT}$ satisfies a LDP (1.18):

$$P(\boldsymbol{A}_{nT} = \boldsymbol{a}) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nTI(a)}, \tag{3.2}$$

with I the LDF, and we introduce its associated SCGF

$$\Gamma(\boldsymbol{\gamma}) = \lim_{n\to\infty} \frac{1}{nT} \ln \mathbb{E}_{\boldsymbol{\pi}(0)} \left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}} \right]. \tag{3.3}$$

In order to study the large deviations of $\boldsymbol{A}_{nT}$ in the long-time limit, let us first look at the observable $\boldsymbol{A}_t[z]$ on the shorter time interval $[0, t]$:

$$\boldsymbol{A}_t[z] \equiv \begin{pmatrix} \frac{1}{t} \sum_{i=0}^{M-1} g_{z_{i+1},z_i}(t_{i+1}) \theta(t - t_{i+1}) \\ \frac{1}{t} \int_0^t \mathrm{d}\tau h_{z(\tau)}(\tau) \end{pmatrix}, \tag{3.4}$$

where θ is the heaviside function defined by

$$\theta(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0. \end{cases} \tag{3.5}$$

As seen in Sec. 1.2, the statistics of $\boldsymbol{A}_t$ is contained in the generating function $G(t, \boldsymbol{\gamma}) = \sum_x \tilde{G}_x(t, \boldsymbol{\gamma})$ defined in Eq. (2.40), where $\tilde{G}_x(t, \boldsymbol{\gamma})$ is the generating function with state x at time t (2.43), and where the vector $\tilde{\boldsymbol{G}}(t, \boldsymbol{\gamma})$ of components $\tilde{G}_x(t, \boldsymbol{\gamma})$ is solution of

$$\frac{\partial}{\partial t} \tilde{\boldsymbol{G}}(t, \boldsymbol{\gamma}) = \boldsymbol{\kappa}(t, \boldsymbol{\gamma}) \tilde{\boldsymbol{G}}(t, \boldsymbol{\gamma}). \tag{3.6}$$

We remind the definition of the biased generator $\boldsymbol{\kappa}$:

$$\kappa_{xy}(t,\boldsymbol{\gamma}) \equiv \begin{cases} k_{xy}(t)\mathrm{e}^{\gamma_1 g_{xy}(t)} & \text{if } x \neq y, \\ -\sum_{y\neq x} k_{xy}(t) + \gamma_2 h_x(t) & \text{if } x = y. \end{cases} \tag{3.7}$$

Notice that, by definition, $\boldsymbol{\kappa}(t+T,\boldsymbol{\gamma}) = \boldsymbol{\kappa}(t,\boldsymbol{\gamma})$, $\forall t$. In the following, we keep in mind that $\boldsymbol{\kappa}$ depends on $\boldsymbol{\gamma}$ and drop $\boldsymbol{\gamma}$ in the notations for clarity. We can formally solve Eq. (3.6) with initial condition $\tilde{\boldsymbol{G}}(0,\boldsymbol{\gamma}) = \boldsymbol{\pi}(0)$ from Eq. (2.43) by writing

$$\tilde{\boldsymbol{G}}(t,\boldsymbol{\gamma}) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,0)\boldsymbol{\pi}(0), \tag{3.8}$$

where the propagator between times t_0 and t is defined by

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,t_0) \equiv \overleftarrow{\exp} \int_{t_0}^{t} \boldsymbol{\kappa}(\tau)\,\mathrm{d}\tau, \tag{3.9}$$

involving the time-ordered exponential $\overleftarrow{\exp}$ defined in Appendix 3.A. The operator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,t_0)$ is the unique solution of the initial matrix value problem $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{\kappa}(t)\boldsymbol{X}(t)$, with initial condition $\boldsymbol{X}(t_0) = \mathbb{1}$ the identity matrix in the state space. Notice that the xy-component of the propagator $\left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,t_0)\right]_{xy}$ is nothing but the biased transition probability $\tilde{G}(x,t \mid y,t_0)$ mentioned in Eq. (2.50). In Appendix 3.A, we sum up the main definitions and properties of the time-ordered exponential.

3.1.2 Setting the problem of conditioning

The LDF $I(\boldsymbol{a})$ of Eq. (3.2) contains the information on the typical values of $\boldsymbol{A}_{nT}$ for the dynamics generated by $\boldsymbol{k}$ as well as its rare fluctuations. When the evolution of the dynamics is constraint-free, the observable $\boldsymbol{A}_{nT}$ converges in probability to its typical values with equal probability. Now, one can be interested in a rare fluctuation of $\boldsymbol{A}_{nT}$. In Refs. [52, 80, 81], the following problem has been addressed: *given the process $z(t)$ conditioned on a rare fluctuation of the observable $\{z(t) \mid \boldsymbol{A}_{nT}[z] = \boldsymbol{a}, \boldsymbol{a} \in \mathbb{R}^2\}$, is there a Markov process which is equivalent to the conditioned process in the long-time limit?* The answer is yes and such process is given by the driven process which has been derived in the case of time-homogeneous processes [80,81]. In the following, we define the driven process for our periodically driven process following two approaches: first using path probabilities, and second from a variational perspective.

3.1.3 Spectral properties of the one-period propagator

In this section, we investigate the spectral properties of the one-period propagator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ and relate them to the generating functions of our observable. We assume from now on that the final time is always nT and omit the subscript nT in the observable $\boldsymbol{A}[z] \equiv \boldsymbol{A}_{nT}[z]$.

The operator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ is positive in the sense of Def. 1 as it is built from the average value of an exponential (2.50). Then, from the Perron-Frobenius theorem (see Appendix 3.B), $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ admits a unique dominant eigenvalue χ_T. Let $\boldsymbol{r}_T$ be its associated

right (column) eigenvector and $\boldsymbol{l}_T$ its associated left (row) eigenvector:

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)\,\boldsymbol{r}_T = \chi_T\,\boldsymbol{r}_T, \tag{3.10}$$

$$\boldsymbol{l}_T\,\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0) = \chi_T\,\boldsymbol{l}_T. \tag{3.11}$$

The eigenvectors $\boldsymbol{r}_T$ and $\boldsymbol{l}_T$ can be chosen up to a multiplicative constant that we set by imposing

$$\mathbf{1}\cdot\boldsymbol{r}_T = 1, \tag{3.12}$$

$$\boldsymbol{r}_T\cdot\boldsymbol{l}_T = 1, \tag{3.13}$$

where $\mathbf{1}$ is the vector whose components are all equal to 1. These spectral elements are found to be connected to the large deviations of $\boldsymbol{A}$, as shown in the following. From Eqs. (2.40, 2.43, 3.8), Property 5 of Appendix 3.A and using the periodicity of $\boldsymbol{\kappa}$, the generating functions $\tilde{G}_x$ and G at time nT reads

$$\tilde{G}_x(nT,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\delta_{x,z(nT)}\right] = \sum_y \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n\right]_{xy}\pi_y(0), \tag{3.14}$$

$$G(nT,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right] = \mathbf{1}\cdot\left(\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n\boldsymbol{\pi}(0)\right) = \sum_{x,y}\left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n\right]_{xy}\pi_y(0). \tag{3.15}$$

The asymptotic expansion of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(nT,0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n$ at large n is dominated by the contribution of its largest eigenvalue:

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n \underset{n\to\infty}{\simeq} (\chi_T)^n\,\boldsymbol{r}_T\,\boldsymbol{l}_T. \tag{3.16}$$

With Eqs. (3.12, 3.13, 3.15), it yields

$$\Gamma(\boldsymbol{\gamma}) = \lim_{n\to\infty}\frac{1}{nT}\ln\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right] = \frac{1}{T}\ln\chi_T. \tag{3.17}$$

Hence, the SCGF Γ is proportional to the logarithm of the largest eigenvalue of the one-period propagator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ [107, 228]. We remind the importance of SCGFs in the study of large deviations of observable as they are related to LDFs via the Gärtner-Ellis theorem (1.27). In our problem, the LDF of $\boldsymbol{A}$ is generally difficult to compute while the SCGF appears to be simply the dominant eigenvalue of the one-period propagator which can generally be obtained at least numerically. Similarly, combining Eqs. (3.14–3.16) and using Eqs. (3.12, 3.13), we find

$$\lim_{n\to\infty}\mathrm{e}^{-nT\Gamma}\mathbb{E}_{x_0}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right] = (l_T)_{x_0}, \tag{3.18}$$

$$\lim_{n\to\infty}\frac{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\delta_{z(nT),x}\right]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right]} = (r_T)_x, \tag{3.19}$$

where $\mathbb{E}_{x_0}$ is the path average over $[z]$ with initial probability $\pi_y(0) = \delta_{y,x_0}$, $\forall y$. Hence, Eqs. (3.18, 3.19) allow writing the eigenvectors of the propagator in terms of path averages. These results are similar to the one stated in Ref. [80] with the difference that the eigenvectors are now defined with respect to the one-period propagator whereas in the time-homogeneous case they correspond to the eigenvectors of the biased generator $\boldsymbol{\kappa}$. Indeed, for a time-homogeneous process, the biased generator is time-independent and the operator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ simplifies to $\mathrm{e}^{T\boldsymbol{\kappa}}$ which has the same eigenvectors as $\boldsymbol{\kappa}$.

3.1.4 Generalized Doob transform

Let $\boldsymbol{M}$ be the generator of a Markov process, $\boldsymbol{v}$ a vector whose elements are strictly positive and $\boldsymbol{f}$ an arbitrary vector function. The *generalized Doob transform* $\boldsymbol{M}^{\boldsymbol{v},f}$ of $\boldsymbol{M}$ associated with $\boldsymbol{v}$ and $\boldsymbol{f}$ is defined by

$$\boldsymbol{M}^{\boldsymbol{v},f} \equiv \mathcal{D}(\boldsymbol{v})\boldsymbol{M}\,\mathcal{D}(\boldsymbol{v})^{-1} - \mathcal{D}(\boldsymbol{f}), \tag{3.20}$$

where $\mathcal{D}(\boldsymbol{v})$ is the diagonal matrix with the components of $\boldsymbol{v}$ on its diagonal. This definition has been introduced in Refs. [52,80,229], generalizing the Doob transform initially introduced by Doob [83] who considered $\boldsymbol{f}$ and $\boldsymbol{v}$ satisfying [80]

$$f_x(t) = \alpha, \quad \forall x \text{ and } t, \tag{3.21}$$

$$\boldsymbol{v}\boldsymbol{M} \leq \alpha \boldsymbol{v}, \tag{3.22}$$

with α a constant. The generator $\boldsymbol{M}^{\boldsymbol{v},f}$ in Eq. (3.20) is norm-conserving if and only if $\boldsymbol{f} = \mathcal{D}(\boldsymbol{v})^{-1}\mathcal{D}(\boldsymbol{v}\boldsymbol{M})$ so that $\sum_x M^{\boldsymbol{v},\boldsymbol{f}}_{xy} = 0$. In the following, we always consider this choice of $\boldsymbol{f}$:

$$\boldsymbol{M}^{\boldsymbol{v}} \equiv \mathcal{D}(\boldsymbol{v})\boldsymbol{M}\,\mathcal{D}(\boldsymbol{v})^{-1} - \mathcal{D}(\boldsymbol{v})^{-1}\mathcal{D}(\boldsymbol{v}\boldsymbol{M}), \tag{3.23}$$

and we call Eq. (3.23) simply a Doob transform by abuse of language. Componentwise, it reads

$$M^{\boldsymbol{v}}_{xy}(t) = v_x(t)M_{xy}(t)v_y^{-1}(t) - v_x^{-1}(t)(\boldsymbol{v}\boldsymbol{M})_x(t)\delta_{xy}. \tag{3.24}$$

The Doob transform is then a tool to build a norm-conserving generator out of an arbitrary one. Notice that if α is a state-independent function, $\boldsymbol{M}^{\alpha\boldsymbol{v}} = \boldsymbol{M}^{\boldsymbol{v}}$. The path probability associated with the Doob transform $\boldsymbol{M}^{\boldsymbol{v}}$ is given by [229]

$$\begin{aligned}\mathbb{P}_{\boldsymbol{M}^{\boldsymbol{v}},\boldsymbol{\pi}(0)}[z] =& \mathbb{P}_{\boldsymbol{M},\boldsymbol{\pi}(0)}[z]\, v_{z_0}^{-1}(0)v_{z(nT)}(nT) \\ & \exp\left[-\int_0^{nT}\left(v_{z(t)}^{-1}(t)(\boldsymbol{v}\boldsymbol{M})_{z(t)}(t) + v_{z(t)}^{-1}(t)\left.\frac{\mathrm{d}\boldsymbol{v}(t)}{\mathrm{d}t}\right|_{z(t)}\right)\mathrm{d}t\right].\end{aligned} \tag{3.25}$$

Eq. (3.25) is obtained by replacing $\boldsymbol{k}$ by $\boldsymbol{M}^{\boldsymbol{v}}$ in Eq. (2.13) and where we used in the second term of the integrand the equality $\ln v_{z_i}(t_{i+1}) - \ln v_{z_i}(t_i) = \int_{t_i}^{t_{i+1}} \mathrm{d}t\, \left.\frac{\mathrm{d}\ln \boldsymbol{v}}{\mathrm{d}t}\right|_{z(t)}$ for each interval of time $[t_i, t_{i+1}]$ between two jumps.

3.1.5 The driven process as the limit of the canonical process

Canonical generator

It has been show in Ref. [80] that the canonical process defined by its path probability (2.89) from the exponential tilting of a time-homogeneous reference process is a norm-conserving Markov process which is not time-homogeneous. We will see in this section that the same result holds for periodically driven processes, namely, the canonical process defined from the exponential tilting of our periodically driven reference process is generated by a norm-conserving Markov generator that is time-dependent but not necessarily time-periodic. To do so, we look for a generator $\boldsymbol{\mathcal{K}}^n$ that satisfies

$$\mathbb{P}_{\boldsymbol{\mathcal{K}}^n,\boldsymbol{\pi}'(0)}[z] \equiv \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z], \tag{3.26}$$

where $\boldsymbol{\pi}'(0)$ is an initial probability that may be different from $\boldsymbol{\pi}(0)$ and where we recall the expression of the canonical path probability

$$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] = \frac{\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right]} = \frac{\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]}{\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right]}. \tag{3.27}$$

We used Eq. (2.51) in the last equality. We want $\boldsymbol{\mathcal{K}}^n$ to generate a norm-conserving process, hence we look for a vector $\boldsymbol{C}^n \equiv \boldsymbol{C}^n(t)$ such that $\boldsymbol{\mathcal{K}}^n$ is built from the Doob transform

$$\boldsymbol{\mathcal{K}}^n \equiv \boldsymbol{\kappa}^{\boldsymbol{C}^n} \equiv \mathcal{D}(\boldsymbol{C}^n)\boldsymbol{\kappa}\,\mathcal{D}(\boldsymbol{C}^n)^{-1} - \mathcal{D}(\boldsymbol{C}^n)^{-1}\mathcal{D}(\boldsymbol{C}^n\boldsymbol{\kappa}). \tag{3.28}$$

From Eq. (3.25), the path probability generated by $\boldsymbol{\mathcal{K}}^n$ reads

$$\mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]\, C^n_{z_0}(0)^{-1}\, C^n_{z(nT)}(nT) \\ \exp\left[-\int_0^{nT}\left(C^n_{z(t)}(t)^{-1}(\boldsymbol{C}^n\boldsymbol{\kappa})_{z(t)}(t) + C^n_{z(t)}(t)^{-1}\left.\frac{\mathrm{d}\boldsymbol{C}^n(t)}{\mathrm{d}t}\right|_{z(t)}\right)\mathrm{d}t\right]. \tag{3.29}$$

For this path probability to coincide with Eq. (3.27), $\boldsymbol{C}^n$ should be chosen so that the time-extensive terms of Eq. (3.29) are all equal to 1. This condition is satisfied for

$$\boldsymbol{C}^n(t) = \mathbf{1}\,\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(nT,t) \tag{3.30}$$

solution of (cf. property 3 of Appendix 3.A)

$$\begin{cases}\dot{\boldsymbol{C}}^n = -\boldsymbol{C}^n\boldsymbol{\kappa},\\ \boldsymbol{C}^n(nT) = \mathbf{1},\end{cases} \tag{3.31}$$

leading to

$$\mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]\,(C^n_{z_0})^{-1}(0), \tag{3.32}$$

or equivalently

$$\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\,\boldsymbol{\pi}(0)\odot\boldsymbol{C}^n(0)}[z], \tag{3.33}$$

where $\odot$ is the Hadamard product: $(\boldsymbol{u}\odot\boldsymbol{v})_x \equiv u_x v_x$. From Eqs. (3.15) and (3.30), we remark that the generating function can be expressed in terms of $\boldsymbol{C}^n$ as

$$\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right] = \boldsymbol{C}^n(0)\cdot\boldsymbol{\pi}(0), \tag{3.34}$$

implying from Eqs. (3.27) and (3.33)

$$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\frac{\boldsymbol{C}^n(0)\odot\boldsymbol{\pi}(0)}{\boldsymbol{\pi}(0)\cdot\boldsymbol{C}^n(0)}}[z]. \tag{3.35}$$

In other words, the canonical path probability is associated with the time-dependent generator $\boldsymbol{\mathcal{K}}^n = \boldsymbol{\kappa}^{\boldsymbol{C}^n}$ with initial probability $\boldsymbol{\pi}'(0) = \frac{\boldsymbol{C}^n(0)\odot\boldsymbol{\pi}(0)}{\boldsymbol{C}^n(0)\cdot\boldsymbol{\pi}(0)}$. It means that the canonical process has a corresponding Markov generator which is norm-conserving contrary to the biased generator. This shows the importance of the normalization factor $\mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right]$ dividing the biased path probability $\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]$ in order to define the canonical path probability $\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$. This normalization at the level of paths translates at the level of generators into a *rectification* of the biased generator by a Doob transform to build a norm-conserving canonical generator. Hence, the canonical process can be seen as a "rectified biased process", and the biased process as an intermediary process used for mathematical calculations. Notice that the canonical generator $\boldsymbol{\mathcal{K}}^n$ depends explicitly on the number of periods n.

Driven process

Let us come back to our original problem, namely finding an equivalent Markov process which is asymptotically equivalent to the process conditioned on a rare fluctuation of the observable. This conditioned process is represented by the microcanonical path ensemble (2.88):

$$\mathbb{P}^{\text{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \equiv \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}\left[z \mid \boldsymbol{A}_{nT}[z] = \mathbf{a}\right]. \tag{3.36}$$

As discussed in Sec. 2.5.3, the canonical process is equivalent to the microcanonical process for specific values of $\boldsymbol{\gamma}$ under convexity conditions on $I(\boldsymbol{a})$, giving a first answer to the question of finding an equivalent Markov process of the conditioned process in the long-time limit. In this section, we make a step further by focusing on the asymptotic dynamics of the canonical process in the limit $n \to \infty$ and by considering the process towards which the canonical process converges at long time. This process is named the "driven process" in Refs. [80, 81].

The driven process is defined as the limit of the canonical process as $n \to \infty$. Since the canonical process is defined from the Doob transform of the tilted operator using the vector $\boldsymbol{C}^n$, the driven process will be built similarly. To determine the vector involved in this Doob transform, let us write the asymptotic expansion of $\boldsymbol{C}^n$ in the limit $n \to \infty$. Using Eqs. (3.12, 3.16, 3.30), the periodicity of $\boldsymbol{\kappa}$ and Property 5, we find that $\boldsymbol{C}^n(\tau)$ for $\tau \in [0, T[$ reads

$$\boldsymbol{C}^n(\tau) = \mathbf{1}\ \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)^n \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1} \underset{n\to\infty}{\sim} (\chi_T)^n \boldsymbol{l}_T \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1}. \tag{3.37}$$

The idea is to define the driven generator as the Doob transform of the biased matrix using a vector $\boldsymbol{l}(\tau)$ corresponding to the limit of $\boldsymbol{C}^n(\tau)$ when $n \to \infty$, but since scalar constants play no role in the Doob transform, it suffices to consider $\boldsymbol{l} \equiv \boldsymbol{l}(\tau)$ defined for all $\tau \in [0, T[$ by

$$\boldsymbol{l}(\tau) \equiv \boldsymbol{l}_T \left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(\tau,0)\right]^{-1}, \tag{3.38}$$

that is by construction the solution of

$$\begin{cases} \dot{\boldsymbol{l}} = -\boldsymbol{l}\boldsymbol{\kappa}, \\ \boldsymbol{l}(0) = \boldsymbol{l}_T. \end{cases} \tag{3.39}$$

For times greater than T, we notice from Eq. (3.11) and the periodicity of $\boldsymbol{\kappa}$ that the vector $\boldsymbol{l}$ satisfies

$$\boldsymbol{l}(\tau + T) = \chi_T^{-1}\, \boldsymbol{l}(\tau). \tag{3.40}$$

We now define the Markov generator $\boldsymbol{K} \equiv \boldsymbol{K}(\tau, \boldsymbol{\gamma})$ of the driven process at any time τ by the Doob transform of the tilted matrix $\boldsymbol{\kappa}$ associated with the vector $\boldsymbol{l}$:

$$\boldsymbol{K} \equiv \boldsymbol{\kappa}^{\boldsymbol{l}} = \mathcal{D}(\boldsymbol{l})\boldsymbol{\kappa}\,\mathcal{D}(\boldsymbol{l})^{-1} - \mathcal{D}(\boldsymbol{l})^{-1}\mathcal{D}(\boldsymbol{l}\boldsymbol{\kappa}). \tag{3.41}$$

Note that the positivity of $\boldsymbol{l}$ at all time is ensured by the positivity of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t,0)$ and the positivity of $\boldsymbol{l}_T$ guaranteed by the PF theorem. By construction, the generator of the driven process is the limit of the canonical transition matrix when $n \to \infty$ as can be shown from Eqs. (3.28) and (3.37):

$$\lim_{n\to\infty} \boldsymbol{\mathcal{K}}^n(\tau, \boldsymbol{\gamma}) = \boldsymbol{\kappa}^{\boldsymbol{l}}(\tau, \boldsymbol{\gamma}) = \boldsymbol{K}(\tau, \boldsymbol{\gamma}). \tag{3.42}$$

These results are consistent with Ref. [80] when considering time-homogeneous processes, i.e. with time-independent generators $\boldsymbol{k}(t) = \boldsymbol{k}$ and $\boldsymbol{\kappa}(t) = \boldsymbol{\kappa}$. In this case, $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)$ simplifies to $\mathrm{e}^{T\boldsymbol{\kappa}}$, and its dominant left eigenvector $\boldsymbol{l}_\star \equiv \boldsymbol{l}_T$ of eigenvalue χ_T becomes T-independent. $\boldsymbol{l}_\star$ is also the dominant left eigenvector of $\boldsymbol{\kappa}$ with eigenvalue $\upsilon \equiv \frac{1}{T}\ln\chi_T$. The function $\boldsymbol{l}(\tau)$ in Eq. (3.38) becomes $\boldsymbol{l}(\tau) = \boldsymbol{l}_\star \mathrm{e}^{-\tau\boldsymbol{\kappa}} = \mathrm{e}^{-\tau\upsilon}\boldsymbol{l}_\star$. Since state-independent functions play no role in the Doob transform, the driven process is the Doob transform of the tilted matrix with respect to $\boldsymbol{l}_\star$, recovering the results of Refs. [52, 80]. We point out that the generators considered in these references are the adjoints of the generators considered in this manuscript. Hence, their right eigenvectors correspond to our left eigenvectors and vice versa.

Let us now discuss some properties of the driven generator. One interesting property of $\boldsymbol{K}$ is its periodicity (contrary to the canonical generator). Indeed, from Eq. (3.40) and the periodicity of $\boldsymbol{\kappa}$, it follows from Eq. (3.41) that $\forall\tau \in [0,T]$:

$$\boldsymbol{K}(\tau + T, \boldsymbol{\gamma}) = \boldsymbol{K}(\tau, \boldsymbol{\gamma}). \tag{3.43}$$

We remark that the TiPS probability for the driven process can be obtained from the solution $\boldsymbol{l}$ of the initial value problem of Eq. (3.39) and from the solution $\boldsymbol{r} \equiv \boldsymbol{r}(\tau)$ of the initial value problem:

$$\begin{cases} \dot{\boldsymbol{r}} = \boldsymbol{\kappa}\boldsymbol{r}, \\ \boldsymbol{r}(0) = \boldsymbol{r}_T, \end{cases} \tag{3.44}$$

or alternatively

$$\boldsymbol{r}(t) \equiv \overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t,0)\boldsymbol{r}_T. \tag{3.45}$$

Using Eq. (3.10) and the periodicity of $\boldsymbol{\kappa}$, the vector $\boldsymbol{r}$ satisfies:

$$\boldsymbol{r}(t+T) = \chi_T \boldsymbol{r}(t). \tag{3.46}$$

The TiPS probability of the driven process $\boldsymbol{\mu} \equiv \boldsymbol{\mu}(t)$, defined as the T-periodic solution of the master equation:

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \boldsymbol{K}\boldsymbol{\mu} \\ \boldsymbol{\mu}(0) = \boldsymbol{\mu}(T), \end{cases} \tag{3.47}$$

is expressed in terms of the vectors $\boldsymbol{l}$ and $\boldsymbol{r}$

$$\boldsymbol{\mu}(t) = \boldsymbol{l}(t) \odot \boldsymbol{r}(t). \tag{3.48}$$

Indeed, Eqs. (3.39, 3.41, 3.44) yield for any state x

$$\begin{aligned} \sum_y K_{xy}(l_y r_y) &= \sum_y \left\{ l_x \kappa_{xy} l_y^{-1} l_y r_y - l_x^{-1} l_y \kappa_{yx} l_x r_x \right\} && (3.49) \\ &= l_x \dot{r}_x + \dot{l}_x r_x && (3.50) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(l_x r_x), && (3.51) \end{aligned}$$

while Eqs. (3.40, 3.46) lead to

$$\boldsymbol{l}(0) \odot \boldsymbol{r}(0) = \boldsymbol{l}(T) \odot \boldsymbol{r}(T), \tag{3.52}$$

which proves that $\boldsymbol{\mu}$ is the solution of Eq. (3.47). Notice that our normalization choice in Eq. (3.13) ensures the normalization of $\boldsymbol{\mu}(0)$.

We saw in Eq. (3.42) that the driven generator is the limit of the canonical generator. This property translates at the path probability level into the fact that the driven and canonical path probabilities are asymptotically equivalent in the sense of (2.91). Using Eqs. (3.17, 3.39–3.41) in Eq. (3.25), we obtain the path probability of the driven process:

$$\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z] = \mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]\, l_{z_{nT}}(0)\, \mathrm{e}^{-nT\Gamma}\, l_{z_0}^{-1}(0). \tag{3.53}$$

Using the definitions of the canonical path probability (3.27) and driven path probability (3.53), we obtain

$$\frac{\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]} = l_{z_{nT}}(0)\, \mathrm{e}^{-nT\Gamma}\, l_{z_0}^{-1}(0)\, \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\right]. \tag{3.54}$$

Finally, using the definition of the SCGF (3.3), we find:

$$\lim_{n\to\infty} \frac{1}{nT} \ln \frac{\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]} = 0, \tag{3.55}$$

showing the logarithmic equivalence between the driven path probability and the canonical path probability for any $\boldsymbol{\gamma}$:

$$\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z] \underset{n\to\infty}{\asymp} \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]. \tag{3.56}$$

Effective process

Given the equivalence between the driven process and the canonical process, we can finally answer our initial question about finding an equivalent process of the microcanonical process conditioned on $\boldsymbol{A} = \boldsymbol{a}$. Let us call this process the *effective process*. As discussed in Sec. 2.5.3, the canonical path probability (for a specific value of $\boldsymbol{\gamma}$) and microcanonical path probability (for $\boldsymbol{A}$ conditioned on $\boldsymbol{a}$) are logarithmically equivalent if the LDF I is strictly convex at $\boldsymbol{a}$ [195]. In this case, and assuming that I is differentiable for simplicity, the equivalence holds for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$, where $\nabla I(\boldsymbol{a})$ is the gradient of I evaluated at $\boldsymbol{a}$. Mathematically, this reads:

$$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \underset{n\to\infty}{\asymp} \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]\,\big|_{\boldsymbol{\gamma}=\nabla I(\boldsymbol{a})}\,. \tag{3.57}$$

When combined with the logarithmic equivalence between the driven and canonical path probabilities of Eq. (3.56), we find that the effective process is the driven process for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$. Mathematically, this reads:

$$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z] \underset{n\to\infty}{\asymp} \mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]\,\big|_{\boldsymbol{\gamma}=\nabla I(\boldsymbol{a})}\,. \tag{3.58}$$

In other words, when $I(\boldsymbol{a})$ is strictly convex at $\boldsymbol{a}$, the effective process is the Markov process of generator $\boldsymbol{K}(\boldsymbol{\gamma} = \nabla I(\boldsymbol{a}))$. When the convexity is not strict, the same value $\boldsymbol{\gamma}$ corresponds to several values $\boldsymbol{a}$ and the SCGF is not differentiable at $\boldsymbol{\gamma}$. Actually, the first derivative of the SCGF at $\boldsymbol{\gamma}$ is discontinuous, implying that the system undergoes a first-order phase transition [230, 231]. This case corresponds physically to a phase coexistence of all the values $\boldsymbol{a}$ associated with the value $\boldsymbol{\gamma}$ [224]. If I is not convex at $\boldsymbol{a}$, there is no Markov process equivalent to the microcanonical process.

3.1.6 Variational approach

The results derived previously can be obtained and understood from an optimisation problem under constraint of LDFs. This variational approach has been investigated for time-homogeneous processes in Refs. [81, 232, 233]. Here, we derive the generator of the driven process as the process minimizing a functional playing the role of entropy in Jaynes' maximum entropy principle of statistical mechanics. This functional is the level 2.5 LDF for empirical occupations and transition probabilities. We find that the driven process is the "most probable" process for which the observable $\boldsymbol{A}$ takes asymptotically a chosen value.

Level 2.5 large deviation function

When dealing with large deviation theory, we can consider different level of descriptions of LDFs with decreasing degree of details [60]. For our Markov jump process, we have:

Level 3: It is the level that describes the probability of the *empirical trajectory occupation* $P(\{x_i, t_i\})[z]$ which measures the fraction of time the system occupies successively the states x_i during the time-intervals $[t_i, t_{i+1}[$ along the path $[z]$ of duration nT. For the stochastic process generated by $\boldsymbol{k}$, the typical behaviour of the empirical trajectory occupation $P(\{x_i, t_i\})[z]$ is given by the path probability $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ (2.13).

Level 2: It is the level that describes the probability of the *empirical occupation* defined for any $\tau \in [0, T[$ by

$$p_x^n[z](\tau) = \frac{1}{n}\sum_{\ell=0}^{n-1} \delta_{x,z(\tau+\ell T)} \tag{3.59}$$

and which represents the percentage of periods in which the system has been in state x at time τ of each period along the path $[z]$ of duration nT.

Level 1: It is the level that describes the probability of observables of the form

$$\boldsymbol{B}_{nT}[z] = \frac{1}{nT}\int_0^{nT} h_{z(t)}(t)\mathrm{d}t. \tag{3.60}$$

Since $\boldsymbol{B}_{nT}$ can be rewritten in terms of the empirical occupation

$$\boldsymbol{B}_{nT}[z] = \frac{1}{T}\int_0^{T} \sum_x p_x^n[z](t) h_x(t)\mathrm{d}t, \tag{3.61}$$

the level 1 LDF can be obtained from the level 2 LDF by the contraction principle (1.35)

$$I_1(\boldsymbol{b}) = \inf_{\boldsymbol{p}|\frac{1}{T}\int_0^T \boldsymbol{p}(t)\cdot\boldsymbol{h}(t)=\boldsymbol{b}} I_2(\boldsymbol{p}). \tag{3.62}$$

For Markov jump processes, it is not possible to derive explicitly the level 2 LDF. However, there exists a more detailed level of description for which the LDF is explicit and which allows describing the fluctuations of observables $\boldsymbol{A}_{nT}$ of the form (2.14): it's the level 2.5 which is the less detailed level such that the LDF is explicit. It describes the large

deviations of the empirical occupation defined in Eq. (3.59) and the *empirical transition probability* $\omega^n_{xy}(\tau)[z]$ defined for each $\tau \in [0, T[$ by

$$\omega^n_{xy}[z](\tau) = \frac{1}{n}\sum_{\ell=0}^{n-1} \frac{1}{\mathrm{d}\tau} \sum_{s\in[\tau,\tau+\mathrm{d}\tau[} \delta_{y,z(s^-+\ell T)}\delta_{x,z(s^++\ell T)}, \tag{3.63}$$

with $\mathrm{d}\tau$ an infinitesimal time. The quantity $n\,\mathrm{d}\tau\,\omega^n_{xy}[z](\tau)$ measures the number of periods in which the transition $y \to x$ occurs at time τ (or more precisely between times τ and $\tau+\mathrm{d}\tau$) of each period along the path $[z]$ of duration nT. This quantity is related to $np^n_x[z]$ as follows

$$np^n_x[z](\tau+\mathrm{d}\tau) - np^n_x[z](\tau) = \sum_{y\neq x} \left[n\,\mathrm{d}\tau\,\omega^n_{xy}[z](\tau) - n\,\mathrm{d}\tau\,\omega^n_{yx}[z](\tau)\right], \tag{3.64}$$

meaning that the variation between times τ and $\tau + \mathrm{d}t$ of each period of the number of periods the system has occupied the state x has two contributions: an additive contribution coming from the transitions having x as arrival state, and a subtractive contribution coming from the transitions having x as departure state. It leads to the continuity equation

$$\frac{\partial p^n_x[z]}{\partial \tau} = \sum_{y\neq x} \left[\omega^n_{xy}[z] - \omega^n_{yx}[z]\right]. \tag{3.65}$$

The empirical occupation and the empirical transition probability are random variables since they are functional of the paths and depend on the number of periods n considered during the total observation time nT. In the limit of large n, the probability to observe the empirical occupation $\boldsymbol{p}^n[z] = \boldsymbol{p}$ and the empirical transition probability $\boldsymbol{\omega}^n[z] = \boldsymbol{\omega}$ satisfies a large deviation principle:

$$P_n(\boldsymbol{\omega}, \boldsymbol{p}) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nTI_{2.5}(\boldsymbol{\omega},\boldsymbol{p})}, \tag{3.66}$$

where the 2.5 LDF is given by [234]

$$I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) = \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_{y,x\neq y} \left[p_y(\tau) \left(k_{xy}(\tau) - \frac{\omega_{xy}(\tau)}{p_y(\tau)}\right) + \omega_{xy}(\tau) \ln \frac{\omega_{xy}(\tau)}{k_{xy}(\tau)p_y(\tau)}\right], \tag{3.67}$$

see Appendix 3.C for an heuristic derivation. This expression holds only for $\boldsymbol{p}(0) = \boldsymbol{p}(T)$ and $\boldsymbol{\omega}(0) = \boldsymbol{\omega}(T)$, for conservative transition probabilities $\dot{p}_x(\tau) = \sum_{y\neq x}[\omega_{xy}(\tau) - \omega_{yx}(\tau)]$, $\forall x$, and normalized occupations $\sum_y p_y(\tau) = 1$, $\forall y$; otherwise I is infinite. Notice that $I(\boldsymbol{\omega}, \boldsymbol{p})$ vanishes for $\boldsymbol{p} = \boldsymbol{\pi}^{\mathrm{TiPS}}$ and $\boldsymbol{\omega} = \boldsymbol{k} \odot \boldsymbol{\pi}^{\mathrm{TiPS}}$. It means that when we consider paths generated by the original process $\boldsymbol{k}$ without any conditioning, the random variable $\boldsymbol{p}^n[z]$ converges in probability to $\boldsymbol{\pi}^{\mathrm{TiPS}}$ in the limit $n \to \infty$ while the random variable $\boldsymbol{\omega}^n$ converges to $\boldsymbol{k} \odot \boldsymbol{\pi}^{\mathrm{TiPS}}$, with $\left[\boldsymbol{k} \odot \boldsymbol{\pi}^{\mathrm{TiPS}}\right]_{xy} \equiv k_{xy}\pi^{\mathrm{TiPS}}_y$.

The level 2.5 LDF is important to study the large deviations of the observable $\boldsymbol{A}$ since the LDF $I(a)$ can be obtained from the 2.5 LDF $I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p})$ using the contraction principle. Indeed, we can rewrite the conditioning observable $\boldsymbol{A}$ in terms of $\boldsymbol{p}^n[z]$ and $\boldsymbol{\omega}^n[z]$ using the periodicity of $\boldsymbol{g}$ and $\boldsymbol{h}$:

$$\boldsymbol{A}\left(\boldsymbol{\omega}^n[z], \boldsymbol{p}^n[z]\right) = \begin{pmatrix} A_1(\boldsymbol{\omega}^n[z]) \\ A_2(\boldsymbol{p}^n[z]) \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_{x,y\neq x} \omega^n_{xy}[z](\tau)\; g_{xy}(\tau) \\ \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_x p^n_x[z](\tau)\; h_x(\tau) \end{pmatrix}. \tag{3.68}$$

Without conditioning, $\boldsymbol{A}$ converges in probability to $\boldsymbol{A}(\boldsymbol{k} \odot \boldsymbol{\pi}^{\mathrm{TiPS}}, \boldsymbol{\pi}^{\mathrm{TiPS}})$ as $n \to \infty$. We are now interested in conditioning our process on the event $\{\boldsymbol{A}(\boldsymbol{\omega}^n[z], \boldsymbol{p}^n[z]) = \boldsymbol{a} \mid \boldsymbol{a} \in \mathbb{R}^2\}$ with $\boldsymbol{a}$ a rare fluctuation. As in Section 3.1.2, we look for the Markov process for which the conditioning value $\boldsymbol{a}$ becomes a typical value of $\boldsymbol{A}$. Here, it amounts to looking for the most probable (or rather least improbable) pair $(\boldsymbol{\omega}, \boldsymbol{p})$ compatible with $\boldsymbol{A}(\boldsymbol{\omega}, \boldsymbol{p}) = \boldsymbol{a}$. It is obtained by minimizing the 2.5 LDF under the following constraints

- **C0:** $\boldsymbol{A}(\boldsymbol{\omega}, \boldsymbol{p}) = \boldsymbol{a}$,
- **C1:** $\sum_y p_y(\tau) = 1, \ \forall \tau \in [0, T]$,
- **C2:** $\dot{p}_x(\tau) = \sum_y \left[\omega_{xy}(\tau) - \omega_{yx}(\tau)\right], \ \forall x$ and $\forall \tau \in [0, T]$,
- **C3:** $\boldsymbol{p}(T) = \boldsymbol{p}(0)$,
- **C4:** $\boldsymbol{\omega}(T) = \boldsymbol{\omega}(0)$.

The LDF of $\boldsymbol{A}$ is then obtained from the contraction principle under the above constraints:

$$I(\boldsymbol{a}) = \inf_{\boldsymbol{p},\boldsymbol{\omega} | \{\mathbf{Ci}\}_{i=0}^4} \Big\{ I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) \Big\}. \tag{3.69}$$

Intuitively, the optimizer $(\boldsymbol{\omega}_{\boldsymbol{a}}, \boldsymbol{p}_{\boldsymbol{a}})$ of Eq. (3.69) is expected to be the typical value of $(\boldsymbol{\omega}^n[z], \boldsymbol{p}^n[z])$ under the dynamics associated with the effective process introduced in Section 3.1.5 and which is equivalent to the microcanonical process as $n \to \infty$. Reciprocally, we can instead consider a dual approach and compute the SCGF defined in Eq. (3.3) and given by the LF transform of the LDF $I(a)$:

$$\Gamma(\boldsymbol{\gamma}) = \sup_{\boldsymbol{p},\boldsymbol{\omega} | \{\mathbf{Ci}\}_{i=1}^4} \left\{ \boldsymbol{\gamma} \cdot \boldsymbol{A}(\boldsymbol{\omega}, \boldsymbol{p}) - I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) \right\}. \tag{3.70}$$

The solution $(\boldsymbol{\omega}_{\boldsymbol{\gamma}}, \boldsymbol{p}_{\boldsymbol{\gamma}})$ of Eq. (3.70) is the typical value of $(\boldsymbol{\omega}^n[z], \boldsymbol{p}^n[z])$ under the canonical path probability [81] and is expected to be associated with the generator of the driven process as we will show. For strictly convex LDF $I(\boldsymbol{a})$, Eqs. (3.69) and (3.70) have the same solutions. More precisely, if I is strictly convex at $\boldsymbol{a}$, $(\boldsymbol{\omega}_{\boldsymbol{a}}, \boldsymbol{p}_{\boldsymbol{a}}) = (\boldsymbol{\omega}_{\boldsymbol{\gamma}}, \boldsymbol{p}_{\boldsymbol{\gamma}})$ for $\boldsymbol{\gamma} = \boldsymbol{\nabla} I(a)$ in agreement with the equivalence between the microcanonical process and the effective process/driven process for $\boldsymbol{\gamma} = \boldsymbol{\nabla} I(a)$ [81]. In the following, we recover this result through direct calculation of the optimum of Eq. (3.70). To take into account the constraints in this calculation, we used the method of Lagrange multipliers (see Appendix 3.D for a review) and we look for the optimum of the functional

$$\begin{aligned} \mathcal{F}(\boldsymbol{\omega}, \boldsymbol{p}) = & -I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) + \gamma_1 \ A_1(\boldsymbol{\omega}) + \gamma_2 \ A_2(\boldsymbol{p}) \\ & - \frac{1}{T} \int_0^T \mathrm{d}\tau \ c(\tau) \ \left[\sum_y p_y(\tau) - 1 \right] \\ & - \frac{1}{T} \int_0^T \mathrm{d}\tau \sum_x u_x(\tau) \ \left[\dot{p}_x(\tau) - \sum_y \left(\omega_{xy}(\tau) - \omega_{yx}(\tau)\right) \right], \end{aligned} \tag{3.71}$$

where c and $\boldsymbol{u}$ are time-dependent Lagrange multipliers respectively associated with the constraints **C1** and **C2**. We assume in addition that $\boldsymbol{u}(T) = \boldsymbol{u}(0)$ in line with the

constraints $\boldsymbol{C3}$ and $\boldsymbol{C4}$. Notice that computing the SCGF in Eq. (3.70) is equivalent to minimizing the LDF $I(\boldsymbol{a})$ with γ_1 and γ_2 the Lagrange multipliers ensuring the constraint $\boldsymbol{C0}$. Functional derivatives with respect to empirical occupations and empirical transition probabilities yield

$$\begin{cases} \frac{\partial \mathcal{F}}{\partial \omega_{xy}(\tau)} = 0 \quad \Rightarrow \quad \ln \frac{\omega_{xy}(\tau)}{k_{xy}(\tau) p_y(\tau)} + (u_y(\tau) - u_x(\tau)) - \gamma_1 \, g_{xy}(\tau) = 0 \qquad \text{for } x \neq y, \\ \frac{\partial \mathcal{F}}{\partial p_y(\tau)} = 0 \quad \Rightarrow \quad \sum_{x \neq y} \left[k_{xy}(\tau) - \frac{\omega_{xy}(\tau)}{p_y(\tau)} \right] + c(\tau) - \gamma_2 \, h_y(\tau) \; - \dot{u}_y(\tau) = 0. \end{cases} \tag{3.72}$$

We transform the first equation of (3.72) into

$$\omega_{xy}(\tau) = K'_{xy}(\tau) \; p_y(\tau), \tag{3.73}$$

with

$$K'_{xy}(\tau) \equiv k_{xy}(\tau) \; \mathrm{e}^{\gamma_1 g_{xy}(\tau)} \; \mathrm{e}^{u_x(\tau) - u_y(\tau)} = \kappa_{xy}(\tau) \mathrm{e}^{u_x(\tau) - u_y(\tau)}, \tag{3.74}$$

for $x \neq y$ and $\tau \in [0, T]$. We define the diagonal elements such that the sum over the lines of any column of $\boldsymbol{K'}(\tau)$ vanishes:

$$K'_{yy}(\tau) \equiv - \sum_{x \neq y} K'_{xy}(\tau), \tag{3.75}$$

so that using condition $\mathbf{C2}$, $\boldsymbol{K'}$ satisfies

$$\dot{\boldsymbol{p}} = \boldsymbol{K'} \boldsymbol{p}. \tag{3.76}$$

From condition $\mathbf{C3}$, $\boldsymbol{p}$ is the TiPS probability associated with $\boldsymbol{K'}$. At this point, it remains to fix the Lagrange multipliers c and $\boldsymbol{u}$. As suggested by the notation, $\boldsymbol{K'}$ will turn out to be the generator $\boldsymbol{K}$ of the driven process defined in (3.41). To prove this, we see that the second equation of Eq. (3.72) becomes

$$c(\tau) = \sum_{x \neq y} K'_{xy}(\tau) - \sum_{x \neq y} k_{xy}(\tau) + \gamma_2 \, h_y(\tau) + \dot{u}_y(\tau), \qquad \forall y. \tag{3.77}$$

Using (2.49, 3.74), we get

$$c = \sum_{x \neq y} \kappa_{xy} \mathrm{e}^{u_x - u_y} + \kappa_{xx} + \dot{u}_y, \tag{3.78}$$

where we dropped the τ-dependency for clarity. Multiplying Eq. (3.78) by e^{u_y}, we finally obtain

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} \left(e^{\boldsymbol{u}} \right) = - (e^{\boldsymbol{u}}) \left(\boldsymbol{\kappa} - c \mathbb{1} \right), \\ e^{\boldsymbol{u}(0)} = e^{\boldsymbol{u}(T)}, \end{cases} \tag{3.79}$$

with $(\mathrm{e}^{\boldsymbol{u}})_x \equiv \mathrm{e}^{u_x}$. The formal solution of (3.79) reads:

$$\mathrm{e}^{\boldsymbol{u}(t)} = \mathrm{e}^{\boldsymbol{u}(T)} \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa} - c\mathbb{1}}(T, t) = \mathrm{e}^{\boldsymbol{u}(0)} \mathrm{e}^{- \int_t^T c} \; \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, t), \tag{3.80}$$

where we used Property 6 of Appendix 3.A in the second equality. Taking $t = 0$, we get

$$\mathrm{e}^{\boldsymbol{u}(0)} \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, 0) = \mathrm{e}^{\int_0^T c} \; \mathrm{e}^{\boldsymbol{u}(0)}. \tag{3.81}$$

Hence, the optimization with respect to $\boldsymbol{p}$ (second equation of (3.72)) leads to a spectral equation. Since the vector $\mathrm{e}^{\boldsymbol{u}(0)}$ has positive components, by Perron-Frobenius theorem it is the unique — up to a multiplicative constant — left eigenvector of $\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)$ associated with its largest eigenvalue χ_T, implying

$$\mathrm{e}^{\int_0^T c} = \chi_T. \tag{3.82}$$

From Eq. (3.17), we find that the SCGF is related to the Lagrange multiplier c by

$$\Gamma = \frac{1}{T}\int_0^T c(\tau)\,\mathrm{d}\tau. \tag{3.83}$$

We recover this result in Appendix 3.E by computing directly Eq. (3.70). Notice that we can rewrite Eq. (3.79) as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{\boldsymbol{u}+\int_t^T c}\right) = -(e^{\boldsymbol{u}+\int_t^T c})\boldsymbol{\kappa}, \tag{3.84}$$

where the vector $e^{\boldsymbol{u}+\int_t^T c}$ is solution of

$$\begin{cases} \dot{\boldsymbol{X}} = -\boldsymbol{X}\boldsymbol{\kappa}, \\ \boldsymbol{X}(T) = \chi_T^{-1}\boldsymbol{X}(0). \end{cases} \tag{3.85}$$

From Eq. (3.39), we conclude that the vector $\boldsymbol{l}$ that appears in the Doob transform used to define the driven generator is related to the Lagrange multipliers through

$$\boldsymbol{l}(t) = e^{\boldsymbol{u}(t)+\int_t^T c}. \tag{3.86}$$

We emphasize that $\boldsymbol{u}(t)$ is set up to an additive and time-dependent function constant in the state space (independent of the states). Indeed, if **C2** is satisfied for all states but one then it is satisfied for all states (in view of **C1**). Then, Eq. (3.86) is a choice for this remaining degree of freedom in $\boldsymbol{u}(t)$.

We now show that the transition rate matrix $\boldsymbol{K'}$ generates the driven process, i.e. we show that $\boldsymbol{K'}$ is expressed as the Doob transform of $\boldsymbol{\kappa}$ associated with the vector $\boldsymbol{l}$. Using Eqs. (2.49, 3.77, 3.86), we transform Eqs. (3.74–3.75) $\forall x, y$ into

$$\begin{aligned}
K'_{xy} &= \kappa_{xy}\mathrm{e}^{u_x-u_y} - [\kappa_{xx} - K'_{xx}]\,\delta_{xy}, && (3.87)\\
&= \kappa_{xy}\mathrm{e}^{u_x-u_y} - [\,(k_{xx}+\gamma_2 h_x) + (c - k_{xx} - \gamma_2\, h_x - \dot{u}_x)\,]\,\delta_{xy}, && (3.88)\\
&= \mathrm{e}^{u_x}\kappa_{xy}\mathrm{e}^{-u_y} - \,[c - \dot{u}_x]\,\delta_{xy}, && (3.89)\\
&= l_x\kappa_{xy}l_x^{-1} + l_x^{-1}\dot{l}_x\,\delta_{xy}, && (3.90)\\
&= l_x\kappa_{xy}l_x^{-1} - l_x^{-1}(\boldsymbol{l}\boldsymbol{\kappa})_x\,\delta_{xy}, && (3.91)\\
&= \kappa^{\boldsymbol{l}}_{xy}. && (3.92)
\end{aligned}$$

Hence $\boldsymbol{K'} = \boldsymbol{K}$ (defined in Eq. (3.41)). It follows from Eq. (3.76) and the constraints $\boldsymbol{C3}-\boldsymbol{C4}$ that the optimum of Eq. (3.70) is reached for $\boldsymbol{p} = \boldsymbol{\mu}$ the TiPS probability of the driven process with generator $\boldsymbol{K}$, and $\boldsymbol{\omega} = \boldsymbol{K}\odot\boldsymbol{p}$ the directional probability current associated with the probability $\boldsymbol{p}$ and rate matrix $\boldsymbol{K}$. In the time-homogeneous case, we recover the results of [81, 233].

To conclude, the driven process is the most probable process that generates the dynamics leading to the imposed value of the conditioning observable. In other words, it is the generator of an "optimal" Markov process for which the conditioning observable takes asymptotically the imposed value as a typical value.

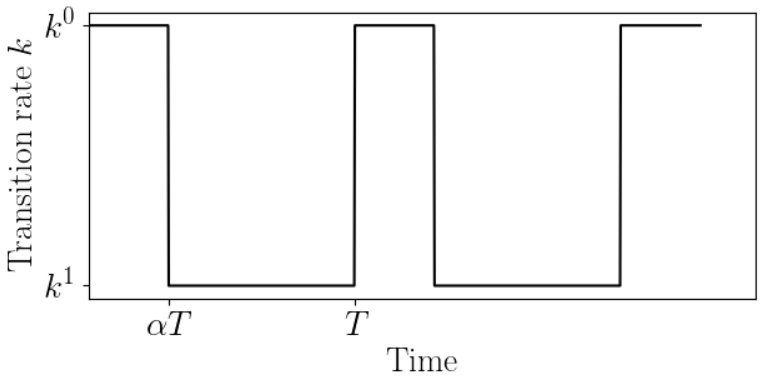

Figure 3.1 – Transition rate $k(t)$ appearing in the generator $\boldsymbol{k}$ of Eq. (3.93).

3.1.7 Applications: modulated two-level system

In this section, we give some intuition about the driven process and the conditioning problem by applying our results on a simple example. We consider a two level system with states denoted by $\mathbf{e}_\pm$. For simplicity, the transition rate matrix is chosen symmetric and piecewise-constant, and we take as conditioning observable a current defined through a time-periodic function. The study of this model is as follows. We compute the SCGF associated with this observable and derive the generators $\boldsymbol{\mathcal{K}}^n$ and $\boldsymbol{K}$ of the canonical and driven processes, respectively. We then study the convergence of the canonical transition rates toward the driven transition rates as the number of periods n grows. Finally, we comment qualitatively on the influence of the conditioning on the transition rates of the driven process.

The transition rate matrix used to model the system reads

$$\boldsymbol{k}(t) = \begin{pmatrix} -k(t) & k(t) \\ k(t) & -k(t) \end{pmatrix}. \tag{3.93}$$

The rate $k(t)$ is a T-periodic and piecewise constant function of time

$$k(t) = \begin{cases} k^0 & \text{for } t \in [0, \alpha T[, \\ k^1 & \text{for } t \in [\alpha T, T[, \end{cases} \tag{3.94}$$

where $k^i > 0$, $i = 0, 1$ are two constants. We chose $k^0 = 1$ to set the time scale. The observable A is the scalar path functional

$$A[z] = \frac{1}{nT} \sum_{t \in [0,nT] \,|\, z(t^+) \neq z(t^-)} g_{z(t^+),z(t^-)}(t), \tag{3.95}$$

where $\boldsymbol{g}$ is antisymmetric: $g_{-+}(t) = -g_{+-}(t) \equiv g(t)$, and piecewise constant:

$$g(t) = \begin{cases} g^0 & \text{for } t \in [0, \alpha T[, \\ g^1 & \text{for } t \in [\alpha T, T[. \end{cases} \tag{3.96}$$

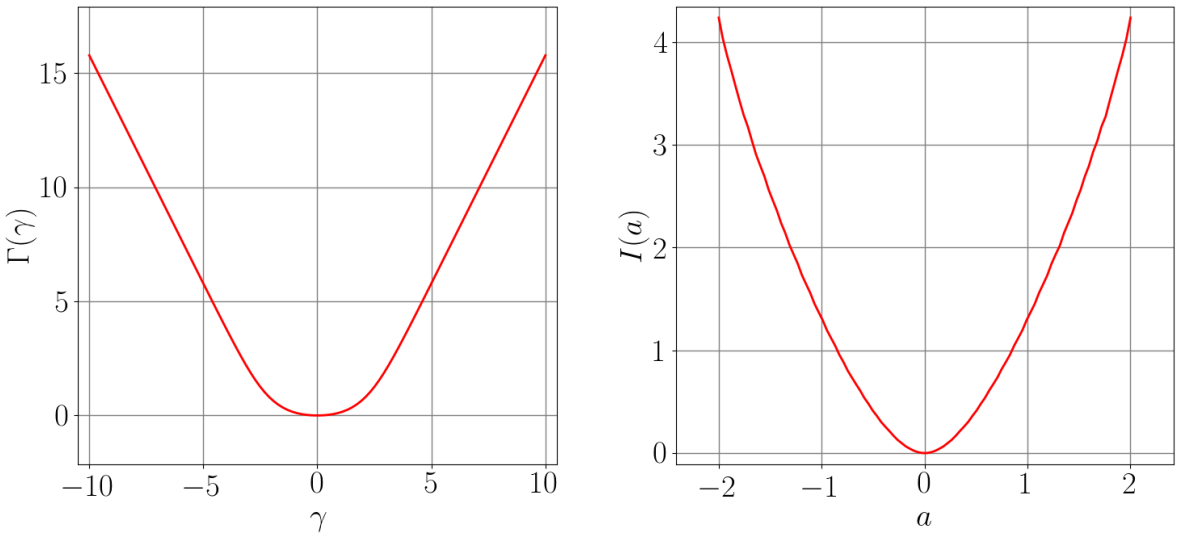

Figure 3.2 – (left) SCGF $\Gamma(\gamma)$ and (right) LDF $I(a)$. The figures are obtained for $\alpha = 0.3$, $T = 1$, $k^0 = 1$, $k^1 = 0.1$, $g^0 = 1$, $g^1 = -1$.

When $g^0 = 1$ and $g^1 = 0$ for instance, nTA counts the net number of transitions $\mathbf{e}_+ \to \mathbf{e}_-$ occuring in the first part of each period. With $\mathbf{e}_+$, $\mathbf{e}_-$ respectively the first and second basis vectors, the tilted operator reads

$$\boldsymbol{\kappa}(t, \boldsymbol{\gamma}) = \begin{pmatrix} -k(t) & k(t)\mathrm{e}^{-\gamma g(t)} \\ k(t)\mathrm{e}^{\gamma g(t)} & -k(t) \end{pmatrix}. \tag{3.97}$$

Our theory relies on the propagator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t, 0)$ that we shall now compute. The tilted operator being piecewise constant, this propagator can be written explicitly since the time-ordered exponential simplifies to a simple exponential on the constant parts. For $t \in [0, \alpha T[$, we obtain

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t, 0) = \mathrm{e}^{-k^0 t} \begin{pmatrix} \cosh(k^0 t) & \mathrm{e}^{-\gamma g^0} \sinh(k^0 t) \\ \mathrm{e}^{\gamma g^0} \sinh(k^0 t) & \cosh(k^0 t) \end{pmatrix}, \tag{3.98}$$

while for $t \in [\alpha T, T[$, and introducing $t^0 \equiv \alpha T$ and $t^1 \equiv t^1(t) \equiv t - \alpha T$, we have

$$\begin{aligned} &\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t, 0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(t^0 + t^1, 0) = \mathrm{e}^{-k^0 t^0 - k^1 t^1} \\ &\times \begin{pmatrix} \prod_i \cosh(k^i t^i) + \prod_i \mathrm{e}^{\gamma(1-2i) g^i} \sinh(k^i t^i) & \sum_i \mathrm{e}^{-\gamma g^i} \sinh(k^i t^i) \cosh(k^{1-i} t^{1-i}) \\ \sum_i \mathrm{e}^{\gamma g^i} \sinh(k^i t^i) \cosh(k^{1-i} t^{1-i}) & \prod_i \cosh(k^i t^i) + \prod_i \mathrm{e}^{-\gamma(1-2i) g^i} \sinh(k^i t^i) \end{pmatrix}, \end{aligned} \tag{3.99}$$

where the sums and products are over $i = 0, 1$. The largest eigenvalue of the propagator over one period reads

$$\chi_T = \frac{1}{2} \left[\mathrm{tr}\ \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, 0) + \sqrt{\left[\mathrm{tr}\ \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, 0) \right]^2 - 4\ \det \overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T, 0)} \right], \tag{3.100}$$

where tr and det stand for the trace and determinant, respectively. Using Eq. (3.17), the SCGF $\Gamma(\gamma)$ follows, see Fig. 3.2 for a numerical computation. The Legendre conjugate

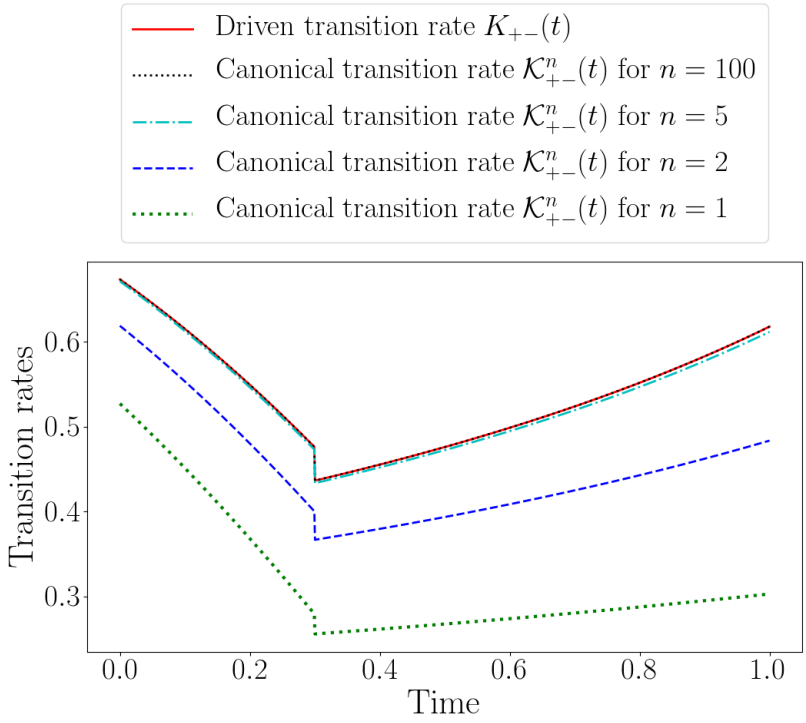

Figure 3.3 – $K_{+-}(t)$ and $\mathcal{K}^n_{+-}(t)$ as a function of time for different number of periods $n = 1, 2, 5, 100$. The figure is obtained for $\alpha = 0.3$, $T = 1$, $k^0 = 1$, $k^1 = 0.1$, $g^0 = 1$, $g^1 = -1$, $a = 0.4$ corresponding to $\gamma = 1.11$.

LDF $I(a)$ is shown on the same figure. Unsurprisingly, I vanishes at $a = 0$ due to the symmetry of the rate matrix $\boldsymbol{k}$: there are asymptotically as many transitions $\mathbf{e}_+ \to \mathbf{e}_-$ as transitions $\mathbf{e}_- \to \mathbf{e}_+$ leading to a vanishing typical value for A.

The generator $\boldsymbol{K}$ being defined as the Doob transform of $\boldsymbol{\kappa}$ based on $\boldsymbol{l}(t) = \boldsymbol{l}(0)\left[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t,0)\right]^{-1}$, we need the left eigenvector $\boldsymbol{l}(0) = \boldsymbol{l}_T$ of the one-period propagator associated with the eigenvalue χ_T:

$$\boldsymbol{l}(0) = \frac{1}{\mathsf{N}} \begin{pmatrix} \prod_i \mathrm{e}^{-k^i t^i} \left[\sum_i \mathrm{e}^{\gamma g^i} \sinh(k^i t^i) \cosh(k^{1-i} t^{1-i}) \right] \\ \rho_T - \prod_i \mathrm{e}^{-k^i t^i} \left[\prod_i \cosh(k^i t^i) + \prod_i \mathrm{e}^{\gamma(1-2i) g^i} \sinh(k^i t^i) \right] \end{pmatrix}^\dagger, \tag{3.101}$$

with N a normalization factor following from Eq. (3.13), and where $\boldsymbol{v}^\dagger$ is the transpose of $\boldsymbol{v}$, for any vector $\boldsymbol{v}$. Inverting the propagators in Eqs. (3.98–3.99), we can compute $\boldsymbol{l}(t)$ at any $t \in [0, T[$. Then, Eq. (3.41) yields an analytic expression for the generator of the driven process from which we have computed numerically one component as shown in Fig. 3.3. Similarly, the generator $\boldsymbol{\mathcal{K}}^n$ is defined as the Doob transform of $\boldsymbol{\kappa}$ based on $\boldsymbol{C}^n(t) = \mathbf{1}\,[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(T,0)]^n[\overleftarrow{\boldsymbol{\mathcal{Q}}}_{\boldsymbol{\kappa}}(t,0)]^{-1}$. Inverting the propagators of Eqs. (3.98–3.99) and taking the n^{th} power of the one-period propagator, we can compute $\boldsymbol{C}^n(t)$ at any $t \in [0, T[$. Then, Eq. (3.28) yields an expression for the generator of the canonical process from which we have computed numerically one component as shown in Fig. 3.3 for $n = 1, 2, 5$ and 100. This figure illustrates the convergence of the canonical generator $\boldsymbol{\mathcal{K}}^n$ towards the driven generator $\boldsymbol{K}$ when $n \to \infty$ as stated in Eq. (3.42). We observe that the two generators $\boldsymbol{K}$

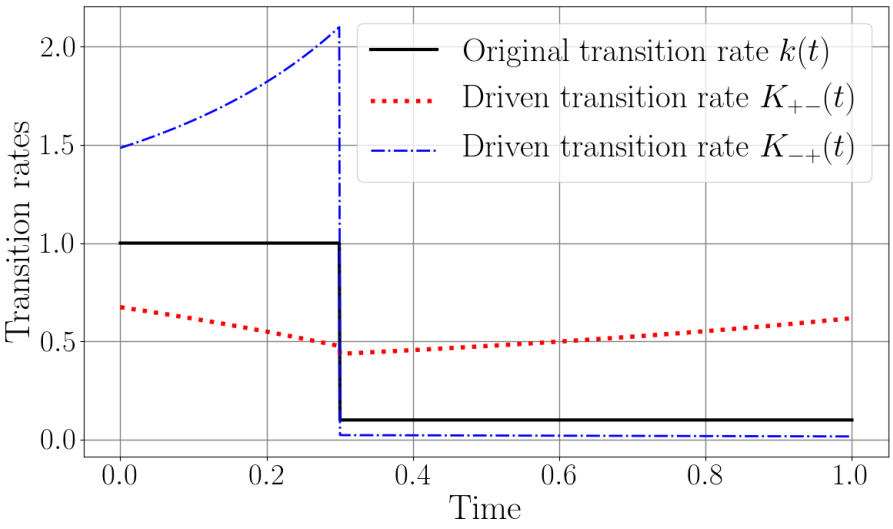

Figure 3.4 – Original transition rate $k(t)$ (solid black line) and driven transition rates $K_{+-}(t)$ (red dotted line) and $K_{-+}(t)$ (blue dash-dotted line). The figure is obtained for $\alpha = 0.3$, $T = 1$, $k^0 = 1$, $k^1 = 0.1$, $g^0 = 1$, $g^1 = -1$, $a = 0.4$ corresponding to $\gamma = 1.11$.

and $\boldsymbol{\mathcal{K}}^n$ are piecewise continuous (with discontinuities at αT and T) and time-dependent even though the original rate matrix $\boldsymbol{k}$ was piecewise constant.

On Fig. 3.4, we plot both driven rates $K_{+-}(t)$ and $K_{-+}(t)$ and original rates $k_{+-}(t) = k_{-+}(t) = k(t)$ to observe qualitatively the effect of the conditioning on our initial Markov process. We chose to impose $a = 0.4$ net transitions from $\mathbf{e}_+$ to $\mathbf{e}_-$ per unit of time, counted positively if they occur on the first part of each cycle ($g^0 = 1$) and negatively on the second part ($g^1 = -1$). In view of the strict convexity of I, the process that has $a = 0.4$ as a typical event is the driven process for $\gamma = 1.11 = I'(0.4)$. In the original process, A is zero on average due to the symmetry of the rate matrix. Hence, imposing $a > 0$ should increase the rate of the driven process for transitions $\mathbf{e}_+ \to \mathbf{e}_-$ on $[0, \alpha T]$ and transitions $\mathbf{e}_- \to \mathbf{e}_+$ on $[\alpha T, T]$. Compared to the original rate k, we see on Fig. 3.4 that indeed $K_{+-} < k^0 < K_{-+}$ on $[0, \alpha T]$ so that transitions $\mathbf{e}_+ \to \mathbf{e}_-$ are preferred on average, and conversely $K_{-+} < k^1 < K_{+-}$ on $[\alpha T, T]$ so that transitions $\mathbf{e}_- \to \mathbf{e}_+$ are preferred on average. Hence, the conditioning has broken the symmetry of the rate matrix and made it fully time-dependent.

On Fig. 3.5, we plot the rate K_{+-} for different values of a (associated to their corresponding γ). We observe that this rate from $\mathbf{e}_- \to \mathbf{e}_+$ deviates more and more from k_{+-} as $|a|$ becomes larger, i.e. as a goes away from the typical value 0 of the original process. The magnitude of change of the driven rate is thus in direct correspondence with the magnitude of the conditioning. However, it is not intuitive to understand the growth of the transition rate. We can just say that the possibility of a time-dependent rate matrix offers a broader dynamical space to explore in the variational calculation compared to case of piecewise-constant rates.

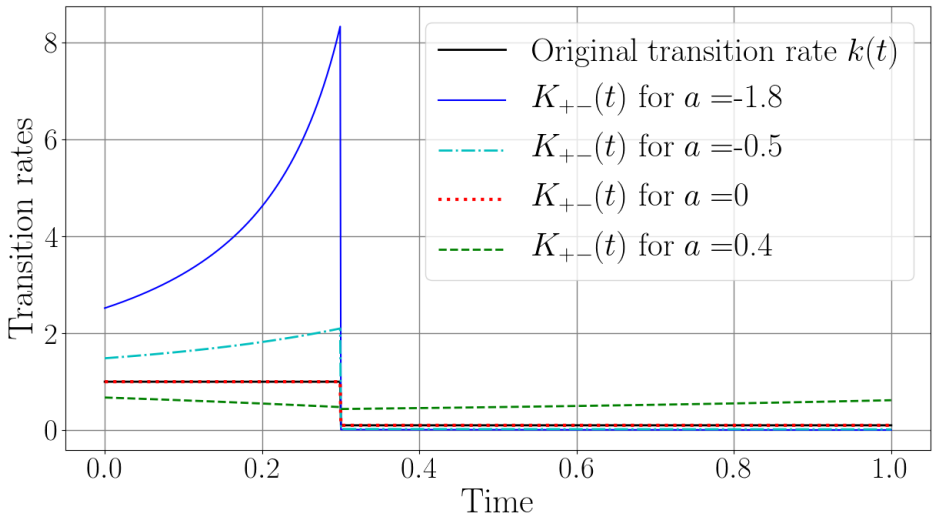

Figure 3.5 – Original transition rate $k(t)$ (solid black line) and driven rate $K_{+-}(t)$ (colored lines) for different values of γ corresponding to different values of the fluctuation a. The figure is obtained for $\alpha = 0.3$, $T = 1$, $k^0 = 1$, $k^1 = 0.1$, $g^0 = 1$, $g^1 = -1$. The values $a = -1.8, -0.5, 0, 0.4$ correspond respectively to $\gamma = -3.59, -1.29, 0, 1.11$. As expected, $\boldsymbol{k} = \boldsymbol{K}$ when conditioning at the typical value $a = 0$.

3.2 Conditioning of periodically driven Markov diffusion processes

In this section, we address the problem of conditioning of periodically driven Markov processes within the framework of diffusion processes.

3.2.1 Periodically driven diffusion processes

We consider a Markov diffusion process z_t evolving according to the Langevin equation of Eq. (2.54) and whose drift b and diffusion coefficient σ are time-periodic of period T:

$$b(x, t+T) = b(x, t), \tag{3.102}$$
$$\sigma(x, t+T) = \sigma(x, t), \tag{3.103}$$

$\forall x$ and t. We set the initial time to $t_0 = 0$ and the final time to $t = nT$ with n the number of periods that have elapsed, and we assume that the solution of the Fokker-Planck equation $\frac{\partial \varrho(x,t)}{\partial t} = (\mathfrak{L}\varrho)(x, t)$ reaches a T-periodic solution ϱ^{TiPS} when $n \to \infty$, i.e. $\varrho^{\mathrm{TiPS}}(t+T) = \varrho^{\mathrm{TiPS}}(t)$. Note that the periodicity of b and σ implies the periodicity of the Fokker-Planck operator:

$$\mathfrak{L}(t+T) = \mathfrak{L}(t), \quad \forall t. \tag{3.104}$$

All the definitions introduced in Sec. 2.4 apply in this section with the difference that the final time is fixed to $t = nT$. As discussed previously, we are interested in the fluctuations of the observable $\boldsymbol{A}_{nT}$, functional of the paths up to the final time nT, defined in Eq. (2.69) by replacing t by nT, and where we consider T-periodic functions g and h:

$g(x,t+T)=g(x,t)$ and $h(x,t+T)=h(x,t)$, $\forall x$ and t. We assume that $\boldsymbol{A}_{nT}$ satisfies a LDP (1.18):

$$P(\boldsymbol{A}_{nT}=\boldsymbol{a}) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nTI(a)}, \tag{3.105}$$

with I the LDF, and we introduce its associated SCGF

$$\Gamma(\boldsymbol{\gamma}) = \lim_{n\to\infty} \frac{1}{nT} \ln \mathbb{E}_{\varrho(0)} \left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}} \right]. \tag{3.106}$$

In order to study the large deviations of $\boldsymbol{A}_{nT}$ in the long-time limit, let us first consider the observable $\boldsymbol{A}_t[z]$ on the shorter time interval $[0,t]$:

$$\boldsymbol{A}_t[z] = \begin{pmatrix} \frac{1}{t}\int_0^t g(z_\tau,\tau)\circ \mathrm{d}z_\tau \\ \frac{1}{t}\int_0^t h(z_\tau,\tau)\mathrm{d}\tau \end{pmatrix}, \tag{3.107}$$

where we remind that $\circ$ refers to the Stratonovich convention. As seen in Sec. 1.2, the statistics of $\boldsymbol{A}_t$ is contained in the generating function $G_{\boldsymbol{\gamma}}(t) = \int \mathrm{d}x \tilde{G}_{\boldsymbol{\gamma}}(x,t)$ defined in Eq. (2.81), where $\tilde{G}_{\boldsymbol{\gamma}}(x,t)$ is the generating function with state x at time t (2.83) solution of

$$\frac{\partial}{\partial t}\tilde{G}_{\boldsymbol{\gamma}}(x,t) = (\Lambda \tilde{G}_{\boldsymbol{\gamma}})(x,t). \tag{3.108}$$

We remind the definition of the biased generator $\Lambda = \Lambda_{\boldsymbol{\gamma}}$ (2.85) defined by its action on an arbitrary function f by:

$$\Lambda_{\boldsymbol{\gamma}} f = (-\nabla + \gamma_1 g)(\hat{b} f) + \frac{1}{2}(-\nabla+\gamma_1 g)\left[D(-\nabla+\gamma_1 g)f\right] + \gamma_2 h f, \tag{3.109}$$

with $\hat{b} = b - \frac{1}{2}\sigma\nabla\sigma$ and $D=\sigma^2$, and where we dropped for clarity the space-time dependence of all the functions. We will often lighten the notations by not mentioning the dependence on x, t and $\boldsymbol{\gamma}$ but one keeps in mind their presence. The adjoint operator $\Lambda^\dagger$ of the biased Fokker-Planck operator acting on an arbitrary function φ reads

$$\Lambda^\dagger \varphi = \hat{b}(\nabla+\gamma_1 g)\varphi + \frac{1}{2}(\nabla+\gamma_1 g)\left[D(\nabla+\gamma_1 g)\varphi\right] + \gamma_2 h\varphi. \tag{3.110}$$

Adjoints operators are related by definition by

$$\varphi\cdot(\Lambda f) = f\cdot(\Lambda^\dagger\varphi), \tag{3.111}$$

for any functions φ, f. The dot stands for the scalar product $\varphi\cdot f \equiv \int_x \varphi(x)f(x)$. By abuse of notation, we use the dot for the scalar product both in the vector-space and in the function-space. Notice that Λ and $\Lambda^\dagger$ are also T-periodic. The formal solution of Eq. (3.108) with initial condition $\tilde{G}_{\boldsymbol{\gamma}}(0)=\varrho(0)$ reads

$$\tilde{G}_{\boldsymbol{\gamma}}(t) = \overleftarrow{\mathcal{Q}}_\Lambda(t,0)\varrho(0), \tag{3.112}$$

to be understood as

$$\tilde{G}_{\boldsymbol{\gamma}}(x,t) = \int \mathrm{d}y \overleftarrow{\mathcal{Q}}_\Lambda(x,t,y,0)\varrho(y,0), \tag{3.113}$$

where the propagator $\overleftarrow{\mathcal{Q}}_\Lambda(x,t,y,0)$ from y at t_0 to x at t is the biased transition probability $\tilde{G}(x,t \mid y,0)$ introduced in Eq. (2.112), and where $\overleftarrow{\mathcal{Q}}_\Lambda(t,0) \equiv \overleftarrow{\exp}\left[\int_0^t \mathrm{d}\tau \Lambda(\tau)\right]$ is the

time-ordered exponential of Λ, defined as the unique solution of the initial value problem $\frac{\mathrm{d}}{\mathrm{d}t}X = \Lambda X$, with $X(0) = \mathbb{1}$ the operator such that $\mathbb{1}f = f$, see Appendix 3.A for more detail.

As in the case of jump processes, we look for an effective Markov process that is equivalent to the microcanonical process in which we conditioned the observable $\boldsymbol{A}_{nT}$ on a rare fluctuation $\boldsymbol{a}$. This process follows from the driven process which is defined as the limit of the canonical process. We will show this result in the formalism of diffusion processes from a path integral perspective first, then from a variational approach, extending the results of Refs. [80,81] for time-homogeneous processes to our periodically driven processes.

3.2.2 Spectral properties of the one-period propagator

From now on, we assume that the final time is always nT and omit the subscript nT for our generic observable $\boldsymbol{A}[z] \equiv \boldsymbol{A}_{nT}[z]$. In this section, we relate the spectral properties of the propagator $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$ to the generating functions of the observable, giving them a physical interpretation.

In 1948, Krein and Rutman extended the Perron-Frobenius theorem to infinite dimensional Banach spaces and compact operators [235, 236]. We will not go into detail and we simply assume that the operator $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$ that we consider satisfies the conditions required by the Krein-Rutman theorem so that it admits a dominant eigenvalue associated with unique positive left and right eigenfunctions, see Theorem 6. Let χ_T be the highest eigenvalue of $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$, r_T its right eigenfunction and l_T its left eigenfunction:

$$\overleftarrow{\mathcal{Q}}_\Lambda(T,0)\, r_T = \chi_T\, r_T, \tag{3.114}$$

$$\overleftarrow{\mathcal{Q}}_\Lambda(T,0)^\dagger\, l_T = \chi_T\, l_T, \tag{3.115}$$

where r_T and l_T are functions of space. The eigenfunctions r_T and l_T can be chosen up to a multiplicative constant that we set by imposing

$$\mathbf{1} \cdot r_T = 1, \tag{3.116}$$

$$l_T \cdot r_T = 1, \tag{3.117}$$

with $\mathbf{1} : x \mapsto 1$. We assume furthermore that

$$l_T \cdot \varrho(0) < \infty. \tag{3.118}$$

At time nT, the generating function $G_{\boldsymbol{\gamma}}$ reads

$$G_{\boldsymbol{\gamma}}(nT) = \mathbf{1} \cdot \tilde{G}_{\boldsymbol{\gamma}}(nT) = \mathbf{1} \cdot \overleftarrow{\mathcal{Q}}_\Lambda(nT,0)\varrho(0) = \int_{x,y} \overleftarrow{\mathcal{Q}}_\Lambda(x,T,y,0)^n \varrho(y,0), \tag{3.119}$$

where we used the periodicity of b and σ — and hence of Λ — in the last equation. The asymptotic expansion of $\overleftarrow{\mathcal{Q}}_\Lambda(nT,0)$ at large n is dominated by the contribution of its dominant eigenvalue:

$$\overleftarrow{\mathcal{Q}}_\Lambda(T,0)^n \underset{n\to\infty}{\simeq} (\chi_T)^n (r_T l_T), \tag{3.120}$$

where $(r_T l_T)$ is to be understood here as the operator $(r_T l_T f)(x) \equiv \int_y r_T(x) l_T(y) f(y)$. From the definition of the SCGF (3.106) and the generating function at time nT (3.119),

and using Eqs. (3.116, 3.118), it follows

$$\Gamma = \frac{1}{T} \ln \chi_T. \tag{3.121}$$

Hence, the SCGF Γ is proportional to the logarithm of the largest eigenvalue of the one-period propagator. It is easy to check that Eqs. (3.18–3.19) obtained in the case of jump processes remain valid for our diffusion process.

3.2.3 Generalized Doob transform

As introduced in Section 3.1.4, the Doob transform of the operator M associated with a positive function v reads

$$M^v \equiv v M v^{-1} - v^{-1}(M^\dagger v), \tag{3.122}$$

with $v^{-1} = \frac{1}{v}$. The action of M^v on an arbitrary function f reads:

$$(M^v f)(x,t) \equiv v(x,t)(M v^{-1} f)(x,t) - v^{-1}(x,t)(M^\dagger v)(x,t) f(x,t). \tag{3.123}$$

M^v is norm-conserving since $\mathbf{1} \cdot M^v = 0$ by definition of the adjoint operator (3.111). The Doob transform allows then building norm-conserving generators.

Doob transform of the Fokker-Planck operator

When M is the Fokker-Planck operator $\mathfrak{L}$, computing explicitly (3.122) yields for any function φ

$$\mathfrak{L}^v \varphi = -\nabla \left[\hat{b}^v \varphi - \frac{1}{2} D \nabla \varphi \right], \tag{3.124}$$

where we introduced

$$b^v \equiv b + D \nabla \ln v, \tag{3.125}$$

$$\hat{b}^v \equiv \hat{b} + D \nabla \ln v. \tag{3.126}$$

Therefore, the Doob transform of a Fokker-Planck operator is a Fokker-Planck operator with the same diffusion coefficient but with a modified drift. The path probability associated with the Doob transform M^v (or equivalently b^v and σ) is obtained by replacing b by b^v in Eq. (2.63) and using the fact that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\ln v(z_t,t)) = v(z_t,t)^{-1} \partial_t v(z_t,t) + \dot{z}_t \nabla (\ln v(z_t,t)). \tag{3.127}$$

It follows

$$\mathbb{P}_{b^v,\sigma,\varrho(0)[z]} = \mathbb{P}_{b,\sigma,\varrho(0)}[z] v(z_{nT}, nT) v(z_0,0)^{-1} \exp\left\{ - \int_0^{nT} \mathrm{d}\tau \left[v(z_\tau,\tau)^{-1} (\mathfrak{L}^\dagger v)(z_\tau,\tau) + v^{-1}(z_\tau,\tau) \frac{\partial v}{\partial t}(z_\tau,\tau) \right] \right\}. \tag{3.128}$$

Doob transform of the biased Fokker-Planck operator

When M is the biased Fokker-Planck operator Λ, computing explicitly (3.122) yields

$$\Lambda^{\upsilon}\varphi = -\nabla\left[\hat{b}^{\upsilon}_{\boldsymbol{\gamma}}\varphi - \frac{1}{2}D\nabla\varphi\right], \tag{3.129}$$

where we introduced

$$b^{\upsilon}_{\boldsymbol{\gamma}} \equiv b + D(\nabla \ln \upsilon + \gamma_1 g), \tag{3.130}$$

$$\hat{b}^{\upsilon}_{\boldsymbol{\gamma}} \equiv \hat{b} + D(\nabla \ln \upsilon + \gamma_1 g). \tag{3.131}$$

The Doob transform of the biased Fokker-Planck operator yields a Fokker-Planck operator of drift $b^{\upsilon}_{\boldsymbol{\gamma}}$ and diffusion coefficient σ that generates a norm-conserving Markov diffusion process. Injecting the expression of $b^{\upsilon}_{\boldsymbol{\gamma}}$ instead of b in Eq. (2.63) yields after lengthy calculations the path probability associated with Λ^{υ}:

$$\begin{aligned}\mathbb{P}_{b^{\upsilon}_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z] = \mathbb{P}_{b,\sigma,\varrho(0)}[z]\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\upsilon(z_{nT},nT)\upsilon(z_0,0)^{-1} \\ \exp\left\{-\int_0^{nT}\mathrm{d}\tau\left[\upsilon(z_\tau,\tau)^{-1}(\Lambda^{\dagger}\upsilon)(z_\tau,\tau) + \upsilon^{-1}(z_\tau,\tau)\frac{\partial\upsilon}{\partial t}(z_\tau,\tau)\right]\right\},\end{aligned} \tag{3.132}$$

where we used

$$\nabla\upsilon = \upsilon\nabla(\ln\upsilon), \tag{3.133}$$

$$\nabla(\upsilon^{-1}) = -\upsilon^{-1}\nabla(\ln\upsilon). \tag{3.134}$$

3.2.4 The driven process as the limit of the canonical process

Canonical generator

As for jump processes, we show that the canonical process defined by its path probability

$$\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z] \equiv \frac{\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\mathbb{P}_{b,\sigma,\varrho(0)}[z]}{\mathbb{E}_{\varrho(0)}[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}}]}, \tag{3.135}$$

is generated by a norm-conserving Markov generator [80]. To do so, we look for a generator $\mathbf{L}_n$ of drift b^c and diffusion coefficient σ^c that satisfy

$$\mathbb{P}_{b^c,\sigma^c,\varrho'(0)}[z] = \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z], \tag{3.136}$$

where $\varrho'(0)$ is an initial probability density that may be different from $\varrho(0)$. Moreover, we want $\mathbf{L}_n$ to generate a norm-conserving process, hence we look for a function C_n such that $\mathbf{L}_n$ is built from the Doob transform

$$\mathbf{L}_n \equiv \Lambda^{C_n} \equiv C_n\mathbf{L}_n\,(C_n)^{-1} - (C_n)^{-1}(\mathbf{L}_n^{\dagger}C_n). \tag{3.137}$$

Using the result obtained in Eqs. (3.130–3.131), the drift and the diffusion coefficient of the canonical generator $\mathbf{L}_n$ are given by

$$b^c = b + D(\nabla \ln C_n + \gamma_1 g), \tag{3.138}$$

$$\hat{b}^c = \hat{b} + D(\nabla \ln C_n + \gamma_1 g), \tag{3.139}$$

$$\sigma^c = \sigma. \tag{3.140}$$

We now need to determine the function C_n. From Eq. (3.132), the path probability generated by the operator $\mathbf{L}_n$ reads

$$\mathbb{P}_{b_{\boldsymbol{\gamma}}^{C_n},\sigma,\varrho(0)}[z] = \mathbb{P}_{b,\sigma,\varrho(0)}[z]\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}C_n(z_{nT},nT)C_n(z_0,0)^{-1} \\ \exp\left\{-\int_0^{nT}\mathrm{d}\tau\left[C_n(z_\tau,\tau)^{-1}(\Lambda^\dagger C_n)(z_\tau,\tau)+C_n^{-1}(z_\tau,\tau)\frac{\partial C_n}{\partial t}(z_\tau,\tau)\right]\right\}. \tag{3.141}$$

For this path probability to coincide with Eq. (3.135), $C_n \equiv C_n(t)$ should be chosen equal to

$$C_n(t) = \left[\overleftarrow{\mathbf{Q}}_\Lambda(nT,t)\right]^\dagger \mathbf{1}, \tag{3.142}$$

solution of

$$\begin{cases}\dot{C}_n = -\Lambda^\dagger C_n,\\ C_n(nT)=\mathbf{1},\end{cases} \tag{3.143}$$

so that the time-extensive terms in the exponential of Eq. (3.141) vanish. Note that the positivity of C_n is ensured by the Feynman-Kac formula[2] [80]. With this choice of C_n, the path probability associated with Λ^{C_n} (3.141) becomes

$$\mathbb{P}_{b^c,\sigma,\varrho(0)}[z] = \mathbb{P}_{b,\sigma,\varrho(0)}[z_t]\,\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]}\,C^n(z_0,0)^{-1}, \tag{3.144}$$

or equivalently

$$\mathbb{P}_{b,\sigma,\varrho(0)}[z]\,\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]} = \mathbb{P}_{b^c,\sigma,\varrho(0)C^n(0)}[z]. \tag{3.145}$$

From Eqs. (3.119) and (3.142) and using the definition of the adjoint operator (3.111), we remark that the generating function can be expressed in terms of C_n as

$$G_{\boldsymbol{\gamma}}(nT) = \mathbb{E}_{\varrho(0)}[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}}] = C_n(0)\cdot\varrho(0), \tag{3.146}$$

implying from Eq. (3.135) and Eq. (3.145) that

$$\mathbb{P}_{b^c,\sigma,\frac{C_n(0)\varrho(0)}{C_n(0)\cdot\varrho(0)}}[z] = \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z]. \tag{3.147}$$

In other words, the canonical path probability is associated with the generator $\mathbf{L}_n = \Lambda^{C_n}$ and the initial probability density $\varrho'(0) = \frac{C_n(0)\varrho(0)}{[C_n(0)\cdot\varrho(0)]}$. This shows that the canonical process has a corresponding Markov generator which is norm-conserving contrary to the biased generator. Notice that the canonical generator $\mathbf{L}_n$ depends explicitly on the number of periods n.

[2] The Feynman-Kac formula [237–239] states that the solution of

$$\begin{cases}\frac{\partial X}{\partial t}+\mathbf{L}X+VX=0,\\ X(t_\mathrm{f},x)=\varphi(x),\end{cases}$$

with $\mathbf{L}$ a Markov operator and V a function of space and time, reads

$$X(t,x) = \mathbb{E}\left[\varphi(z_t)\mathrm{e}^{\int_t^{t_\mathrm{f}}V(z_\tau,\tau)\mathrm{d}\tau}\middle| z_t = x\right].$$

Driven process

In this section, we focus on the asymptotic dynamics of the canonical process in the limit $n \to \infty$. From Eq. (3.142) and using Eqs. (3.116–3.120), we find that $C_n(\tau)$ for $\tau \in [0, T[$ is given asymptotically by

$$C_n(\tau) = \left[\overleftarrow{\mathfrak{Q}}_\Lambda(T,0)^n \overleftarrow{\mathfrak{Q}}_\Lambda(\tau,0)^{-1}\right]^\dagger \mathbf{1} \underset{n\to\infty}{\sim} (\chi_T)^n \left[\overleftarrow{\mathfrak{Q}}_\Lambda(\tau,0)^{-1}\right]^\dagger l_T. \tag{3.148}$$

We want to define the driven generator as the Doob transform of the biased operator using a function $l(\tau)$ corresponding to the limit of $C_n(\tau)$ when $n \to \infty$. Since scalar constants play no role in the Doob transform, it suffices to consider $l \equiv l(\tau)$ defined for $\tau \in [0, T[$ by

$$l(\tau) \equiv \left[\overleftarrow{\mathfrak{Q}}_\Lambda(\tau,0)^{-1}\right]^\dagger l_T, \tag{3.149}$$

which is by construction the solution of

$$\begin{cases} \dot{l} = -\Lambda^\dagger l, \\ l(0) = l_T. \end{cases} \tag{3.150}$$

Using Eq. (3.115) and the periodicity of Λ, we notice that the function l satisfies

$$l(\tau + T) = \chi_T^{-1} l(\tau). \tag{3.151}$$

We define the Markov generator $\mathcal{L} \equiv \mathcal{L}(\tau, \boldsymbol{\gamma})$ of the driven process at any time τ by the Doob transform of the tilted generator Λ associated with the function l:

$$\mathcal{L} \equiv \Lambda^l = l\Lambda l^{-1} - l^{-1}(\Lambda^\dagger l). \tag{3.152}$$

Note that the positivity of $l(t)$ at all t is ensured by the positivity of C_n, of $\overleftarrow{\mathfrak{Q}}_\Lambda(\tau,0)$ and of l_T (Krein–Rutman Theorem). By construction from Eqs. (3.137) and (3.148), the generator of the driven process is the limit of the canonical operator as $n \to \infty$:

$$\lim_{n\to\infty} \mathbf{L}_n = \mathcal{L}. \tag{3.153}$$

Using Eqs. (3.130–3.131), it follows that the driven generator $\mathcal{L}$ is a Fokker-Planck operator of diffusion coefficient σ and drift

$$\begin{aligned} B_{\boldsymbol{\gamma}} &\equiv b + D(\nabla \ln l + \gamma_1 g), \\ \hat{B}_{\boldsymbol{\gamma}} &\equiv \hat{b} + D(\nabla \ln l + \gamma_1 g). \end{aligned} \tag{3.154}$$

In the following, we discuss the properties of the driven generator. First of all, the operator $\mathcal{L}$ is T-periodic. Indeed, using Eq. (3.151) and the periodicity of Λ in Eq. (3.152), it follows for all τ

$$\mathcal{L}(\tau + T) = \mathcal{L}(\tau). \tag{3.155}$$

Secondly, the TiPS probability of the driven process can be obtained from the solution l of the initial value problem of Eq. (3.150) and the solution $r \equiv r(\tau)$ of the initial value problem

$$\begin{cases} \dot{r} = \Lambda r, \\ r(0) = r_T, \end{cases} \tag{3.156}$$

or alternatively

$$r(t) \equiv \overleftarrow{\mathfrak{Q}}_\Lambda(t,0) r_T. \tag{3.157}$$

Using Eq. (3.114) and the periodicity of Λ, the function r satisfies:

$$r(t+T) = \chi_T r(t). \tag{3.158}$$

The TiPS probability of the driven process $\mu \equiv \mu(t)$, defined as the T-periodic solution of the Fokker-Planck equation:

$$\begin{cases} \frac{\mathrm{d}\mu}{\mathrm{d}t} = \Lambda\mu \\ \mu(0) = \mu(T), \end{cases} \tag{3.159}$$

is expressed in terms of the functions l and r

$$\mu(x,t) = l(x,t) r(x,t), \quad \forall x, t. \tag{3.160}$$

Indeed, Eqs. (3.150, 3.152, 3.156) yield for all x and τ

$$\begin{aligned} (\mathcal{L} l r)(x,\tau) &= l(x,\tau)(\Lambda l^{-1} l r)(x,\tau) - l^{-1}(x,\tau)(\Lambda^\dagger l)(x,\tau) l(x,\tau) r(x,\tau) && (3.161)\\ &= l(x,\tau)(\Lambda r)(x,\tau) - (\Lambda^\dagger l)(x,\tau) r(x,\tau) && (3.162)\\ &= l(x,\tau)\dot{r}(x,\tau) + \dot{l}(x,\tau) r(x,\tau) && (3.163)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[l(x,\tau) r(x,\tau) \right], && (3.164) \end{aligned}$$

while Eqs. (3.151, 3.158) lead to

$$l(0) r(0) = l(T) r(T), \tag{3.165}$$

which proves that μ is the solution of Eq. (3.159). Notice that our normalization choice in Eq. (3.117) ensures the normalization of $\mu(0)$.

Let us now discuss the asymptotic equivalence between the canonical and driven path probabilities. Using Eqs. (3.121, 3.150–3.152) in Eq. (3.132), the path probability of the driven process reads

$$\mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z] = \mathbb{P}_{b,\sigma,\varrho(0)}[z] \mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}[z]} l(z_{nT},0) \mathrm{e}^{-nT\Gamma} l(z_0,0)^{-1}. \tag{3.166}$$

From the definitions of the canonical path probability (3.135) and the driven path probability (3.166), we obtain

$$\frac{\mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z]} = l(z_{nT},0)\, \mathrm{e}^{-nT\Gamma}\, l(z_0,0)^{-1}\, \mathbb{E}_{\varrho(0)}[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}}]. \tag{3.167}$$

Using the definition of the SCGF (3.106), we finally find:

$$\lim_{n\to\infty} \frac{1}{nT} \ln \frac{\mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z]} = 0. \tag{3.168}$$

Hence, the driven path probability and the canonical path probability are logarithmically equivalent:

$$\mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z] \underset{n\to\infty}{\asymp} \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z]. \tag{3.169}$$

Given this equivalence, and using the fact that the canonical path probability (for a specific value of $\boldsymbol{\gamma}$) and the microcanonical path probability (for $\boldsymbol{A}$ conditioned on $\boldsymbol{a}$) are logarithmically equivalent if the LDF I is strictly convex at $\boldsymbol{a}$, we conclude that the effective process is equivalent to the microcanonical process conditioned on $\boldsymbol{A} = \boldsymbol{a}$ is given by the driven process for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$ (provided that I is strictly convex and differentiable). Mathematically, this reads:

$$\mathbb{P}^{\text{micro}}_{\mathbf{a},\varrho(0)}[z] \underset{n\to\infty}{\asymp} \mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z]\,\big|_{\boldsymbol{\gamma}=\nabla I(\boldsymbol{a})}\,. \tag{3.170}$$

3.2.5 Variational approach

In this section, we extend the results of Ref. [81] to periodically driven diffusion processes and derive the driven process from an optimisation problem under constraint of the 2.5 LDF of the empirical occupation density $\rho^n[z]$:

$$\rho^n[z](x,\tau) = \frac{1}{n}\sum_{\ell=1}^{n-1}\delta(z_{\tau+\ell T} - x), \tag{3.171}$$

and the empirical current $j[z]$:

$$j^n[z](x,\tau) = \frac{1}{n}\sum_{\ell=1}^{n-1}\delta(z_{\tau+\ell T} - x)\circ\dot{z}_{\tau+\ell T}. \tag{3.172}$$

The quantity $\rho^n[z](x,t)\mathrm{d}x$ measures the fraction of periods in which the system has occupied a state in the infinitesimal interval $[x, x+\mathrm{d}x[$ at time τ of each period along the path $[z]$ of duration nT, while the quantity $n\mathrm{d}\tau j^n[z](x,t)$ measures the number of periods in which the system performed a displacement between x and $x+\mathrm{d}x$ at time τ of each period along the trajectory $[z]$ of duration nT. These two quantities are related by

$$n\rho^n[z](x,t+\mathrm{d}\tau)\mathrm{d}x - n\rho^n[z](x,t)\mathrm{d}x = n\mathrm{d}\tau j^n[z](x,t) - n\mathrm{d}\tau j^n[z](x+\mathrm{d}x,t), \tag{3.173}$$

meaning that the variation in the number of periods in which the system occupies a state in $[x, x+\mathrm{d}x[$ between times τ and $\tau+\mathrm{d}t$ has two contributions: an additive contribution coming from a displacement of the system between x and $x+\mathrm{d}x$, and a subtractive contribution coming from a displacement of the system out of the interval $[x, x+\mathrm{d}x[$. This leads to the continuity equation:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\rho^n[z](x,\tau) = -\nabla j^n[z](x,\tau). \tag{3.174}$$

The empirical occupation density and the empirical current are random variables since they are functional of the paths and depend on the number of periods n considered during the total observation time nT. In the limit of large n, the probability to observe the empirical occupation density $\rho^n[z] = \rho$ and the empirical current $j^n[z] = j$ satisfies a large deviation principle:

$$P_n(j,\rho) \underset{n\to\infty}{\asymp} \mathrm{e}^{nTI_{2.5}(j,\rho)}, \tag{3.175}$$

where the 2.5 LDF is given by [240]

$$I_{2.5}(j,\rho) = \frac{1}{T}\int_0^T \mathrm{d}\tau \int \mathrm{d}x \left[\frac{1}{2\rho D}(j - J^\rho)^2\right], \tag{3.176}$$

with $J^\rho = \hat{b}\rho - \frac{1}{2}D\nabla\rho$, see Appendix 3.C for an heuristic derivation. This expression holds only for $\rho(0) = \rho(T)$ and $j(0) = j(T)$, and for conservative currents $\frac{\partial}{\partial t}\rho(x,\tau) = -\nabla j(x,\tau)$, $\forall x, \tau$, and normalized occupations $\int \mathrm{d}x \rho(x,\tau) = 1$, otherwise $I_{2.5}$ is infinite. Notice that $I_{2.5}(j,\rho)$ vanishes for $\rho = \varrho^{\mathrm{TiPS}}$ and $j = J^{\varrho^{\mathrm{TiPS}}}$. Hence, without any conditioning on the paths generated by our original process of drift b and diffusion coefficient σ, the random variable $\rho^n[z]$ converges in probability to ϱ^{TiPS} in the limit $n \to \infty$ while the random variable $j^n[z]$ converges in probability to $J^{\varrho^{\mathrm{TiPS}}}$.

The level 2.5 LDF is appropriate for studying the large deviations of the observable $\boldsymbol{A}$ since the LDF $I(\boldsymbol{a})$ can be obtained from the 2.5 LDF $I_{2.5}(j,\rho)$ using the contraction principle. Indeed, using the periodicity of g and h, we can rewrite the conditioning observable $\boldsymbol{A}$ of Eq. (3.107) with final time $t = nT$ in terms of $\rho^n[z]$ and $j^n[z]$:

$$\boldsymbol{A}(j^n[z],\rho^n[z]) = \begin{pmatrix} A_1(j^n[z]) \\ A_2(\rho^n[z]) \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{T}\int_0^T \mathrm{d}\tau \int \mathrm{d}x g(x,\tau) j^n[z](x,\tau) \\ \frac{1}{T}\int_0^T \mathrm{d}\tau \int \mathrm{d}x h(x,\tau)\rho^n[z](x,\tau) \end{pmatrix}. \tag{3.177}$$

Without conditioning, the observable $\boldsymbol{A}$ converges to $\boldsymbol{A}(J^{\varrho^{\mathrm{TiPS}}}, \varrho^{\mathrm{TiPS}})$ as $n \to \infty$. As before, we are interested in conditioning our process on the event $\{\boldsymbol{A}(j,\rho) = \boldsymbol{a} \mid \boldsymbol{a} \in \mathbb{R}^2\}$. We look for the most probable pair (j,ρ) compatible with $\boldsymbol{A}(j,\rho) = \boldsymbol{a}$. It is obtained by minimizing the 2.5 LDF under the following constraints

- **C0:** $\boldsymbol{A}(j,\rho) = \boldsymbol{a}$,
- **C1:** $\int \mathrm{d}x \rho(x,\tau) = 1, \forall \tau \in [0,T]$,
- **C2:** $\frac{\partial}{\partial \tau}\rho(x,\tau) = -\nabla j(x,\tau), \forall x$ and $\forall \tau \in [0,T]$,
- **C3:** $\rho(T) = \rho(0)$,
- **C4:** $j(T) = j(0)$.

This optimization problem amounts to computing the LDF of $\boldsymbol{A}$ and reads mathematically

$$I(\boldsymbol{a}) = \inf_{\rho,j|\{\mathbf{Ci}\}_{i=0}^4} \Big\{ I_{2.5}(j,\rho) \Big\}. \tag{3.178}$$

Intuitively, the optimizer $(j_{\boldsymbol{a}}, \rho_{\boldsymbol{a}})$ is expected to be the typical value of $(j^n[z], \rho^n[z])$ under the dynamics associated with the effective process and which is equivalent to the microcanonical process as $n \to \infty$. Reciprocally, we can instead consider a dual approach and compute the SCGF defined in Eq. (3.106) and given by the LF transform of the LDF $I(\boldsymbol{a})$

$$\Gamma(\boldsymbol{\gamma}) = \sup_{\rho,j|\{\mathbf{Ci}\}_{i=1}^4} \left\{ \boldsymbol{\gamma} \cdot \boldsymbol{A}(j,\rho) - I_{2.5}(j,\rho) \right\}. \tag{3.179}$$

The solution $(j_{\boldsymbol{\gamma}}, \rho_{\boldsymbol{\gamma}})$ of Eq. (3.179) is the typical value of $(j[z], \rho[z])$ under the canonical path probability [81]. It is expected to be associated with the generator of the driven process as will be shown. For strictly convex LDF, Eqs. (3.178) and (3.179) have the same solutions, i.e. if I is strictly convex at $\boldsymbol{a}$, $(j_{\boldsymbol{a}}, \rho_{\boldsymbol{a}}) = (j_{\boldsymbol{\gamma}}, \rho_{\boldsymbol{\gamma}})$ for $\boldsymbol{\gamma} = \boldsymbol{\nabla} I(a)$. This is in agreement with the equivalence of the microcanonical process and the effective process/driven process for $\boldsymbol{\gamma} = \boldsymbol{\nabla} I(a)$ [81]. In the following, we recover this result through

direct calculation of the optimum of

$$\begin{aligned}\mathcal{F}(j,\rho) &= -I_{2.5}(j,\rho) + \gamma_1\ A_1(j) + \gamma_2\ A_2(\rho) \\ &- \frac{1}{T}\int_0^T \mathrm{d}\tau\ c(\tau)\ \left[\ \int \mathrm{d}x \rho(x,\tau) - 1\ \right] \\ &- \frac{1}{T}\int_0^T \mathrm{d}\tau \int \mathrm{d}x u(x,t)\left[\dot{\rho}(x,\tau) + \nabla j(x,\tau)\ \right], \end{aligned} \tag{3.180}$$

where c and u are the Lagrange multipliers associated with the constraints **C1** and **C2** respectively. We assume in addition that $u(T) = u(0)$ in line with the constraints **C3** and **C4**. Functional derivatives with respect to the empirical occupation and empirical current yield

$$\begin{cases} \frac{\delta \mathcal{F}}{\delta j(x,\tau)} = 0 \Rightarrow -\frac{1}{T}\frac{1}{\rho D}(j - J^\rho) + \frac{1}{T}\gamma_1 g + \frac{1}{T}\nabla u = 0, \\ \frac{\delta \mathcal{F}}{\delta \rho(x,\tau)} = 0 \Rightarrow \frac{1}{T}\frac{1}{2\rho^2 D}(j - J^\rho)^2 - \int_x \frac{1}{2\rho D}\frac{\delta}{\delta \rho}(j - J^\rho)^2 - \frac{1}{T}c + \frac{1}{T}\partial_t u + \frac{1}{T}\gamma_2 h = 0, \end{cases} \tag{3.181}$$

where we dropped for clarity the (x,τ) dependence of the functions in the right-hand side of Eq. (3.181). The first equation leads to

$$j = \hat{B}'_{\boldsymbol{\gamma}}\rho - \frac{1}{2}D\nabla\rho, \tag{3.182}$$

where we introduced

$$\hat{B}'_{\boldsymbol{\gamma}} \equiv \hat{b} + D(\gamma_1 g + \nabla u). \tag{3.183}$$

The current j is then associated with an operator $\mathcal{L}' \equiv -\nabla\left[\hat{B}'_{\boldsymbol{\gamma}} - \frac{1}{2}D\nabla\right]$ so that $\dot{\rho} = \mathcal{L}'\rho$ by condition **C2**. From condition **C3**, ρ is the TiPS probability associated with $\mathcal{L}'$. As suggested by the notation, $\mathcal{L}'$ will turn out to be the generator $\mathcal{L}$ of the driven process defined in Eq. (3.152).

Using an integration by parts and the fact that $j - J^\rho = D\rho(\gamma_1 g + \nabla u)$, the second equation of (3.181) becomes

$$-\dot{u} = -c + \gamma_2 h + \frac{1}{2}D(\gamma_1 g + \nabla u)^2 + \hat{b}(\gamma_1 g + \nabla u) + \frac{1}{2}\nabla\left[D(\gamma_1 g + \nabla u)\right]. \tag{3.184}$$

After computing explicitly the action of $\Lambda^\dagger$ of Eq. (3.110) on e^u:

$$\Lambda^\dagger \mathrm{e}^u = \hat{b}(\gamma_1 g + \nabla u)\mathrm{e}^u + \frac{1}{2}D(\gamma_1 g + \nabla u)^2 \mathrm{e}^u + \frac{1}{2}\nabla\left[D(\gamma_1 g + \nabla u)\right]\mathrm{e}^u + \gamma_2 h \mathrm{e}^u, \tag{3.185}$$

we remark that Eq. (3.184) reads

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\left(e^u\right) = -\left(\Lambda - c\right)^\dagger\left(e^u\right), \\ e^{u(0)} = e^{u(T)}, \end{cases} \tag{3.186}$$

with $(\mathrm{e}^u)(x,t) \equiv \mathrm{e}^{u(x,t)}$. Using the definitions of Appendix. 3.A, the formal solution of Eq. (3.186) reads:

$$\mathrm{e}^{u(t)} = \left[\overleftarrow{\mathbf{Q}}_{\Lambda - c}(T,t)\right]^\dagger \mathrm{e}^{u(T)} = \left[\overleftarrow{\mathbf{Q}}_{\Lambda}(T,t)\right]^\dagger \mathrm{e}^{-\int_t^T c}\ \mathrm{e}^{u(0)}, \tag{3.187}$$

where we used Property 6 in the second equality. Taking $t = 0$, it follows

$$\left[\overleftarrow{\mathfrak{Q}}_\Lambda(T,0)\right]^\dagger \mathrm{e}^{u(0)} = \mathrm{e}^{\int_0^T c}\, \mathrm{e}^{u(0)}. \tag{3.188}$$

Thus, the optimization with respect to ρ leads to a spectral equation. Since $\mathrm{e}^{u(0)}$ is a positive function, by Krein-Rutman theorem it is the unique — up to a multiplicative constant — left eigenfunction of $\overleftarrow{\mathfrak{Q}}_\Lambda(T,0)$ associated with its largest eigenvalue $\mathrm{e}^{\int_0^T c} = \chi_T$. From Eq. (3.121), we find that the SCGF reads $\Gamma = \frac{1}{T}\int_0^T c$. Notice that we can rewrite Eq. (3.186) as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{u+\int_t^T c}\right) = -\Lambda^\dagger(e^{u+\int_t^T c}). \tag{3.189}$$

Hence the vector $e^{u+\int_t^T c}$ is solution of

$$\begin{cases} \dot{X} = -\Lambda^\dagger X, \\ X(T) = \chi_T^{-1} X(0). \end{cases} \tag{3.190}$$

From Eqs. (3.150–3.151), we conclude that the function l that appears in the Doob transform leading to the driven generator is related to the Lagrange multipliers through

$$l(t) = e^{u(t)+\int_t^T c}. \tag{3.191}$$

We now show that the operator $\mathcal{L}'$ generates the driven process, i.e. we show that $\mathcal{L}'$ is given by the Doob transform of Λ associated with the function l. To do so, we compute explicitly $\Lambda^l f$ for an arbitrary function f in the one hand, and $\mathcal{L}' f$ on the other hand, and using Eqs. (3.182–3.184), we obtain after long calculation that

$$\mathcal{L}' f = l(\Lambda l^{-1} f) - l^{-1}(\Lambda^\dagger l) f. \tag{3.192}$$

Hence, $\mathcal{L}'$ is obtained from the Doob transform of Λ using l and we conclude that $\mathcal{L}' = \mathcal{L}$ is the generator of the driven process defined in Eq. (3.152). Another way to deduce this result is to notice that the generator $\mathcal{L}' = -\nabla\left[\hat{B}'_{\boldsymbol{\gamma}} - \frac{1}{2}D\nabla\right]$ is a Fokker-Planck operator with diffusion coefficient σ and drift (3.183)

$$\hat{B}'_{\boldsymbol{\gamma}} = \hat{b} + D(\gamma_1 g + \nabla u) = \hat{b} + D(\gamma_1 g + \nabla \ln l), \tag{3.193}$$

where we used Eq. (3.191) in the last equality, implying that $\hat{B}'_{\boldsymbol{\gamma}} = \hat{B}_{\boldsymbol{\gamma}}$ is indeed the drift of the driven process (3.154). It follows that the optimum of Eq. (3.179) is reached for $\rho = \mu$ the TiPS probability density of the driven process of generator $\mathcal{L}$, and for $j = \hat{B}_{\boldsymbol{\gamma}}\mu - \frac{1}{2}D\nabla\mu$ the probability current associated with the probability density μ and operator $\mathcal{L}$, both related by $\dot{\mu} = -\nabla j = \mathcal{L}\mu$. In the time-homogeneous case, we recover the results of [81].

To conclude this section, the driven process is the most probable process that generates the dynamics leading to the imposed value of the conditioning observable. In other words, it is the generator of the Markov process for which the conditioning observable takes asymptotically the imposed value as a typical value.

3.3 Conclusion

In this chapter, we addressed the problem of process conditioning for observables defined through periodic functions in the framework of Markov processes with time-periodic generators. We took the period of these functions equal to the period of the generator, with no loss of generality compared to the case of commensurable periods. Starting from nonequilibrium path probabilities generalizing the canonical and microcanonical ensembles, we defined the Markov generator of the canonical process and its asymptotic limit after a large number of periods. The latter is the driven generator obtained from the Doob transform involving the left eigenvector of the one-period propagator for the tilted operator (and its time evolution). This is consistent with the time-homogeneous theory where the eigenvector of the tilted matrix is involved instead. Finally, the effective process for which $\boldsymbol{A}$ takes asymptotically the microcanonical value $\boldsymbol{a}$ follows from the driven process. This result requires the equivalence between microcanonical and canonical path ensembles which is guaranteed by the convexity of the LDF $I(\boldsymbol{a})$.

We can sum up the problem of conditioning in the following sentence: *"The rare fluctuation of one process is the typical value of another process"* [241]. The problem of conditioning is interesting from the fundamental and theoretical point of view as it provides a generalization of equilibrium concepts. It defines nonequilibrium microcanonical and canonical path ensembles and studies their equivalence. This equivalence relies on the convexity of the LDF for $\boldsymbol{A}$, in straight connection with the concavity of the entropy for the equivalence of equilibrium ensembles. This analogy between entropy and LDF is broader than the question of ensemble equivalence. In the same way that the canonical state probability follows from Jaynes' maximum entropy principle in equilibrium statistical mechanics, the driven process follows from a constrained optimization problem on the 2.5 LDF. Another nonequilibrium generalization is mentioned when the LDF is convex but not strictly at one point (the LDF has linear parts, or several points with the same slope). This case of partial equivalence is identified with a phase transition corresponding to the coexistence of multiple values $\boldsymbol{a}$ of the observable for one value of $\boldsymbol{\gamma}_c = \boldsymbol{\nabla} I(\boldsymbol{a})$ [230, 231]. For this reason, there is not, strictly speaking, an effective driven process $\boldsymbol{K}(\boldsymbol{\gamma}_c)$ for which $\boldsymbol{A}$ will have one identified typical value corresponding to the conditioning value of the microcanonical process. Yet, when varying the parameter $\boldsymbol{\gamma}$ across $\boldsymbol{\gamma}_c$, the set of typical values of the driven process will "jump" from one typical value for $\boldsymbol{A}$ to another (given the strict convexity of I at $\boldsymbol{\gamma} \neq \boldsymbol{\gamma}_c$), with a phase coexistence of typical values at $\boldsymbol{\gamma} = \boldsymbol{\gamma}_c$.

Appendices

"L'absence de preuve n'est pas preuve de l'absence."

William Wright

3.A Time-ordered exponential: definitions and properties

Definitions

Let $\boldsymbol{M}$ be a linear operator (e.g.: transition rate matrix or Fokker-Planck operator). The ordered exponential $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0) \equiv \overleftarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$ is the unique solution of the initial value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{M}(t)\boldsymbol{X}(t), \quad \text{with } \boldsymbol{X}(0) = \mathbb{1}, \tag{3.194}$$

that has the integral form

$$\boldsymbol{X}(t) = \mathbb{1} + \int_0^t \boldsymbol{M}(t')\boldsymbol{X}(t')\mathrm{d}t'. \tag{3.195}$$

Inserting this integral form into itself, one obtains the series expansion of the time ordered exponential

$$\begin{aligned}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0) &= \mathbb{1} + \int_0^t \boldsymbol{M}(t_1)\mathrm{d}t_1 + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, \boldsymbol{M}(t_1)\boldsymbol{M}(t_2) \\ &\quad + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3\, \boldsymbol{M}(t_1)\boldsymbol{M}(t_2)\boldsymbol{M}(t_3) + \ldots \end{aligned} \tag{3.196}$$

Notice that the arrow on the exponential specifies the ordering of the product of $\boldsymbol{M}$ in the expansion for increasing time from right to left.

The reverse-ordered exponential $\overrightarrow{\mathcal{Q}}_{\boldsymbol{M}}(0,t) \equiv \overrightarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$ is unique solution of the initial value problem:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{X}(t)\boldsymbol{M}(t), \quad \text{with } \boldsymbol{X}(0) = \mathbb{1}, \tag{3.197}$$

that has the integral form

$$\boldsymbol{X}(t) = \mathbb{1} + \int_0^t \boldsymbol{X}(t')\boldsymbol{M}(t')\mathrm{d}t'. \tag{3.198}$$

Inserting this integral form into itself, one obtains the series expansion of the reverse-ordered exponential

$$\begin{aligned}\overrightarrow{\mathcal{Q}}_{\boldsymbol{M}}(0,t) &= \mathbb{1} + \int_0^t \boldsymbol{M}(t_1)\mathrm{d}t_1 + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2\, \boldsymbol{M}(t_2)\boldsymbol{M}(t_1) \\ &\quad + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3\, \boldsymbol{M}(t_3)\boldsymbol{M}(t_2)\boldsymbol{M}(t_1) + \dots \end{aligned} \tag{3.199}$$

Notice that the arrow on the exponential specifies the ordering of the product of $\boldsymbol{M}$ in the expansion for increasing time from left to right.

Properties of the time-ordered exponential

For the convenience of the reader, we recall useful properties on linear differential equations. See Ref. [242] for a full description of the theory.

Property 1 (Adjoint of a propagator) *The adjoint of a propagator based on the operator $\boldsymbol{M}$ is the reverse-ordered propagator based on the adjoint operator $\boldsymbol{M}^\dagger$:*

$$\left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0)\right]^\dagger = \overrightarrow{\mathcal{Q}}_{\boldsymbol{M}^\dagger}(0,t). \tag{3.200}$$

Note that if the operator $\boldsymbol{M}$ is a matrix, the adjoint is the transpose. This property follows from the definitions via the series expansions and the fact that the adjoint of a product of two linear operators is the product of the two adjoint operators taken in the reverse order.

Property 2 (Inverse of a propagator) *The inverse of a propagator based on the operator $\boldsymbol{M}$ is the reverse-ordered propagator based on the opposite operator $-\boldsymbol{M}$:*

$$\left[\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0)\right]^{-1} = \overrightarrow{\mathcal{Q}}_{-\boldsymbol{M}}(0,t). \tag{3.201}$$

Indeed, starting from $\dot{\boldsymbol{X}} = \boldsymbol{M}\boldsymbol{X}$ and using $\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{X}\boldsymbol{X}^{-1}) = \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}\boldsymbol{X}^{-1} + \boldsymbol{X}\frac{\mathrm{d}\boldsymbol{X}^{-1}}{\mathrm{d}t} = 0$, it follows that $\frac{\mathrm{d}\boldsymbol{X}^{-1}}{\mathrm{d}t} = -\boldsymbol{X}^{-1}\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t}\boldsymbol{X}^{-1} = -\boldsymbol{X}^{-1}\boldsymbol{M}$. Note that the invertibility of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0)$ is ensured by the Abel-Jacobi-Liouville identity which states that the determinant of $\boldsymbol{X}(t)$ solution of (3.194) is given by

$$\det \boldsymbol{X}(t) = \det \boldsymbol{X}(0)\, \mathrm{e}^{\int_0^t \mathrm{tr}M(t')\mathrm{d}t'}. \tag{3.202}$$

Since $\boldsymbol{X}(0) = \mathbb{1}$, it follows that the determinant of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0)$ is nonzero.

Property 3 (First relation between final and initial value problems) *The solution of the final value problem*

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = -\boldsymbol{X}(t)\boldsymbol{M}(t), \quad \text{with } \boldsymbol{X}(t_\mathrm{f}) = \mathbb{1}, \tag{3.203}$$

is given by $\boldsymbol{X}(t) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_\mathrm{f}, t)$.

Indeed, one can check directly that

$$\frac{\mathrm{d}}{\mathrm{d}t}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t) = \lim_{a\to 0}\frac{\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t+a) - \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t+a)\,\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t+a,t)}{a} \tag{3.204}$$

$$= \lim_{a\to 0}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t+a)\,\frac{\mathbb{1}-\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t+a,t)}{a} \tag{3.205}$$

$$= \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t)\Big[-\frac{\mathrm{d}}{\mathrm{d}s}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(s,t)\mid_{s=t}\Big] \tag{3.206}$$

$$= -\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t)\;\boldsymbol{M}\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,t) \tag{3.207}$$

$$= -\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t)\;\boldsymbol{M}. \tag{3.208}$$

Property 4 (Second relation between final and initial value problems) *The solution of the final value problem:*

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{X}(t) = \boldsymbol{M}(t)\boldsymbol{X}(t), \quad \text{with } \boldsymbol{X}(t_{\mathrm{f}}) = \mathbb{1}, \tag{3.209}$$

is given by $\boldsymbol{X}(t) = \Big[\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_{\mathrm{f}},t)\Big]^{-1} = \overrightarrow{\mathcal{Q}}_{-\boldsymbol{M}}(t,t_{\mathrm{f}})$.

This follows from combining properties 2 and 3.

Property 5 (Multiplicative property of propagators) *For any* $t_1 \in [t_0,t]$,

$$\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,t_0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,t_1)\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t_1,t_0). \tag{3.210}$$

Property 6 (Time-ordered exponential of the sum of commuting operators) *If* $\boldsymbol{M}(t_1)$ *and* $\boldsymbol{N}(t_2)$ *commute for any* t_1,$t_2 \in \mathbb{R}$, *then* $\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}+\boldsymbol{N}}(t,t_0) = \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,t_0)\overleftarrow{\mathcal{Q}}_{\boldsymbol{N}}(t,t_0)$.

Let us denote the left-hand side of the equality by $\boldsymbol{X}(t)$ and the right-hand side by $\boldsymbol{Y}(t)$. On the one hand, $\dot{\boldsymbol{X}} = (\boldsymbol{M}+\boldsymbol{N})\boldsymbol{X}$. On the other hand, $\dot{\boldsymbol{Y}} = \boldsymbol{M}\boldsymbol{Y} + \overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,t_0)\boldsymbol{N}\overleftarrow{\mathcal{Q}}_{\boldsymbol{N}}(t,t_0) = (\boldsymbol{M}+\boldsymbol{N})\boldsymbol{Y}$ since $\boldsymbol{M}$ and $\boldsymbol{N}$ commute for any time. Thus, the operators $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfy the same differential equation. Besides, $\boldsymbol{X}(t_0) = \boldsymbol{Y}(t_0) = \mathbb{1}$, hence $\boldsymbol{X}(t) = \boldsymbol{Y}(t)$, $\forall t \in \mathbb{R}$.

3.B Perron-Frobenius theorems

Definitions

Let $\boldsymbol{M} = (M_{ij})$ be a square matrix of dimension N defined on the set of real numbers.

Definition 1 (Positive matrices) $\boldsymbol{M}$ *is positive if* $M_{ij} > 0$, $\forall i,j$.

Definition 2 (Nonnegative matrices) $\boldsymbol{M}$ *is nonnegative if* $M_{ij} \geq 0$, $\forall i,j$.

Definition 3 (Irreducible matrices) *Suppose* $\boldsymbol{M}$ *is nonnegative.* $\boldsymbol{M}$ *is said to be irreducible if* $(\mathbb{1}+\boldsymbol{M})^{n-1}$ *is a positive matrix.*

Definition 4 (Primitive matrices) *Suppose* $\boldsymbol{M}$ *is nonnegative.* $\boldsymbol{M}$ *is said to be primitive if there exists* $\ell \in \mathbb{N}$ *such that* $\boldsymbol{M}^{\ell}$ *is positive.*

Definition 5 (Metzler matrices) *The matrix* $\boldsymbol{M}$ *is said to be a Metzler matrix if its off-diagonal entries are nonnegative:* $M_{ij} \geq 0$, $\forall i \neq j$.

Perron-Frobenius (PF) theorems

Theorem 1 (PF theorem for positive matrices) *Let $\boldsymbol{M}$ be a positive matrix. Then [243, 244]:*

- *$\boldsymbol{M}$ admits a positive dominant eigenvalue χ called the Perron-Frobenius eigenvalue. Thus, χ is the spectral radius of M: $\mid s \mid < \chi$ for any other eigenvalue s of $\boldsymbol{M}$.*
- *χ is a simple eigenvalue: the eigenspace associated with χ is one-dimensional and is generated by one positive (right) eigenvector $\boldsymbol{r}$ called the Perron-Frobenius eigenvector.*
- *There is no positive eigenvector of $\boldsymbol{M}$ other than positive multiples of $\boldsymbol{r}$.*
- *Since $\boldsymbol{M}$ and its adjoint $\boldsymbol{M}^\dagger$ have the same characteristic polynomial $\det(s\mathbb{1} - \boldsymbol{M})$, the same results apply to $\boldsymbol{M}^\dagger$. Hence $\boldsymbol{M}$ has a unique positive left eigenvector $\boldsymbol{l}$ associated with the eigenvalue χ.*

Theorem 2 (PF theorem for irreducible nonnegative matrices) *Suppose M is a nonnegative matrix. If $\boldsymbol{M}$ is irreducible, then Theorem 1 holds [243].*

Theorem 3 (PF theorem for primitive matrices) *Suppose $\boldsymbol{M}$ a is primitive matrix. Then Theorem 1 applies to $\boldsymbol{M}$ [243].*

Theorem 4 (PF for general nonnegative matrices) *If $\boldsymbol{M}$ is a nonnegative matrix then [243, 244]:*

- *$\boldsymbol{M}$ admits a nonnegative eigenvalue $\chi \geq 0$ such that $\mid s \mid \leq \chi$ for any other eigenvalue s of $\boldsymbol{M}$.*
- *χ is associated with a nonnegative right eigenvector $\boldsymbol{r}$: $\boldsymbol{r} \geq 0$ with $\boldsymbol{r} \neq 0$.*
- *χ is associated with a nonnegative left eigenvector $\boldsymbol{l}$: $\boldsymbol{l} \geq 0$ with $\boldsymbol{l} \neq 0$.*
- *The eigenvectors $\boldsymbol{l}$ and $\boldsymbol{r}$ are not necessarily unique or positive.*

Theorem 5 (PF theorem for Metzler matrices) *Suppose $\boldsymbol{M}$ is a Metzler matrix. Then we have [244]:*

- *$\mathrm{e}^{\boldsymbol{M}t}$ is nonnegative.*
- *There exists $\alpha \in \mathbb{R}$ such that $\boldsymbol{M} + \alpha I$ is nonnegative.*

Hence, applying the Perron-Frobenius theorem for nonnegative matrices on $\boldsymbol{M} + \alpha I$ yields:

- *There exists a real dominant eigenvalue χ of $\boldsymbol{M}$ such that $Re(s) \leq \chi$ for any other eigenvalue s of $\boldsymbol{M}$, with $Re(s)$ the real part of s.*
- *The real dominant eigenvalue χ of $\boldsymbol{M}$ is associated with nonnegative (and non-zero) left and right eigenvectors.*
- *If $\boldsymbol{r}$ is a positive eigenvector of $\boldsymbol{M}$ then $\boldsymbol{r}$ is associated with χ.*

If we assume furthermore that $\boldsymbol{M}$ is irreducible, then [244]:

- χ is a simple eigenvalue.
- χ is associated with a unique positive right eigenvector $\boldsymbol{r}$ and a unique positive left eigenvector $\boldsymbol{l}$.

Theorem 6 (Krein–Rutman theorem) *The results above are valid for finite spaces. For infinite spaces, one refers to the Krein–Rutman theorem [235, 236] which generalizes the Perron–Frobenius theorem to infinite-dimensional Banach spaces. Let $\boldsymbol{M}$ be an operator with positive spectral radius χ and satisfying the conditions stated in the Krein–Rutman theorem. Then (Theorem 19.3 of Ref. [245])*

- *χ is a simple eigenvalue of $\boldsymbol{M}$ with positive eigenvector.*
- *There is no other eigenvalue with positive eigenvector.*
- *$|s| < \chi$ for any other eigenvalue s.*

3.C Derivation of level 2.5 large deviation functions

The following derivations are done for periodically driven processes implying time-periodic generators. We present the derivation for both jump processes and diffusion processes.

Level 2.5 large deviation function of jump processes

In the following, we derive heuristically the 2.5 LDF of empirical transition probability and empirical occupation of Eq. (3.67). For a more rigorous derivation, see Ref. [234]. The probability of observing an empirical transition probability $\boldsymbol{\omega}^n[z] = \boldsymbol{\omega}$ and an empirical occupation $\boldsymbol{p}^n[z] = \boldsymbol{p}$ for a our reference process of generator $\boldsymbol{k}$ is given by:

$$P_n(\boldsymbol{\omega}, \boldsymbol{p}) = \int \mathfrak{D}[z] \mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] \delta(\boldsymbol{p}^n[z] - \boldsymbol{p}) \delta(\boldsymbol{\omega}^n[z] - \boldsymbol{\omega}), \tag{3.211}$$

where the path probability (2.13) can be rewritten as:

$$\begin{aligned}
\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z] &= \pi_{z_0}(0) \exp\left[\sum_{t\in[0,nT]|z(t^+)\neq z(t^-)} \ln\left(k_{z(t^+)z(t^-)}(t)\delta t\right) - \int_0^{nT} \sum_{x\neq z(t)} k_{xz(t)}(t) \mathrm{d}t \right] \\
&= \pi_{z_0}(0) \exp \sum_{\ell=0}^{n-1} \left\{ \sum_{t\in[\ell T,(\ell+1)T]|z(t^+)\neq z(t^-)} \ln\left(k_{z(t^+)z(t^-)}(t)\delta t\right) \right. \\
&\qquad \left. - \int_{\ell T}^{(\ell+1)T} \mathrm{d}t \sum_{x\neq z(t)} k_{xz(t)}(t) \right\} \\
&= \pi_{z_0}(0) \exp \sum_{\ell=0}^{n-1} \left\{ \sum_{\tau\in[0,T]|z(\tau^+ + \ell T)\neq z(\tau^- + \ell T)} \ln\left(k_{z(\tau^+ + \ell T)z(\tau^- + \ell T)}(\tau)\delta t\right) \right. \\
&\qquad \left. - \int_0^{T} \mathrm{d}\tau \sum_{x\neq z(\tau+\ell T)} k_{xz(\tau+\ell T)}(\tau) \right\},
\end{aligned}$$

where we used the periodicity of $\boldsymbol{k}$ in the last equation. From Eq. (3.65), $\boldsymbol{p}$ and $\boldsymbol{\omega}$ need to be related by

$$\dot{p}_x = \sum_{y \neq x} [\omega_{xy} - \omega_{yx}], \tag{3.212}$$

otherwise $P_n(\boldsymbol{\omega}, \boldsymbol{p}) = 0$; and in the limit of large number of periods, Eq. (3.59) and Eq. (3.63) require by continuity that $\boldsymbol{p}(T) = \boldsymbol{p}(0)$ and $\boldsymbol{\omega}(T) = \boldsymbol{\omega}(0)$. Let us introduce an auxiliary process whose T-periodic generator $\boldsymbol{k}'$ will be chosen such that $\boldsymbol{\omega}$ and $\boldsymbol{p}$ are typical for the dynamics it generates. The path probability $\mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z]$ associated with $\boldsymbol{k}'$ is obtained by replacing $\boldsymbol{k}$ by $\boldsymbol{k}'$ in Eq. (3.212). We can rewrite Eq. (3.211) as

$$P_n(\boldsymbol{\omega}, \boldsymbol{p}) = \int \mathfrak{D}[z] \mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z] \; \mathrm{e}^{-\mathcal{A}[z]} \delta(\boldsymbol{p}^n[z] - \boldsymbol{p}) \delta(\boldsymbol{\omega}^n[z] - \boldsymbol{\omega}), \tag{3.213}$$

where we introduced

$$\mathcal{A}[z] \equiv \ln \frac{\mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z]}{\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]}. \tag{3.214}$$

It follows from Eq. (3.212):

$$\begin{aligned}
\mathcal{A}[z] &= \sum_{\ell=0}^{n-1} \Bigg\{ \sum_{\tau \in [0,T] | z(\tau^+ + \ell T) \neq z(\tau^- + \ell T)} \ln \frac{k'_{z(\tau^+ + \ell T) z(\tau^- + \ell T)}(\tau)}{k_{z(\tau^+ + \ell T) z(\tau^- + \ell T)}(\tau)} \mathrm{d}t \\
&\qquad - \int_0^T \mathrm{d}\tau \sum_{x \neq z(\tau + \ell T)} \left[k'_{x z(\tau + \ell T)}(\tau) - k_{x z(\tau + \ell T)}(\tau) \right] \Bigg\} \\
&= \sum_{\ell=0}^{n-1} \Bigg\{ \sum_{\tau \in [0,T] | z(\tau^+ + \ell T) \neq z(\tau^- + \ell T)} \sum_{x,y} \delta_{x, z(\tau^+ + \ell T)} \delta_{y, z(\tau^- + \ell T)} \ln \frac{k'_{xy}(\tau)}{k_{xy}(\tau)} \\
&\qquad - \int_0^T \mathrm{d}\tau \sum_y \delta_{y, z(\tau + \ell T)} \sum_{x \neq y} \left[k'_{xy}(\tau) - k_{xy}(\tau) \right] \Bigg\} \\
&= \sum_{\ell=0}^{n-1} \Bigg\{ \int_0^T \mathrm{d}\tau \sum_{s \in [\tau, \tau + \mathrm{d}\tau [| z(s^+ + \ell T) \neq z(s^- + \ell T)} \frac{1}{\mathrm{d}\tau} \sum_{x,y} \delta_{x, z(s^+ + \ell T)} \delta_{y, z(s^- + \ell T)} \ln \frac{k'_{xy}(\tau)}{k_{xy}(\tau)} \\
&\qquad - \int_0^T \mathrm{d}\tau \sum_{y, x \neq y} p^n_y[z] \left[k'_{xy}(\tau) - k_{xy}(\tau) \right] \Bigg\} \\
&= n \int_0^T \mathrm{d}\tau \sum_{y, x \neq y} \left[\omega^n_{xy}[z](\tau) \ln \frac{k'_{xy}(\tau)}{k_{xy}(\tau)} + p^n_y[z](\tau) \; \left(k_{xy}(\tau) - k'_{xy}(\tau) \right) \right].
\end{aligned} \tag{3.215}$$

where we used $\sum_{\tau \in [0,T] | z(\tau^+ + \ell T) \neq z(\tau^- + \ell T)} = \int_0^T \sum_{s \in [\tau, \tau + \mathrm{d}\tau [| z(s^+ + \ell T) \neq z(s^- + \ell T)}$ and the definitions of the empirical occupation (3.59) and empirical transition probability (3.63) in the third and last equations. We now choose $\boldsymbol{k}'$ for which $\boldsymbol{p}$ and $\boldsymbol{\omega}$ are typical in the long-time limit, i.e. such that

$$\boldsymbol{p}^n[z] \xrightarrow[n \to \infty]{\boldsymbol{k}'} \boldsymbol{p}, \tag{3.216}$$

$$\boldsymbol{\omega}^n[z] \xrightarrow[n \to \infty]{\boldsymbol{k}'} \boldsymbol{\omega}, \tag{3.217}$$

where the arrows refer to the convergence in probability with respect to $\mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z]$. This limit is verified for $\boldsymbol{k}'$ satisfying $\boldsymbol{\omega} = \boldsymbol{k}' \odot \boldsymbol{p}$. We finally get :

$$P_n(\boldsymbol{\omega}, \boldsymbol{p}) = \mathrm{e}^{-nTI_{2.5}(\boldsymbol{\omega},\boldsymbol{p})} \int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z]\delta(\boldsymbol{p}^n[z] - \boldsymbol{p})\delta(\boldsymbol{\omega}^n[z] - \boldsymbol{\omega}), \tag{3.218}$$

where we defined $I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p})$ by

$$I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p}) = \frac{1}{T}\int_0^T \mathrm{d}\tau \sum_{y,x\neq y} \left[p_y(\tau)\left(k_{xy}(\tau) - \frac{\omega_{xy}(\tau)}{p_y(\tau)}\right) + \omega_{xy}(\tau)\ln\frac{\omega_{xy}(\tau)}{k_{xy}(\tau)p_y(\tau)}\right]. \tag{3.219}$$

Since $\boldsymbol{k}'$ has been chosen such that Eqs. (3.216–3.217) are satisfied, we have

$$\int \mathfrak{D}[z]\mathbb{P}_{\boldsymbol{k}',\boldsymbol{\pi}(0)}[z]\delta(\boldsymbol{p}^n[z] - \boldsymbol{p})\delta(\boldsymbol{\omega}^n[z] - \boldsymbol{\omega}) \underset{n\to\infty}{\asymp} 1 \tag{3.220}$$

leading to

$$P_n(\boldsymbol{\omega}, \boldsymbol{p}) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nTI_{2.5}(\boldsymbol{\omega},\boldsymbol{p})} \tag{3.221}$$

when Eq. (3.212) is satisfied. Hence, $I_{2.5}(\boldsymbol{\omega}, \boldsymbol{p})$ is the 2.5 LDF of empirical occupation and empirical transition probability.

Level 2.5 large deviation function of diffusion processes

In the following, we derive heuristically the 2.5 LDF of empirical current and empirical occupation density of Eq. (3.176) [240].
The probability of observing an empirical current $j^n[z] = j$ and an empirical occupation density $\rho^n[z] = \rho$ for a our reference process of generator $\mathfrak{L}$ — or equivalently of drift b and diffusion coefficient σ — is given by:

$$P_n(j, \rho) = \int \mathfrak{D}[z]\mathbb{P}_{b,\sigma,\varrho(0)}[z]\delta(\rho^n[z] - \rho)\delta(j^n[z] - j), \tag{3.222}$$

where the path probability (2.63) can be rewritten as:

$$\begin{aligned}
\mathbb{P}_{b,\sigma,\varrho(0)}[z] &= \varrho(z_0, 0)\exp\left\{-\int_0^{nT}\mathrm{d}t\left[\frac{1}{2D(z_t,t)}\left(\dot{z}_t - \hat{b}(z_t,t)\right)^2 + \frac{1}{2}\nabla b(z_t,t)\right]\right\}. \\
&= \varrho(z_0, 0)\exp\sum_{\ell=0}^{n-1}\left\{-\int_{\ell T}^{(\ell+1)T}\mathrm{d}t\left[\frac{1}{2D(z_t,t)}\left(\dot{z}_t - \hat{b}(z_t,t)\right)^2 + \frac{1}{2}\nabla b(z_t,t)\right]\right\} \\
&= \varrho(z_0, 0)\exp\sum_{\ell=0}^{n-1}\left\{-\int_0^{T}\mathrm{d}\tau\left[\frac{1}{2D(z_{\tau+\ell T},\tau)}\left(\dot{z}_{\tau+\ell T} - \hat{b}(z_{\tau+\ell T},\tau)\right)^2\right.\right. \\
&\qquad\qquad \left.\left. + \frac{1}{2}\nabla b(z_{\tau+\ell T},\tau)\right]\right\},
\end{aligned} \tag{3.223}$$

where we used the periodicity of b and σ in the last equation. From Eq. (3.174), ρ and j need to be related by

$$\dot{\rho} = -\nabla j, \tag{3.224}$$

otherwise $P_n(j,\rho) = 0$; and in the limit of large number of periods, Eq. (3.171) and Eq. (3.172) require by continuity that $\rho(T) = \rho(0)$ and $j(T) = j(0)$. Let us introduce an auxiliary process whose T-periodic Fokker-Planck generator $\mathcal{L}'$ of drift B' and diffusion coefficient σ will be chosen such that j and ρ are typical for the dynamics it generates. The path probability $\mathbb{P}_{B',\sigma,\varrho(0)}[z]$ associated with $\mathcal{L}'$ is obtained by replacing b by B' in Eq. (3.223). We can rewrite Eq. (3.222) as

$$P_n(j,\rho) = \int \mathfrak{D}[z]\mathbb{P}_{B',\sigma,\varrho(0)}[z]\ \mathrm{e}^{-\mathcal{A}[z]}\delta(\rho^n[z]-\rho)\delta(j^n[z]-j), \tag{3.225}$$

where we introduced

$$\mathcal{A}[z] \equiv \ln \frac{\mathbb{P}_{B',\sigma,\varrho(0)}[z]}{\mathbb{P}_{b,\sigma,\varrho(0)}[z]}. \tag{3.226}$$

It follows from Eq. (3.223):

$$\begin{aligned}
\mathcal{A}[z] &= -\sum_{\ell=0}^{n-1}\int_0^T \mathrm{d}\tau \left\{ \frac{1}{2D(z_{\tau+\ell T},\tau)} \left[\left(\dot{z}_{\tau+\ell T} - \hat{B}'(z_{\tau+\ell T},\tau)\right)^2 - \left(\dot{z}_{\tau+\ell T} - \hat{b}(z_{\tau+\ell T},\tau)\right)^2 \right] \right. \\
&\qquad \left. + \frac{1}{2}\left[\nabla B'(z_{\tau+\ell T},\tau) - \nabla b(z_{\tau+\ell T},\tau)\right] \right\} \\
&= -\sum_{\ell=0}^{n-1}\int_0^T \mathrm{d}\tau \left\{ \frac{1}{2D(z_{\tau+\ell T},\tau)} \left[\left(2\dot{z}_{\tau+\ell T} - \hat{b}(z_{\tau+\ell T},\tau) - \hat{B}'(z_{\tau+\ell T},\tau)\right) \right.\right. \\
&\qquad \left.\left. \times \left(\hat{b}(z_{\tau+\ell T},\tau) - \hat{B}'(z_{\tau+\ell T},\tau)\right) \right] + \frac{1}{2}\left[\nabla B'(z_{\tau+\ell T},\tau) - \nabla b(z_{\tau+\ell T},\tau)\right] \right\} \\
&= -n\int_0^T \mathrm{d}\tau \int \mathrm{d}x \left\{ \frac{1}{2D(x,\tau)} \left[2j^n[z](x,\tau)\left(\hat{b}(x,\tau) - \hat{B}'(x,\tau)\right) \right.\right. \\
&\qquad \left.\left. - \rho^n[z](x,\tau)\left(\hat{b}(x,\tau)^2 - \hat{B}'(x,\tau)^2\right) \right] + \frac{1}{2}\left[\nabla B'(x,\tau) - \nabla b(x,\tau)\right]\rho^n[z](x,\tau) \right\},
\end{aligned} \tag{3.227}$$

where we used in the last equation the definitions of the empirical occupation density and empirical current:

$$\rho[z](x,\tau) = \frac{1}{n}\sum_{\ell=1}^{n-1}\delta(z_{\tau+\ell T} - x), \tag{3.228}$$

$$j^n[z](x,\tau) = \frac{1}{n}\sum_{\ell=1}^{n-1}\delta(z_{\tau+\ell T} - x)\circ \mathrm{d}\dot{z}_{\tau+\ell T}. \tag{3.229}$$

We now choose B' such that ρ and j are typical in the long-time limit for the dynamics generated by $\mathcal{L}'$, i.e. such that

$$\rho^n[z] \xrightarrow[n\to\infty]{\mathcal{L}'} \rho, \tag{3.230}$$

$$j^n[z] \xrightarrow[n\to\infty]{\mathcal{L}'} j, \tag{3.231}$$

where the arrows refer to the convergence in probability with respect to $\mathbb{P}_{B',\sigma,\varrho(0)}[z]$. This limit is verified for B' satisfying $j = \hat{B}'(x,t)\rho(x,t) - \frac{1}{2}D(x,t)\nabla\rho(x,t)$, and we have

$$\hat{B}'(x,t)\rho(x,t) - \hat{b}\rho(x,t) = j(x,t) - J^{\rho}(x,t), \tag{3.232}$$

where J^{ρ} is defined in Eq. (2.66). It yields:

$$P_n(j,\rho) = \mathrm{e}^{-nTI_{2.5}(j,\rho)} \int \mathfrak{D}[z]\mathbb{P}_{B',\sigma,\varrho(0)}[z]\delta(\rho^n[z]-\rho)\delta(j^n[z]-j), \tag{3.233}$$

where we have from Eq. (3.227) and Eq. (3.232) and using an integration by parts:

$$\begin{aligned}
TI_{2.5}(j,\rho) &= -\int_{\tau\in[0,T]}\int_x \left\{ \frac{1}{2D\rho}\left[2j\left(\hat{b}\rho - \hat{B}'\rho\right) - \rho^2\left(\hat{b}^2 - \hat{B}'^2\right)\right] - \frac{1}{2}\left[B'-b\right]\nabla\rho \right\} \\
&= -\int_{\tau\in[0,T]}\int_x \left\{ \frac{1}{2D\rho}\left[-2j\left(j-J^{\rho}\right) - \left((\hat{b}\rho)^2 - (\hat{B}'\rho)^2\right) - D\left[B'\rho - b\rho\right]\nabla\rho\right] \right\} \\
&= \int_{\tau\in[0,T]}\int_x \left\{ \frac{1}{2D\rho}\left[2j\left(j-J^{\rho}\right) - \left(j-J^{\rho}\right)\left(j+J^{\rho}+D\nabla\rho\right) + D\left(j-J^{\rho}\right)\nabla\rho\right] \right\} \\
&= \int_{\mathrm{d}\tau\in[0,T]}\int_x \left\{ \frac{1}{2D\rho}\left[\left(j-J^{\rho}\right)^2\right] \right\},
\end{aligned} \tag{3.234}$$

where we dropped for clarity the dependence on x and τ of all the functions in the integral. Since B' has been chosen such that Eqs. (3.230–3.231) are satisfied, we have $\int \mathfrak{D}[z]\mathbb{P}_{B',\sigma,\varrho(0)}[z]\delta(\rho^n[z]-\rho)\delta(j^n[z]-j) \underset{n\to\infty}{\asymp} 1$, leading to $P_n(j,\rho) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nTI_{2.5}(j,\rho)}$ when Eq. (3.224) is satisfied. Hence, $I_{2.5}(j,\rho)$ is the 2.5 LDF of empirical occupation density and empirical current.

Remark For diffusion processes of dimension $d > 1$, the previous calculation leads to the 2.5 LDF

$$I_{2.5}(\boldsymbol{j},\rho) = \frac{1}{T}\int_{\mathrm{d}\tau\in[0,T]}\int_{\boldsymbol{x}}\left[\frac{1}{2\rho}\,{}^t(\boldsymbol{j}-\boldsymbol{J}^{\rho})\boldsymbol{D}^{-1}(\boldsymbol{j}-\boldsymbol{J}^{\rho})\right], \tag{3.235}$$

where $\boldsymbol{x}$, $\boldsymbol{j}$, $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ are now d-dimensional vectors, $\boldsymbol{D} \equiv \boldsymbol{\sigma}\boldsymbol{\sigma}^{\dagger}$ is a $d \times d$ matrix and ∇ is the gradient vector.

3.D Lagrange multipliers

The method of Lagrange multipliers allows finding the extrema (maxima or minima) of a function under constraints. To proceed, the constraint on the variables of the function is converted into a constraint on new variables called *Lagrange multipliers*, in the same way the constraint on the energy in the equilibrium microcanonical ensemble translates into a constraint on the temperature in the equilibrium canonical ensemble. The method is summarized as follows [189]:

1. We want to find the optimizer of a function $F(v_1,\ldots,v_N)$ under M constraints

$$C_{\ell}(v_1,\ldots,v_N) = 0, \qquad \ell = 1,\ldots,M. \tag{3.236}$$

2. We introduce a new function $\tilde{F}(v_1, \dots, v_N; \gamma_1, \cdots \gamma_M)$:

$$\tilde{F}(v_1, \dots, v_N; \gamma_1, \cdots \gamma_M) \equiv F(v_1, \dots, v_N) + \sum_{\ell=1}^{M} \gamma_\ell C_\ell(v_1, \dots, v_N), \tag{3.237}$$

where the new variable γ_ℓ is the Lagrange multiplier associated with the constraint C_ℓ. $\tilde{F}$ is called the *Lagrange function*.

3. We compute the optimizer $\{v_1^*(\gamma_1, \dots, \gamma_M), \dots, v_N^*(\gamma_1, \dots, \gamma_M)\}$ of $\tilde{F}$ solutions of

$$\begin{cases} \frac{\partial \tilde{F}}{v_1}(v_1^*, \dots, v_N^*; \gamma_1, \cdots \gamma_M) = 0 \\ \vdots \\ \frac{\partial \tilde{F}}{v_N}(v_1^*, \dots, v_N^*; \gamma_1, \cdots \gamma_M) = 0 \end{cases} \tag{3.238}$$

4. We inject the optimizer $\{v_1^*(\gamma_1, \dots, \gamma_M), \dots, v_N^*(\gamma_1, \dots, \gamma_M)\}$ into the set of constraints $\{C_\ell\}$, leading to a set of equations on the Lagrange multipliers of solution $\{\gamma_1^*, \dots \gamma_M^*\}$.

5. The final solution of our problem is $\{v_1^*(\gamma_1^*, \dots, \gamma_M^*), \dots, v_N^*(\gamma_1^*, \dots, \gamma_M^*)\}$.

An example of application of the method of Lagrange multipliers is the calculation of the equilibrium microcanonical probability distribution. It follows from the maximization of the Shannon entropy under the constraint of probability normalization.

3.E Direct calculation of the SCGF for jump processes

We recover the SCGF of the observable $\boldsymbol{A}$ by evaluating the 2.5 LDF (3.67) at the optimum $(\boldsymbol{\omega}, \boldsymbol{p})$ of our variational problem stated in Eq. (3.70). By definition of the second component of our observable $\boldsymbol{A}$ in Eq. (3.68) and using Eq. (3.77) we find

$$\gamma_2 \, A_2(\boldsymbol{p}) = \frac{1}{T} \int_0^T \mathrm{d}t \sum_y p_y \gamma_2 h_y, \tag{3.239}$$

$$= \frac{1}{T} \int_0^T \mathrm{d}t \left[c + \sum_y p_y \left(\sum_{x \neq y} k_{xy} - \sum_{x \neq y} K'_{xy} \right) - \sum_y p_y \dot{u}_y \right], \tag{3.240}$$

$$= \frac{1}{T} \int_0^T \mathrm{d}t \left[c + \sum_y p_y \left(\sum_{x \neq y} k_{xy} - \sum_{x \neq y} K'_{xy} \right) + \sum_y \dot{p}_y u_y \right], \tag{3.241}$$

where we used **C3** in the integration by parts. Using Eqs. (3.73–3.74) and **C2**, the LDF at the optimum ($\boldsymbol{\omega} = \boldsymbol{K'} \odot \boldsymbol{p}, \boldsymbol{p}$) reads

$$I(\boldsymbol{\omega}, \boldsymbol{p}) = \frac{1}{T}\int_0^T \mathrm{d}t \left[\sum_y p_y \left(\sum_{x\neq y} k_{xy} - \sum_{x\neq y} K'_{xy}\right) + \sum_{x,y\neq x} \omega_{xy}\left[u_x - u_y\right] + \gamma_1 \sum_{x,y\neq x} \omega_{xy}\, g_{xy}\right] \tag{3.242}$$

$$= \frac{1}{T}\int_0^T \mathrm{d}t \left[\sum_y p_y \left(\sum_{x\neq y} k_{xy} - \sum_{x\neq y} K'_{xy}\right) + \sum_{x,y\neq x} u_x\left[\omega_{xy} - \omega_{yx}\right] + \gamma_1 \sum_{x,y\neq x} \omega_{xy}\, g_{xy}\right] \tag{3.243}$$

$$= \frac{1}{T}\int_0^T \mathrm{d}t \left[\sum_y p_y \left(\sum_{x\neq y} k_{xy} - \sum_{x\neq y} K'_{xy}\right) + \sum_x u_x\, \dot{p}_x\right] + \gamma_1 A_1(\boldsymbol{\omega}). \tag{3.244}$$

Combining Eqs. (3.241) and (3.244), we finally obtain

$$\gamma_1\ A_1(\boldsymbol{K'} \odot \boldsymbol{p}) + \gamma_2\ A_2(\boldsymbol{p}) - I_{2.5}(\boldsymbol{K'} \odot \boldsymbol{p}, \boldsymbol{p}) = \frac{1}{T}\int_0^T c(\tau)\mathrm{d}\tau. \tag{3.245}$$

We recognize the SCGF Γ in the left-hand side of Eq. (3.245) as the Legendre transform of the LDF. It follows that the SCGF is the time-average over a period of the Lagrange multiplier used to ensure the normalization of the empirical occupation, recovering the result stated in Eq. (3.83). The variational calculation of the SCGF is similar in many ways to the calculation of the equilibrium canonical probability via the maximum entropy principle in which the SCGF (free energy) also follows from the Lagrange multiplier that imposes probability normalization.

Chapter 4

Conditioning nonlinear Markov processes on large deviations

"Tu peux avoir tout l'argent du monde. Mais y'a un truc que tu n'auras jamais... Un dinosaure."

Homer Simpson

Previously, we extended the work of R. Chétrite and H. Touchette of conditioning and biasing time-homogeneous Markov processes to periodically driven Markov processes. This problem of conditioning relies on the theory of large deviations in the limit of large times. It has been shown that the driven process is obtained from the biased process by using the so-called generalized Doob transform. This transformation leads to a norm-conserving Markov generator starting from a biased generator that is not norm-conserving. We say that the generalized Doob transform *rectifies* the biased generator. For the purpose of this chapter, we call *linear operator formalism* the framework in which the latter results have been derived, as opposed to the *Hamiltonian/Lagrangian formalism* that will be introduced in the following [109–123].

In this chapter, we consider Markov processes for which the large deviations are studied in the limit of a large parameter that is not time, for instance volume or number of particles. In this limit, the generator of the Markov process becomes a function of an intensive variable (concentration, density, etc.). When it is linear in this variable, as for $\mathcal{N}$ independent Markov jump or diffusion processes, we talk about *linear Markov processes*. On the contrary, when the Markov generator is not linear in the intensive variable as it happens for population processes or interacting many-body Markov processes, we talk about *nonlinear Markov processes*.

In the following, we aim to generalize the Doob transform and the driven process considered in the linear operator formalism to general nonlinear Markov processes using the Hamiltonian framework. Considering this class of processes is motivated by its applications in a wide range of fields. Examples include biology [246, 247], biochemistry [141, 248–255], sociophysics [256–259] and psychology [260, 261].

This generalization relies on the theory of large deviations in which the scaling parameter is a size-type parameter. We will see how concepts of the linear operator formalism such as biasing, Doob transform (rectification) and Perron-Frobenius theorem translates within

the Hamiltonian framework.

The outline of this chapter is as follows. In Sections 4.1, 4.2 and 4.3, we deal only with time-independent generators. The case of time-periodic generator will be broached in Section 4.4 as an opening. In Section 4.1, we derive explicitly the Lagrangian/Hamiltonian formalism for the specific case of population processes starting from the well-known results of the linear operator formalism for time-homogeneous Markov jump processes [123, 124]. This derivation leads to the definition of Lagrangians and Hamiltonians describing both the non-biased and biased processes. The novelty consists in the derivation of the Lagrangians and Hamiltonians associated with the driven process, generalizing the Doob transform/rectification to nonlinear processes. On the basis of these results, we propose in Section 4.2 a general theory of the rectification of time-homogeneous nonlinear Markov processes within the Lagrangian/Hamiltonian formalism. In Section 4.3, we illustrate this theory with two specific models of population processes: the Brownian Donkey [119, 120, 262] and a chemical reaction network [124]. The main content of this chapter has been published in Ref. [263].

4.1 Markov population processes

A population process is a Markov jump process in which a state refers to the number of individuals in a population, and a transition between states amounts to adding or removing individuals from the population. For instance, a chemical reaction network can be modeled by a population process, the states being the number of molecules of each species and the transitions being the occurrence of a chemical reaction. Population processes find application in a large number of domains such as chemical kinetics, biology, ecology, finance, epidemiology, demography, queueing theory, etc. In the following, we review and adapt the Lagrangian/Hamiltonian formalism considered in Ref. [124] for chemical networks to general nonlinear rates. We then go a step further by describing within this formalism the driven process and the Doob transform from which it derives. The procedure leading to the derivation of the driven process in the Hamiltonian formalism is generically referred to as *the rectification*.

The following section is a pedagogical springboard to the general theory of Section 4.2 on the rectification of nonlinear Markov processes. For this reason, some elements are not dealt with in depth as they are in the general theory of Section 4.2. The interest of this section is to gradually introduce new concepts that might have seemed to come out of nowhere if we had started this chapter with the general theory. Here, we derive each object of the Lagrangian/Hamiltonian formalism from the linear operator formalism, thus giving a mapping between the two frameworks and some intuition to the definitions of the general theory.

We consider a many-body system modeled by a time-homogeneous Markov jump process defined on an infinite lattice. We denote by $\{X\}$ the states occupied by each particle and $\boldsymbol{N}$ the population state vector whose component $N_X \in \mathbb{N}$ is the number of particles in state X. We denoted by $\{\alpha\}$ the set of allowed transitions. For instance, it may represent the set of reactions in a given chemical system. For an initial state $\boldsymbol{N}$ and a transition α, the final state $\boldsymbol{N}'$ is then constrained by

$$\boldsymbol{N}' = \boldsymbol{N} + \mathfrak{D}_\alpha, \tag{4.1}$$

where we introduced the vector $\boldsymbol{\mathfrak{D}}_\alpha$ whose component $\mathfrak{D}_{X,\alpha}$ is the variation of the number of particles in state X when the transition α occurs. For example, consider the system of reactions $\alpha = \{\pm 1, \pm 2\}$:

$$2A \underset{\alpha=-1}{\overset{\alpha=+1}{\rightleftharpoons}} 3B, \tag{4.2}$$

$$B \underset{\alpha=-2}{\overset{\alpha=+2}{\rightleftharpoons}} A + C, \tag{4.3}$$

where A, B and C are three chemical species. The states in this model are $\{A, B, C\}$, the population state vector is $\boldsymbol{N} = (N_A, N_B, N_C)$, with N_X ($X = A, B, C$) the number of species X, and we have

$$\begin{array}{lll} \mathfrak{D}_{A,+1} = -2, & \mathfrak{D}_{B,+1} = 3, & \mathfrak{D}_{C,+1} = 0, \\ \mathfrak{D}_{A,-1} = 2, & \mathfrak{D}_{B,-1} = -3, & \mathfrak{D}_{C,-1} = 0, \\ \mathfrak{D}_{A,+2} = 1, & \mathfrak{D}_{B,+2} = -1, & \mathfrak{D}_{C,+2} = 1, \\ \mathfrak{D}_{A,-2} = -1, & \mathfrak{D}_{B,-2} = 1, & \mathfrak{D}_{C,-2} = -1. \end{array}$$

Because of the constraint (4.1), the transition rate $k_{\boldsymbol{N}'\boldsymbol{N}}$ from $\boldsymbol{N}$ to $\boldsymbol{N}'$ depends only on the initial state $\boldsymbol{N}$ and the transition α. Hence, we write them $k_{\alpha,\boldsymbol{N}} \equiv k_{\boldsymbol{N}+\boldsymbol{\mathfrak{D}}_\alpha,\boldsymbol{N}}$. Note that the transition rates are time-independent by definition of a time-homogeneous process. We now assume that the transition rate $k_{\alpha,\boldsymbol{N}}$ scales linearly with a large size-type parameter $\mathcal{N}$ (volume, total number of particle, etc.) and that the state vector $\boldsymbol{N}$ is of order $\mathcal{N}$: $\boldsymbol{N} = O(\mathcal{N})$. In the limit of large $\mathcal{N}$, the appropriate state variable is the empirical density (or concentration) at any time t defined by

$$\boldsymbol{z}(t) \equiv \frac{\boldsymbol{N}(t)}{\mathcal{N}}, \tag{4.4}$$

where $\boldsymbol{N}(t)$ is the state vector at time t. Note that we made an abuse of notation by denoting by $\boldsymbol{N}$ both the state of the system and the empirical state vector $\boldsymbol{N}(t)$ which is a stochastic process. From the constraint (4.1), the number of particles in state X at time $t + \delta t$ is given by the number of particles in state X at time t plus the number of particles added or removed because of the transitions occurred between t and $t + \delta t$:

$$N_X(t + \delta t) = N_X(t) + \sum_\alpha [\Omega_\alpha]_t^{t+\delta t} \mathfrak{D}_{X,\alpha}, \tag{4.5}$$

where $[\Omega_\alpha]_t^{t+\delta t} \equiv \sum_{s\in[t,t+\mathrm{d}t[} \delta_{\alpha(s),\alpha}$ is the number of transitions α that have occurred between times t and $t + \delta t$. Introducing the empirical particle current due to transition α

$$\lambda_\alpha(t) \equiv \frac{1}{\mathcal{N}\delta t} [\Omega_\alpha]_t^{t+\delta t}, \tag{4.6}$$

and taking the limit $\delta t \to 0$, we obtain the continuity equation relating $\dot{\boldsymbol{z}}$ to the vector $\boldsymbol{\lambda}$ of component λ_α:

$$\dot{\boldsymbol{z}} = \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}, \tag{4.7}$$

with $\boldsymbol{\mathfrak{D}}$ the matrix of component $\mathfrak{D}_{X,\alpha}$. In our chemical example, $\mathcal{N}$ is the volume, $\boldsymbol{z} = (z_A, z_B, z_C)$ is the concentration vector of species A, B and C and $\boldsymbol{\lambda}$ is the empirical

chemical current. Since we assumed that the transition rate $k_{\alpha,\boldsymbol{N}}$ scales linearly with $\mathcal{N}$, we can define the intensive transition rate $\boldsymbol{k}(\boldsymbol{z})$ by:

$$k_\alpha(\boldsymbol{z}) \equiv \lim_{\mathcal{N}\to\infty} \frac{k_{\alpha,\boldsymbol{N}}}{\mathcal{N}}, \tag{4.8}$$

and assume the existence of such limit. In general, $k_\alpha(\boldsymbol{z})$ is not linear in $\boldsymbol{z}$, justifying the designation of *nonlinear Markov processes*. Contrary to Section 3.1 where we looked at the large deviations **in time** of empirical observables, we are interested here in the dynamical large deviations **in size** of the empirical current $\boldsymbol{\lambda}$ and empirical density $\boldsymbol{z}$ during an infinitesimal time interval δt when $\mathcal{N}$ goes to infinity. To do so, we derive the $\mathcal{N}$-large deviation function associated with these observables by computing the transition probability $P_{\delta t}(\boldsymbol{N}_\mathrm{f}, t+\delta t \mid \boldsymbol{N}_\mathrm{i}, t)$ from $\boldsymbol{N}(t) = \boldsymbol{N}_\mathrm{i}$ at time t to $\boldsymbol{N}(t+\delta t) = \boldsymbol{N}_\mathrm{f}$ at time $t+\delta t$ in the continuous limit defined by

$$\begin{cases} \mathcal{N} \to \infty, \\ \delta t \to 0, \\ \mathcal{N}\delta t \to \infty. \end{cases} \tag{4.9}$$

Since our Markov process is time-homogeneous, the transition probability depends only on the difference δt between initial and final times and we write $P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$ the transition probability from $\boldsymbol{N}_\mathrm{i}$ to $\boldsymbol{N}_\mathrm{f}$ after a time δt. In Appendix 4.A, we reproduce the derivation done in the appendix of Ref. [124] for rates satisfying a mass-action law, but the proof remains essentially the same for our general rates (4.8). We find that in the continuous limit, the transition probability satisfies a LDP

$$P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\delta t \mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z}_\mathrm{i})} \delta(\dot{\boldsymbol{z}} - \mathfrak{D}\boldsymbol{\lambda}), \tag{4.10}$$

with $\boldsymbol{z}_\mathrm{i} \equiv \frac{\boldsymbol{N}_\mathrm{i}}{\mathcal{N}}$ and $\boldsymbol{z}_\mathrm{f} \equiv \frac{\boldsymbol{N}_\mathrm{f}}{\mathcal{N}} = \boldsymbol{z}_\mathrm{i} + \delta t \dot{\boldsymbol{z}}$, and where the LDF or *detailed Lagrangian* is given by

$$\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}) = \sum_\alpha \left[\lambda_\alpha \ln\left(\frac{\lambda_\alpha}{k_\alpha(\boldsymbol{z})}\right) - \lambda_\alpha + k_\alpha(\boldsymbol{z}) \right]. \tag{4.11}$$

The transition probability $P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$ becomes in the continuous limit the transition probability $P_{\delta t}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i})$, and because the final density is constrained by $\boldsymbol{z}_\mathrm{f} = \boldsymbol{z}_\mathrm{i} + \delta t \mathfrak{D}\boldsymbol{\lambda}$, the transition probability $P_{\delta t}(\boldsymbol{z}_\mathrm{i} + \delta t \mathfrak{D}\boldsymbol{\lambda} \mid \boldsymbol{z}_\mathrm{i})$ can be interpreted as the conditional probability $P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}_\mathrm{i})$ to observe during δt an empirical current $\boldsymbol{\lambda}$ given the initial density $\boldsymbol{z}_\mathrm{i}$. In order to have a normalized probability, the detailed Lagrangian is always nonnegative and has to vanish for at least one $\boldsymbol{\lambda}$ for each $\boldsymbol{z}$. Note that the path probability $\mathbb{P}_t[\boldsymbol{N} \mid \boldsymbol{N}_\mathrm{i}]$ of the path $[\boldsymbol{N}] = (\boldsymbol{N}_\tau)_{\tau\in[0,t]}$ along the time interval $[0,t]$ given that the system was in $\boldsymbol{N}_\mathrm{i}$ at $t=0$ is obtained from the product of transition probabilities over M steps of time δt with $M \to \infty$, $\delta t \to 0$ and $t = M\delta t$:

$$\mathbb{P}_t[\boldsymbol{N} \mid \boldsymbol{N}_\mathrm{i}] = \prod_{\ell=1}^{M} P_{\delta t}(\boldsymbol{N}(t_\ell) \mid \boldsymbol{N}(t_{\ell-1})), \tag{4.12}$$

where $t_\ell \equiv \ell\delta t$. In the continuous limit, it leads to the path probability of $[\boldsymbol{z}]$ given the initial state $\boldsymbol{z}_\mathrm{i}$:

$$\mathbb{P}_t[\boldsymbol{z} \mid \boldsymbol{z}_\mathrm{i}] \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\int_0^t \mathrm{d}\tau \mathscr{L}(\boldsymbol{\lambda}_\tau, \boldsymbol{z}_\tau)}. \tag{4.13}$$

On the other hand, the integrated transition probability $P_t(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$ to observe $\boldsymbol{N}_\mathrm{f}$ after a time t given the system was in $\boldsymbol{N}_\mathrm{i}$ reads:

$$P_t(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) = \sum_{\substack{\{\boldsymbol{N}(t_\ell)\} \\ \ell \in [\![1, M-1]\!]}} \prod_{\ell=1}^{M-1} P_{\delta t}(\boldsymbol{N}(t_\ell) \mid \boldsymbol{N}(t_{\ell-1})). \tag{4.14}$$

In the continuous limit, it leads to:

$$P_t(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i}) \underset{n\to\infty}{\asymp} \int \mathfrak{D}[\boldsymbol{\lambda}, \boldsymbol{z}] \mathrm{e}^{-\mathcal{N}\int_0^t \mathrm{d}\tau \mathscr{L}(\boldsymbol{\lambda}_\tau, \boldsymbol{z}_\tau)} \delta(\dot{\boldsymbol{z}} - \mathfrak{D}\boldsymbol{\lambda}). \tag{4.15}$$

Note the difference between $\mathbb{P}_t[\boldsymbol{z} \mid \boldsymbol{z}_\mathrm{i}]$ which is the probability of the **path** $[\boldsymbol{z}]$ of duration t given the initial state $\boldsymbol{z}_\mathrm{i}$, and $P_t(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i})$ which is the **transition probability** from $\boldsymbol{z}_\mathrm{i}$ to $\boldsymbol{z}_\mathrm{f}$ after a time t (independently of the path followed in between).

We define the *detailed Hamiltonian* as the SCGF of the observable $\Omega_\alpha = \mathcal{N}\delta t \lambda_\alpha$, giving the number of each transition α during δt:

$$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z}) \equiv \lim_{\mathcal{N}\to\infty} \frac{1}{\mathcal{N}\delta t} \ln \left(\int \mathrm{d}\boldsymbol{\lambda} P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}) \mathrm{e}^{\mathcal{N}\delta t \boldsymbol{f}\cdot\boldsymbol{\lambda}} \right), \tag{4.16}$$

and corresponding to the LF transform of the detailed Lagrangian [124]

$$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z}) = \sup_{\boldsymbol{\lambda}} \{\boldsymbol{f}\cdot\boldsymbol{\lambda} - \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})\}, \tag{4.17}$$

where $\boldsymbol{f}$ is conjugated to $\boldsymbol{\lambda}$ and where the central dot $\cdot$ stands for the scalar product $\boldsymbol{f}\cdot\boldsymbol{\lambda} \equiv \sum_\alpha \lambda_\alpha f_\alpha$. Note that $\mathscr{H}$ is convex in $\boldsymbol{f}$ since it follows from a LF transform with respect to $\boldsymbol{\lambda}$ (see Appendix 1.A). Note also that the Lagrangian and Hamiltonian do not depend explicitly on time as a consequence of the time-homogeneity of our process. Computing explicitly the LF transform of (4.17), we obtain that the detailed Hamiltonian is given by

$$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z}) = \sum_\alpha k_\alpha(\boldsymbol{z}) \left[\mathrm{e}^{f_\alpha} - 1\right]. \tag{4.18}$$

For $\boldsymbol{f} = 0$, the value of the Hamiltonian is zero for any $\boldsymbol{z}$. This property translates the fact that $\mathscr{H}$ is associated with a norm-conserving Markov generator, and hence a normalized transition probability, in line with the fact that the Lagrangian has to be nonnegative and to have zero as a minimal value.

The detailed Lagrangian contains the information on every individual current, which represents a surplus of information for what is usually needed to study the system. We can consider a coarse-grained description of the model by looking at the $\mathcal{N}$-large deviations of the empirical density flux $\dot{\boldsymbol{z}}$. This information is encoded in the *standard Lagrangian* L obtained by contracting the detailed Lagrangian over the empirical currents under the constraint (4.7):

$$L(\dot{\boldsymbol{z}}, \boldsymbol{z}) \equiv \inf_{\boldsymbol{\lambda} | \dot{\boldsymbol{z}} = \mathfrak{D}\boldsymbol{\lambda}} \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}). \tag{4.19}$$

Contrary to the detailed Lagrangian, the standard Lagrangian is usually difficult to compute explicitly. The corresponding *standard Hamiltonian* is defined by

$$H(\boldsymbol{p}, \boldsymbol{z}) \equiv \sup_{\dot{\boldsymbol{z}}} \{\boldsymbol{p}\cdot\dot{\boldsymbol{z}} - L(\dot{\boldsymbol{z}}, \boldsymbol{z})\}, \tag{4.20}$$

where $\boldsymbol{p}$ is conjugated to $\dot{\boldsymbol{z}}$ and $\boldsymbol{p}\cdot\dot{\boldsymbol{z}} \equiv \sum_X p_X \dot{z}_X$. Contrary to the standard Lagrangian which is difficult to obtain in general, the standard Hamiltonian is simply obtained from the detailed Hamiltonian:

$$H(\boldsymbol{p},\boldsymbol{z}) = \sup_{\dot{\boldsymbol{z}}} \{\boldsymbol{p}\cdot\dot{\boldsymbol{z}} - L(\dot{\boldsymbol{z}},\boldsymbol{z})\} \tag{4.21}$$

$$= \sup_{\dot{\boldsymbol{z}}} \left\{ \boldsymbol{p}\cdot\dot{\boldsymbol{z}} - \inf_{\boldsymbol{\lambda}|\dot{\boldsymbol{z}}=\mathfrak{D}\boldsymbol{\lambda}} [\mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z})] \right\} \tag{4.22}$$

$$= \sup_{\dot{\boldsymbol{z}}} \left\{ \sup_{\boldsymbol{\lambda}|\dot{\boldsymbol{z}}=\mathfrak{D}\boldsymbol{\lambda}} [\boldsymbol{p}\cdot\dot{\boldsymbol{z}} - \mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z})] \right\} \tag{4.23}$$

$$= \sup_{\boldsymbol{\lambda}} \{\boldsymbol{p}\cdot(\mathfrak{D}\boldsymbol{\lambda}) - \mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z})\} \tag{4.24}$$

$$= \sup_{\boldsymbol{\lambda}} \left\{(\mathfrak{D}^\dagger\boldsymbol{p})\cdot\boldsymbol{\lambda} - \mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z})\right\} \tag{4.25}$$

$$= \mathscr{H}(\boldsymbol{f} = \mathfrak{D}^\dagger\boldsymbol{p},\boldsymbol{z}), \tag{4.26}$$

where $\mathfrak{D}^\dagger$ is the adjoint of $\mathfrak{D}$ and where we used Eq. (4.19). It yields

$$H(\boldsymbol{p},\boldsymbol{z}) = \sum_\alpha k_\alpha(\boldsymbol{z}) \left[\mathrm{e}^{\boldsymbol{p}\cdot\mathfrak{D}_\alpha} - 1\right]. \tag{4.27}$$

Since Eq. (4.26) allows explicit formulation of standard Hamiltonians, we expect that the Hamiltonian framework is more convenient for analytical computations than the Lagrangian framework.

Illustration with a toy model

We illustrate the results derived above with a simple linear population process. Even if looking at a linear model does not seem interesting, it remains a pedagogical toy model as it allows us to compute analytical expressions and to understand the dynamics of the system. Nonlinear models will be studied in Section 4.3. We are interested here in a linear chemical reaction. In general, chemical reaction networks are represented by a set of chemical reactions. The r^{th} reaction reads:

$$\sum_X \nu_X^{+\mathsf{r}} X \underset{\mathfrak{K}_{-\mathsf{r}}}{\overset{\mathfrak{K}_{+\mathsf{r}}}{\rightleftharpoons}} \sum_X \nu_X^{-\mathsf{r}} X, \tag{4.28}$$

where $\mathfrak{K}_{\epsilon\mathsf{r}}$ is the kinetic constant of the reaction $\epsilon\mathsf{r}$, where the sum runs over the chemical species X, and where $\nu_X^{\epsilon\mathsf{r}}$ is the stoichiometry of species X in the reaction $\epsilon\mathsf{r}$, with $\epsilon = \pm$. Chemical reaction networks are a particular case of population processes where $\mathcal{N}$ is a volume, the variable $\boldsymbol{z}$ is the concentration vector whose component z_X is the concentration of species X, and the variable $\boldsymbol{\lambda}$ is the chemical current such that $N\delta t\lambda_{\epsilon\mathsf{r}}$ is the number of reactions $\epsilon\mathsf{r}$ occurring during an infinitesimal time δt. Both variables are related by $\dot{\boldsymbol{z}} = \mathfrak{D}\boldsymbol{\lambda}$, where $\mathfrak{D}$ is the matrix whose component $\mathfrak{D}_{X,\epsilon\mathsf{r}} = \nu_X^{-\epsilon\mathsf{r}} - \nu_X^{\epsilon\mathsf{r}}$ is the variation of the number of species X when the reaction $\epsilon\mathsf{r}$ occurs. We choose the intensive rates of Eq. (4.8) according to the mass-action law [124, 264]:

$$k_{\epsilon\mathsf{r}}(\boldsymbol{z}) \equiv \mathfrak{K}_{\epsilon\mathsf{r}}\boldsymbol{z}^{\boldsymbol{\nu}^{\epsilon\mathsf{r}}}, \tag{4.29}$$

with $\boldsymbol{z}^{\boldsymbol{\nu}^{\epsilon\mathfrak{r}}} \equiv \prod_X {z_X}^{\nu_X^{\epsilon\mathfrak{r}}}$. Now, we are interested in the following chemical reaction

$$A \underset{\mathfrak{K}_{-1}}{\overset{\mathfrak{K}_{+1}}{\rightleftharpoons}} X \underset{\mathfrak{K}_{-2}}{\overset{\mathfrak{K}_{+2}}{\rightleftharpoons}} B, \tag{4.30}$$

where species A and B are chemostatted, i.e. they have constant concentrations a and b, respectively. We call x the concentration of species X and we set $k_{-1} = 1$, fixing the time scale in numerical plots. Since a and b are constant, the variable $\boldsymbol{z}$ reduces to x and the matrix $\mathfrak{D}$ becomes the vector of components

$$\begin{aligned} \mathfrak{D}_{+1} &= +1, \quad \mathfrak{D}_{+2} = -1, \\ \mathfrak{D}_{-1} &= -1, \quad \mathfrak{D}_{-2} = +1. \end{aligned} \tag{4.31}$$

The transition rates $k_{\epsilon\mathfrak{r}}(\boldsymbol{z})$ for this system are given by

$$k_{+1} = \mathfrak{K}_{+1} a, \tag{4.32}$$
$$k_{-1} = \mathfrak{K}_{-1} x, \tag{4.33}$$
$$k_{+2} = \mathfrak{K}_{+2} x, \tag{4.34}$$
$$k_{-2} = \mathfrak{K}_{-2} b. \tag{4.35}$$

From Eqs. (4.18, 4.27), the detailed and standard Hamiltonians for this model are given by:

$$\mathscr{H}(\boldsymbol{f}, x) = \mathfrak{K}_{+1} a \left[\mathrm{e}^{f_{+1}} - 1\right] + \mathfrak{K}_{-1} x \left[\mathrm{e}^{f_{-1}} - 1\right] + \mathfrak{K}_{+2} x \left[\mathrm{e}^{f_{+2}} - 1\right] + \mathfrak{K}_{-2} b \left[\mathrm{e}^{f_{-2}} - 1\right], \tag{4.36}$$

$$H(p, x) = (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b)[\mathrm{e}^{p} - 1] + (\mathfrak{K}_{-1} + \mathfrak{K}_{+2}) x [\mathrm{e}^{-p} - 1], \tag{4.37}$$

both related by $\mathscr{H}(\boldsymbol{f} = \mathfrak{D}^\dagger p, x) = H(p, x)$. For this model, the standard Lagrangian can be explicitly computed by contracting the detailed Lagrangian (4.11) according to Eq. (4.19). It yields

$$\begin{aligned} L(\dot{x}, x) = \dot{x} \ln \frac{\dot{x} + \sqrt{\dot{x}^2 + 4(\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b)(\mathfrak{K}_{-1} + \mathfrak{K}_{+2}) x}}{2(\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b)} \\ - \sqrt{\dot{x}^2 + 4(\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b)(\mathfrak{K}_{-1} + \mathfrak{K}_{+2}) x} + (\mathfrak{K}_{-1} + \mathfrak{K}_{+2}) x + (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b). \end{aligned} \tag{4.38}$$

The same result can be obtained by taking the LF transform of Eq. (4.37) with respect to p since for differentiable Hamiltonians, the LF transform of Eq. (4.20) is involutive. When we compare the expressions of standard Lagrangian and Hamiltonian, it is obvious that the Hamiltonian formalism is more convenient for explicit computation. A fortiori, when dealing with differentiable Hamiltonians or convex Lagrangians, the two formalisms are strictly equivalent since the LF transform is involutive. In the rest of this manuscript, we assume that it is always the case.

4.1.1 Biased process

Similarly to the linear operator formalism, we are interested in the fluctuations of the observable

$$\boldsymbol{A}_t \equiv \begin{pmatrix} \frac{1}{t} \sum_{t' \in [0,t]} [\boldsymbol{\Omega}]_{t'}^{t'+\mathrm{d}t'} \\ \frac{1}{t} \int_0^t \boldsymbol{N}(t') \mathrm{d}t' \end{pmatrix} = \mathcal{N} \begin{pmatrix} \frac{1}{t} \int_0^t \boldsymbol{\lambda}(t') \mathrm{d}t' \\ \frac{1}{t} \int_0^t \boldsymbol{z}(t') \mathrm{d}t' \end{pmatrix}, \tag{4.39}$$

where $[\boldsymbol{\Omega}]_{t'}^{t'+\mathrm{d}t'}$ is the vector function of component $[\Omega_\alpha]_{t'}^{t'+\mathrm{d}t'}$. Hence, the first component of $\boldsymbol{A}_t$ counts the number of each transition occurring during the time interval $[0,t]$, and the second component of $\boldsymbol{A}_t$ is the time average of the number of individuals at each state during the interval $[0,t]$. Using the results of Section 2.3.4, the biased transition matrix $\boldsymbol{\kappa}$ ruling the time-evolution of the generating function $\tilde{G}_{\boldsymbol{N}}(t,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{N}_\mathrm{i}}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\delta_{\boldsymbol{N}(t),\boldsymbol{N}}\right]$ — where the average is done with respect to the path probability $\mathbb{P}_t[\boldsymbol{N} \mid \boldsymbol{N}_\mathrm{i}]$ — reads in the Dirac notation

$$\boldsymbol{\kappa} = \sum_{\alpha,\boldsymbol{N}} k_{\alpha,\boldsymbol{N}}\mathrm{e}^{\gamma_\alpha^1}\left|\boldsymbol{N}+\boldsymbol{\mathfrak{D}}_\alpha\right\rangle\left\langle\boldsymbol{N}\right| - k_{\alpha,\boldsymbol{N}}\left|\boldsymbol{N}\right\rangle\left\langle\boldsymbol{N}\right| + \sum_{\boldsymbol{N}}\boldsymbol{\gamma}_2\cdot\boldsymbol{N}\left|\boldsymbol{N}\right\rangle\left\langle\boldsymbol{N}\right|, \tag{4.40}$$

where the vectors $\boldsymbol{\gamma}_1$ of component γ_α^1 and $\boldsymbol{\gamma}_2$ of component γ_X^2 are conjugated to the first and second component of $\boldsymbol{A}_t$, respectively. Once again, we aim to describe the $\mathcal{N}$-LDF of the observables $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ but now with respect to the biased process. To do so, we compute the biased transition probability $\tilde{G}_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) = \langle\boldsymbol{N}_\mathrm{f}|\,\mathrm{e}^{\delta t\boldsymbol{\kappa}}\,|\boldsymbol{N}_\mathrm{i}\rangle$ from an initial state $\boldsymbol{N}_\mathrm{i}$ to a final state $\boldsymbol{N}_\mathrm{f}$ after a time δt [124]. Following the same procedure used to derive the detailed Lagrangian in Appendix 4.A, we obtain in the continuous limit that the biased transition probability satisfies a LDP

$$\tilde{G}_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\delta t\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda},\boldsymbol{z})}\delta(\dot{\boldsymbol{z}} - \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}), \tag{4.41}$$

where the *detailed biased Lagrangian* is given by:

$$\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda},\boldsymbol{z}) = \sum_\alpha\left[\lambda_\alpha\ln\left(\frac{\lambda_\alpha}{k_\alpha(\boldsymbol{z})}\right) - \lambda_\alpha + k_\alpha(\boldsymbol{z}) - \gamma_\alpha^1\lambda_\alpha\right] - \boldsymbol{\gamma}_2\cdot\boldsymbol{z}. \tag{4.42}$$

The LF transform of Eq. (4.42) with respect to $\boldsymbol{\lambda}$ defines the *detailed biased Hamiltonian*

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f},\boldsymbol{z}) = \sum_\alpha k_\alpha(\boldsymbol{z})\left[\mathrm{e}^{f_\alpha+\gamma_\alpha^1} - 1\right] + \boldsymbol{\gamma}_2\cdot\boldsymbol{z} = \mathscr{H}(\boldsymbol{f}+\boldsymbol{\gamma}_1,\boldsymbol{z}) + \boldsymbol{\gamma}_2\cdot\boldsymbol{z}, \tag{4.43}$$

where $\boldsymbol{f}$ is once again the conjugate variable of $\boldsymbol{\lambda}$.

As discussed in Chapter 3, the biased matrix $\boldsymbol{\kappa}$ generates a Markov process that is not norm-conserving, which implies that $\tilde{G}_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$ is not a proper transition probability as it is not normalized. In the Lagrangian/Hamiltonian formalism, this translates into the fact that the biased Lagrangian needs not to vanish at a given $\boldsymbol{\lambda}$ for each $\boldsymbol{z}$ and that the biased Hamiltonian is not zero anymore when $\boldsymbol{f} = 0$, as can be noticed from Eqs. (4.42–4.43).

Again, we can consider a coarse-grained description of the model by looking at the $\mathcal{N}$-large deviations of the empirical density flux $\dot{\boldsymbol{z}}$ given $\boldsymbol{z}$, and we define the *standard biased Lagrangian* $L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}},\boldsymbol{z})$ and *standard biased Hamiltonian* $H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z})$ in the same way as for the non-biased case:

$$L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}},\boldsymbol{z}) \equiv \inf_{\boldsymbol{\lambda}|\dot{\boldsymbol{z}}=\boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}}\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda},\boldsymbol{z}), \tag{4.44}$$

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z}) \equiv \sup_{\dot{\boldsymbol{z}}}\{\boldsymbol{p}\cdot\dot{\boldsymbol{z}} - L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}},\boldsymbol{z})\} = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} = \boldsymbol{\mathfrak{D}}^\dagger\boldsymbol{p},\boldsymbol{z}). \tag{4.45}$$

In the following, we may skip the attributes *detailed* and *standard* as the notations used for the Lagrangians and Hamiltonians are distinguishing, namely calligraphic letters are used for detailed Lagrangians ($\mathscr{L}$) and Hamiltonians ($\mathscr{H}$) while straight letters are for standard Lagrangians (L) and Hamiltonians (H).

Illustration with the toy model

Let us go back to our toy model (4.30). Using Eqs. (4.43, 4.45), the (standard) biased Hamiltonian reads

$$H_{\boldsymbol{\gamma}}(p,x) = \alpha \mathrm{e}^{p} + \beta x \mathrm{e}^{-p} - \delta x - (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b), \tag{4.46}$$

where we introduced for clarity

$$\alpha \equiv \mathfrak{K}_{+1} a \mathrm{e}^{\gamma^1_{+1}} + \mathfrak{K}_{-2} b \mathrm{e}^{\gamma^1_{-2}}, \tag{4.47}$$

$$\beta \equiv \mathfrak{K}_{-1} \mathrm{e}^{\gamma^1_{-1}} + \mathfrak{K}_{+2} \mathrm{e}^{\gamma^1_{+2}}, \tag{4.48}$$

$$\delta \equiv \mathfrak{K}_{-1} + \mathfrak{K}_{+2} - \gamma_2. \tag{4.49}$$

We clearly see that for $\boldsymbol{\gamma} \neq 0$, $H_{\boldsymbol{\gamma}}(p,x) \neq 0$. Computing the LF transform of $H_{\boldsymbol{\gamma}}$ leads to the biased Lagrangian

$$L_{\boldsymbol{\gamma}}(\dot{x},x) = \dot{x} \ln \frac{\dot{x} + \sqrt{\dot{x}^2 + 4\alpha\beta x}}{2\alpha} - \sqrt{\dot{x}^2 + 4\alpha\beta x} + \delta x + (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b). \tag{4.50}$$

Notice that for $\boldsymbol{\gamma} = 0$, we recover the non-biased Hamiltonian and Lagrangian of Eqs. (4.37–4.38).

4.1.2 Equations of motion

What follows includes the non-biased formalism by taking $\boldsymbol{\gamma} = 0$. We are interested in the typical behavior of our system during the time-interval $[0, \mathcal{T}]$. The biased transition probability of observing $\boldsymbol{z}_{\mathrm{f}} = \boldsymbol{z}_{\mathrm{i}} + \int_0^{\mathcal{T}} \dot{\boldsymbol{z}}_t \mathrm{d}t = \boldsymbol{z}_{\mathrm{i}} + \int_0^{\mathcal{T}} \mathfrak{D}\boldsymbol{\lambda}_t \mathrm{d}t$ at $t = \mathcal{T}$ given the initial state $\boldsymbol{z}_{\mathrm{i}}$ at $t = 0$ satisfies

$$\tilde{G}_{\mathcal{T}}(\boldsymbol{z}_{\mathrm{f}} \mid \boldsymbol{z}_{\mathrm{i}}) \underset{\mathcal{N}\to\infty}{\asymp} \int \mathfrak{D}[\boldsymbol{\lambda}, \boldsymbol{z}] \mathrm{e}^{-\mathcal{N} \int_0^{\mathcal{T}} \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}_t, \boldsymbol{z}_t) \mathrm{d}t} \delta(\dot{\boldsymbol{z}} - \mathfrak{D}\boldsymbol{\lambda}) \underset{\mathcal{N}\to\infty}{\asymp} \int \mathfrak{D}[\dot{\boldsymbol{z}}, \boldsymbol{z}] \mathrm{e}^{-\mathcal{N} \int_0^{\mathcal{T}} L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}_t, \boldsymbol{z}_t) \mathrm{d}t}, \tag{4.51}$$

where we used Laplace's approximation and the definition of the standard biased Lagrangian (4.19) in the last equality. There is a family of trajectories $\{(\boldsymbol{z}_t)^{\varepsilon}\}$ indexed by ε connecting $\boldsymbol{z}_{\mathrm{i}}$ to $\boldsymbol{z}_{\mathrm{f}}$ during a time $\mathcal{T}$. The typical trajectory followed by the system is the one minimizing the action $S[\dot{\boldsymbol{z}}, \boldsymbol{z}]_0^{\mathcal{T}} \equiv \int_0^{\mathcal{T}} \mathrm{d}t L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}_t, \boldsymbol{z}_t)$, which appears to be the one solving the Euler-Lagrange equation for any $t \in [0, \mathcal{T}]$:

$$\frac{\partial L_{\boldsymbol{\gamma}}}{\partial \boldsymbol{z}} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L_{\boldsymbol{\gamma}}}{\partial \dot{\boldsymbol{z}}} \right) = 0, \tag{4.52}$$

with boundary conditions $\boldsymbol{z}_0 = \boldsymbol{z}_{\mathrm{i}}$ and $\boldsymbol{z}_{\mathcal{T}} = \boldsymbol{z}_{\mathrm{f}}$, see Appendix 4.B for the derivation. Assuming that the Hamiltonian is differentiable, the Lagrangian is given by the Legendre transform

$$L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}, \boldsymbol{z}) = \boldsymbol{p} \cdot \dot{\boldsymbol{z}} - H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}), \tag{4.53}$$

with $\boldsymbol{p} = \frac{\partial L_{\boldsymbol{\gamma}}}{\partial \dot{\boldsymbol{z}}}$. The biased transition probability can then be written from a Hamiltonian perspective

$$\tilde{G}_{\mathcal{T}}(\boldsymbol{z}_{\mathrm{f}} \mid \boldsymbol{z}_{\mathrm{i}}) \underset{\mathcal{N}\to\infty}{\asymp} \int \mathfrak{D}[\boldsymbol{p}, \boldsymbol{z}] \mathrm{e}^{-\mathcal{N} \int_0^{\mathcal{T}} [\boldsymbol{p}_t \cdot \dot{\boldsymbol{z}}_t - H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t)] \mathrm{d}t}. \tag{4.54}$$

As before, we look for the typical trajectory $(\boldsymbol{z}_t)^\varepsilon$ connecting $\boldsymbol{z}_\mathrm{i}$ to $\boldsymbol{z}_\mathrm{f}$ after a time $\mathcal{T}$. This trajectory is the one minimizing the action $S[\boldsymbol{p},\boldsymbol{z}]_0^{\mathcal{T}} \equiv \int_0^{\mathcal{T}} [\boldsymbol{p}_t \cdot \dot{\boldsymbol{z}}_t - H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t)]\,\mathrm{d}t$ and appears to be the solution of Hamilton's equations for any $t \in [0,\mathcal{T}]$:

$$\begin{cases} \dot{\boldsymbol{z}}_t = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t), \\ \dot{\boldsymbol{p}}_t = -\partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t), \end{cases} \tag{4.55}$$

with the same boundary conditions $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$, see Appendix 4.B for the derivation, and where $\partial_{\boldsymbol{p}}$ and $\partial_{\boldsymbol{z}}$ are the gradients with respect to $\boldsymbol{p}$ and $\boldsymbol{z}$, respectively. Along this solution, the value of the Hamiltonian is a constant of time. Indeed, using Eq. (4.55) and the fact that the Hamiltonian is not explicitly time-dependent, we have

$$\frac{\mathrm{d}H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z})}{\mathrm{d}t} = \dot{\boldsymbol{p}}\,\partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z}) + \dot{\boldsymbol{z}}\,\partial_{\boldsymbol{z}} H(\boldsymbol{p},\boldsymbol{z}) = 0. \tag{4.56}$$

Hence, the value of the Hamiltonian depends essentially on the boundary conditions. The biased transition probability satisfies then

$$\tilde{G}_{\mathcal{T}}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{\mathcal{N}\left[H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t^*,\boldsymbol{z}_t^*)\mathcal{T} - \int_0^{\mathcal{T}} \boldsymbol{p}_t^*\cdot\dot{\boldsymbol{z}}_t^*\,\mathrm{d}t\right]}, \tag{4.57}$$

for $\boldsymbol{p}_t^*$ and $\boldsymbol{z}_t^*$ solutions of Hamilton's equations (4.55). We recognize in the second term of the exponential the so-called *reduced action*

$$S_\mathrm{r}(\mathcal{T}) \equiv \int_0^{\mathcal{T}} \boldsymbol{p}_t^* \cdot \dot{\boldsymbol{z}}_t^* \mathrm{d}t. \tag{4.58}$$

Hamilton-Jacobi equation

An alternative description of the dynamics can be obtained by considering *Hamilton's principal function* — also called *Jacobi's action* — defined as the action evaluated along the solutions of Hamilton's equations (or equivalently the Euler-Lagrange equation) with initial state $\boldsymbol{z}_\mathrm{i}$ and arrival state $\boldsymbol{z}$ at time t:

$$S(\boldsymbol{z},\boldsymbol{z}_\mathrm{i},t) \equiv \int_0^t \mathrm{d}\tau L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}_\tau,\boldsymbol{z}_\tau), \tag{4.59}$$

with $(\dot{\boldsymbol{z}}_\tau,\boldsymbol{z}_\tau)$ solution of Eq. (4.52). The action S contains all the information on the dynamics of the system (see Ref. [265] for more details). It can be obtained by solving a partial differential equation called *Hamilton-Jacobi (HJ) equation*:

$$\frac{\partial S}{\partial t} + H_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} S, \boldsymbol{z}) = 0. \tag{4.60}$$

When the Hamiltonian is time-independent, hence a constant along a solution of Hamilton's equations $H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z}) = E$, it is convenient to consider *Hamilton's characteristic function* defined as the Legendre transform of S with respect to time:

$$W(\boldsymbol{z},\boldsymbol{z}_\mathrm{i},E) = Et + S(\boldsymbol{z},\boldsymbol{z}_\mathrm{i},t), \tag{4.61}$$

where the *eigenrate* E has the dimension of an inverse time and replaces "energy" in the HJ equation of analytical mechanics. The Hamilton-Jacobi equation for W reads

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} W, \boldsymbol{z}) = E. \tag{4.62}$$

Unless specified otherwise, the terminology "Hamilton-Jacobi equation" always refers to Eq. (4.62) in this manuscript.

Representation of the trajectories

From now on, we work only from the Hamiltonian perspective as standard Hamiltonians are more convenient than standard Lagrangians. We saw that the trajectory $(\boldsymbol{p}_t, \boldsymbol{z}_t)$ likely followed by the system is the one satisfying Hamilton's equations. These solutions depend on the boundary conditions. If $\partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}$ and $\partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}$ are smooth enough (which is assumed to be always the case in this manuscript), then the solutions of Hamilton's equations exist and are unique given the boundary conditions $\boldsymbol{z}_\mathrm{i}$ at initial time $t = 0$ and $\boldsymbol{z}_\mathrm{f}$ at final time $t = \mathcal{T}$ (or any other conditions) [266].

We represent the ensemble of solutions of Hamilton's equations in the phase space $(\boldsymbol{p}, \boldsymbol{z})$ in which the dependence in time is made implicit (see Fig. 4.1). Each point of the phase space represents the value of $\boldsymbol{z}$ and $\boldsymbol{p}$ at a given time t_0, and each line represents one trajectory $(\boldsymbol{p}^*_t, \boldsymbol{z}^*_t)$ solution of Hamilton's equations and is a line of constant Hamiltonian. Starting from the point $(\boldsymbol{p}_{t_0}, \boldsymbol{z}_{t_0})$, the system will evolve to another point $(\boldsymbol{p}_{t_0+\tau}, \boldsymbol{z}_{t_0+\tau})$ after a time τ by following the corresponding trajectory. The direction followed by the system along the trajectory is indicated by an arrow. For an initial state $\boldsymbol{z}_\mathrm{i}$ and a final state $\boldsymbol{z}_\mathrm{f}$, several trajectories are possible (red trajectories in Fig. 4.1). If now we impose that $\boldsymbol{z}_\mathrm{f}$ is reached after a specific time $\mathcal{T}$, then only one trajectory is admitted according to the uniqueness of the solutions of Hamilton's equations. This uniqueness translates into the fact that the trajectories in the phase space never intersect. If two trajectories intersected, the system could follow two different trajectories starting from the same point in the phase space (the crossing point), meaning that there would be two solutions of Hamilton's equations, which would contradict the uniqueness theorem [266]. An important remark is that in Fig. 4.1, two trajectories seem to intersect in the red point (fixed point). In fact, it is not the case because the system takes an infinite time to reach or leave that point, and trajectories only approach it infinitesimally closely but will never reach it (see the next sections for more details). Thereafter, we look for the trajectory followed by the system when the final condition $\boldsymbol{z}_\mathrm{f}$ is imposed in the limit of long final time $\mathcal{T} \to \infty$.

Critical manifolds

There is a particular class of trajectories that we call *critical manifolds*[1], defined as an ensemble of compact trajectories (in the sense that they are entirely included in some compact[2] set of the phase space) and such that at least one other trajectory converges towards it forward or backward in time. We name these trajectories that lead to a critical manifold forward (resp. backward) in time *relaxing (resp. fluctuating) transient trajectories*. The simplest critical manifolds are *fixed points*, whose phase space coordinates solve the stationary Hamilton's equations

$$\begin{cases} \dot{\boldsymbol{z}} = 0, \\ \dot{\boldsymbol{p}} = 0. \end{cases} \tag{4.63}$$

The fixed point in Fig. 4.1 (red point) is a critical manifold. Note that centers (e.g. the magenta point in Fig. 4.7), are fixed points but not critical manifolds as no other trajectory

[1] A manifold is informally defined as a geometrical space generalizing the notion of curve or surface to arbitrary dimensions. For instance, a one-dimensional manifold is a curve and includes lines and circles. A two-dimensional manifold is a surface and includes plans, spheres and tori.

[2] In $\mathbb{R}^n$, a set is compact if and only if it is closed and bounded.

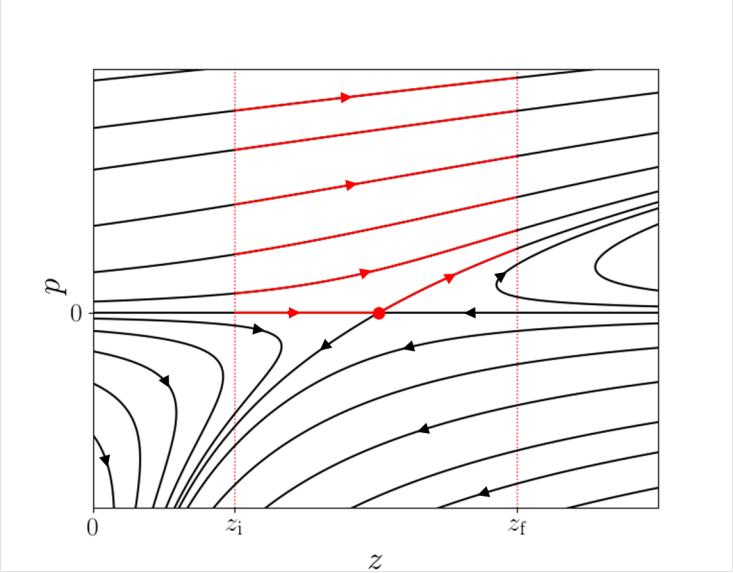

Figure 4.1 – Trajectories in the phase space of the toy model (4.30). Each line corresponds to a value of H. The arrows indicate the direction of each trajectory. All the red trajectories start at $\boldsymbol{z}_\mathrm{i}$ at $t = 0$ and end at $\boldsymbol{z}_\mathrm{f}$ at different final times. Imposing the final time $\mathcal{T}$ such that $\boldsymbol{z}_\mathcal{T} = \boldsymbol{z}_\mathrm{f}$ selects a unique trajectory compatible with $\boldsymbol{z}_0 = \boldsymbol{z}_\mathrm{i}$.

converges to them. Another type of critical manifolds is *limit cycles* which are closed critical manifolds of dimension 1. Limit cycles are nonlinear phenomena and cannot occur in linear systems. Note that circles (periodic trajectories encircling a center, see Fig. 4.7) are not critical manifolds as no other trajectory leads to them. Other examples of critical manifolds include tori or complex geometric structures called *strange attractors*[3]. When the system is at a critical manifold, it will take an infinite time to leave it via a fluctuating transient trajectory. On the contrary, it takes an infinite time for the system to reach a critical manifold via a relaxing transient trajectory.

Long-time limit dynamics

We are interested in characterizing the typical (or dominant) trajectory $(\boldsymbol{p}_t^*, \boldsymbol{z}_t^*)$ when $\mathcal{T} \to \infty$. It means that we look for a trajectory connecting $\boldsymbol{z}_\mathrm{i}$ to the final state $\boldsymbol{z}_\mathrm{f}$ after an infinitely long time $\mathcal{T}$. Intuitively, we sense that the only way to connect two finite positions of the phase space during an infinite amount of time is to follow transient trajectories. In Section 4.2.3, we will confirm this intuition and see that for any boundary conditions $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$, the dominant trajectory in the long-time limit approaches asymptotically transient trajectories along which the Hamiltonian equals the SCGF:

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t^*, \boldsymbol{z}_t^*) = \bar{\Gamma}, \tag{4.64}$$

where we defined the SCGF $\bar{\Gamma}$ scaled in time and size:

$$\bar{\Gamma} \equiv \frac{\Gamma}{\mathcal{N}} = \frac{1}{\mathcal{N}} \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \ln \mathbb{E}_{\boldsymbol{z}_\mathrm{i}} \left[\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_\mathcal{T}} \right], \tag{4.65}$$

[3] We made here an abuse of language as *attractor* means that all trajectories converge toward it forward in time. Here, the strange attractor may be stable for some trajectories (attractor) and unstable for others (repeller).

where the average is done with respect to $\mathbb{P}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$:

$$\mathbb{E}_{\boldsymbol{z}_{\mathrm{i}}}\left[\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\right] = \int \mathfrak{D}[\boldsymbol{z}]\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\mathbb{P}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}] \underset{\mathcal{N}\to\infty}{\asymp} \int \mathfrak{D}[\boldsymbol{\lambda}, \boldsymbol{z}]\, \mathrm{e}^{-\mathcal{N}\int_0^{\mathcal{T}} \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}_t, \boldsymbol{z}_t)\mathrm{d}t}\delta(\dot{\boldsymbol{z}} - \mathfrak{D}\boldsymbol{\lambda}). \quad (4.66)$$

We used Eqs. (4.13), (4.39) and (4.42) in Eq. (4.66). In particular, in the non-biased case ($\boldsymbol{\gamma} = 0$), $\bar{\Gamma} = 0$ implying that the trajectories at $\boldsymbol{p} = 0$ are dominant in the long-time limit since $H(\boldsymbol{p} = 0, \boldsymbol{z}) = 0$, $\forall \boldsymbol{z}$. For now, we try to give some intuition to this prediction using our toy model (4.30).

Illustration with the toy model

Let us investigate the dynamics governed by the biased Hamiltonian. Hamilton's Equations read

$$\begin{cases} \dot{x} = \frac{\partial H_{\boldsymbol{\gamma}}}{\partial p} = \alpha \mathrm{e}^{p} - \beta x \mathrm{e}^{-p}, \\ \dot{p} = -\frac{\partial H_{\boldsymbol{\gamma}}}{\partial x} = -\beta \mathrm{e}^{-p} + \delta. \end{cases} \quad (4.67)$$

The critical manifolds of this system consist of one fixed point of coordinates

$$\begin{cases} x_\star = \frac{\alpha\beta}{\delta^2}, \\ p_\star = \ln \frac{\beta}{\delta}. \end{cases} \quad (4.68)$$

Note that the allowed values of γ_2 are such that $\delta > 0$ to ensure the existence of $p_\star$. At this fixed point, the biased Hamiltonian is given by

$$H_{\boldsymbol{\gamma}}(p_\star, x_\star) = \frac{\alpha\beta}{\delta} - (\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b). \quad (4.69)$$

The trajectories (p_t^*, x_t^*), solutions of Hamilton's equations, are represented in Fig. 4.2 both in the non-biased and biased cases. There are four transient trajectories in both cases: two trajectories leading to the fixed point *forward* in time (relaxing) and two trajectories leading to the fixed point *backward* in time (fluctuating). When biasing, the original fixed point is shifted from its position x_{NB} (where NB stands for "Non-Biased") to a new concentration $x_\star$ fixed by the value $\boldsymbol{\gamma}$, but with $p_\star \neq 0$ translating the fact that the biased process is not norm-conserving.

In the following, we illustrate analytically and/or numerically the following points:

1. The dominant trajectory in the long-time limit approaches the transient trajectory connecting the boundary conditions.

2. Along a relaxing transient trajectory, it takes an infinite time for the system to reach the fixed point. Conversely, when the system is at the fixed point, it takes an infinite time to leave it via a fluctuating transient trajectory.

3. The value of the Hamiltonian along the transient trajectories is equal to the SCGF $\bar{\Gamma}(\boldsymbol{\gamma})$. In particular, this value is 0 for the non-biased Hamiltonian.

To show these points, we compute analytically the solutions of Eq. (4.67) with initial condition x_{i} at $t = 0$ and final condition x_{f} at $t = \mathcal{T}$. We obtain

$$\begin{cases} p_t^* = \ln\left[\left(\mathrm{e}^{p_{\mathrm{f}}} - \mathrm{e}^{p_\star}\right)\mathrm{e}^{\delta(t-\mathcal{T})} + \mathrm{e}^{p_\star}\right], \\ x_t^* = \left[\left(x_{\mathrm{i}}\mathrm{e}^{-p_0^*} - x_\star \mathrm{e}^{-p_\star}\right)\mathrm{e}^{-\delta t} + x_\star \mathrm{e}^{-p_\star}\right]\mathrm{e}^{p_t^*}. \end{cases} \quad (4.70)$$

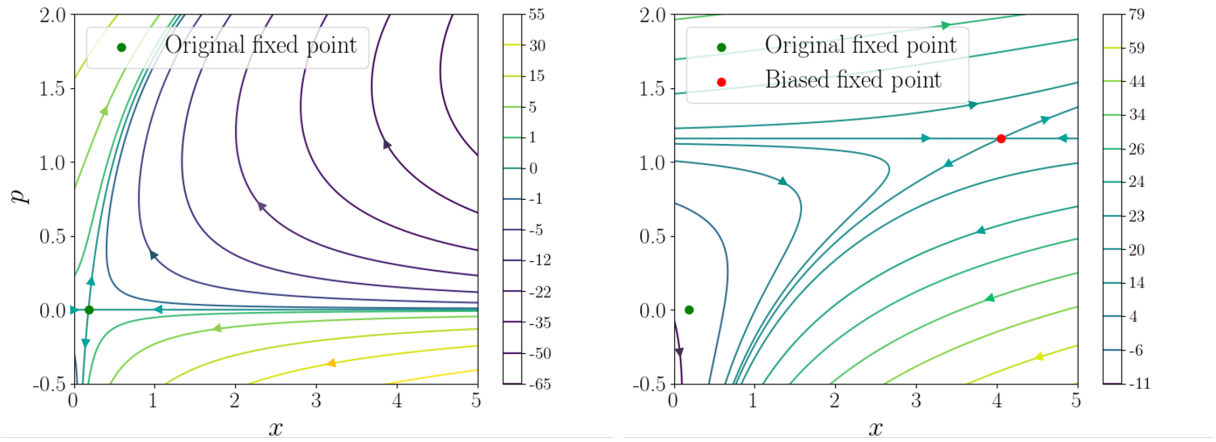

Figure 4.2 – (Left) Trajectories of the non-biased Hamiltonian ($\boldsymbol{\gamma} = 0$). The coordinates of the fixed point (green point) are $(x_{\text{NB}} = (\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b)/(\mathfrak{K}_{-1} + \mathfrak{K}_{+2}), p_{\text{NB}} = 0)$. (Right) Trajectories of the biased Hamiltonian. The coordinates of the fixed point (red point) are $(x_\star = \alpha\beta/\delta^2, p_\star = \ln(\beta/\delta))$.
The figures are obtained for $\mathfrak{K}_{+1}a = 3$, $\mathfrak{K}_{-1} = 1$, $\mathfrak{K}_{+2} = 3$, $\mathfrak{K}_{-2}b = 2$, $\gamma^1_{+1} = 1$, $\gamma^1_{-1} = -2$, $\gamma^1_{+2} = 2$, $\gamma^1_{-2} = -1$, $\gamma_2 = -3$ leading to $x_{\text{NB}} = 0.18$, $x_\star = 4.05$ and $p_\star = 1.16$.

For convenience, we imposed in Eq. (4.70) the final condition p_{f} instead of x_{f}, both being related by $x_{\text{f}} = \left[(x_{\text{i}}\mathrm{e}^{-p_0} - x_\star \mathrm{e}^{-p_\star})\,\mathrm{e}^{-\delta\mathcal{T}} + x_\star \mathrm{e}^{-p_\star}\right]\mathrm{e}^{p_{\text{f}}}$. The solutions (p^*_t, x^*_t) are represented in Fig. 4.3. We see that for each value of the final time $\mathcal{T}$, there is a unique trajectory connecting x_{i} to x_{f} (right figure). The greater $\mathcal{T}$ is, the closer the trajectory is to the transient trajectories. In the limit $\mathcal{T} \to \infty$ (with $\mathcal{T} < \infty$), (p^*_t, x^*_t) approaches asymptotically the transient trajectories (but never reach them!). We can see this result analytically from Eq. (4.70). For large final time $\mathcal{T}$, we have for t sufficiently away from $\mathcal{T}$

$$\begin{cases} p^*_t \underset{\mathcal{T}\to\infty}{\longrightarrow} p_\star, \\ x^*_t \underset{\mathcal{T}\to\infty}{\longrightarrow} (x_{\text{i}} - x_\star)\mathrm{e}^{-\delta t} + x_\star \underset{t\gtrsim\frac{1}{\delta}}{\simeq} x_\star, \end{cases} \tag{4.71}$$

showing that the system is infinitesimally close to the fixed point (but not strictly at it) (left figure), and then that the dominant trajectory in the long-time limit converges to the transient trajectories (point 1).

We now prove the point 2. We show that along the relaxing transient trajectory ($p^*_t = p_\star$, $\forall t$), it takes an infinite time to reach the fixed point. Starting from $x_{\text{i}} \neq x_\star$, the solutions for all t are

$$\begin{cases} p^*_t = p_\star, \\ x^*_t = (x_{\text{i}} - x_\star)\mathrm{e}^{-\delta t} + x_\star. \end{cases} \tag{4.72}$$

Hence, $x_\star$ is reached theoretically at $t = \infty$. Conversely, let us show that if the system is initially at the fixed point, it takes an infinite time to leave it. Taking $t = 0$ in Eq. (4.70), we obtain

$$\begin{cases} p^*_0 = \ln\left[(\mathrm{e}^{p_{\text{f}}} - \mathrm{e}^{p_\star})\,\mathrm{e}^{-\delta\mathcal{T}} + \mathrm{e}^{p_\star}\right], \\ x^*_0 = x_{\text{i}}. \end{cases} \tag{4.73}$$

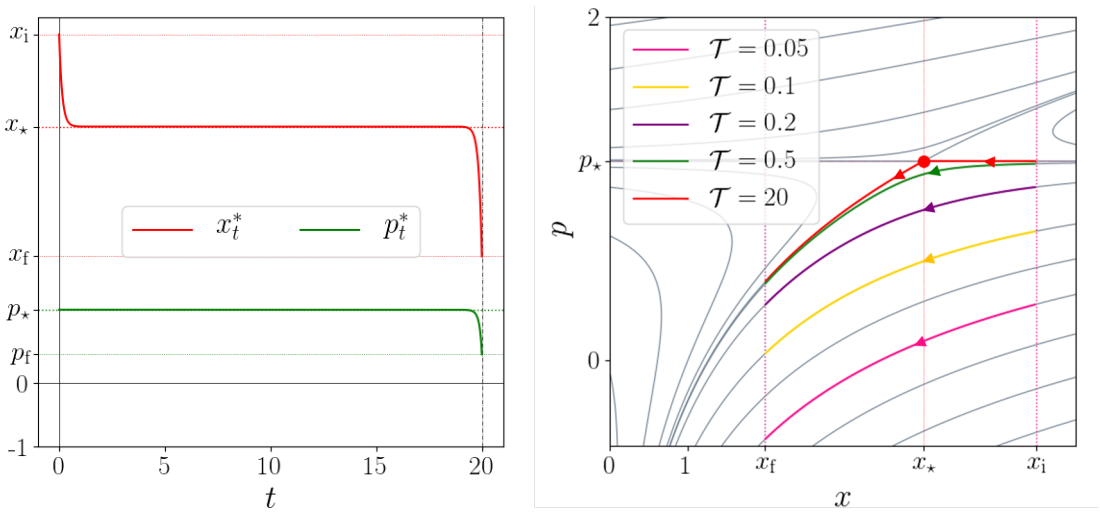

Figure 4.3 – (Left) p_t^* and x_t^* solutions of Hamilton's equations with boundary conditions $x_0^* = x_\mathrm{i}$ and $p_\mathcal{T}^* = p_\mathrm{f}$ for large $\mathcal{T}$. These solutions leave their initial values to approach closely the fixed point $(x_\star, p_\star)$, then reach their final values at $t = \mathcal{T}$ (red trajectory in the right figure). Note that the fixed point is not reached, but the solution is infinitesimally close to it ($x_t^* - x_\star \sim 10^{-15}$ for t sufficiently away from 0 and $\mathcal{T}$).
(Right) Trajectories followed by the system from x_i to x_f at time $\mathcal{T}$ for different values of $\mathcal{T}$. When we increase $\mathcal{T}$, the trajectory followed by the system approaches the two transient trajectories connecting x_i to x_f. When $\mathcal{T} \to \infty$, the trajectory followed by the system (red trajectory) is infinitesimally close to the two transient trajectories.
Both figures are obtained for $x_\mathrm{i} = 5.5$, $x_\mathrm{f} = 2$, $\mathfrak{K}_{+1}a = 3$, $\mathfrak{K}_{-1} = 1$, $\mathfrak{K}_{+2} = 3$, $\mathfrak{K}_{-2}b = 2$, $\gamma^1_{+1} = 1$, $\gamma^1_{-1} = -2$, $\gamma^1_{+2} = 2$, $\gamma^1_{-2} = -1$, $\gamma_2 = -3$ leading to $x_\star = 4.05$ and $p_\star = 1.16$. In the left figure, $\mathcal{T} = 20$.

We impose the final condition $p_\mathrm{f} \neq p_\star$ at $t = \mathcal{T}$. It means that if the system is initially at the fixed point $(x_\mathrm{i} = x_\star, p_0^* = p_\star)$, we have from the first equality of Eq. (4.73) that the final time $\mathcal{T}$ is necessarily infinite. Hence, the system is at the fixed point at any finite time t and leaves the fixed point only for t close to $\mathcal{T} = \infty$ to reach $(p_\mathrm{f}, x_\mathrm{f})$ theoretically at $t = \mathcal{T}$. This shows point 2.

Let us now discuss the point 3. Injecting the solutions of Eq. (4.70) in the expression of the biased Hamiltonian (4.46) leads to

$$H_{\boldsymbol{\gamma}}(p_t^*, x_t^*) = \alpha \left[\left(\mathrm{e}^{p_\mathrm{f}} - \frac{\beta}{\delta} \right) \mathrm{e}^{-\delta \mathcal{T}} + \frac{\beta}{\delta} \right] + \frac{\beta x_\mathrm{i}}{\left(\mathrm{e}^{p_\mathrm{f}} - \frac{\beta}{\delta} \right) \mathrm{e}^{-\delta \mathcal{T}} + \frac{\beta}{\delta}} - \delta x_\mathrm{i} - (\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b). \quad (4.74)$$

Notice that indeed $H_{\boldsymbol{\gamma}}(p_t^*, x_t^*)$ does not depend on time along the solution of Hamilton's and is equal to $H_{\boldsymbol{\gamma}}(p_0^*, x_0^*)$. In the long-time limit, we obtain

$$H_{\boldsymbol{\gamma}}(p_t^*, x_t^*) \xrightarrow[\mathcal{T} \to \infty]{} \frac{\alpha\beta}{\delta} - (\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b) = H_{\boldsymbol{\gamma}}(x_\star, p_\star), \quad (4.75)$$

for any boundary conditions x_i and p_f. This is expected since the dominant trajectory in the long-time limit converges to the trajectories leading to the fixed point. Let us now

prove that the SCGF $\bar{\Gamma}$ coincides with $H_{\boldsymbol{\gamma}}(p_\star, x_\star)$. From Eqs. (4.51, 4.57, 4.66), we have

$$\bar{\Gamma} = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{N}\mathcal{T}} \ln \mathbb{E}_{x_\mathrm{i}} \left[\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_t} \right] = \left\{ H_{\boldsymbol{\gamma}}(p_t^*, x_t^*) - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} p_t^* \dot{x}_t^* \mathrm{d}t \right\}, \tag{4.76}$$

for the final condition x_f maximizing the action. In the limit $\mathcal{T} \to \infty$, the system follows asymptotically the relaxing transient trajectory ($p = p_\star$) from x_i to $x_\star$, then it follows asymptotically the fluctuating transient trajectory from $x_\star$ to x_f (red trajectory in Fig. 4.3). The reduced action $S_\mathrm{r} = \int_0^{\mathcal{T}} p_t^* \dot{x}_t^* \mathrm{d}t$ along those trajectories reads then

$$S_\mathrm{r}(\mathcal{T}) = \int_{[x_t^*]_0^{\mathcal{T}}} p^*(x)\mathrm{d}x \underset{\mathcal{T}\to\infty}{\simeq} \int_{x_\mathrm{i}}^{x_\star} p^*(x)\mathrm{d}x + \int_{x_\star}^{x_\mathrm{f}} p^*(x)\mathrm{d}x, \tag{4.77}$$

with $p^*(x)$ the solution of the HJ equation corresponding to the trajectory (p_t^*, x_t^*) and where we used the fact that at the fixed point, the reduced action is zero since $\dot{x}_\star = 0$. Eq. (4.77) implies that the reduced action is finite in the limit $\mathcal{T} \to \infty$, hence

$$\lim_{\mathcal{T}\to\infty} \frac{S_\mathrm{r}(\mathcal{T})}{\mathcal{T}} = 0. \tag{4.78}$$

It follows from Eqs. (4.75, 4.76, 4.78) that $\bar{\Gamma} = H_{\boldsymbol{\gamma}}(p_\star, x_\star)$, proving the point 3.

Some remarks are worth to be done. For the non-biased Hamiltonian ($\boldsymbol{\gamma} = 0$), one can check that the SCGF $\bar{\Gamma}$ is zero as expected. We also have $p_\star(\boldsymbol{\gamma} = 0) = p_{\mathrm{NB}} = 0$, which is expected since the non-biased Hamiltonian has to vanish at $p = 0$ to ensure the normalization of the transition probability. When biasing ($\boldsymbol{\gamma} \neq 0$), the biased transition probability is not normalized anymore and $p_\star$ is shifted from 0 while $x_\star$ takes a new value which depends on $\boldsymbol{\gamma}$. Hence, biasing defines a new process with a new fixed point (corresponding to a new typical trajectory) but that is not associated with a normalized transition probability ($H_{\boldsymbol{\gamma}}(p = 0, x) \neq 0$). In the next section, we will see how to *rectify* the biased process in order to build a norm-conserving process. This *rectified process* corresponds to the driven process in the linear operator formalism.

4.1.3 Rectified process

In the linear operator formalism, the driven generator for time-homogeneous Markov processes is given by the Doob transform (3.23) of the biased generator $\boldsymbol{\kappa}$ using its dominant left eigenvector [80]:

$$\boldsymbol{K} \equiv \boldsymbol{\kappa}^{\mathrm{e}^{\boldsymbol{U}}} = \mathcal{D}(\mathrm{e}^{\boldsymbol{U}})\boldsymbol{\kappa}\,\mathcal{D}(\mathrm{e}^{-\boldsymbol{U}}) - \mathcal{D}(\mathrm{e}^{-\boldsymbol{U}})\mathcal{D}(\mathrm{e}^{\boldsymbol{U}}\boldsymbol{\kappa}), \tag{4.79}$$

where the vector $\mathrm{e}^{\boldsymbol{U}}$ of component $(\mathrm{e}^{\boldsymbol{U}})_x \equiv \mathrm{e}^{U_x}$ is the dominant left eigenvector of $\boldsymbol{\kappa}$ whose highest eigenvalue coincides with the SCGF $\Gamma = \mathcal{N}\bar{\Gamma}$ defined in Eq. (4.65). Note that the positivity of $\mathrm{e}^{\boldsymbol{U}}$ is ensured by the PF theorem.

Spectral problem in the continuous limit

Before deriving the Lagrangian and Hamiltonian associated with the dynamics generated by $\boldsymbol{K}$, let us first investigate in the continuous limit the spectral equation

$$\mathrm{e}^{\boldsymbol{U}}\boldsymbol{\kappa} = \mathcal{N}\bar{\Gamma}\mathrm{e}^{\boldsymbol{U}}. \tag{4.80}$$

To do so, assume that for $\boldsymbol{N} = O(\mathcal{N})$, $\mathcal{N} \to \infty$, there exists for any α a function $W_{\mathrm{s}}(\boldsymbol{z})$ such that

$$U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}} \simeq \mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \tag{4.81}$$

where the scalar product is performed over the states: $\mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} W_{\mathrm{s}} \equiv \sum_X \mathfrak{D}_{X,\alpha} \partial_{z_X} W_{\mathrm{s}}$. We investigate the nature of the function W_{s} by writing the spectral relation between $\boldsymbol{U}$ and $\boldsymbol{\kappa}$ (4.80):

$$\sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha}} \kappa_{\alpha,\boldsymbol{N}} + \mathrm{e}^{U_{\boldsymbol{N}}} \kappa_{\boldsymbol{N},\boldsymbol{N}} = \mathcal{N}\bar{\Gamma} \mathrm{e}^{U_{\boldsymbol{N}}} \tag{4.82}$$

$$\sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}}} \kappa_{\alpha,\boldsymbol{N}} + \kappa_{\boldsymbol{N},\boldsymbol{N}} = \mathcal{N}\bar{\Gamma} \tag{4.83}$$

$$\sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}}} k_{\alpha,\boldsymbol{N}} \mathrm{e}^{\gamma^1_\alpha} - \sum_\alpha k_{\alpha,\boldsymbol{N}} + \boldsymbol{\gamma}_2 \cdot \boldsymbol{N} = \mathcal{N}\bar{\Gamma}, \tag{4.84}$$

where we used Eq. (4.40) in the last equation. Taking the continuous limit (4.9) and using the assumption (4.81), we finally obtain

$$\sum_\alpha k_\alpha(\boldsymbol{z}) \left[\mathrm{e}^{\mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} W_{\mathrm{s}} + \gamma^1_\alpha} - 1 \right] + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z} = \bar{\Gamma}. \tag{4.85}$$

We recognize the biased Hamiltonian (4.43) in the left-hand side of Eq. (4.85):

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} = \mathfrak{D}^\dagger \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) = \bar{\Gamma}, \tag{4.86}$$

or equivalently

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) = \bar{\Gamma}. \tag{4.87}$$

The function W_{s} appears to be Hamilton's characteristic function, solution of the HJ equation (4.62) for the eigenrate $E = \bar{\Gamma}$. One keeps in mind that W_{s} depends on $\boldsymbol{\gamma}$. Then, the spectral problem in the linear operator formalism amounts to solving a Hamilton-Jacobi equation in the continuous limit. Following the same derivation for the dominant right eigenvector of $\boldsymbol{\kappa}$ leads to the same equation (4.87). Hence, Eq. (4.87) admits two solutions $W_{\mathrm{s,u}}$ corresponding to the left and right Perron-Frobenius eigenvectors of $\boldsymbol{\kappa}$. By convention, we call W_{s} the solution corresponding to the left eigenvector (this choice of notation will become clearer in Section 4.2). In the non-biased case ($\boldsymbol{\gamma} = 0$), $\bar{\Gamma} = 0$ and $W_{\mathrm{s}}(\boldsymbol{\gamma} = 0) = 0$, in line with the fact that the dominant eigenvalue and left eigenvector of the non-biased generator $\boldsymbol{k}$ are respectively 0 and $\mathrm{e}^{\boldsymbol{U}(\boldsymbol{\gamma}=0)} = \mathbf{1}$, implying $\boldsymbol{U}(\boldsymbol{\gamma} = 0) = 0$.

Rectified process

We now derive the Lagrangian associated with the driven generator $\boldsymbol{K}$. From Eq. (4.79) and using Eq. (4.40) and Eq. (4.80), $\boldsymbol{K}$ reads in the Dirac notation

$$\boldsymbol{K} = \sum_{\alpha,\boldsymbol{N}} \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}}} k_{\alpha,\boldsymbol{N}} \mathrm{e}^{\gamma^1_\alpha} \left| \boldsymbol{N} + \mathfrak{D}_\alpha \boldsymbol{N} \right\rangle \left\langle \boldsymbol{N} \right| - \sum_{\boldsymbol{N}} \left(\sum_\alpha k_{\alpha,\boldsymbol{N}} - \boldsymbol{\gamma}_2 \cdot \boldsymbol{N} + \mathcal{N}\bar{\Gamma} \right) \left| \boldsymbol{N} \right\rangle \left\langle \boldsymbol{N} \right|. \tag{4.88}$$

Replacing $\boldsymbol{k}$ by $\boldsymbol{K}$ in the procedure used to derive the original Lagrangian (Appendix 4.A), we obtain that the transition probability $P_{\boldsymbol{K},\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) \equiv \langle \boldsymbol{N}_\mathrm{f} \mid \mathrm{e}^{\delta t \boldsymbol{K}} \mid \boldsymbol{N}_\mathrm{i} \rangle$ associated with the driven process satisfies a LDP:

$$P_{\boldsymbol{K},\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\delta t \mathcal{N} \mathscr{L}^\mathrm{r}(\boldsymbol{\lambda},\boldsymbol{z};\boldsymbol{\gamma})} \delta(\dot{\boldsymbol{z}} - \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}), \tag{4.89}$$

where the *detailed rectified Lagrangian* is given by

$$\mathscr{L}^\mathrm{r}(\boldsymbol{\lambda},\boldsymbol{z};\boldsymbol{\gamma}) = \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda},\boldsymbol{z}) - \boldsymbol{\lambda}\cdot\boldsymbol{\mathfrak{D}}^\dagger \partial_{\boldsymbol{z}} W_\mathrm{s} + \bar{\Gamma}. \tag{4.90}$$

The corresponding *detailed rectified Hamiltonian* is obtained by taking the LF transform of Eq. (4.90) with respect to $\boldsymbol{\lambda}$ and it yields

$$\mathscr{H}^\mathrm{r}(\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}) = \sum_\alpha k_\alpha(\boldsymbol{z}) \left[\mathrm{e}^{f_\alpha + \gamma_\alpha^1 + \boldsymbol{\mathfrak{D}}_\alpha\cdot\partial_{\boldsymbol{z}} W_\mathrm{s}} - 1\right] + \boldsymbol{\gamma}_2\cdot\boldsymbol{z} - \bar{\Gamma}, \tag{4.91}$$

or equivalently

$$\mathscr{H}^\mathrm{r}(\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} + \boldsymbol{\mathfrak{D}}^\dagger\partial_{\boldsymbol{z}} W_\mathrm{s},\boldsymbol{z}) - \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{\mathfrak{D}}^\dagger\partial_{\boldsymbol{z}} W_\mathrm{s},\boldsymbol{z}), \tag{4.92}$$

where we used Eqs. (4.43, 4.86). Again, we can consider a coarse-grained description of the model by looking at the $\mathcal{N}$-large deviations of the empirical density flux $\dot{\boldsymbol{z}}$, and we define the *standard rectified Lagrangian* $L^\mathrm{r}(\dot{\boldsymbol{z}},\boldsymbol{z};\boldsymbol{\gamma})$ and the *standard rectified Hamiltonian* $H^\mathrm{r}(\boldsymbol{p},\boldsymbol{z};\boldsymbol{\gamma})$ similarly to the non-biased case:

$$L^\mathrm{r}(\dot{\boldsymbol{z}},\boldsymbol{z};\boldsymbol{\gamma}) \equiv \inf_{\boldsymbol{\lambda}\mid\dot{\boldsymbol{z}}=\boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}} \mathscr{L}^\mathrm{r}(\boldsymbol{\lambda},\boldsymbol{z};\boldsymbol{\gamma}), \tag{4.93}$$

$$H^\mathrm{r}(\boldsymbol{p},\boldsymbol{z};\boldsymbol{\gamma}) \equiv \sup_{\dot{\boldsymbol{z}}}\{\boldsymbol{p}\cdot\dot{\boldsymbol{z}} - L^\mathrm{r}(\dot{\boldsymbol{z}},\boldsymbol{z};\boldsymbol{\gamma})\} = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} = \boldsymbol{\mathfrak{D}}^\dagger\boldsymbol{p},\boldsymbol{z};\boldsymbol{\gamma}) \tag{4.94}$$

$$= H^\mathrm{r}(\boldsymbol{p} + \partial_{\boldsymbol{z}} W_\mathrm{s},\boldsymbol{z}) - H_{\boldsymbol{\gamma}}(\partial_{\boldsymbol{z}} W_\mathrm{s},\boldsymbol{z}). \tag{4.95}$$

As expected for Hamiltonians deriving from norm-conserving Markov generators, the rectified Hamiltonian vanishes at $\boldsymbol{p} = 0$.

Illustration with the toy model

Let us compute the rectified Hamiltonian for our toy model (4.30):

$$H^\mathrm{r}(p,x;\boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(p + \partial_x W_\mathrm{s},x) - H_{\boldsymbol{\gamma}}(\partial_x W_\mathrm{s},x) \tag{4.96}$$

To do so, we first solve the Hamilton-Jacobi equation

$$H_{\boldsymbol{\gamma}}(p = \partial_x W,x) = \bar{\Gamma}. \tag{4.97}$$

This equation admits exactly two solutions

$$W_\mathrm{s}(x) \equiv p_\star x \tag{4.98}$$

$$W_\mathrm{u}(x) \equiv x\ln\left(\frac{x\delta}{\alpha}\right) - x, \tag{4.99}$$

up to an additive constant that we set equal to 0 since only the derivative $\partial_x W$ appears in the definition of the rectified Hamiltonian. Note that W_s is the solution associated with

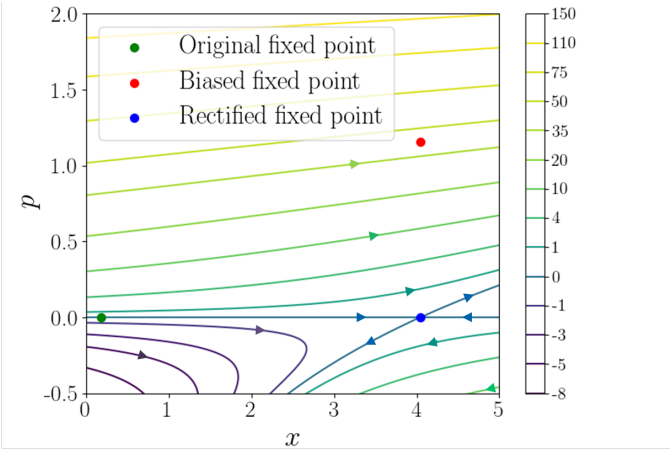

Figure 4.4 – Trajectories of rectified biased Hamiltonian. The coordinates of the fixed point (blue point) are $(x_\star, p = 0)$.
The figures are obtained for $\mathfrak{K}_{+1}a = 3$, $\mathfrak{K}_{-1} = 1$, $\mathfrak{K}_{+2} = 3$, $\mathfrak{K}_{-2}b = 2$, $\gamma^1_{+1} = 1$, $\gamma^1_{-1} = -2$, $\gamma^1_{+2} = 2$, $\gamma^1_{-2} = -1$, $\gamma_2 = -3$ leading to $x_\star = 4.05$ and $p_\star = 1.16$.

the dominant left eigenvector of the biased generator as it satisfies $W_{\mathrm{s}}(\boldsymbol{\gamma} = 0) = 0$. In Section 4.2, we give a more rigorous criterion for the choice of W_{s}. Using the expression of W_{s} in Eq. (4.97), we obtain the rectified Hamiltonian

$$H^{\mathrm{r}}(p, x; \boldsymbol{\gamma}) = \delta x_\star \left(\mathrm{e}^{p} - 1\right) + \delta x \left(\mathrm{e}^{-p} - 1\right). \tag{4.100}$$

Contrary to the biased Hamiltonian, the rectified Hamiltonian is a proper Hamiltonian ($H(p = 0, x) = 0$) associated with a norm-conserving Marvov process, which guarantees that the concentrations following from this Hamiltonian are physical concentrations. Unsurprisingly, it has the same structure as the original non-biased Hamiltonian $H(p, x) = (\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b)[\mathrm{e}^{p} - 1] + (\mathfrak{K}_{-1} + \mathfrak{K}_{+2})x[\mathrm{e}^{-p} - 1]$, meaning that the rectified Hamiltonian also describes the chemical reaction network of Eq. (4.30) but with modified kinetic constants $\mathfrak{K}^{\mathrm{r}}_{\epsilon\mathrm{r}}$ that we control through $\boldsymbol{\gamma}$. The rectified dynamics admits a fixed point with the same concentration $x_\star$ as for the biased Hamiltonian with the difference that the associated variable p is now 0, in line with the rectification procedure (see Fig. 4.4 for an illustration).

Discussion on the equivalence with the microcanonical process

In the linear operator formalism, the driven process (with the appropriate value of $\boldsymbol{\gamma}$) is equivalent to the microcanonical process, i.e. the process conditioned on one value of the observable $\boldsymbol{A}$. Similarly, we expect the typical trajectory of the rectified Hamiltonian H^{r} in the long-time limit to be such that $\boldsymbol{A}$ takes a new typical value according to the value of $\boldsymbol{\gamma}$. This assertion is backed by the derivation of the rectified Hamiltonian from the driven generator. In Section 4.2.4, we will show this result directly from the logarithmic equivalence of path probabilities independently of the linear operator formalism.

4.2 Nonlinear Markov processes: general formalism

On the basis of the results derived in the previous section, we propose a general and model-free theory of the rectification of time-independent Hamiltonians. This section may seem redundant with the previous section on certain points but it provides an abstract framework for studying nonlinear Markov processes without specifying their generators. Moreover, the theory is dealt here more rigorously and more deeply than previously. Population processes appear then to be an application of the results developed in this section. We made the choice to expose first the derivations done on population processes to justify the following general theory in hopes that it enlightens the reader.

4.2.1 Lagrangian/Hamiltonian formalism

We consider time-homogeneous Markov processes characterized by a large size-type parameter $\mathcal{N}$ (number of particles, volume, etc.) and described by two empirical observables: a current variable $\boldsymbol{\lambda}$ and a state variable $\boldsymbol{z}$. These variables will have precise definitions in specific contexts. For instance, in the case of chemical reaction networks (4.28), the variable $\boldsymbol{\lambda}$ is the empirical chemical current and the variable $\boldsymbol{z}$ is the empirical concentration vector. The dynamics of $\boldsymbol{z}$ is determined by the currents $\boldsymbol{\lambda}$ through a conservation law:

$$\dot{\boldsymbol{z}} = \mathcal{D}\boldsymbol{\lambda}, \tag{4.101}$$

where $\mathcal{D}$ stands for a generalized differential operator that will have precise definitions in specific contexts. For example, in the case of chemical network reactions (4.28), $\mathcal{D}$ becomes the matrix $\mathfrak{D}$ whose component $\mathfrak{D}_{X,\epsilon\mathsf{r}} = \nu_X^{-\epsilon\mathsf{r}} - \nu_X^{\epsilon\mathsf{r}}$ is the variation of the number of species X when reaction $\epsilon\mathsf{r}$ occurs.

We are interested in the transition probability $P(\boldsymbol{z}', t+\delta t \mid \boldsymbol{z}, t)$ of observing $\boldsymbol{z}_{t+\delta t} = \boldsymbol{z}'$ at time $t+\delta t$ given $\boldsymbol{z}_t = \boldsymbol{z}$ at time t, with δt an infinitesimal time. Since we consider time-homogeneous Markov processes, the transition probability depends only on the difference δt between final and initial times and we write $P_{\delta t}(\boldsymbol{z}' \mid \boldsymbol{z})$ the conditional probability to observe $\boldsymbol{z}'$ after a time δt given that the system was in $\boldsymbol{z}$. From Eq. (4.101), observing $\boldsymbol{z}'$ after δt given $\boldsymbol{z}$ is entirely determined by the data of $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ since $\boldsymbol{z}_{t+\delta t} = \boldsymbol{z}_t + \delta t \mathcal{D}\boldsymbol{\lambda}$. We can thus equivalently consider $P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z})$ the conditional probability of the current variable $\boldsymbol{\lambda}$ given the state variable $\boldsymbol{z}$ during the infinitesimal time interval δt. We assume that this probability satisfies a LDP whose LDF is the detailed Lagrangian $\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})$:

$$\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}) \equiv -\lim_{\mathcal{N}\to\infty} \frac{1}{\delta t \mathcal{N}} \ln P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}), \tag{4.102}$$

with $\delta t \to 0$, $\delta t \mathcal{N} \to \infty$, and we write

$$P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\delta t \mathcal{N} \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})}. \tag{4.103}$$

For $P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z})$ to be indeed a propability, the detailed Lagrangian must satisfy for all $\boldsymbol{z}$

$$\begin{cases} \forall \boldsymbol{\lambda},\ \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}) \geq 0, \\ \exists\, \boldsymbol{\lambda}_*(\boldsymbol{z}),\ \mathscr{L}(\boldsymbol{\lambda}_*(\boldsymbol{z}), \boldsymbol{z}) = 0. \end{cases} \tag{4.104}$$

We can also consider a more coarse-grained description of the system by considering the standard Lagrangian $L(\dot{\boldsymbol{z}}, \boldsymbol{z})$ obtained from the contraction principle (1.35) over $\boldsymbol{\lambda}$ under the constraint (4.101) by:

$$L(\dot{\boldsymbol{z}}, \boldsymbol{z}) \equiv \inf_{\boldsymbol{\lambda}|\dot{\boldsymbol{z}}=\mathcal{D}\boldsymbol{\lambda}} \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}). \tag{4.105}$$

This is the usual Lagrangian used in analytical mechanics. This Lagrangian is usually difficult to compute explicitly, contrary to the detailed Lagrangian which has an explicit formula for a wide number of systems, such as systems modeled by diffusive processes or by Markov jump processes.

Reciprocally, we can consider the detailed Hamiltonian $\mathscr{H}(\boldsymbol{f}, \boldsymbol{z})$ corresponding to the scaled cumulant generating function for $\boldsymbol{\lambda}$ obtained from the Legendre-Fenchel transform of $\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})$:

$$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z}) \equiv \sup_{\boldsymbol{\lambda}} \{\boldsymbol{f} \cdot \boldsymbol{\lambda} - \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})\}, \tag{4.106}$$

where the central dot $\cdot$ denotes the scalar product in current space and $\boldsymbol{f}$ is conjugated to $\boldsymbol{\lambda}$. Note that $\mathscr{H}$ is convex in $\boldsymbol{f}$ since it follows from a LF transform with respect to $\boldsymbol{\lambda}$. Hamiltonians associated with proper stochastic processes must satisfy, $\forall \boldsymbol{z}$,

$$\mathscr{H}(\boldsymbol{f} = 0, \boldsymbol{z}) = 0 \tag{4.107}$$

to ensure that $P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z})$ is a propability. Indeed, condition (4.104) and

$$\mathscr{H}(\boldsymbol{f} = 0, \boldsymbol{z}) = \sup_{\boldsymbol{\lambda}} \{-\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})\} = -\inf_{\boldsymbol{\lambda}} \{\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})\} \tag{4.108}$$

imply Eq. (4.107). Similarly, the standard Hamiltonian $H(\boldsymbol{p}, \boldsymbol{z})$ is obtained from the LF transform of the standard Lagrangian $L(\dot{\boldsymbol{z}}, \boldsymbol{z})$:

$$H(\boldsymbol{p}, \boldsymbol{z}) \equiv \sup_{\dot{\boldsymbol{z}}} \{\boldsymbol{p} \cdot \dot{\boldsymbol{z}} - L(\dot{\boldsymbol{z}}, \boldsymbol{z})\}. \tag{4.109}$$

Following the same calculation as Eqs. (4.21–4.26), standard and detailed Hamiltonians are simply related by:

$$\mathscr{H}(\boldsymbol{f} = \mathcal{D}^\dagger \boldsymbol{p}, \boldsymbol{z}) = H(\boldsymbol{p}, \boldsymbol{z}), \tag{4.110}$$

where $\mathcal{D}^\dagger$ is the adjoint of $\mathcal{D}$. Since the Hamiltonian framework is more convenient than the Lagrangian framework as glimpsed in Section 4.1, we will essentially use the Hamiltonian framework in the remaining of this chapter.

Biased Lagrangian and Hamiltonian

We are now interested in the fluctuations in the limit of large parameter $\mathcal{N}$ of the two-component observable $\bar{\boldsymbol{A}}_t$ defined by

$$\bar{\boldsymbol{A}}_t \equiv \frac{1}{t} \begin{pmatrix} \int_0^t \mathrm{d}t' \boldsymbol{\lambda}(t') \\ \int_0^t \mathrm{d}t' \boldsymbol{z}(t') \end{pmatrix}. \tag{4.111}$$

We use an overbar to emphasize that the observable is scaled with $\mathcal{N}$ such that $\mathcal{N}\bar{\boldsymbol{A}}$ is an extensive observable. The statistics of $\bar{\boldsymbol{A}}_t$ is contained in the generating function

$$G_{\boldsymbol{\gamma}}(t) = \mathbb{E}_{\boldsymbol{z}_\mathrm{i}} \left[\mathrm{e}^{\mathcal{N} t \boldsymbol{\gamma} \cdot \bar{\boldsymbol{A}}_t} \right], \tag{4.112}$$

where $\mathbb{E}_{\boldsymbol{z}_\mathrm{i}}[\dots]$ is the path average with respect to the path probability $\mathbb{P}_t[\boldsymbol{z} \mid \boldsymbol{z}_\mathrm{i}]$ of the trajectory $[\boldsymbol{z}]$ up to time t given the initial state $\boldsymbol{z}_\mathrm{i}$ at time $t = 0$:

$$G_{\boldsymbol{\gamma}}(t) = \int \mathfrak{D}[\boldsymbol{\lambda}, \boldsymbol{z}] \mathrm{e}^{\mathcal{N} t \boldsymbol{\gamma} \cdot \bar{\boldsymbol{A}}_t} \mathbb{P}_t[\boldsymbol{z} \mid \boldsymbol{z}_\mathrm{i}] \delta(\dot{\boldsymbol{z}} - \mathcal{D}\boldsymbol{\lambda}). \tag{4.113}$$

Using the fact that

$$\mathbb{P}_t[\boldsymbol{z} \mid \boldsymbol{z}_\mathrm{i}] = \prod_{\ell=0}^{M} P_{\delta t}(\boldsymbol{\lambda}_{\tau_\ell} \mid \boldsymbol{z}_{\tau_\ell}), \tag{4.114}$$

where the product runs over times $\tau_\ell \equiv \ell \delta t$ with initial time $\tau_0 = 0$ and final time $\tau_M = t$, it follows from Eq. (4.113) that

$$G_{\boldsymbol{\gamma}}(t) = \int \prod_{\ell=0}^{M} \mathrm{d}\boldsymbol{\lambda}_{\tau_\ell} \mathrm{d}\boldsymbol{z}_{\tau_\ell} G_{\delta t}(\boldsymbol{\lambda}_{\tau_\ell} \mid \boldsymbol{z}_{\tau_\ell}) \delta(\boldsymbol{z}_{\tau_{\ell+1}} - \boldsymbol{z}_{\tau_\ell} - \delta t \mathcal{D} \boldsymbol{\lambda}_{\tau_\ell}), \tag{4.115}$$

where we introduced the biased transition probability $G_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z})$ during the infinitesimal time δt

$$G_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}) \equiv P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}) \mathrm{e}^{\mathcal{N} \delta t (\boldsymbol{\gamma}_1 \cdot \boldsymbol{\lambda} + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z})}. \tag{4.116}$$

From Eqs. (4.103, 4.116), we find that the biased transition probability is associated with the detailed biased Lagrangian $\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z})$:

$$G_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z}) \underset{\mathcal{N} \to \infty}{\asymp} \mathrm{e}^{-\mathcal{N} \delta t \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z})}, \tag{4.117}$$

with

$$\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z}) \equiv \mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}) - \boldsymbol{\gamma}_1 \cdot \boldsymbol{\lambda} - \boldsymbol{\gamma}_2 \cdot \boldsymbol{z}. \tag{4.118}$$

The detailed biased Hamiltonian $\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{z})$ is given by the LF transform of the detailed biased Lagrangian:

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{z}) \equiv \sup_{\boldsymbol{\lambda}} \{\boldsymbol{f} \cdot \boldsymbol{\lambda} - \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z})\} = \mathscr{H}(\boldsymbol{f} + \boldsymbol{\gamma}_1, \boldsymbol{z}) + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z}, \tag{4.119}$$

where we used in the second equality Eqs. (4.106, 4.118) and the fact that $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ are independent in Eq. (4.116). Note that the standard biased Lagrangian $L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}, \boldsymbol{z})$ and Hamiltonian $H_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{z})$ follow from the detailed ones as in the non-biased case, namely

$$L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}, \boldsymbol{z}) \equiv \inf_{\boldsymbol{\lambda} | \dot{\boldsymbol{z}} = \mathcal{D}\boldsymbol{\lambda}} \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z}), \tag{4.120}$$

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}) \equiv \sup_{\dot{\boldsymbol{z}}} \{\boldsymbol{p} \cdot \dot{\boldsymbol{z}} - L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}, \boldsymbol{z})\} = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} = \mathcal{D}^\dagger \boldsymbol{p}, \boldsymbol{z}). \tag{4.121}$$

The biased Lagrangian and Hamiltonian are not associated with a norm-conserving Markov process as they do not respect conditions (4.104) and (4.107), respectively. In the next section, we define a transformation on the biased Lagrangian and Hamiltonian that restores these conditions. This transformation translates in the Lagrangian/Hamiltonian language the generalized Doob transform used to build the driven generator in the linear operator formalism. This requires to investigate the nonlinear counterpart of the Perron-Frobenius theorem and the dominant eigenvalue and eigenvectors of the biased generator on which is based the definition of the driven generator.

Evolution of the dynamics

As discussed in Section 4.1.2, the trajectory typically followed by the system given an initial state $\boldsymbol{z}_\mathrm{i}$ and a final state $\boldsymbol{z}_\mathcal{T} = \boldsymbol{z}_\mathrm{f}$ is the one solving Hamilton's equations:

$$\dot{\boldsymbol{z}}_t = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t), \qquad \dot{\boldsymbol{p}}_t = -\partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t), \tag{4.122}$$

or equivalently the HJ equation

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} W, \boldsymbol{z}) = E \tag{4.123}$$

with an eigenrate E compatible with the boundary conditions. We say that a solution $W(\boldsymbol{z}, E)$ of this equation is *global* if it is defined and analytic for all $\boldsymbol{z}$. We define the corresponding *reduced dynamics* describing the evolution of the state variable only by the equation

$$\dot{\boldsymbol{z}} = \left.\frac{\partial H_{\boldsymbol{\gamma}}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=\partial_{\boldsymbol{z}} W, \boldsymbol{z}}. \tag{4.124}$$

This dynamics is said to be *globally stable* (respectively *globally unstable*) if there exists a compact set C in $\boldsymbol{z}$-space such that all trajectories of the reduced dynamics converge to (respectively exit from) C, i.e.

$$\forall \boldsymbol{z}_\mathrm{i}, \exists t^\star \in \mathbb{R}^\pm, \forall t \lessgtr t^\star, \boldsymbol{z}_t \in C. \tag{4.125}$$

We will see that such stability conditions can guarantee the existence of critical manifolds, defined in Section 4.1.2.

4.2.2 Spectral properties of the Hamiltonian

The rectification procedure in the linear operator formalism relies heavily on the Perron-Frobenius theorem since it ensures the non-degeneracy of the largest eigenvalue of the biased generator and the positivity of its dominant left eigenvector, both used in the definition of the driven generator. In order to extend the rectification to nonlinear processes, one needs to translate the Perron-Frobenius theorem in the Hamiltonian framework in which the spectral problem is expressed by a Hamilton-Jacobi equation. Given the difficulty of such a generalization, we instead propose a conjecture based on physically reasonable assumptions on the structure of the Hamiltonian under consideration.

Assumptions on statistical Hamiltonians

In the following, we make a series of assumptions on the properties of the Hamiltonians we consider. We assume these properties to be generically preserved under biasing, if not we restrain $\boldsymbol{\gamma}$ to the values for which it is the case. We call the class of Hamiltonians satisfying the following properties *statistical Hamiltonians*. Without loss of generality, we focus on the biased Hamiltonian $H_{\boldsymbol{\gamma}}$ (the non-biased case follows from $\boldsymbol{\gamma} = 0$) and illustrate numerically each assumption on the nonlinear model called "Brownian Donkey" that will be studied in Section 4.3.1. In the remaining of this manuscript, we assume $\boldsymbol{z}$ and $\boldsymbol{p}$ to be defined on $\mathbb{R}^\mathrm{n}$ or open sets of $\mathbb{R}^\mathrm{n}$ (with n an integer).

First, $H_{\boldsymbol{\gamma}}$ is convex at $\boldsymbol{p}$ for any $\boldsymbol{z}$ since it follows from a LF transform. We assume in addition that it is *strictly convex* as well as *coercive*, i.e. for any $\boldsymbol{z}$, $H(\boldsymbol{p}, \boldsymbol{z}) \to \infty$ when

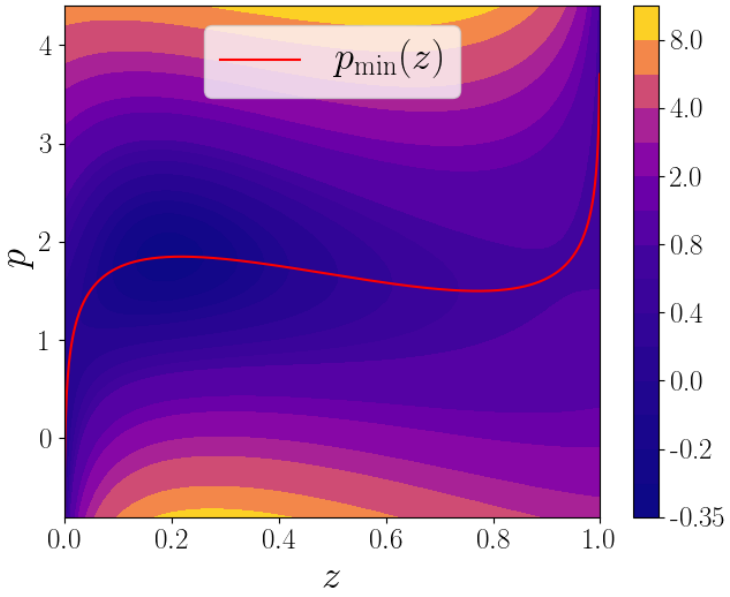

Figure 4.5 – $H_{\boldsymbol{\gamma}}(p, z)$ and $p_{\min}(z)$ for the model of the Brownian Donkey. $H_{\boldsymbol{\gamma}}(p, z)$ is strictly convex in p and diverges for infinite values of p for any z (except near the edges of z where the assumptions of the model are not verified anymore), implying the existence of a unique minimum $p_{\min}(z)$ as illustrated on the figure.

$|\boldsymbol{p}| \to \infty$, where $|\boldsymbol{p}|$ is the Euclidean norm of $\boldsymbol{p}$. Given these assumptions, there is for any $\boldsymbol{z}$ a unique value $\boldsymbol{p} = \boldsymbol{p}_{\min}(\boldsymbol{z})$ that minimizes $H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z})$:

$$\partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\min}(\boldsymbol{z}), \boldsymbol{z}) = 0 \quad \text{and} \quad \partial^2_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\min}(\boldsymbol{z}), \boldsymbol{z}) > 0, \tag{4.126}$$

see Figure 4.5. From the first equation of Eq. (4.122), the minimizer $\boldsymbol{p}_{\min}(\boldsymbol{z})$ is associated with a stopping point for $\boldsymbol{z}$, i.e. $\dot{\boldsymbol{z}} = 0$. We define, for future use, the minimal value of $H_{\boldsymbol{\gamma}}$ for each $\boldsymbol{z}$:

$$H_{\min}(\boldsymbol{z}) \equiv H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\min}(\boldsymbol{z}), \boldsymbol{z}). \tag{4.127}$$

Second, there exists a compact set B such that $H_{\min}(\boldsymbol{z})$ admits at least one maximum inside B and no extrema outside. Note that the extrema $\{\boldsymbol{z}_\star\}$ of $H_{\min}(\boldsymbol{z})$ are the positions of the fixed points $\{(\boldsymbol{p}_\star = \boldsymbol{p}_{\min}(\boldsymbol{z}_\star), \boldsymbol{z}_\star)\}$ of the Hamiltonian dynamics since at an extremum $\boldsymbol{z}_\star$, we have

$$\dot{\boldsymbol{p}} = \partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\min}(\boldsymbol{z}_\star), \boldsymbol{z}_\star) = 0. \tag{4.128}$$

We label $\boldsymbol{z}^\star_\ell$ ($\ell = 0, 1, \dots$) the positions of the maxima of $H_{\boldsymbol{\gamma}}$ on the manifold $\boldsymbol{p} = \boldsymbol{p}_{\min}(\boldsymbol{z})$ (which we assume to be countable for the sake of simplicity) and we introduce $\boldsymbol{p}^\star_\ell \equiv \boldsymbol{p}_{\min}(\boldsymbol{z}^\star_\ell)$. We define $H^\star_\ell(\boldsymbol{\gamma}) \equiv H_{\min}(\boldsymbol{z}^\star_\ell) = H(\boldsymbol{p}^\star_\ell, \boldsymbol{z}^\star_\ell)$ and choose the indices of $\boldsymbol{z}^\star_\ell$ such that $H^\star_0(\boldsymbol{\gamma}) \geq H^\star_1(\boldsymbol{\gamma}) \geq H^\star_2(\boldsymbol{\gamma}) \geq \dots$, so that $H^\star_0(\boldsymbol{\gamma}) = \max_\ell H^\star_\ell(\boldsymbol{\gamma})$, see Figure 4.6. The corresponding fixed point $(\boldsymbol{p}^\star_0, \boldsymbol{z}^\star_0)$ is particularly important and will be called the *dominant fixed point* for reasons that will be explained in section 4.2.3.

Finally, we assume that the absolute maximum of $H_{\min}(\boldsymbol{z})$ is non-degenerate, i.e. $H^\star_0(\boldsymbol{\gamma}) > H^\star_1(\boldsymbol{\gamma})$, in order to avoid first-order phase transitions (see Appendix 4.D for

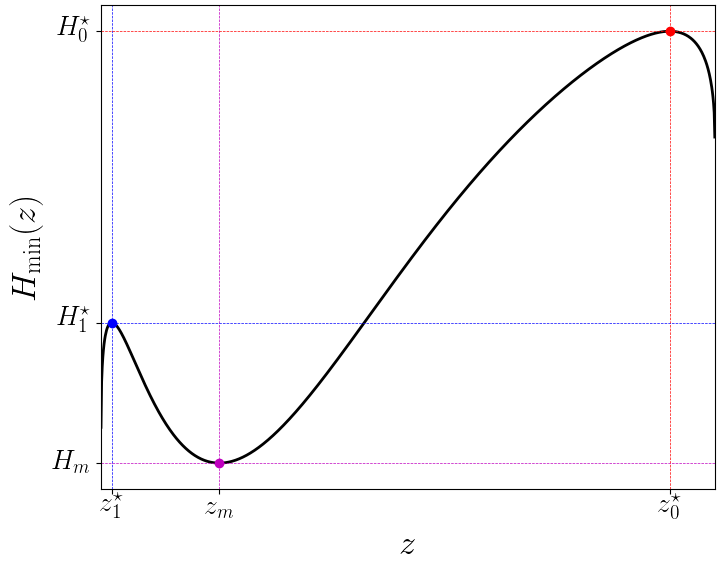

Figure 4.6 – $H_{\min}(z)$ vs z for the model of the Brownian Donkey. It admits three extrema corresponding to the positions of the fixed points: two maxima $z_0^\star$ and $z_1^\star$ and one minimum z_m. All the extrema are contained in $B = [z_1^\star, z_0^\star]$.

an example). This implies that

$$H_0^\star(\boldsymbol{\gamma}) = \max_{\boldsymbol{z}} \min_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}), \tag{4.129}$$

which will be an important object in the following conjecture.

Conjecture for a nonlinear generalization of the Perron-Frobenius theorem

Under the assumptions of the previous section, we make the following conjecture concerning the solutions of the HJ equation (4.62):

Conjecture 1 *There exists a value $E^\star(\boldsymbol{\gamma})$ of $H_{\boldsymbol{\gamma}}$ such that*

1. *For $E > E^\star(\boldsymbol{\gamma})$, all trajectories tend towards the boundaries of the system forward and backward in time, so that none of them contain or reach a critical manifold (fixed points, limit cycles, strange attractors, etc.).*

2. *For $E < E^\star(\boldsymbol{\gamma})$, there is no global solution to the HJ equation, and the reduced action of any solution W along any bounded trajectory (such as closed trajectories or strange attractors) is nonnegative: $\int \partial_{\boldsymbol{z}} W \cdot \mathrm{d}\boldsymbol{z} \geq 0$.*

3. *For $E = E^\star(\boldsymbol{\gamma})$, the HJ equation admits at least two global solutions (up to an additive constant). Among these solutions, there is exactly one globally stable solution $W_{\mathrm{s}}(\boldsymbol{z}, \boldsymbol{\gamma})$, and one globally unstable solution $W_{\mathrm{u}}(\boldsymbol{z}, \boldsymbol{\gamma})$. These two solutions $W_{\mathrm{s}}(\boldsymbol{z}, \boldsymbol{\gamma})$ and $W_{\mathrm{u}}(\boldsymbol{z}, \boldsymbol{\gamma})$ coincide on each of their critical manifolds.*

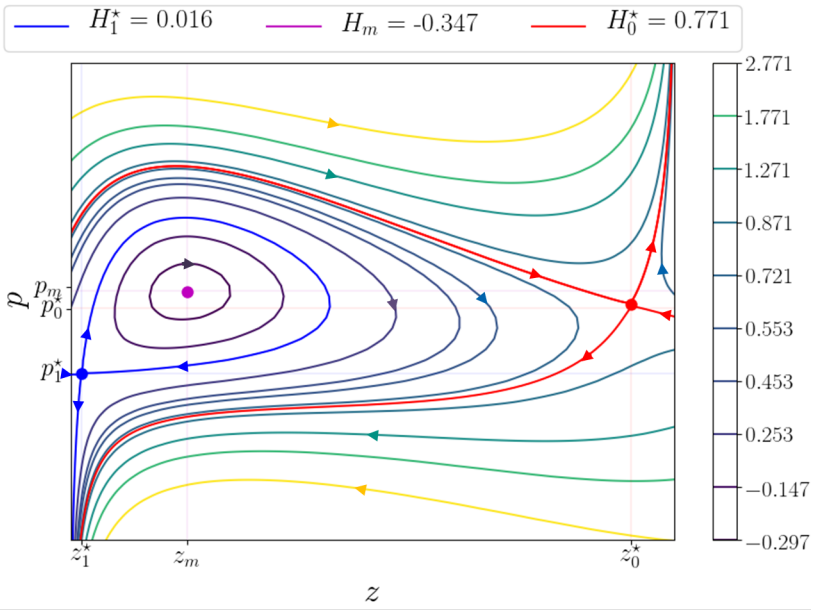

Figure 4.7 – Trajectories in the phase space associated with $H_{\boldsymbol{\gamma}}(p, z)$ for the model of the Brownian donkey. The red, mangenta and blue points correspond respectively to the fixed points of positions $z_0^\star$, z_m and $z_1^\star$ given by the extrema of $H_{\min}(z)$ as illustrated in Figure 4.6. In red, the trajectory of eigenrate $E = H_0^\star$.

4. *The dominant fixed point* $(\boldsymbol{p}_0^\star, \boldsymbol{z}_0^\star)$ *is contained in both the globally stable solution* $W_{\mathrm{s}}(\boldsymbol{z}, \boldsymbol{\gamma})$ *and the globally unstable solution* $W_{\mathrm{u}}(\boldsymbol{z}, \boldsymbol{\gamma})$*. The critical value* $E^\star(\boldsymbol{\gamma})$ *of the eigenrate can therefore be obtained by a max-min formula:*

$$E^\star(\boldsymbol{\gamma}) = H_0^\star(\boldsymbol{\gamma}) = \max_{\boldsymbol{z}} \min_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}). \tag{4.130}$$

By analogy with the PF theorem, $E^\star(\boldsymbol{\gamma})$ corresponds to the dominant eigenvalue, $W_{\mathrm{s}}(\boldsymbol{z}, \boldsymbol{\gamma})$ corresponds to the dominant left eigenvector, which is the solution that vanishes when $\boldsymbol{\gamma} = 0$, and $W_{\mathrm{u}}(\boldsymbol{z}, \boldsymbol{\gamma})$ corresponds to the dominant right eigenvector, and determines the stationary distribution of $\boldsymbol{z}$ when $\boldsymbol{\gamma} = 0$. Fig. 4.7 provides an illustration of this conjecture. We see that for values E of the Hamiltonian smaller than $H_0^\star$ (trajectories between the magenta point and the red trajectory), there are intervals of z for which the equation $H_{\boldsymbol{\gamma}}(p, z) = E$ does not admit solutions for p. Starting from $E = H_0^\star$, we see that $H_{\boldsymbol{\gamma}}(p, z) = E$ admits two solutions for p for any z. For $E > H_0^\star$, all trajectories tend to the boundaries of the state-space forward and backward in time. The reader may refer to Ref. [263] for additional remarks about this conjecture that go beyond the framework of this manuscript.

4.2.3 Long-time limit and SCGF

As in the linear operator formalism, we are interested in finding an equivalent process to the conditioned process in the long-time limit. Finding this process relies on the Perron-Frobenius theorem in the linear operator formalism. In the Lagrangian/Hamiltonian

framework, we use instead our conjecture to obtain similar information. More specifically, the globally stable and unstable solutions mentioned in points 3 and 4 contain the long-time dynamics of the system in the sense that, for any choice of boundary conditions, the trajectories that dominate the action in the long-time limit are included in those two manifolds. Under the assumptions of Sec. 4.2.2 and using the previous conjecture, we have the following result:

In the long-time limit $\mathcal{T} \to \infty$, the trajectory $(\boldsymbol{p}_t^*, \boldsymbol{z}_t^*)_{t\in[0,\mathcal{T}]}$ dominating the path integral is such that $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t) \to E^\star(\boldsymbol{\gamma})$ and approaches the trajectory with the following structure:

- *A relaxation phase in which the system follows a transient trajectory of the stable manifold (corresponding to $\boldsymbol{p} = \partial_{\boldsymbol{z}} W_{\mathrm{s}}$) from $\boldsymbol{z}_{\mathrm{i}}$ to the associated critical manifold.*
- *A stationary phase (or switching phase) in which the system remains at the critical manifold (or alternates between multiple critical manifolds through the trajectories connecting them).*
- *A fluctuation phase in which the system leaves the critical manifold to reach $\boldsymbol{z}_{\mathrm{f}}$ via a transient trajectory of the unstable manifold (corresponding to $\boldsymbol{p} = \partial_{\boldsymbol{z}} W_{\mathrm{u}}$).*

Moreover, the SCGF is given by

$$\bar{\Gamma}(\boldsymbol{\gamma}) = H_0^\star(\boldsymbol{\gamma}) = E^\star(\boldsymbol{\gamma}). \tag{4.131}$$

This result can be proven when there is a single critical manifold at $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$ (i.e. the dominant fixed point $(\boldsymbol{p}_0^\star, \boldsymbol{z}_0^\star)$ by point 4). The proof, which we present in the following, relies on first showing that such a trajectory exists, and then that any trajectory of higher or lower eigenrate has necessarily a lower Jacobi's action. For complex cases with more that one dominant critical manifold, the statement above is presented as a conjecture. The reader can lean on Fig. 4.7 to illustrate each argument.

Let us assume that the dominant fixed point is the only critical manifold at $H_{\boldsymbol{\gamma}} = E^\star(\boldsymbol{\gamma})$. We first show that there exists a trajectory connecting the initial and final conditions according to the description above. Using point 3 of our conjecture, for $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$, the HJ equation admits one globally stable solution W_{s} and one globally unstable solution W_{u}. Hence, all the trajectories of the manifold $\boldsymbol{p} = \partial_{\boldsymbol{z}} W_{\mathrm{s}}$ converge toward the dominant fixed point, by definition of the global stability. Conversely, all the trajectories of the manifold $\boldsymbol{p} = \partial_{\boldsymbol{z}} W_{\mathrm{u}}$ leave the dominant fixed point, by definition of the global instability. Thus, for any initial condition $\boldsymbol{z}_{\mathrm{i}}$, there exists a relaxing transient trajectory connecting $\boldsymbol{z}_{\mathrm{i}}$ to the $(\boldsymbol{p}_0^\star, \boldsymbol{z}_0^\star)$. Similarly, for any final condition $\boldsymbol{z}_{\mathrm{f}}$, there exists a fluctuating transient trajectory connecting $(\boldsymbol{p}_0^\star, \boldsymbol{z}_0^\star)$ to $\boldsymbol{z}_{\mathrm{f}}$. The boundary conditions are therefore connected through a trajectory with the correct eigenrate and of infinite time duration due to the presence of the fixed point. Note that any trajectory with $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$ that might connect the boundary conditions without passing through $(\boldsymbol{p}_0^\star, \boldsymbol{z}_0^\star)$ must be of finite duration and is then not a candidate for the infinite time limit.
For further use, we compute the value of the opposite of the scaled action

$$-\frac{1}{\mathcal{T}} S[\boldsymbol{p}, \boldsymbol{z}]_0^{\mathcal{T}} = H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t, \boldsymbol{z}_t) - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \boldsymbol{p}_t \cdot \dot{\boldsymbol{z}}_t \mathrm{d}t \tag{4.132}$$

along the trajectory described above. The reduced action is expressed in terms of the global solutions $W_{\mathrm{s,u}}$ according to:

$$\int_0^{\mathcal{T}} \boldsymbol{p}_t \cdot \dot{\boldsymbol{z}}_t \mathrm{d}t \underset{\mathcal{T}\to\infty}{\simeq} \int_{\boldsymbol{z}_\mathrm{i}}^{\boldsymbol{z}_0^\star} \partial_{\boldsymbol{z}} W_\mathrm{s} \cdot \mathrm{d}\boldsymbol{z} + \int_{\boldsymbol{z}_0^\star}^{\boldsymbol{z}_\mathrm{f}} \partial_{\boldsymbol{z}} W_\mathrm{u} \cdot \mathrm{d}\boldsymbol{z} = W_\mathrm{u}(\boldsymbol{z}_\mathrm{f}) - W_\mathrm{u}(\boldsymbol{z}_0^\star) + W_\mathrm{s}(\boldsymbol{z}_0^\star) - W_\mathrm{s}(\boldsymbol{z}_\mathrm{i}). \tag{4.133}$$

Hence, it is not time-extensive, leading to

$$\lim_{\mathcal{T}\to\infty} -\frac{1}{\mathcal{T}} S[\boldsymbol{p},\boldsymbol{z}]_0^{\mathcal{T}} = H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma}). \tag{4.134}$$

We now need to exclude possible trajectories at other values of $H_{\boldsymbol{\gamma}}$. Let us first look at the case $H_{\boldsymbol{\gamma}} > E^\star(\boldsymbol{\gamma})$. From point 1, all the corresponding trajectories tend towards the boundaries of the domain of $\boldsymbol{z}$ when $\mathcal{T} \to \pm\infty$ so that any trajectory connecting $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$ is necessarily of finite duration.

Finally, let us look at the case $H_{\boldsymbol{\gamma}} < E^\star(\boldsymbol{\gamma})$. From point 2 of the conjecture, there are no global solutions of the HJ equation so that some boundary conditions $\boldsymbol{z}_\mathrm{i}$ to $\boldsymbol{z}_\mathrm{f}$ are not connected by trajectories, but some may be. Given $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$, we distinguish three cases:

- There is no trajectory connecting $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$.
- There is a trajectory connecting $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$ but which is neither periodic nor leads to a critical manifold so that the trajectory between $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$ is necessarily of finite duration.
- There is a trajectory connecting $\boldsymbol{z}_\mathrm{i}$ and $\boldsymbol{z}_\mathrm{f}$ that is either periodic or leads to a critical manifold. Since the trajectory dominating the dynamics is the one maximizing the opposite of the action, we need to compare the value of $-\frac{1}{\mathcal{T}} S[\boldsymbol{p},\boldsymbol{z}]_0^{\mathcal{T}}$ along that trajectory with the value $E^\star(\boldsymbol{\gamma})$ found above. On the one hand, the value the Hamiltonian term is smaller than $E^\star(\boldsymbol{\gamma})$. On the other hand, we have conjectured in point 2 that the reduced action is nonnegative so that it reduces the value of $-\frac{1}{\mathcal{T}} S[\boldsymbol{p},\boldsymbol{z}]_0^{\mathcal{T}}$ even more. This implies that such a trajectory will be exponentially less likely than the one found for $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$.

In all cases, the dominant trajectory in the long-time limit is the one corresponding to $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$. It follows from Eq. (4.57) and point 4 that for any initial and final boundary conditions,

$$\tilde{G}_{\mathcal{T}}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i}) \underset{\mathcal{T}\to\infty}{\asymp} \mathrm{e}^{\mathcal{N}\mathcal{T}H_0^\star(\boldsymbol{\gamma})}. \tag{4.135}$$

Using Eqs. (4.51, 4.66, 4.135), it follows that the SCGF $\bar{\Gamma} = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}\mathcal{N}} \ln \mathbb{E}_{\boldsymbol{z}_\mathrm{i}} \left[\mathrm{e}^{\mathcal{N}\mathcal{T}\boldsymbol{\gamma}\cdot\bar{\boldsymbol{A}}_t}\right]$ is given by

$$\bar{\Gamma}(\boldsymbol{\gamma}) = H_0^\star(\boldsymbol{\gamma}). \tag{4.136}$$

Hence, the Hamiltonian converges to the SCGF $\bar{\Gamma}$ in the long-time limit. For the model of the Brownian Donkey, this corresponds to the red trajectory in Figure 4.7. Eq. (4.136) is the nonlinear analogue of the result relating the SCGF of a Markov process to the largest eigenvalue of the biased generator in the linear operator formalism.

In the case where there is more than one critical manifold at $H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t) = E^\star(\boldsymbol{\gamma})$, we conjecture that it is always possible to connect one critical manifold to another via

trajectories of the stable and unstable manifolds. Combined with point 3, it follows that there exists a relaxing transient trajectory of the globally stable manifold connecting $\boldsymbol{z}_\mathrm{i}$ to a first critical manifold, a fluctuating transient trajectory of the globally unstable manifold connecting a second (or eventually the same) critical manifold to $\boldsymbol{z}_\mathrm{f}$, and in between there exists trajectories connecting the first and second critical manifolds (switching phase). Along any of those critical manifolds, the reduced action is non-extensive in time since $S_\mathrm{r} = \int_0^\mathcal{T} \partial_{\boldsymbol{z}} W_\mathrm{s} \cdot \mathrm{d}\boldsymbol{z} = W_\mathrm{s}(\boldsymbol{z}_\mathrm{f}, \boldsymbol{\gamma}) - W_\mathrm{s}(\boldsymbol{z}_\mathrm{i}, \boldsymbol{\gamma})$. Subsequently, the rest of the proof done for the case of a single fixed point holds, concluding our reasoning.

Some remarks on the non-biased case

In the non-biased case ($\boldsymbol{\gamma} = 0$), the SCGF vanishes ($\bar{\Gamma} = 0$) by definition, implying $H_0^\star(\boldsymbol{\gamma} = 0) = 0$. It follows that $\boldsymbol{p} = 0$ is always a solution of the HJ equation in the non-biased case by virtue of Eq. (4.107) and Eq. (4.26). Moreover, for the transition probability $P_\mathcal{T}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i})$ to be normalizable, this solution has to correspond to the globally stable solution $W_\mathrm{s} = 0$. Indeed, the dependence of $P_\mathcal{T}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i})$ comes from the globally unstable solution W_u in the reduced action term as $\boldsymbol{z}_\mathrm{f}$ is reached via the unstable manifold:

$$P_\mathcal{T}(\boldsymbol{z}_\mathrm{f} \mid \boldsymbol{z}_\mathrm{i}) \underset{\mathcal{T}\to\infty}{\asymp} \mathrm{e}^{W_\mathrm{u}(\boldsymbol{z}_\mathrm{f}) - W_\mathrm{s}(\boldsymbol{z}_\mathrm{i})}. \tag{4.137}$$

If W_u was zero, the transition probability would be a constant of $\boldsymbol{z}_\mathrm{f}$, hence non-normalized. So $\boldsymbol{p} = 0$ needs to correspond to the globally stable manifold.

4.2.4 Rectification on Hamiltonians

Relying on our conjecture for the nonlinear counterpart of the Perron-Frobenius and our description of the long-time behavior of statistical Hamiltonians, we now propose a procedure turning the biased dynamics into a rectified dynamics. This rectification translates in the Hamiltonian formalism the generalized Doob transform (3.23, 3.122) of the linear operator formalism and allows us to build the nonlinear counterpart of the driven process whose dynamics is equivalent to the microcanonical dynamics in the long-time limit.

Structure of the rectification

Let $\mathbf{H}(\boldsymbol{p}, \boldsymbol{z})$ be a Hamiltonian which does not necessarily satisfy Eq. (4.107) and $\boldsymbol{\varpi}$ a function of the same nature as $\boldsymbol{p}$. We define the *rectification* $\mathbf{H}^{\boldsymbol{\varpi}}(\boldsymbol{p}, \boldsymbol{z})$ of $\mathbf{H}(\boldsymbol{p}, \boldsymbol{z})$ associated with $\boldsymbol{\varpi}$ by

$$\mathbf{H}^{\boldsymbol{\varpi}}(\boldsymbol{p}, \boldsymbol{z}) \equiv \mathbf{H}(\boldsymbol{p} + \boldsymbol{\varpi}, \boldsymbol{z}) - \mathbf{H}(\boldsymbol{\varpi}, \boldsymbol{z}). \tag{4.138}$$

By definition, $\mathbf{H}^{\boldsymbol{\varpi}}(\boldsymbol{p} = 0, \boldsymbol{z}) = 0$, as wanted. This definition is purely mathematical and is the analogue of the Doob transform in the linear operator formalism. Yet, in order to build physically relevant Hamiltonians, one needs to choose $\boldsymbol{\varpi}$ in a way that keeps the Hamiltonian structure, i.e. whose dynamics satisfies Hamilton's equations. To guarantee the latter condition, $\mathbf{H}^{\boldsymbol{\varpi}}(\boldsymbol{p}, \boldsymbol{z})$ has to derive from a canonical transformation $(\boldsymbol{Z}, \boldsymbol{P}) \to (\boldsymbol{z}, \boldsymbol{p})$ (see Appendix 4.B.3). In addition to this constraint, we want to build a rectified Hamiltonian whose long-time limit dynamics is equivalent to the microcanonical dynamics, imposing a specific choice of $\boldsymbol{\varpi}$ as discussed in the following.

Rectified Hamiltonian

For the purpose of this section, we denote by $\boldsymbol{Z}$ and $\boldsymbol{P}$ the variables of the biased Hamiltonian $H_{\boldsymbol{\gamma}}$. We aim to derive the rectified Hamiltonian from the rectification transform of the biased Hamiltonian and while keeping the Hamiltonian structure, i.e. whose dynamics satisfies Hamilton's equations. To guarantee the latter condition, we define the rectified Hamiltonian from a canonical transformation $(\boldsymbol{Z}, \boldsymbol{P}) \to (\boldsymbol{z}, \boldsymbol{p})$ associated with the generating function of the second type

$$F_2(\boldsymbol{p}, \boldsymbol{Z}, t) = \boldsymbol{Z} \cdot \boldsymbol{p} + W_{\mathrm{s}}(\boldsymbol{Z}) - \bar{\Gamma} t, \tag{4.139}$$

see Appendix 4.B.3 for a review. The new variables $\boldsymbol{z}$, $\boldsymbol{p}$ are obtained from the rules of transformation associated with the generating function F_2:

$$\boldsymbol{z} = \partial_{\boldsymbol{p}} F_2 = \boldsymbol{Z} \tag{4.140}$$

$$\boldsymbol{P} = \partial_{\boldsymbol{Z}} F_2 = \boldsymbol{p} + \partial_{\boldsymbol{Z}} W_{\mathrm{s}}, \tag{4.141}$$

with $\partial_{\boldsymbol{P}}$ the gradient operator with respect to $\boldsymbol{P}$, and the new Hamiltonian follows from

$$H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(\boldsymbol{P}, \boldsymbol{Z}) + \partial_t F_2 = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) - \bar{\Gamma}, \tag{4.142}$$

leading to

$$H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) - H_{\boldsymbol{\gamma}}(\partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}), \tag{4.143}$$

where we used the fact that W_{s} is solution of the HJ equation (4.62) with $E = \bar{\Gamma}$. The rectified Hamiltonian derives indeed from the rectification transform (4.138) of the biased Hamiltonian with $\boldsymbol{\varpi} = \partial_{\boldsymbol{z}} W_{\mathrm{s}}$. The detailed rectified Hamiltonian follows using (4.121):

$$\mathscr{H}^{\mathrm{r}}(\boldsymbol{f}, \boldsymbol{z}; \boldsymbol{\gamma}) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} + \mathcal{D}^\dagger \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) - \mathscr{H}_{\boldsymbol{\gamma}}(\mathcal{D}^\dagger \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}). \tag{4.144}$$

Assuming $\mathscr{H}^{\mathrm{r}}$ is everywhere differentiable in $\boldsymbol{f}$, the Legendre-Fenchel transform is involutive and the rectified detailed Lagrangian reads

$$\mathscr{L}^{\mathrm{r}}(\boldsymbol{\lambda}, \boldsymbol{z}; \boldsymbol{\gamma}) = \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z}) - \boldsymbol{\lambda} \cdot \mathcal{D}^\dagger \partial_{\boldsymbol{z}} W_{\mathrm{s}} + \bar{\Gamma}(\boldsymbol{\gamma}). \tag{4.145}$$

The rectified standard Lagrangian follows immediately from

$$L^{\mathrm{r}}(\dot{\boldsymbol{z}}, \boldsymbol{z}; \boldsymbol{\gamma}) = \inf_{\boldsymbol{\lambda} | \dot{\boldsymbol{z}} = \mathcal{D}\boldsymbol{\lambda}} \mathscr{L}^{\mathrm{r}}(\boldsymbol{\lambda}, \boldsymbol{z}; \boldsymbol{\gamma}). \tag{4.146}$$

Note that these definitions are consistent with the Lagrangians and Hamiltonians derived from the driven generator for the case of population processes (Section 4.1.3).

Properties of the rectified Hamiltonian

Let us show that the rectified Hamiltonian is proper, i.e. it has the properties of a statistical Hamiltonian on the one hand, and the properties of a non-biased Hamiltonian on the other hand (see Section 4.2.2).

We first show that the rectified Hamiltonian $H^{\mathrm{r}}(\boldsymbol{\gamma})$ is a statistical Hamiltonian given that the biased Hamiltonian $H_{\boldsymbol{\gamma}}$ is a statistical Hamiltonian. First, the strict convexity in

$\boldsymbol{p}$ of H^{r} is inherited from the strict convexity in $\boldsymbol{p}$ of the biased Hamiltonian $H_{\boldsymbol{\gamma}}$. Indeed, $H_{\boldsymbol{\gamma}}$ is strictly convex if and only if for all $\boldsymbol{p}, \boldsymbol{p}'$, $\boldsymbol{p} \neq \boldsymbol{p}'$,

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}) > H_{\boldsymbol{\gamma}}(\boldsymbol{p}', \boldsymbol{z}) + \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{z}) \cdot (\boldsymbol{p} - \boldsymbol{p}'). \tag{4.147}$$

It follows from Eqs. (4.143) and (4.147):

$$H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}; \boldsymbol{\gamma}) > \left[H_{\boldsymbol{\gamma}}(\boldsymbol{p}' + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) - \bar{\Gamma}\right] + \partial_{\boldsymbol{p}} \left[H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) - \bar{\Gamma}\right] \cdot (\boldsymbol{p} - \boldsymbol{p}') \tag{4.148}$$

$$> H^{\mathrm{r}}(\boldsymbol{p}', \boldsymbol{z}; \boldsymbol{\gamma}) + \partial_{\boldsymbol{p}} H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}; \boldsymbol{\gamma}) \cdot (\boldsymbol{p} - \boldsymbol{p}'), \tag{4.149}$$

proving the strict convexity of H^{r}. Note that the coercivity of H^{r} follows immediately from the coercivity of $H_{\boldsymbol{\gamma}}$. It follows that H^{r} admits for each $\boldsymbol{z}$ a unique minimum reached for $\boldsymbol{p} = \boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}) = \boldsymbol{p}_{\mathrm{min}}(z) - \partial_{\boldsymbol{z}} W_{\mathrm{s}}(\boldsymbol{z})$, with $\boldsymbol{p}_{\mathrm{min}}$ the minimizer of $H_{\boldsymbol{\gamma}}$:

$$\partial_{\boldsymbol{p}} H^{\mathrm{r}}(\boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}; \boldsymbol{\gamma}) = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}) + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}) = 0, \tag{4.150}$$

$$\partial^2_{\boldsymbol{p}} H^{\mathrm{r}}(\boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}; \boldsymbol{\gamma}) = \partial^2_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}) + \partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}) = \partial^2_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}) > 0, \tag{4.151}$$

where we used Eq. (4.126). The minimal value of H^{r} is then related to the minimal value of $H_{\boldsymbol{\gamma}}$ by:

$$H^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}) \equiv H^{\mathrm{r}}(\boldsymbol{p}^{\mathrm{r}}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}) = H_{\boldsymbol{\gamma}}(\boldsymbol{p}_{\mathrm{min}}(\boldsymbol{z}), \boldsymbol{z}) - \bar{\Gamma} = H_{\mathrm{min}}(\boldsymbol{z}) - \bar{\Gamma}. \tag{4.152}$$

Consequently, the extrema of $H^{\mathrm{r}}_{\mathrm{min}}$ are given by the extrema of $H_{\mathrm{min}}(\boldsymbol{z})$ shifted by the constant $\bar{\Gamma}$, both reached at the same positions $\boldsymbol{z}^\star_\ell$. In particular, it implies the non-degeneracy of the absolute maximum of H^{r}. Finally, we remind that the rectified dynamics satisfies the solutions of Hamilton's equations since the rectified Hamiltonian derives from a canonical transformation.

We now show that the rectified Hamiltonian has the properties of a *non-biased* statistical Hamiltonian. First, we have by construction that $H^{\mathrm{r}}(\boldsymbol{p} = 0, \boldsymbol{z}) = 0$, as required for a non-biased Hamiltonian. Second, the absolute maximum of $H^{\mathrm{r}}_{\mathrm{min}}$ is zero by virtue of Eqs. (4.136) and (4.152). It remains to show that the solution $\boldsymbol{p} = 0$ is the globally stable solution of the HJ equation for the eigenrate $E = 0$. From Eq. (4.124) and (4.143), we have:

$$\dot{\boldsymbol{z}} = \left.\frac{\partial H^{\mathrm{r}}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=0, \boldsymbol{z}} = \left.\frac{\partial H_{\boldsymbol{\gamma}}}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=\partial_{\boldsymbol{z}} W_{\mathrm{s}}, \boldsymbol{z}}, \tag{4.153}$$

meaning that the reduced dynamics at $\boldsymbol{p} = 0$ of the rectified Hamiltonian corresponds to the reduced dynamics of the biased Hamiltonian along its globally stable manifold. Hence, all the trajectories for the rectified dynamics at $\boldsymbol{p} = 0$ converge to a compact set, showing the global stability of the manifold $\boldsymbol{p} = 0$ for the rectified dynamics. This explain why it is necessary to perform the rectification with respect to W_{s} rather than W_{u} or any other characteristic function. Notice that the corresponding globally unstable solution is given by $\boldsymbol{p} = \partial_{\boldsymbol{z}} (W_{\mathrm{u}} - W_{\mathrm{s}})$.

Equivalence of microcanonical, rectified and canonical processes

In this section, we show that the rectified path probability $\mathbb{P}^{\mathrm{r}}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$ of the trajectory $[\boldsymbol{z}]$ of duration $\mathcal{T}$ given the initial state $\boldsymbol{z}_{\mathrm{i}}$ is asymptotically equivalent in the long-time limit to the canonical path probability

$$\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma}, \mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}] \equiv \frac{\mathrm{e}^{\mathcal{T} \boldsymbol{\gamma} \cdot \boldsymbol{A}_{\mathcal{T}}} \mathbb{P}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]}{\mathbb{E}_{\boldsymbol{z}_{\mathrm{i}}}\left[\mathrm{e}^{\mathcal{T} \boldsymbol{\gamma} \cdot \boldsymbol{A}_{\mathcal{T}}}\right]}, \tag{4.154}$$

with $\boldsymbol{A}_\mathcal{T} = \mathcal{N}\bar{\boldsymbol{A}}_\mathcal{T}$. We know by construction that the rectified and biased Lagrangians satisfy

$$\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}] \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\int_0^\mathcal{T} \mathrm{d}\tau \mathscr{L}^{\mathrm{r}}(\boldsymbol{\lambda}_\tau, \boldsymbol{z}_t)}, \tag{4.155}$$

$$\mathbb{P}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_\mathcal{T}} \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\int_0^\mathcal{T} \mathrm{d}\tau \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}_\tau, \boldsymbol{z}_t)}, \tag{4.156}$$

with $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ related by Eq. (4.101). Combined with Eqs. (4.65, 4.145), it follows

$$\frac{\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]} \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\int_0^\mathcal{T} \mathrm{d}\tau \boldsymbol{\lambda}_\tau\cdot\mathfrak{D}^\dagger\partial_{\boldsymbol{z}} W_{\mathrm{s}}(\boldsymbol{z}_\tau)} \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\int_0^\mathcal{T} \mathrm{d}\tau \dot{\boldsymbol{z}}_\tau\cdot\partial_{\boldsymbol{z}} W_{\mathrm{s}}(\boldsymbol{z}_t)}, \tag{4.157}$$

where we used Eq. (4.101) in the last equality. It follows

$$\frac{\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]} \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}[W_{\mathrm{s}}(\boldsymbol{z}_\mathcal{T}) - W_{\mathrm{s}}(\boldsymbol{z}_{\mathrm{i}})]}, \tag{4.158}$$

leading to:

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \ln \frac{\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]}{\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]} = 0. \tag{4.159}$$

Hence, the rectified path probability and the canonical path probability are logarithmically equivalent:

$$\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}] \underset{\mathcal{T}\to\infty}{\asymp} \mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]. \tag{4.160}$$

Finally, the equivalence between the rectified path probability $\mathbb{P}^{\mathrm{r}}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$ and the microcanonical path probability $\mathbb{P}^{\mathrm{micro}}_{\boldsymbol{a},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}] = \mathbb{P}_\mathcal{T}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}, \boldsymbol{A}_\mathcal{T} = \boldsymbol{a}]$ follows from the equivalence between the canonical and microcanonical path probabilities for $\boldsymbol{\gamma} = \nabla I(\boldsymbol{a})$, with I the LDF (in time) of $\boldsymbol{A}_\mathcal{T}$ [195].

4.2.5 Fluctuation relations

We say that the Hamiltonian $\mathscr{H}$ satisfies a fluctuation relation if there exists a quantity $\boldsymbol{F}$, called *affinity*, such that

$$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z}) = \mathscr{H}(\boldsymbol{F} + \theta\boldsymbol{f}, \boldsymbol{z}), \tag{4.161}$$

where the current-reversal operator θ is an involutive linear operator acting on $\boldsymbol{f}$, i.e. $\theta^2\boldsymbol{f} = \boldsymbol{f}$. For instance, for overdamped diffusion processes, θ is equal to minus the identity: $\theta\boldsymbol{f} = -\boldsymbol{f}$, while for jump processes θ is the operator that exchanges initial and final states of a jump: $\theta f_{nm} = f_{mn}$. The affinity $\boldsymbol{F}$ is such that $\theta\boldsymbol{F} = -\boldsymbol{F}$ and may depend on $\boldsymbol{z}$. For example, we have $F_{nm} = \ln\left(\tilde{k}_{nm}\mu_m/\tilde{k}_{mn}\mu_n\right)$ for independent many-body jump processes (Appendix 4.C.1) and $F(x) = -\frac{2J^\rho(x)}{D(x)\rho(x)}$ for independent many-body diffusion processes (Appendix 4.C.2).

In this section, we investigate the inheritance of the fluctuation relation by the biased and rectified Hamiltonians given the fluctuation relation (4.161) for the original Hamiltonian $\mathscr{H}$. From the definition of the biased Hamiltonian (4.119) and using Eq. (4.161), we have

$$\begin{aligned}
\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{z}) &= \mathscr{H}(\boldsymbol{f} + \boldsymbol{\gamma}_1, \boldsymbol{z}) + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z} && (4.162)\\
&= \mathscr{H}(\boldsymbol{F} + \theta(\boldsymbol{f} + \boldsymbol{\gamma}_1), \boldsymbol{z}) + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z} && (4.163)\\
&= \mathscr{H}(\boldsymbol{F} + (\theta - 1)\boldsymbol{\gamma}_1 + \theta\boldsymbol{f} + \boldsymbol{\gamma}_1, \boldsymbol{z}) + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z}, && (4.164)
\end{aligned}$$

leading to the fluctuation relation for the biased Hamiltonian

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f},\boldsymbol{z}) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{F}_{\boldsymbol{\gamma}} + \theta\boldsymbol{f},\boldsymbol{z}), \tag{4.165}$$

where we introduced the *biased affinity*

$$\boldsymbol{F}_{\boldsymbol{\gamma}} \equiv \boldsymbol{F} + (\theta-1)\boldsymbol{\gamma}_1 \tag{4.166}$$

that satisfies $\theta\boldsymbol{F}_{\boldsymbol{\gamma}} = -\boldsymbol{F}_{\boldsymbol{\gamma}}$. Similarly, from the definition of the rectified Hamiltonian (4.144) and the fluctuation relation for $\mathscr{H}_{\boldsymbol{\gamma}}$ (4.165), the rectified Hamiltonian satisfies

$$\begin{aligned}
\mathscr{H}^{\mathrm{r}}(\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}) &= \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} + \mathcal{D}^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}},\boldsymbol{z}) - \mathscr{H}_{\boldsymbol{\gamma}}(\mathcal{D}^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}},\boldsymbol{z}) && (4.167)\\
&= \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{F}_{\boldsymbol{\gamma}} + \theta(\boldsymbol{f} + \mathcal{D}^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}}),\boldsymbol{z}) - \mathscr{H}_{\boldsymbol{\gamma}}(\mathcal{D}^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}},\boldsymbol{z}) && (4.168)\\
&= \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{F}_{\boldsymbol{\gamma}} + (\theta-1)\mathcal{D}^\dagger\partial_z W_{\mathrm{s}} + \theta\boldsymbol{f} + \mathcal{D}^\dagger\partial_z W_{\mathrm{s}},z) - \mathscr{H}_{\boldsymbol{\gamma}}(\mathcal{D}^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}},\boldsymbol{z}) && (4.169)\\
&= \mathscr{H}^{\mathrm{r}}(\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} + \theta\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}), && (4.170)
\end{aligned}$$

leading to the fluctuation relation

$$\mathscr{H}^{\mathrm{r}}(\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}) = \mathscr{H}^{\mathrm{r}}(\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} + \theta\boldsymbol{f},\boldsymbol{z};\boldsymbol{\gamma}), \tag{4.171}$$

where we introduced the *rectified affinity*

$$\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} \equiv \boldsymbol{F} + (\theta-1)(\boldsymbol{\gamma}_1 + D^\dagger\partial_{\boldsymbol{z}}W_{\mathrm{s}}) \tag{4.172}$$

that satisfies $\theta\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} = -\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}}$. Hence, the non-biased, biased and rectified Hamiltonians have a similar fluctuation symmetry with different affinities given respectively by $\boldsymbol{F}$, $\boldsymbol{F}_{\gamma}$ and $\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}}$.

4.3 Applications

In Section 4.1, we considered general population processes with unspecified nonlinear rates. In the following section, we apply our general theory to specific nonlinear models of population processes: the Brownian Donkey and a nonlinear chemical reaction network. In Appendix 4.C, we deal with the linear case of independent many-body Markov processes, both for jump and diffusion processes.

4.3.1 Brownian Donkey

We apply the formalism stated in Section 4.1 to the model of the Brownian Donkey introduced by B. Cleuren and C. Van den Broeck [262]. This model is made of $\mathcal{N}$ interacting unicyclic machines. Each machine consists of a two-level system jumping between a lower state of energy 0 and a higher state of energy $\mathcal{E} > 0$ via two heat reservoirs labeled by $\nu = 1, 2$ and of inverse temperature β^ν. Two machines interact via an interaction energy $\frac{V}{\mathcal{N}}$ only when they are in different states. We denote by N the number of machines in the high energy state. When the global system is in state N, it has a total energy

$$U_N = N\mathcal{E} + N(\mathcal{N}-N)\frac{V}{\mathcal{N}}. \tag{4.173}$$

We assume that during an infinitesimal time δt, two machines cannot jump at the same time and only one transition occurs so that jumping from an initial state N to a final state N' imposes that

$$N' = N \pm 1. \tag{4.174}$$

The probability p_N to be in state N satisfies the master equation $\dot{p}_N = \sum_{\epsilon=\pm1,\nu} k^\nu_{N,\epsilon+N} p_{N+\epsilon}$, where the transition rate $k^\nu_{\epsilon+N,N}$ from state N to state $N+\epsilon$, with $\epsilon = \pm 1$, through reservoir ν is given by

$$k^\nu_{N+\epsilon,N} = \mathcal{N}\left(\frac{1+\epsilon}{2} - \epsilon\frac{N}{\mathcal{N}}\right) \mathrm{e}^{-\frac{\beta^\nu}{2}(E_a + U_{N+\epsilon} - U_N + \epsilon(-1)^\nu F)}, \tag{4.175}$$

with E_a an activation energy and F a non-conservative force. The Brownian Donkey is a particular model from the class of population processes introduced in Sec. 4.1 where we have the following connections:

$$\boldsymbol{N} \to N, \tag{4.176}$$
$$\alpha \to (\epsilon, \nu), \tag{4.177}$$
$$k_{\alpha,\boldsymbol{N}} \to k^\nu_{N+\epsilon,N}, \tag{4.178}$$
$$\boldsymbol{\mathfrak{D}}_\alpha \to \mathfrak{D}^\nu_\epsilon \equiv \epsilon, \tag{4.179}$$

for any $\nu = 1, 2$ and any $\epsilon = \pm 1$. The empirical density $\boldsymbol{z}$ is now the density of machines in the high energy state $z \equiv \frac{N}{\mathcal{N}}$ and the current $\boldsymbol{\lambda}$ of component λ^ν_ϵ is now the current density of machines passing from the high (resp. low) energy level to the low (resp. high) energy level when $\epsilon = +1$ (resp. $\epsilon = -1$) via channel ν. Both are related by $\dot{z} = \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda} = \sum_{\epsilon,\nu} \epsilon\lambda^\nu_\epsilon$, where $\boldsymbol{\mathfrak{D}}$ is here the line vector operator of components $\mathfrak{D}^\nu_\epsilon = \epsilon$, $\forall\nu$. The transition rate $\boldsymbol{k}(z)$ in the continuous limit defined in Eq. (4.8) reads

$$k^\nu_\epsilon(z) = \left(\frac{1+\epsilon}{2} - \epsilon z\right) \mathrm{e}^{-\frac{\beta^\nu}{2}(E_a + \epsilon\mathcal{E} + \epsilon V(1-2z) + \epsilon(-1)^\nu F)}. \tag{4.180}$$

From Eqs. (4.42–4.43), the detailed biased Lagrangian and Hamiltonian read:

$$\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, z) = \sum_{\epsilon,\nu}\left[\lambda^\nu_\epsilon \ln\left(\frac{\lambda^\nu_\epsilon}{k^\nu_\epsilon(z)}\right) - \lambda^\nu_\epsilon + k^\nu_\epsilon(z)\right] - \boldsymbol{\gamma}_1 \cdot \boldsymbol{\lambda} - \gamma_2 z, \tag{4.181}$$

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, z) = \sum_{\epsilon,\nu} k^\nu_\epsilon(z)\left[\mathrm{e}^{f^\nu_\epsilon + \gamma^\nu_{1,\epsilon}} - 1\right] + \gamma_2 z, \tag{4.182}$$

with $\boldsymbol{\gamma}_1 \cdot \boldsymbol{\lambda} = \sum_{\epsilon,\nu} \gamma^\nu_{1,\epsilon}\lambda^\nu_\epsilon$. We recall that the non-biased Lagrangian and Hamiltonian are obtained from the biased ones by taking $\boldsymbol{\gamma} = 0$. For this model, we can compute explicitly the standard Lagrangian using Eq. (4.19). Using the method of Lagrange multipliers (Appendix 3.D) to ensure the constraint $\dot{z} = \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}$, we obtain after optimizing the functional

$$\tilde{\mathscr{L}}_{\boldsymbol{\gamma}}(\dot{z}, z) \equiv \mathscr{L}_{\boldsymbol{\gamma}}(\dot{z}, z) - \xi[\dot{z} - \boldsymbol{\mathfrak{D}}\boldsymbol{\lambda}], \tag{4.183}$$

with ξ a Lagrange multiplier, that the standard biased Lagrangian is given by

$$L_{\boldsymbol{\gamma}}(\dot{z}, z) = -\sqrt{\dot{z}^2 + \varphi(z, \boldsymbol{\gamma}_1)} + \sum_{\epsilon,\nu} k^\nu_\epsilon(z) - \dot{z}\ln\left[\frac{-\dot{z} + \sqrt{\dot{z}^2 + \varphi(z, \boldsymbol{\gamma}_1)}}{2\sum_\nu k^\nu_- \mathrm{e}^{\gamma^\nu_{1,-}}}\right] - \gamma_2 z, \tag{4.184}$$

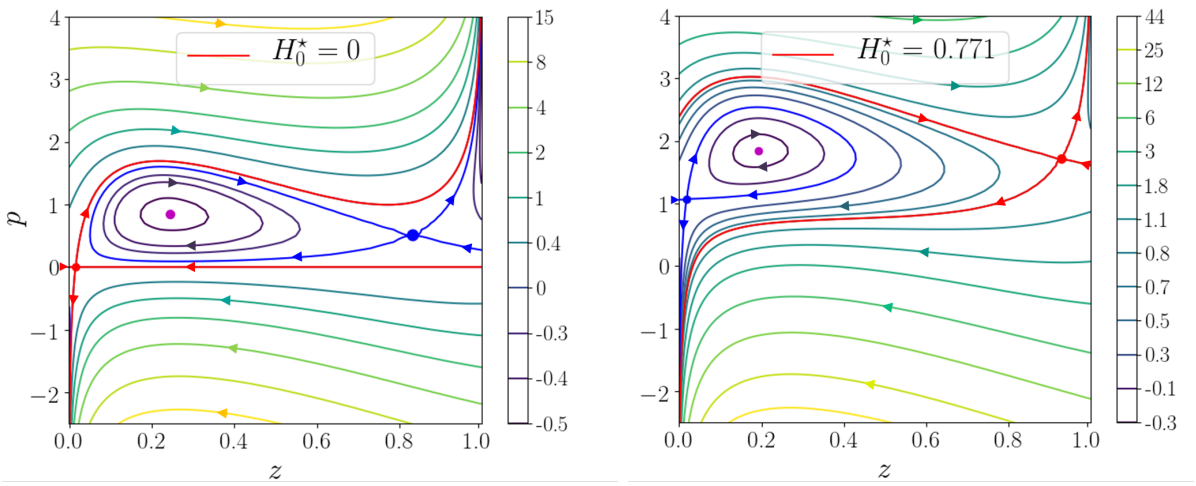

Figure 4.8 – (Left) Trajectories of the original Hamiltonian ($\boldsymbol{\gamma} = 0$). (Right) Trajectories of the biased Hamiltonian.
In both figures, there are three fixed points represented by the three colored points (extrema of $H_{\min}(z)$ in Fig. 4.6). The red trajectory is associated with the max-min value of the Hamiltonian. As expected, this value is equal to 0 in the non-biased case, while it is different from 0 in the biased case.
Both figures are obtained for $\mathcal{E} = 0.8$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$. In the right figure, we have $\gamma^\nu_{1,\epsilon} = \epsilon$ ($\epsilon = \pm 1$ and $\nu = 1, 2$) and $\gamma_2 = 1$.

with $\varphi(z, \boldsymbol{\gamma}_1) \equiv 4 \prod_\epsilon \sum_\nu k^\nu_\epsilon(z) \mathrm{e}^{\gamma^\nu_{1,\epsilon}}$, recovering the result of Ref. [119]. Taking the LF transform of Eq. (4.184), we obtain the standard biased Hamiltonian:

$$H_{\boldsymbol{\gamma}}(p, z) = \sum_{\epsilon,\nu} k^\nu_\epsilon(z) \left[\mathrm{e}^{\epsilon p + \gamma^\nu_{1,\epsilon}} - 1\right] + \gamma_2 z. \tag{4.185}$$

Note that that the detailed Hamiltonian (4.182) and the standard Hamiltonian (4.185) are indeed related by $H_{\boldsymbol{\gamma}}(p, z) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} = \mathfrak{D}^\dagger p, z)$. For this model, the solutions of the implicit equation

$$H_{\boldsymbol{\gamma}}(p(z, \boldsymbol{\gamma}), z) = \bar{\Gamma} \tag{4.186}$$

can be explicitly computed provided that $\left(\bar{\Gamma} - \gamma_2 z + \sum_{\epsilon,\nu} k^\nu_\epsilon(z)\right)^2 \geq \varphi(z, \boldsymbol{\gamma}_1)$. In this case, Eq. (4.186) admits two solutions $p_\pm \equiv p_\pm(z, \boldsymbol{\gamma})$ with

$$p_\pm(z, \boldsymbol{\gamma}) \equiv \ln \left[\frac{\bar{\Gamma} - \gamma_2 z + \sum_{\epsilon,\nu} k^\nu_\epsilon(z) \pm \sqrt{\left(\bar{\Gamma} - \gamma_2 z + \sum_{\epsilon,\nu} k^\nu_\epsilon(z)\right)^2 - \varphi(z, \boldsymbol{\gamma}_1)}}{2 \sum_\nu k^\nu_+(z) \mathrm{e}^{\gamma^\nu_{1,+}}} \right]. \tag{4.187}$$

Note that the solution p_- exists only if $\bar{\Gamma} - \gamma_2 z + \sum_{\epsilon,\nu} k^\nu_\epsilon(z) \geq 0$, defining a domain of validity for γ_2. The SCGF $\bar{\Gamma}$ coincides with the value $H_0^\star = \max_z \min_p H_{\boldsymbol{\gamma}}(p, z)$ of the biased Hamiltonian at the dominant fixed point $(p_0^\star, z_0^\star)$, represented by a red point in Fig. 4.8.

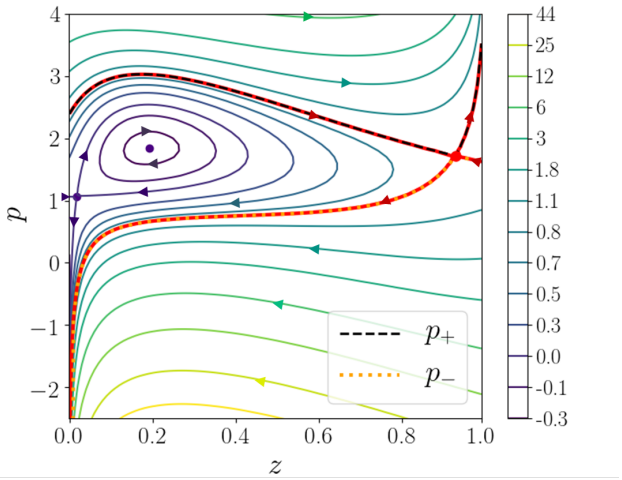

Figure 4.9 – Trajectories of the biased Hamiltonian. The dashed black line corresponds to the solution p_+ and the gold dotted line corresponds to the solution p_-.
The figure is obtained for $\mathcal{E} = 0.8$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$, $\gamma^{\nu}_{1,\epsilon} = \epsilon$ ($\epsilon = \pm 1$ and $\nu = 1, 2$) and $\gamma_2 = 1$.

Stability of the solutions

As illustrated in Fig. 4.9, the solutions $p_\pm$ of Eq. (4.187) have the following stability properties:

- For $z \in [0, z_0^\star]$:

$$\begin{cases} \text{The branch } p_+ \text{ is stable,} \\ \text{The branch } p_- \text{ is unstable} \end{cases} \tag{4.188}$$

- For $z \in [z_0^\star, 1]$:

$$\begin{cases} \text{The branch } p_+ \text{ is unstable,} \\ \text{The branch } p_- \text{ is stable.} \end{cases} \tag{4.189}$$

In Appendix 4.D, we illustrate numerically how the stability of these solutions changes with the biasing parameter $\boldsymbol{\gamma}$ and when the system undergoes a phase transition. It follows from Eqs. (4.188–4.189) that the HJ equation at $E = \bar{\Gamma}$ admits one globally stable solution $p_{\rm s} \equiv \partial_z W_{\rm s}$ and one globally unstable solution $p_{\rm u} \equiv \partial_z W_{\rm u}$ with

$$\partial_z W_{\rm s} \equiv \begin{cases} p_+ & \text{if } z \in [0, z_0^\star] \\ p_- & \text{if } z \in [z_0^\star, 1] \end{cases} , \tag{4.190}$$

$$\partial_z W_{\rm u} \equiv \begin{cases} p_- & \text{if } z \in [0, z_0^\star] \\ p_+ & \text{if } z \in [z_0^\star, 1] \end{cases} . \tag{4.191}$$

The globally stable (resp. unstable) manifold $p = p_{\rm s}$ (resp. $p = p_{\rm u}$) contains two relaxing (resp. fluctuating) transient trajectories that converge to (resp. exit from) the red fixed point as illustrated in Fig. 4.9. In the non-biased case, the globally stable manifold is $p_{\rm s}(z, \boldsymbol{\gamma} = 0) = 0$, as required (see Fig. 4.8).

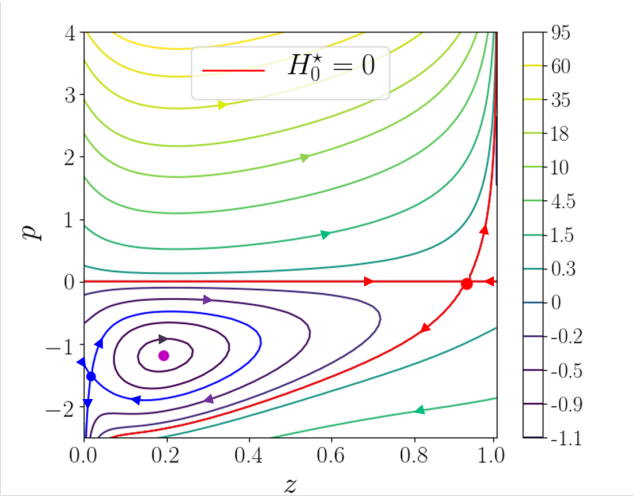

Figure 4.10 – Trajectories of rectified Hamiltonian. There are three fixed points represented by the three colored dot. The red trajectory is associated with the max-min value of the rectified Hamiltonian. As expected, this value is zero.
The figure is obtained for $\mathcal{E} = 0.8$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$, $\gamma^\nu_{1,\epsilon} = \epsilon$ ($\epsilon = \pm 1$ and $\nu = 1, 2$) and $\gamma_2 = 1$.

Rectification

The rectified Lagrangian is given by

$$\mathscr{L}^{\mathrm{r}}(\boldsymbol{\lambda}, z; \boldsymbol{\gamma}) = \mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, z) - \boldsymbol{\lambda} \cdot \mathfrak{D}^\dagger p_{\mathrm{s}} + \bar{\Gamma}, \tag{4.192}$$

with $\left(\mathfrak{D}^\dagger p_{\mathrm{s}}\right)^\nu_\epsilon \equiv \epsilon p_{\mathrm{s}}$, and the rectified Hamiltonians by

$$\mathscr{H}^{\mathrm{r}}(\boldsymbol{f}, z; \boldsymbol{\gamma}) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f} + \mathfrak{D}^\dagger p_{\mathrm{s}}, z) - \bar{\Gamma}, \tag{4.193}$$

$$H^{\mathrm{r}}(p, z; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(p + p_{\mathrm{s}}, z) - \bar{\Gamma}. \tag{4.194}$$

For this model, we can compute explicitly the standard rectified Lagrangian by using the relation (4.105) on Eq. (4.192), and it yields

$$L^{\mathrm{r}}(\dot{z}, z; \boldsymbol{\gamma}) = -\sqrt{\dot{z}^2 + \varphi(z, \boldsymbol{\gamma}_1 + \mathfrak{D}^\dagger p_{\mathrm{s}})} + \sum_{\epsilon,\nu} k^\nu_\epsilon - \dot{z} \ln \left[\frac{-\dot{z} + \sqrt{\dot{z}^2 + \varphi(z, \boldsymbol{\gamma}_1 + \mathfrak{D}^\dagger p_{\mathrm{s}})}}{2 \sum_\nu k^\nu_- \mathrm{e}^{\gamma^\nu_{1,-} - p_{\mathrm{s}}}} \right] - \gamma_2 z + \bar{\Gamma}, \tag{4.195}$$

which can be written as

$$L^{\mathrm{r}}(\dot{z}, z; \boldsymbol{\gamma}) = L_{\boldsymbol{\gamma}^{\mathrm{r}}}(\dot{z}, z) + \bar{\Gamma}, \tag{4.196}$$

with $\boldsymbol{\gamma}^{\mathrm{r}} \equiv (\boldsymbol{\gamma}_1 + \mathfrak{D}^\dagger p_{\mathrm{s}}, \gamma_2)$.

Fluctuation symmetry

The biased Hamiltonian has a fluctuation symmetry:

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, z) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{F}_{\boldsymbol{\gamma}} + \theta \boldsymbol{f}, z), \tag{4.197}$$

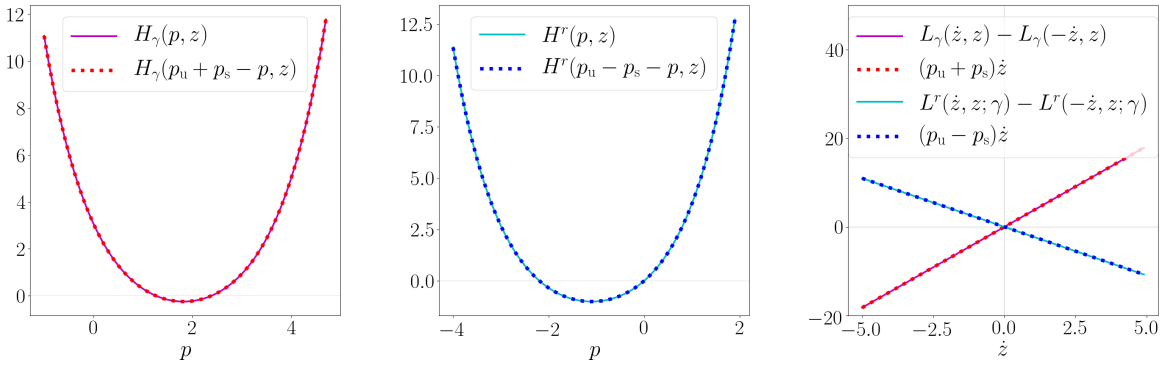

Figure 4.11 – (Left) Fluctuation symmetry for the biased Hamiltonian (4.202). (Middle) Fluctuation symmetry for the rectified Hamiltonian (4.203–4.205). (Right) Fluctuation symmetries for the biased and rectified Lagrangians (4.204).
The figures are obtained for $z = 0.3$, $\mathcal{E} = 0.8$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$, $\gamma^\nu_{1,\epsilon} = \epsilon$ ($\epsilon = \pm 1$ and $\nu = 1, 2$) and $\gamma_2 = 1$.

with

$$(\boldsymbol{F}_{\boldsymbol{\gamma}})^\nu_\epsilon \equiv \epsilon \ln \frac{k^\nu_- \mathrm{e}^{\gamma^\nu_{1,-}}}{k^\nu_+ \mathrm{e}^{\gamma^\nu_{1,+}}}, \tag{4.198}$$

$$\theta f^\nu_\epsilon \equiv f^\nu_{-\epsilon}. \tag{4.199}$$

As seen in Sec. 4.2.5, this symmetry is inherited by the rectified Hamiltonian through

$$\mathscr{H}^{\mathrm{r}}(\boldsymbol{f}, z; \boldsymbol{\gamma}) = \mathscr{H}^{\mathrm{r}}(\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} + \theta \boldsymbol{f}, z; \boldsymbol{\gamma}), \tag{4.200}$$

with

$$\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}} \equiv \boldsymbol{F}_{\boldsymbol{\gamma}} + (\theta - 1)\mathfrak{D}^\dagger p_{\mathrm{s}}. \tag{4.201}$$

The fluctuation symmetry reads for the standard biased and rectified Hamiltonians:

$$H_{\boldsymbol{\gamma}}(p, z) = H_{\boldsymbol{\gamma}}(p_{\mathrm{u}} + p_{\mathrm{s}} - p, z), \tag{4.202}$$

$$H^{\mathrm{r}}(p, z; \boldsymbol{\gamma}) = H^{\mathrm{r}}(p_{\mathrm{u}} - p_{\mathrm{s}} - p, z; \boldsymbol{\gamma}), \tag{4.203}$$

where we used $\theta\mathfrak{D}^\dagger = -\mathfrak{D}^\dagger$. We check numerically this symmetries in Fig. 4.11 (left and middle). We close this remark by pointing at the fact that these symmetries translate for the Lagrangians into:

$$L_{\boldsymbol{\gamma}}(\dot z, z) - L_{\boldsymbol{\gamma}}(-\dot z, z) = (p_{\mathrm{u}} + p_{\mathrm{s}})\dot z, \tag{4.204}$$

$$L^{\mathrm{r}}(\dot z, z; \boldsymbol{\gamma}) - L^{\mathrm{r}}(-\dot z, z; \boldsymbol{\gamma}) = (p_{\mathrm{u}} - p_{\mathrm{s}})\dot z, \tag{4.205}$$

as shown numerically in Fig. 4.11 (right).

4.3.2 Nonlinear chemical reaction

We now apply the results stated in Section 4.1 to a chemical system modeled by the following chemical reaction:

$$A \underset{\mathfrak{K}_{-1}}{\overset{\mathfrak{K}_{+1}}{\rightleftharpoons}} 2X \underset{\mathfrak{K}_{-2}}{\overset{\mathfrak{K}_{+2}}{\rightleftharpoons}} B. \tag{4.206}$$

The species A and B are chemostatted and have constant concentrations a and b, while the concentration x of the species X is allowed to fluctuate. The observables $\boldsymbol{z}$ and $\boldsymbol{\lambda}$ represent respectively the concentration x and the chemical current, both related by $\dot{x} = \mathcal{D}\boldsymbol{\lambda}$ where $\mathcal{D}$ is the line vector whose component $\mathcal{D}_{\epsilon \mathsf{r}} \equiv -2\epsilon(-1)^{\mathsf{r}}$ is the variation of the number of species X when the reaction $\epsilon \mathsf{r}$ occurs. From Eq. (4.29), the transition rates $k_{\epsilon \mathsf{r}}(\boldsymbol{z})$ read

$$
\begin{aligned}
k_{+1} &= \mathfrak{K}_{+1} a, && (4.207)\\
k_{-1} &= \mathfrak{K}_{-1} x^2, && (4.208)\\
k_{+2} &= \mathfrak{K}_{+2} x^2, && (4.209)\\
k_{-2} &= \mathfrak{K}_{-2} b. && (4.210)
\end{aligned}
$$

For simplicity, the observable $\boldsymbol{A}_t$ is chosen to be the chemical current $\frac{1}{t}\int_0^t \boldsymbol{\lambda}(t')\mathrm{d}t'$ of conjugate variable $\boldsymbol{\gamma} \equiv (\{\gamma_{\epsilon \mathsf{r}}\})$. The standard biased Hamiltonian reads then

$$
H_{\boldsymbol{\gamma}}(p, x) = \sum_{\epsilon \mathsf{r}} k_{\epsilon \mathsf{r}}(x) \left[\mathrm{e}^{-2(-1)^{\mathsf{r}} \epsilon p + \gamma_{\epsilon \mathsf{r}}} - 1 \right]. \tag{4.211}
$$

For clarity we introduce:

$$
\begin{aligned}
\alpha &\equiv \mathfrak{K}_{+1} a \mathrm{e}^{\gamma_{+1}} + \mathfrak{K}_{-2} b \mathrm{e}^{\gamma_{-2}}, && (4.212)\\
\beta &\equiv \mathfrak{K}_{-1} \mathrm{e}^{\gamma_{-1}} + \mathfrak{K}_{+2} \mathrm{e}^{\gamma_{+2}}, && (4.213)\\
\delta &\equiv \mathfrak{K}_{-1} + \mathfrak{K}_{+2}, && (4.214)
\end{aligned}
$$

so that the biased Hamiltonian (4.211) simplifies to

$$
H_{\boldsymbol{\gamma}}(p, x) = \alpha \mathrm{e}^{2p} + \beta x^2 \mathrm{e}^{-2p} - \delta x^2 - (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b). \tag{4.215}
$$

Hamilton's equations for the biased Hamiltonian are given by

$$
\begin{cases}
\dot{x} = \frac{\partial H_{\boldsymbol{\gamma}}}{\partial p} = 2\alpha \mathrm{e}^{2p} - 2\beta x^2 \mathrm{e}^{-2p}, \\
-\dot{p} = \frac{\partial H_{\boldsymbol{\gamma}}}{\partial x} = 2\beta x \mathrm{e}^{-2p} - 2\delta x.
\end{cases} \tag{4.216}
$$

This system admits one critical manifold consisting of a fixed point of coordinates

$$
\begin{cases}
x_0^\star = \sqrt{\frac{\alpha\beta}{\delta^2}}, \\
p_0^\star = \frac{1}{2} \ln \frac{\beta}{\delta}.
\end{cases} \tag{4.217}
$$

In Fig 4.12, we plot the phase portrait both in the non-biased and biased cases. The dominant trajectory in the long-time limit corresponds to $H_{\boldsymbol{\gamma}}(p, x) = \bar{\Gamma}$ with

$$
\bar{\Gamma} = H_{\boldsymbol{\gamma}}(p_0^\star, x_0^\star) = \frac{\alpha\beta}{\delta} - (\mathfrak{K}_{+1} a + \mathfrak{K}_{-2} b). \tag{4.218}
$$

It is easy to check that $\bar{\Gamma}(\boldsymbol{\gamma} = 0) = 0$, as required in the non-biased case. Biasing modifies the position of the fixed point from a concentration x_{NB} to a new concentration $x_0^\star$ and shifts the variable p of the fixed point from 0 to $p_0^\star \neq 0$, which is consistent with the fact that the SCGF is non-zero when $\boldsymbol{\gamma} \neq 0$ and that the biased Hamiltonian does not vanish at $p = 0$. The implicit equation

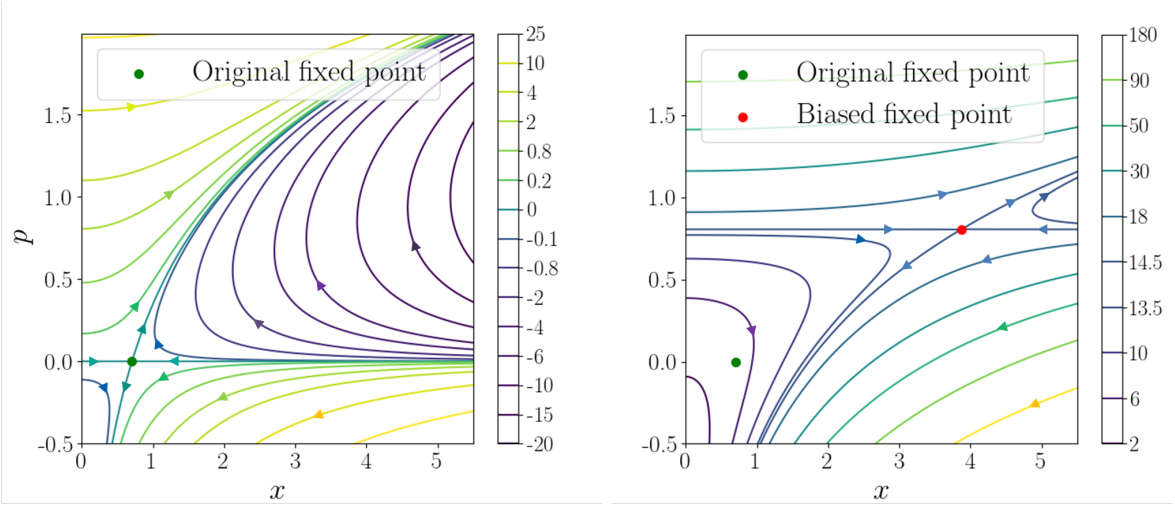

Figure 4.12 – (Left) Trajectories of the original Hamiltonian ($\boldsymbol{\gamma} = 0$). The coordinates of the fixed point (green point) are $\left(x_{\text{NB}} = \sqrt{(\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b)/\delta}, p_{\text{NB}} = 0\right)$. (Right) Trajectories of the biased Hamiltonian. The coordinates of the fixed point (red point) are $\left(x_0^\star = \sqrt{\alpha\beta/\delta^2}, p_0^\star = (1/2)\ln(\beta/\delta)\right)$.
The left figure is obtained for $\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b = 0.5$, $\delta = 1$, for which $x_{\text{NB}} = 0.707$. The right figure is obtained for $\alpha = 3$, $\beta = 5$, $\delta = 1$, $\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b = 0.5$ for which $x_0^\star = 3.873$ and $p_0^\star = 0.805$.

$$H_{\boldsymbol{\gamma}}(p(x, \boldsymbol{\gamma}), x) = \bar{\Gamma} \tag{4.219}$$

admits two solutions $p_{\text{s,u}} \equiv p_{\text{s,u}}(x, \boldsymbol{\gamma})$ with:

$$\begin{cases} p_{\text{s}}(x, \boldsymbol{\gamma}) \equiv p_0^\star = \frac{1}{2}\ln\frac{\beta}{\delta}, \\ p_{\text{u}}(x, \boldsymbol{\gamma}) \equiv \frac{1}{2}\ln\frac{x^2}{x_0^{\star 2}\mathrm{e}^{-2p_\star}} = \frac{1}{2}\ln\frac{\delta x^2}{\alpha}. \end{cases} \tag{4.220}$$

The solution p_{s} corresponds to the globally stable solution and contains two relaxing transient trajectories, and the solution p_{u} corresponds to the globally unstable solution and contains two fluctuating transient trajectories (see Fig. 4.12). Notice that in the non-biased case, we have indeed that $p_{\text{s}}(\boldsymbol{\gamma} = 0) = 0$. Finally, we can compute the rectified Hamiltonian from Eq. (4.143):

$$H^{\text{r}}(p, x; \boldsymbol{\gamma}) = \delta x_0^{\star 2}\left(\mathrm{e}^{2p} - 1\right) + \delta x^2\left(\mathrm{e}^{-2p} - 1\right), \tag{4.221}$$

which can be rewritten as

$$H^{\text{r}}(p, x; \boldsymbol{\gamma}) = \sum_{\epsilon\text{r}} K_{\epsilon\text{r}}(x)\left[\mathrm{e}^{-2(-1)^{\text{r}}\epsilon p} - 1\right], \tag{4.222}$$

where $K_{\epsilon\text{r}}(x) \equiv k_{\epsilon\text{r}}(x)\mathrm{e}^{\gamma_{\epsilon\text{r}} - 2(-1)^{\text{r}}\epsilon p_{\text{s}}}$ is the rectified intensive rate obtained from the Doob transform of the biased transition rate in the linear operator formalism. As expected, the rectified Hamiltonian respects the structure of the original non-biased Hamiltonian with new rates.

Figs. 4.12 and 4.13 offer a visualization of the effect of biasing and rectification on the fixed point. Starting from the original Hamiltonian with fixed point ($x_{\text{NB}}, p_{\text{NB}} = 0$), we

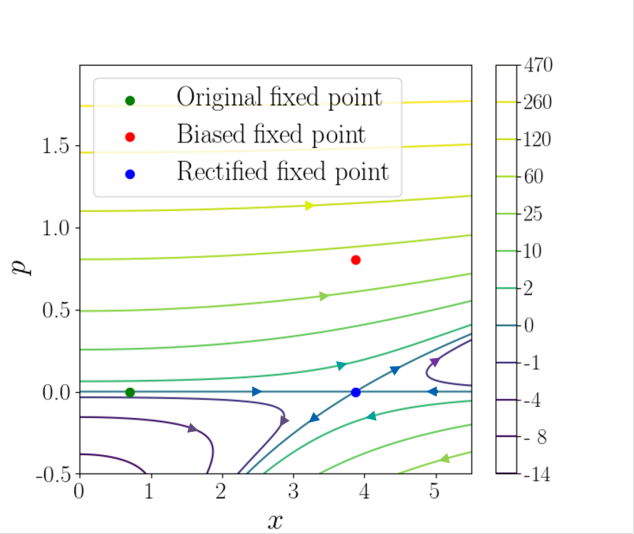

Figure 4.13 – Trajectories of the rectified Hamiltonian. The coordinates of the fixed point (blue point) are $\left(x_0^\star = \sqrt{\alpha\beta/\delta^2}, p = 0\right)$. The figure is obtained for $\alpha = 3$, $\beta = 5$, $\delta = 1$, $\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b = 0.5$.

bias the dynamics via the parameter $\boldsymbol{\gamma}$ to impose a new typical value of $\bar{\boldsymbol{A}}$ in the long-time limit. This translates into a new dominant trajectory, hence a new fixed point $(x_0^\star, p_0^\star)$. Yet, the biased Hamiltonian does not vanish at $p = 0$, and the globally stable $p_\mathrm{s} = p_0^\star$ is not zero, as required for a norm-conserving Markov process. The rectification allows one to build a proper statistical Hamiltonian by shifting the globally stable manifold p_s to 0, leading to a new fixed point at $p = 0$ while remaining at the same position $x_0^\star$, compatible with the imposed value of $\bar{\boldsymbol{A}}$.

Fluctuation symmetry

The biased Hamiltonian has a fluctuation symmetry:

$$H_{\boldsymbol{\gamma}}(p, x) = H_{\boldsymbol{\gamma}}(p_\mathrm{u} + p_\mathrm{s} - p, x), \tag{4.223}$$

which implies a fluctuation symmetry for the rectified Hamiltonian

$$H^\mathrm{r}(p, x; \boldsymbol{\gamma}) = H^\mathrm{r}(p_\mathrm{u} - p_\mathrm{s} - p, x; \boldsymbol{\gamma}). \tag{4.224}$$

We check numerically this symmetries in Fig. 4.14 (right and middle). For this system, the standard biased Lagrangian can be computed explicitly and we obtain

$$L_{\boldsymbol{\gamma}}(\dot{x}, x) = \frac{1}{2}\dot{x}\ln\left(\frac{\dot{x} + \sqrt{\dot{x}^2 + 16\alpha\beta x^2}}{4\alpha}\right) - \frac{1}{2}\sqrt{\dot{x}^2 + 16\alpha\beta x^2} + \delta x^2 + \mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b, \tag{4.225}$$

while the standard rectified Lagrangian reads

$$L^\mathrm{r}(\dot{x}, x) = \frac{1}{2}\dot{x}\ln\left(\frac{\dot{x} + \sqrt{\dot{x}^2 + 16(\delta x x_0^\star)^2}}{4\delta x_0^{\star 2}}\right) - \frac{1}{2}\sqrt{\dot{x}^2 + 16(\delta x x_0^\star)^2} + \delta(x^2 + x_0^{\star 2}), \tag{4.226}$$

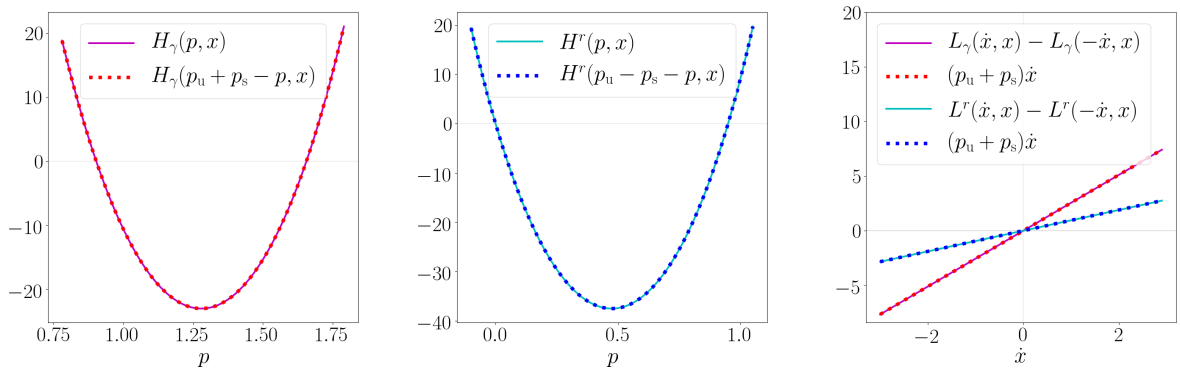

Figure 4.14 – (Left) Fluctuation symmetry for the biased Hamiltonian (4.223). (Middle) Fluctuation symmetry for the rectified Hamiltonian (4.224). (Right) Fluctuation symmetries for the biased and rectified Lagrangians (4.227–4.228).
The figures are obtained for $\alpha = 3$, $\beta = 5$, $\delta = 1$, $\mathfrak{K}_{+1}a + \mathfrak{K}_{-2}b = 0.5$ and $x = 10$.

Both Lagrangians satisfy the following fluctuation symmetry:

$$L_{\boldsymbol{\gamma}}(\dot{x},x) - L_{\boldsymbol{\gamma}}(-\dot{x},x) = (p_\mathrm{u}+p_\mathrm{s})\dot{x}, \tag{4.227}$$
$$L^\mathrm{r}(\dot{x},x;\boldsymbol{\gamma}) - L^\mathrm{r}(-\dot{x},x;\boldsymbol{\gamma}) = (p_\mathrm{u}-p_\mathrm{s})\dot{x}, \tag{4.228}$$

as shown numerically in Fig 4.14 (right).

4.4 Opening on periodically driven Markov processes

We expect our results to be transposable to periodically driven processes with the appropriate changes. In the following, we present a first investigation of these changes for periodically driven population processes. All the notations, definitions and expressions obtained in Section 4.1 for the non-biased and biased processes hold for periodically driven population processes with the difference that the generators $\boldsymbol{k}$ and $\boldsymbol{\kappa}$, and eventually the parameter $\boldsymbol{\gamma}$, are now T-periodic. It implies that the Lagrangians and Hamiltonians become explicitly time-dependent and periodic with period T. We aim to derive the rectified Hamiltonian associated with the generator of the driven process. According to the results of Section 3.1.5, the driven generator for periodically-driven Markov processes reads

$$\boldsymbol{K} \equiv \boldsymbol{\kappa}^{\mathrm{e}^{\boldsymbol{U}}} = \mathcal{D}(\mathrm{e}^{\boldsymbol{U}})\boldsymbol{\kappa}\,\mathcal{D}(\mathrm{e}^{-\boldsymbol{U}}) - \mathcal{D}(\mathrm{e}^{-\boldsymbol{U}})\mathcal{D}(\mathrm{e}^{\boldsymbol{U}}\boldsymbol{\kappa}), \tag{4.229}$$

where $\mathrm{e}^{\boldsymbol{U}}$ is now the solution of

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\boldsymbol{U}} = -\mathrm{e}^{\boldsymbol{U}}\boldsymbol{\kappa}, \\ \boldsymbol{U}(0) = \boldsymbol{U}(T) + T\Gamma, \end{cases} \tag{4.230}$$

and where the SCGF $\Gamma = \mathcal{N}\bar{\Gamma}$ is related to the dominant eigenvalue χ_T of the single-period propagator $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$ via $\Gamma = \frac{1}{T}\ln\chi_T$.

Differential equation in the continuous limit

Before deriving the Lagrangian and Hamiltonian associated with the dynamics generated by $\boldsymbol{K}$, let us first investigate in the continuous limit the differential equation (4.230). To do so, assume that for $\boldsymbol{N} = O(\mathcal{N})$, $\mathcal{N} \to \infty$, there exists a function $S(\boldsymbol{z}, t)$ such that

$$U_{\boldsymbol{N}}(t) = \mathcal{N} S(\boldsymbol{z}, t) + o(\mathcal{N}), \tag{4.231}$$

so that

$$\begin{aligned} \dot{U}_{\boldsymbol{N}}(t) &\simeq \mathcal{N} \frac{\partial S}{\partial t}(\boldsymbol{z}, t), \\ U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}} &\simeq \mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} S, \end{aligned} \tag{4.232}$$

for any α. From Eq. (4.230), we have:

$$-\dot{U}_{\boldsymbol{N}} \mathrm{e}^{U_{\boldsymbol{N}}} = \sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha}} \kappa_{\alpha,\boldsymbol{N}} + \mathrm{e}^{U_{\boldsymbol{N}}} \kappa_{\boldsymbol{N},\boldsymbol{N}} \tag{4.233}$$

$$-\dot{U}_{\boldsymbol{N}} = \sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}}} \kappa_{\alpha,\boldsymbol{N}} + \kappa_{\boldsymbol{N},\boldsymbol{N}} \tag{4.234}$$

$$-\dot{U}_{\boldsymbol{N}} = \sum_\alpha \mathrm{e}^{U_{\boldsymbol{N}+\mathfrak{D}_\alpha} - U_{\boldsymbol{N}}} k_{\alpha,\boldsymbol{N}} \mathrm{e}^{\gamma^1_\alpha} - \sum_\alpha k_{\alpha,\boldsymbol{N}} + \boldsymbol{\gamma}_2 \cdot \boldsymbol{N}, \tag{4.235}$$

where we used Eq. (4.40) in the last equation. Taking the continuous limit (4.9) and using the assumption of Eq. (4.232), we finally obtain

$$\frac{\partial S}{\partial t} + \sum_\alpha k_\alpha(\boldsymbol{z}, t) \left[\mathrm{e}^{\mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} S + \gamma^1_\alpha} - 1 \right] + \boldsymbol{\gamma}_2 \cdot \boldsymbol{z} = 0. \tag{4.236}$$

We recognize the biased Hamiltonian (4.43) in the left-hand side of Eq. (4.236), leading to

$$\begin{cases} \frac{\partial S}{\partial t} + \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} S, \boldsymbol{z}, t) = 0, \\ S(0) = S(T) + T\bar{\Gamma}. \end{cases} \tag{4.237}$$

Hence, S appears to be Hamilton's principal function, solution of the time-dependent Hamilton-Jacobi equation Eq. (4.60).

Rectified process

Following the procedure presented in Appendix 4.A, we derive the rectified Hamiltonian associated with the driven generator $\boldsymbol{K}$ and we obtain

$$H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}, t; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} S, \boldsymbol{z}, t) + \frac{\partial S}{\partial t} = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} S, \boldsymbol{z}, t) - H_{\boldsymbol{\gamma}}(\partial_{\boldsymbol{z}} S, \boldsymbol{z}, t), \tag{4.238}$$

which vanishes at $\boldsymbol{p} = 0$ as required. In the case of time-independent Hamiltonians, S is the Legendre transform of W according to Eq. (4.61): $S(\boldsymbol{z}, t) = -Et + W(\boldsymbol{z}, E)$, and we have for any $\boldsymbol{z}$: $S(\boldsymbol{z}, 0) = S(\boldsymbol{z}, T) + ET$. From the boundary condition of Eq. (4.237), it leads to $E = \bar{\Gamma}$ and $S(\boldsymbol{z}, t) = -\bar{\Gamma} t + W(\boldsymbol{z}, E)$. Finally, using the identities $\partial_{\boldsymbol{z}} S = \partial_{\boldsymbol{z}} W$ and $\frac{\partial S}{\partial t} = -\bar{\Gamma}$ in Eq. (4.238), we recover the definition of the rectified Hamiltonian in the time-homogeneous case.

To finish, we remark that the rectified Hamiltonian (4.238) stems from a canonical transformation $(\boldsymbol{Z}, \boldsymbol{P}) \to (\boldsymbol{z}, \boldsymbol{p})$ starting from the biased Hamiltonian $H_{\boldsymbol{\gamma}}(\boldsymbol{P}, \boldsymbol{Z})$. Indeed, consider the type-2 generating function

$$F_2(\boldsymbol{p}, \boldsymbol{Z}, t) = \boldsymbol{Z} \cdot \boldsymbol{p} + S(\boldsymbol{Z}, t). \tag{4.239}$$

Then, the new variables $\boldsymbol{z}$, $\boldsymbol{p}$ follow from

$$\boldsymbol{z} = \partial_{\boldsymbol{p}} F_2 = \boldsymbol{Z}, \tag{4.240}$$
$$\boldsymbol{P} = \partial_{\boldsymbol{Z}} F_2 = \boldsymbol{p} + \partial_{\boldsymbol{Z}} S, \tag{4.241}$$

and the new Hamiltonian follows from

$$H^{\mathrm{r}}(\boldsymbol{p}, \boldsymbol{z}; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(\boldsymbol{P}, \boldsymbol{Z}) + \partial_t F_2 = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \partial_{\boldsymbol{z}} S, \boldsymbol{z}) + \frac{\partial S}{\partial t}, \tag{4.242}$$

which is the result of Eq. (4.238).

These derivations give a first idea on the rectification procedure for periodically driven nonlinear processes. However, a whole investigation remains to be done in order to construct a proper theory, especially with regard to the characterization of the solutions of the time-dependent HJ equation (4.237) and the behavior of the typical dynamics of the system whose representation is made difficult by the time-dependence of the Hamiltonian, implying that it is not conservative anymore along the solutions of Hamilton's equations.

4.5 Conclusion

In this Chapter, we generalized the problem of conditioning time-homogeneous Markov processes developed in Refs. [80, 81] (linear operator formalism) to nonlinear Markov processes by adapting the Lagrangian/Hamiltonian formalism introduced in Ref. [124] for chemical reaction networks to general nonlinear Markov processes. In Ref. [124], this formalism was used to describe the non-biased and biased processes. In this Chapter, we went a step further by describing the driven process in the Hamiltonian formalism. In the linear operator formalism, the generator of the driven process is obtained from the generalized Doob transform of the biased generator using its dominant left eigenvector, and whose dominant eigenvalue coincides with the SCGF. In the Hamiltonian formalism, this amounts to rectifying the biased Hamiltonian by performing a canonical transformation that involves the globally stable solution of the Hamilton-Jacobi equation with eigenrate equal to the SCGF. These results have been explicitly derived for population processes by taking the continuous limit of the linear operator formalism. On this basis, we built a general and abstract theory for the rectification of biased Hamiltonians. This theory relies on two points:

- The existence of a unique globally stable solution and a unique globally unstable solution of the Hamilton-Jacobi equation with eigenrate equal to the max-min value of the Hamiltonian (analogue of the PF theorem for Markov generators ensuring the existence of a unique left eigenvector and a unique right eigenvector associated with their largest eigenvalue).

- The max-min value of the Hamiltonian coincides with the SCGF (in time and size) (analogue to the fact that the dominant eigenvalue of the biased generator coincides with the SCGF (in time)). This result follows from the convergence of the Hamiltonian along the long-time typical trajectory to its max-min value.

For now, the first point is presented as a conjecture and the second point is proven only in the case of a unique fixed point in the long-time dynamics. Thereby, investigating rigorous demonstrations for these two results remains to be done in order to complete our theory. Nevertheless, we point out that our conjectures have been verified in all our examples and are backed by their derivation from the linear operator formalism for the case of population processes.

To conclude, we hope that the results presented in this chapter will be the starting point for a complete theory on the conditioning/rectification of nonlinear Markov processes, both for time-homogeneous and periodically driven processes.

Appendices

"Le meilleur moyen de réaliser l'impossible est de croire que c'est possible."

Lewis Carroll

4.A Derivation of the Lagrangian of population processes

In this Appendix, we compute the transition probability $P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$ in the continuous limit defined by

$$\begin{cases} \mathcal{N} \to \infty, \\ \delta t \to 0, \\ \mathcal{N}\delta t \to \infty, \end{cases} \tag{4.243}$$

by following the procedure described in the appendix of Ref. [124]. As seen in Sec. 1.1.3, the transition probabilities are solutions of the master equation with transition matrix $\boldsymbol{k}$ that reads in the Dirac notation

$$\boldsymbol{k} = \sum_{\alpha,\boldsymbol{N}} \left(k_{\alpha,\boldsymbol{N}} \left| \boldsymbol{N} + \boldsymbol{\mathfrak{D}}_\alpha \right\rangle \left\langle \boldsymbol{N} \right| - k_{\alpha,\boldsymbol{N}} \left| \boldsymbol{N} \right\rangle \left\langle \boldsymbol{N} \right| \right). \tag{4.244}$$

It follows

$$P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) = \left\langle \boldsymbol{N}_\mathrm{f} \right| \mathrm{e}^{\delta t \boldsymbol{k}} \left| \boldsymbol{N}_\mathrm{i} \right\rangle. \tag{4.245}$$

For clarity, we introduce the following operators

$$A_\alpha \equiv \sum_{\boldsymbol{N}} k_{\alpha,\boldsymbol{N}} \left| \boldsymbol{N} + \boldsymbol{\mathfrak{D}}_\alpha \right\rangle \left\langle \boldsymbol{N} \right|, \tag{4.246}$$

$$B_\alpha \equiv -\sum_{\boldsymbol{N}} k_{\alpha,\boldsymbol{N}} \left| \boldsymbol{N} \right\rangle \left\langle \boldsymbol{N} \right|, \tag{4.247}$$

so that the transition matrix reads $\boldsymbol{k} = \sum_\alpha (A_\alpha + B_\alpha)$. In order to compute the exponential of these operators, we need to investigate their commutation properties. Note that $[B_\alpha, B_\beta] = 0$ for any transitions α, β since both operators are diagonal. Let us compute

the other commutators:

$$[A_\alpha, A_\beta] = \sum_{\boldsymbol{N}} \left[k_{\alpha,\boldsymbol{N}+\mathfrak{D}_\beta} k_{\beta,\boldsymbol{N}} - k_{\beta,\boldsymbol{N}+\mathfrak{D}_\alpha} k_{\alpha,\boldsymbol{N}}\right] |\boldsymbol{N} + \mathfrak{D}_\alpha + \mathfrak{D}_\beta\rangle \langle \boldsymbol{N}| \,, \tag{4.248}$$

$$[A_\alpha, B_\beta] = \sum_{\boldsymbol{N}} \left[k_{\beta,\boldsymbol{N}+\mathfrak{D}_\alpha} k_{\alpha,\boldsymbol{N}} - k_{\alpha,\boldsymbol{N}} k_{\beta,\boldsymbol{N}}\right] |\boldsymbol{N} + \mathfrak{D}_\alpha\rangle \langle \boldsymbol{N}| \,. \tag{4.249}$$

Using the definitions of the density (4.4) $\boldsymbol{z} = \frac{\boldsymbol{N}}{\mathcal{N}}$ and the intensive rate (4.8) $k_{\alpha,\boldsymbol{N}} = \mathcal{N} k_\alpha\left(\frac{\boldsymbol{N}}{\mathcal{N}}\right) + o(\mathcal{N})$, we can expand the rate $k_{\beta,\boldsymbol{N}+\mathfrak{D}_\alpha}$ for any transition α, β and for any population vector state $\boldsymbol{N}$ of order $\mathcal{N}$, i.e. such that $N_X = O(\mathcal{N})$ for any state X:

$$k_{\beta,\boldsymbol{N}+\mathfrak{D}_\alpha} = \mathcal{N} k_\beta\left(\frac{\boldsymbol{N}+\mathfrak{D}_\alpha}{\mathcal{N}}\right) + o(\mathcal{N}) \approx \mathcal{N}\left[k_\beta(\boldsymbol{z}) + \frac{1}{\mathcal{N}}\mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} k_\beta(\boldsymbol{z})\right], \tag{4.250}$$

where we used the fact that $\frac{\mathfrak{D}_\alpha}{\mathcal{N}} \xrightarrow[\mathcal{N}\to\infty]{} 0$. Using Eq. (4.250), and assuming that the accessible population states $\boldsymbol{N}$ are all of order $\mathcal{N}$, i.e. we dismiss the case where only few particles occupy a given state, the commutators of Eqs. (4.248–4.249) read

$$[A_\alpha, A_\beta] = \sum_{\boldsymbol{N}} \mathcal{N}\left[\mathfrak{D}_\beta \cdot \partial_{\boldsymbol{z}} k_\alpha(\boldsymbol{z}) k_\beta(\boldsymbol{z}) - \mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} k_\beta(\boldsymbol{z}) k_\alpha(\boldsymbol{z})\right] |\boldsymbol{N} + \mathfrak{D}_\alpha + \mathfrak{D}_\beta\rangle \langle \boldsymbol{N}| \,, \tag{4.251}$$

$$[A_\alpha, B_\beta] = \sum_{\boldsymbol{N}} \mathcal{N}\left[\mathfrak{D}_\alpha \cdot \partial_{\boldsymbol{z}} k_\beta(\boldsymbol{z}) k_\alpha(\boldsymbol{z})\right] |\boldsymbol{N} + \mathfrak{D}_\alpha\rangle \langle \boldsymbol{N}| \,. \tag{4.252}$$

Hence, these commutators are of order $\mathcal{N}$ and it will be the case of any higher order commutator. Using Baker-Cambell-Hausdorff formula, we can write

$$\mathrm{e}^{\delta t(A_\alpha + A_\beta)} \simeq \mathrm{e}^{\delta t A_\alpha} \mathrm{e}^{\delta t A_\beta} \mathrm{e}^{-\frac{1}{2}\delta t^2 [A_\alpha, A_\beta]}. \tag{4.253}$$

Since, A_α, A_β and their commutator are all of the same order, we can neglect the last term in the limit $\delta t \to 0$ implying the commutation of A_α and A_β in the continuous limit. The same reasoning leads to the commutation of A_α and B_β. Let us now compute the term $\langle \boldsymbol{N}_\mathrm{f}| \mathrm{e}^{\delta t(A_\alpha + B_\alpha)} |\boldsymbol{N}_\mathrm{i}\rangle$:

$$\begin{aligned}
\langle \boldsymbol{N}_\mathrm{f}| \mathrm{e}^{\delta t(A_\alpha + B_\alpha)} |\boldsymbol{N}_\mathrm{i}\rangle &\simeq \langle \boldsymbol{N}_\mathrm{f}| \mathrm{e}^{\delta t A_\alpha} \mathrm{e}^{\delta t B_\alpha} |\boldsymbol{N}_\mathrm{i}\rangle && (4.254)\\
&= \mathrm{e}^{-\delta t k_{\alpha,\boldsymbol{N}_\mathrm{i}}} \langle \boldsymbol{N}_\mathrm{f}| \mathrm{e}^{\delta t A_\alpha} |\boldsymbol{N}_\mathrm{i}\rangle && (4.255)\\
&= \mathrm{e}^{-\delta t k_{\alpha,\boldsymbol{N}_\mathrm{i}}} \sum_{\ell=0}^{\infty} \langle \boldsymbol{N}_\mathrm{f}| \frac{\delta t^\ell}{\ell!} (A_\alpha)^\ell |\boldsymbol{N}_\mathrm{i}\rangle \,. && (4.256)
\end{aligned}$$

where we used the fact that B_α is diagonal in the second equality and where we expanded the exponential in the third equality. The ℓ^{th} power of A_α is given by

$$(A_\alpha)^\ell = \sum_{\boldsymbol{N}} \left[\prod_{m=1}^{\ell} k_{\alpha,\boldsymbol{N}+(m-1)\mathfrak{D}_\alpha}\right] |\boldsymbol{N} + \ell\,\mathfrak{D}_\alpha\rangle \langle \boldsymbol{N}| \,, \tag{4.257}$$

leading to

$$\langle \boldsymbol{N}_\mathrm{f}| \mathrm{e}^{\delta t(A_\alpha + B_\alpha)} |\boldsymbol{N}_\mathrm{i}\rangle = \mathrm{e}^{-\delta t k_{\alpha,\boldsymbol{N}_\mathrm{i}}} \sum_{\ell=0}^{\infty} \frac{(\delta t \mathcal{N})^\ell}{\ell!} \prod_{m=1}^{\ell} \frac{k_{\alpha,\boldsymbol{N}_\mathrm{i}+(m-1)\mathfrak{D}_\alpha}}{\mathcal{N}} \delta(\boldsymbol{N}_\mathrm{f} - \boldsymbol{N}_\mathrm{i} - \ell\,\mathfrak{D}_\alpha). \tag{4.258}$$

For a large factor X, and using Stirling's formula, we find that the function $\ell \mapsto \frac{X^\ell}{\ell!}$ reaches its maximum for $\ell_* \simeq X$. Applying this result for $X = \delta t\mathcal{N}$, we can approximate the sum over ℓ by its dominant contribution given by $\ell_* \propto \delta t\mathcal{N}$ in the limit $\delta t\mathcal{N} \to \infty$. The proportionality factor between ℓ_* and $\delta t\mathcal{N}$ has to be the empirical current λ_α so that the constraint appearing in the Dirac delta function $\boldsymbol{N}_\mathrm{f} - \boldsymbol{N}_\mathrm{i} = \ell\,\mathfrak{D}_\alpha$ is consistent with the constraint of Eq. (4.5), hence $\ell_* = \delta t\mathcal{N}\lambda_\alpha$. It yields

$$\langle \boldsymbol{N}_\mathrm{f}|\, \mathrm{e}^{\delta t(A_\alpha+B_\alpha)}\, |\boldsymbol{N}_\mathrm{i}\rangle \underset{\mathcal{N}\to\infty}{\simeq} \mathrm{e}^{-\delta t k_{\alpha,\boldsymbol{N}_\mathrm{i}}} \frac{(\delta t\mathcal{N})^{\ell_*}}{\ell_*!} \prod_{m=1}^{\ell_*} \frac{k_{\alpha,\boldsymbol{N}_\mathrm{i}+(m-1)\mathfrak{D}_\alpha}}{\mathcal{N}} \delta(\boldsymbol{N}_\mathrm{f} - \boldsymbol{N}_\mathrm{i} - \ell_*\,\mathfrak{D}_\alpha). \quad (4.259)$$

Since $\frac{\ell_*}{\mathcal{N}} = \delta t\lambda_\alpha \xrightarrow[\delta t\to 0]{} 0$, we have from the expansion in Eq. (4.250):

$$\prod_{m=1}^{\ell_*} \frac{k_{\alpha,\boldsymbol{N}_\mathrm{i}+(m-1)\mathfrak{D}_\alpha}}{\mathcal{N}} = \prod_{m=1}^{\ell_*} [k_\alpha(\boldsymbol{z}_\mathrm{i}) + o(1)] \simeq k_\alpha(\boldsymbol{z}_\mathrm{i})^{\ell_*}, \quad (4.260)$$

with $\boldsymbol{z}_\mathrm{i} \equiv \frac{\boldsymbol{N}_\mathrm{i}}{\mathcal{N}}$. It follows

$$\langle \boldsymbol{N}_\mathrm{f}|\, \mathrm{e}^{\delta t(A_\alpha+B_\alpha)}\, |\boldsymbol{N}_\mathrm{i}\rangle \underset{\mathcal{N}\to\infty}{\simeq} \mathrm{e}^{-\delta t\mathcal{N}k_\alpha(\boldsymbol{z}_\mathrm{i})} \frac{(\delta t\mathcal{N})^{\ell_*}}{\ell_*!} k_\alpha(\boldsymbol{z}_\mathrm{i})^{\ell_*} \delta(\boldsymbol{z}_\mathrm{f} - \boldsymbol{z}_\mathrm{i} - \delta t\,\mathfrak{D}_\alpha\lambda_\alpha). \quad (4.261)$$

Using Stirling's Formula, we finally obtain

$$\langle \boldsymbol{N}_\mathrm{f}|\, \mathrm{e}^{\delta t(A_\alpha+B_\alpha)}\, |\boldsymbol{N}_\mathrm{i}\rangle \underset{\mathcal{N}\to\infty}{\simeq} \mathrm{e}^{-\delta t\mathcal{N}\left[\lambda_\alpha \ln\left(\frac{\lambda_\alpha}{k_\alpha(\boldsymbol{z}_\mathrm{i})}\right) - \lambda_\alpha + k_\alpha(\boldsymbol{z}_\mathrm{i})\right]} \delta(\boldsymbol{z}_\mathrm{f} - \boldsymbol{z}_\mathrm{i} - \delta t\,\mathfrak{D}_\alpha\lambda_\alpha). \quad (4.262)$$

We remind that this result is valid within the continuous limit (4.243) and for population state $\boldsymbol{N}$ of order $\mathcal{N}$. We can now compute the transition probability $\langle \boldsymbol{N}_\mathrm{f}|\, \mathrm{e}^{\delta t\boldsymbol{k}}\, |\boldsymbol{N}_\mathrm{i}\rangle = \langle \boldsymbol{N}_\mathrm{f}|\, \prod_\alpha \mathrm{e}^{\delta t(A_\alpha+B_\alpha)}\, |\boldsymbol{N}_\mathrm{i}\rangle$ by injecting the identity $\mathbb{1} = \sum_\ell |\boldsymbol{N}_\ell\rangle\langle \boldsymbol{N}_\ell|$ between each term of the product. We finally obtain

$$P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i}) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\delta t\mathcal{N}\mathscr{L}(\boldsymbol{\lambda},\boldsymbol{z}_\mathrm{i})} \delta(\dot{\boldsymbol{z}} - \mathfrak{D}\boldsymbol{\lambda}), \quad (4.263)$$

with $\dot{\boldsymbol{z}} \equiv \frac{\boldsymbol{z}_\mathrm{f}-\boldsymbol{z}_\mathrm{i}}{\delta t}$, recovering the continuity equation (4.7), and where we introduced the *detailed Lagrangian*:

$$\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z}) = \sum_\alpha \left[\lambda_\alpha \ln\left(\frac{\lambda_\alpha}{k_\alpha(\boldsymbol{z})}\right) - \lambda_\alpha + k_\alpha(\boldsymbol{z})\right]. \quad (4.264)$$

4.B Some elements of analytical mechanics

The results presented here come mainly from Ref. [265].

4.B.1 Euler-Lagrange equation

We call $\{\varepsilon\}$ the ensemble of trajectories connecting $\boldsymbol{z}_\mathrm{i}$ to $\boldsymbol{z}_\mathrm{f}$ after a time $\mathcal{T}$ and denote by $\frac{\delta \boldsymbol{z}_t^\varepsilon}{\delta\varepsilon}$ the variation of $\boldsymbol{z}_t$ between two trajectories at fixed time t as represented in Fig. 4.15.

The fact that all the trajectories have the same starting and arrival point is expressed by the boundary conditions

$$\frac{\delta \boldsymbol{z}^\varepsilon_{t=0}}{\delta\varepsilon} = \frac{\delta \boldsymbol{z}^\varepsilon_{t=\mathcal{T}}}{\delta\varepsilon} = 0. \tag{4.265}$$

We look for the trajectory ε that minimizes the action $S[\dot{\boldsymbol{z}}, \boldsymbol{z}]_0^{\mathcal{T}}$ defined in Section 4.1.2. Mathematically, it reads

$$\frac{\delta S}{\delta\varepsilon}[\dot{\boldsymbol{z}}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t]_0^{\mathcal{T}} = \int_0^{\mathcal{T}} \left[\frac{\partial L_\gamma}{\partial \boldsymbol{z}}(\dot{\boldsymbol{z}}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t)\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} + \frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}(\dot{\boldsymbol{z}}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t)\frac{\delta \dot{\boldsymbol{z}}^\varepsilon_t}{\delta\varepsilon}\right] \mathrm{d}t = 0, \tag{4.266}$$

where the variation of the action is with respect to an infinitesimal variation of the trajectory $\delta\varepsilon$. It follows

$$\begin{aligned}
\frac{\delta S}{\delta\varepsilon} &= \int_0^{\mathcal{T}} \left[\frac{\partial L_\gamma}{\partial \boldsymbol{z}}\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} + \frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}\right)\right] \mathrm{d}t && (4.267)\\
&= \int_0^{\mathcal{T}} \left[\frac{\partial L_\gamma}{\partial \boldsymbol{z}}\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}\right)\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}\right] \mathrm{d}t + \left[\frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}\right]_0^{\mathcal{T}} && (4.268)\\
&= \int_0^{\mathcal{T}} \left[\frac{\partial L_\gamma}{\partial \boldsymbol{z}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}\right)\right]\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}\mathrm{d}t && (4.269)\\
&= 0, && (4.270)
\end{aligned}$$

where we used in Eq. (4.267) an integration by parts, and in Eq. (4.268) the condition (4.265) to remove the boundary term. We want Eq. (4.269) to be 0 for any $\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}$. The only way to guarantee this is to have at all times

$$\frac{\partial L_\gamma}{\partial \boldsymbol{z}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_\gamma}{\partial \dot{\boldsymbol{z}}}\right) = 0. \tag{4.271}$$

Eq. (4.271) is known as the *Euler-Lagrange equation*. It means that the trajectory minimizing the action, hence the trajectory typically followed by the system, is the one satisfying the Euler-Lagrange equation.

4.B.2 Hamilton's equations

We look for the trajectory ε that minimizes the action $S[\boldsymbol{p}, \boldsymbol{z}]_0^{\mathcal{T}}$ defined in Section 4.1.2. Mathematically, it reads

$$\frac{\delta S}{\delta\varepsilon}[\boldsymbol{p}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t]_0^{\mathcal{T}} = \int_0^{\mathcal{T}} \left[\boldsymbol{p}^\varepsilon_t \cdot \frac{\delta \dot{\boldsymbol{z}}^\varepsilon_t}{\delta\varepsilon} + \frac{\delta \boldsymbol{p}^\varepsilon_t}{\delta\varepsilon}\cdot \dot{\boldsymbol{z}}^\varepsilon_t - \frac{\partial H_\gamma}{\partial \boldsymbol{z}}(\boldsymbol{p}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t)\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} - \frac{\partial H_\gamma}{\partial \boldsymbol{p}}(\boldsymbol{p}^\varepsilon_t, \boldsymbol{z}^\varepsilon_t)\frac{\delta \boldsymbol{p}^\varepsilon_t}{\delta\varepsilon}\right] \mathrm{d}t = 0. \tag{4.272}$$

It follows

$$\begin{aligned}
\frac{\delta S}{\delta\varepsilon} &= \int_0^{\mathcal{T}} \left[\boldsymbol{p}^\varepsilon_t \cdot \frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} + \frac{\delta \boldsymbol{p}^\varepsilon_t}{\delta\varepsilon}\cdot \dot{\boldsymbol{z}}^\varepsilon_t - \frac{\partial H_\gamma}{\partial \boldsymbol{z}}\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon} - \frac{\partial H_\gamma}{\partial \boldsymbol{p}}\frac{\delta \boldsymbol{p}^\varepsilon_t}{\delta\varepsilon}\right] \mathrm{d}t && (4.273)\\
&= \int_0^{\mathcal{T}} \left[\dot{\boldsymbol{z}}^\varepsilon_t - \frac{\partial H_\gamma}{\partial \boldsymbol{p}}\right]\frac{\delta \boldsymbol{p}^\varepsilon_t}{\delta\varepsilon}\mathrm{d}t - \int_0^{\mathcal{T}} \left[\dot{\boldsymbol{p}}^\varepsilon_t + \frac{\partial H_\gamma}{\partial \boldsymbol{z}}\right]\frac{\delta \boldsymbol{z}^\varepsilon_t}{\delta\varepsilon}\mathrm{d}t, && (4.274)\\
&= 0, && (4.275)
\end{aligned}$$

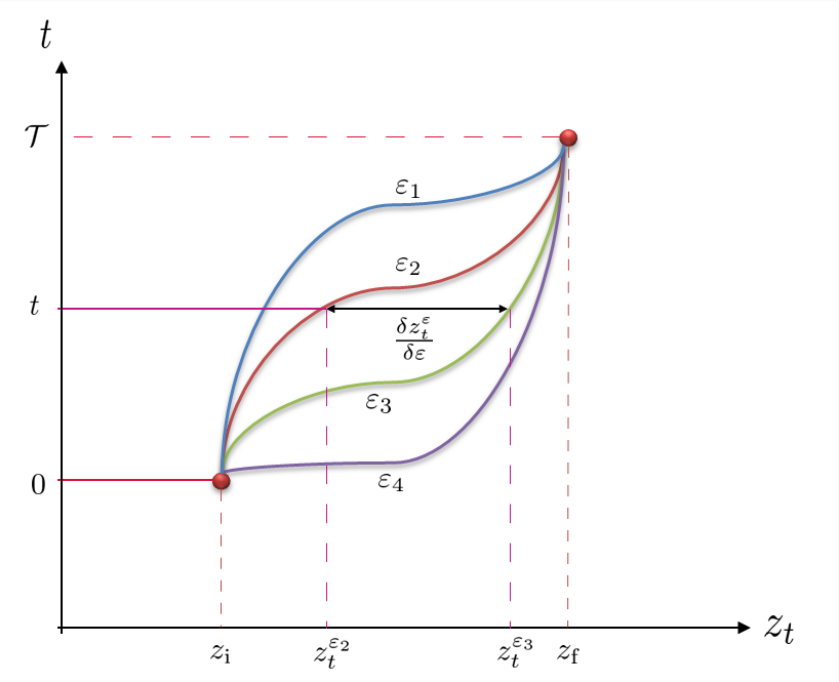

Figure 4.15 – The trajectories connecting $\boldsymbol{z}_{\rm i}$ at time $t = 0$ to $\boldsymbol{z}_{\rm f}$ at time $t = \mathcal{T}$ are here $\{\varepsilon\} = \{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$. A time t, $\frac{\delta \boldsymbol{z}_t^\varepsilon}{\delta \varepsilon}$ represents the variation of $\boldsymbol{z}_t$ between trajectories ε_2 and ε_3.

where we used an integration by parts and Eq. (4.265) to remove the boundary term. We want Eq. (4.274) to be 0 for any $\frac{\delta \boldsymbol{z}_t^\varepsilon}{\delta \varepsilon}$ and $\frac{\delta \boldsymbol{p}_t^\varepsilon}{\delta \varepsilon}$. The only way to guarantee this is to have at all times

$$\begin{cases} \dot{\boldsymbol{z}}_t^\varepsilon = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t^\varepsilon, \boldsymbol{z}_t^\varepsilon), \\ \dot{\boldsymbol{p}}_t^\varepsilon = -\partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t^\varepsilon, \boldsymbol{z}_t^\varepsilon). \end{cases} \tag{4.276}$$

Eq. (4.276) is known as *Hamilton's equations*. It means that the trajectory minimizing the action, hence the trajectory typically followed by the system, is the one satisfying Hamilton's equations.

4.B.3 Canonical transformation

Let us consider some Hamiltonian $H(\boldsymbol{P}, \boldsymbol{Z}, t)$ whose dynamics is ruled by Hamilton's equations:

$$\begin{cases} \dot{\boldsymbol{Z}} = \partial_{\boldsymbol{P}} H(\boldsymbol{P}, \boldsymbol{Z}), \\ \dot{\boldsymbol{P}} = -\partial_{\boldsymbol{Z}} H(\boldsymbol{P}, \boldsymbol{Z}). \end{cases} \tag{4.277}$$

A *canonical transformation* is a change of variable $(\boldsymbol{P}, \boldsymbol{Z}) \to (\boldsymbol{p}(\boldsymbol{P}, \boldsymbol{Z}, t), \boldsymbol{z}(\boldsymbol{P}, \boldsymbol{Z}, t))$ such that there exists a new Hamiltonian $\tilde{H}(\boldsymbol{p}, \boldsymbol{z})$ whose dynamics satisfies Hamilton's equations with respect to these new variables:

$$\begin{cases} \dot{\boldsymbol{z}} = \partial_{\boldsymbol{p}} \tilde{H}(\boldsymbol{p}, \boldsymbol{z}), \\ \dot{\boldsymbol{p}} = -\partial_{\boldsymbol{z}} \tilde{H}(\boldsymbol{p}, \boldsymbol{z}). \end{cases} \tag{4.278}$$

In practice, such a transformation can be obtained from a function $F_1(\boldsymbol{Z}, \boldsymbol{z}, t)$ called a *generating function of type* 1 via the following relations:

$$\boldsymbol{P} = \partial_{\boldsymbol{Z}} F_1, \quad \boldsymbol{p} = -\partial_{\boldsymbol{z}} F_1, \quad \tilde{H}(\boldsymbol{p}, \boldsymbol{z}, t) = H(\boldsymbol{P}, \boldsymbol{Z}, t) + \partial_t F_1. \tag{4.279}$$

It is possible to define three other generating functions from the Legendre transforms of the F_1. For instance, the type-2 generating function $F_2(\boldsymbol{Z}, \boldsymbol{p}, t)$ is defined by

$$F_2(\boldsymbol{Z}, \boldsymbol{p}, t) \equiv F_1(\boldsymbol{Z}, \boldsymbol{z}, t) + \boldsymbol{z} \cdot \boldsymbol{p}, \tag{4.280}$$

leading to the transformation rules:

$$\boldsymbol{P} = \partial_{\boldsymbol{Z}} F_2, \quad \boldsymbol{z} = \partial_{\boldsymbol{p}} F_2, \quad \tilde{H}(\boldsymbol{p}, \boldsymbol{z}, t) = H(\boldsymbol{P}, \boldsymbol{Z}, t) + \partial_t F_2. \tag{4.281}$$

Similarly, one can define the type-3 and type-4 generating functions $F_3(\boldsymbol{P}, \boldsymbol{z})$ and $F_4(\boldsymbol{P}, \boldsymbol{p})$, see [265] for more details.

4.C Independent many-body Markov processes

Nonlinear processes allow modeling a large range of systems. An important example of nonlinear processes is interacting many-body Markov jump [267] and diffusion processes [268] of which independent many-body Markov processes appear as a particular linear case. In this appendix, we illustrate our nonlinear theory developed in Section 4.2 in the case of $\mathcal{N}$ independent jump and diffusion processes.

4.C.1 Independent many-body Markov jump processes

We consider $\mathcal{N}$ independent and identical systems, each one modeled by a time-homogeneous Markov jump process of time-independent generator $\tilde{\boldsymbol{k}}$. We denote by $\nu \in \{1, 2, \dots \mathcal{N}\}$ the ν^{th} system and by $n^\nu \in \{1, \dots M\}$ the state it occupies. The microstate vector $\boldsymbol{n} \equiv \begin{pmatrix} n^1 & \dots & n^\nu & \dots & n^{\mathcal{N}} \end{pmatrix}^\dagger$ denotes the state of the global system and informs on the state of each subsystem. The probability $p_{\boldsymbol{n}}$ that the global system is in state $\boldsymbol{n}$ satisfies the master equation

$$\dot{p}_{\boldsymbol{n}} = \sum_{\boldsymbol{m}} \tilde{k}_{\boldsymbol{n}\boldsymbol{m}} p_{\boldsymbol{m}}, \tag{4.282}$$

where we introduced (with a slight abuse of notation) the transition rate from $\boldsymbol{m}$ to $\boldsymbol{n}$:

$$\tilde{k}_{\boldsymbol{n}\boldsymbol{m}} \equiv \sum_{\nu=1}^{N} \tilde{k}_{n^\nu m^\nu} (1 - \delta_{n^\nu, m^\nu}), \tag{4.283}$$

meaning that the transition $\boldsymbol{m} \to \boldsymbol{n}$ at the level of the global system corresponds to only one transition $m^\nu \to n^\nu$ performed by one of the systems ν. We would like to describe the same system at a more coarse-grained level. To do so, we introduce the mesostate vector $\boldsymbol{N}(\boldsymbol{n}) \equiv \begin{pmatrix} N_1 & \dots & N_n & \dots & N_M \end{pmatrix}^\dagger$ whose component $N_n \equiv \sum_\nu \delta_{n,n^\nu}$ gives the number of systems in state n given the microstate $\boldsymbol{n}$. We are interested in the probability $P_{\boldsymbol{N}}$ that the global system is in state $\boldsymbol{N}$:

$$P_{\boldsymbol{N}} = \sum_{\boldsymbol{n} | \boldsymbol{N}(\boldsymbol{n}) = \boldsymbol{N}} p_{\boldsymbol{n}} \equiv \sum_{\boldsymbol{n}_{\boldsymbol{N}}} p_{\boldsymbol{n}_{\boldsymbol{N}}}, \tag{4.284}$$

where the last sum is over the ensemble $\{\boldsymbol{n}_{\boldsymbol{N}}\}$ of microstates compatible with the mesostate $\boldsymbol{N}$, implying that $p_{\boldsymbol{n}_{\boldsymbol{N}}}$ is the joint probability to be in $\boldsymbol{n}$ and $\boldsymbol{N}$. The master equation satisfied by $P_{\boldsymbol{N}}$ reads then

$$\dot{P}_{\boldsymbol{N}} = \sum_{\boldsymbol{N}'} k_{\boldsymbol{N}\boldsymbol{N}'} P_{\boldsymbol{N}'}, \tag{4.285}$$

where the mesoscopic transition matrix $\boldsymbol{k}$ reads in the Dirac notation:

$$\boldsymbol{k} \equiv \sum_{\boldsymbol{N},\boldsymbol{N}'\neq\boldsymbol{N}} \sum_{n,m} \tilde{k}_{nm} N'_m \delta_{N_n,N'_n+1} \delta_{N_m,N'_m-1} \left|\boldsymbol{N}\right\rangle \left\langle\boldsymbol{N}'\right| - \sum_{\boldsymbol{N}'} \sum_{n,m\neq n} \tilde{k}_{nm} N'_m \left|\boldsymbol{N}'\right\rangle \left\langle\boldsymbol{N}'\right|, \tag{4.286}$$

with $\sum_{\boldsymbol{N}} k_{\boldsymbol{N}\boldsymbol{N}'} = 0$, and where the sum $\sum_{\boldsymbol{N},\boldsymbol{N}'\neq\boldsymbol{N}}$ is implied to run over the mesostates $\boldsymbol{N}$ and $\boldsymbol{N}'$ such that they differ by only one microscopic transition:

$$\begin{cases} N_n = N'_n + 1, \\ N_m = N'_m - 1, \\ N_l = N'_l, \quad \forall l \neq n, m. \end{cases} \tag{4.287}$$

A particular linear case of population processes

Eq. (4.286) means that the transition probability to jump from $\boldsymbol{N}'$ to $\boldsymbol{N}$ is given by the probability of the microscopic transition $m \to n$ implied in the transition $\boldsymbol{N}' \to \boldsymbol{N}$ performed by any of the N'_m systems occupying the state m, and we can write:

$$k_{\boldsymbol{N}\boldsymbol{N}'} = k_{nm,\boldsymbol{N}'} \equiv N'_m \tilde{k}_{nm}. \tag{4.288}$$

Eq. (4.287) can be rewritten as

$$\boldsymbol{N} = \boldsymbol{N}' + \boldsymbol{D}_{nm}, \tag{4.289}$$

where $\boldsymbol{D}_{nm}$ is the vector of component

$$D_{l,nm} = \begin{cases} 1 & \text{if } l = n, \\ -1 & \text{if } l = m, \\ 0 & \text{otherwise,} \end{cases} \tag{4.290}$$

so that the matrix $\boldsymbol{D}$ of components $D_{l,nm}$ is the incidence matrix. Hence, the $\mathcal{N}$ independent Markov processes appear to be a particular linear case of population processes of Section 4.1 with the following connections:

$$\alpha \to (nm), \tag{4.291}$$

$$k_{\alpha,\boldsymbol{N}'} \to k_{nm,\boldsymbol{N}'}, \tag{4.292}$$

$$\mathfrak{D}_\alpha \to \boldsymbol{D}_{nm}, \tag{4.293}$$

and where the observables $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ are respectively the empirical transition current $\boldsymbol{\omega}$ and the empirical density $\boldsymbol{\mu}$ defined by

$$\omega_{lm}(t) \equiv \frac{1}{\mathcal{N}\delta t} \sum_{\nu=1}^{\mathcal{N}} \sum_{s\in[t,t+\delta t[} \delta_{l,n^\nu_{s^+}} \delta_{m,n^\nu_{s^-}} = \frac{1}{\mathcal{N}\mathrm{d}t} \sum_{s\in[t,t+\mathrm{d}t[} \left[\delta_{N_l(s^+),N_l(s^-)+1} \delta_{N_m(s^+),N_m(s^-)-1}\right], \tag{4.294}$$

$$\mu_m(t) \equiv \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \delta_{n^\nu_t,m} = \frac{N_m(t)}{\mathcal{N}}, \tag{4.295}$$

where δt is an infinitesimal time interval, $\boldsymbol{n}_t$ the empirical microstate vector of the global system at time t and $\boldsymbol{N}(t) \equiv \boldsymbol{N}(\boldsymbol{n}_t)$. We assume that the typical time scale during which a single process performs one transition is δt, so that during δt the global system undergoes typically $\mathcal{N}$ transitions. The component $\mathcal{N}\mathrm{d}t\,\omega_{lm}(t)$ counts the number of systems performing the transition $m \to l$ between times t and $t + \delta t$, and the component $\mu_m(t)$ counts the fraction of systems being in state m at time t. From Eqs. (4.294–4.295), the two variables are related by

$$\dot{\boldsymbol{\mu}} = \boldsymbol{D}\boldsymbol{\omega}. \tag{4.296}$$

From Eq. (4.288), the intensive transition rate (4.8) reads for this model

$$k_{nm}(\boldsymbol{\mu}) = \tilde{k}_{nm}\mu_m, \tag{4.297}$$

which is indeed linear in $\boldsymbol{\mu}$. Now, the different Lagrangians and Hamiltonians simply follow from the ones derived in Section 4.1 with the appropriate changes:

$$\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \sum_{n,m\neq n} \left[\omega_{nm} \ln\left(\frac{\omega_{nm}}{k_{nm}\mu_m}\right) - \omega_{nm} + k_{nm}\mu_m\right] - \boldsymbol{\gamma}_1 \cdot \boldsymbol{\omega} - \boldsymbol{\gamma}_2 \cdot \boldsymbol{\mu}, \tag{4.298}$$

$$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{\mu}) = \sum_{n,m\neq n} \tilde{k}_{nm}\mu_m \left[\mathrm{e}^{f_{nm}+\gamma^1_{nm}} - 1\right] + \boldsymbol{\gamma}_2 \cdot \boldsymbol{\mu}, \tag{4.299}$$

$$H_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{\mu}) = \mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{D}^\dagger \boldsymbol{p}, \boldsymbol{\mu}) = \sum_{n,m\neq n} \tilde{k}_{nm}\mu_m \left[\mathrm{e}^{p_n - p_m + \gamma^1_{nm}} - 1\right] + \boldsymbol{\gamma}_2 \cdot \boldsymbol{\mu}, \tag{4.300}$$

with $\boldsymbol{\gamma}_1 \cdot \boldsymbol{\omega} \equiv \sum_{n,m\neq n} \gamma^1_{nm}\omega_{nm}$ and $\boldsymbol{\gamma}_2 \cdot \boldsymbol{\mu} \equiv \sum_n \gamma^2_n \mu_n$. The non-biased Lagrangian and Hamiltonian are recovered when taking $\boldsymbol{\gamma} = 0$. Note that the standard Lagrangian cannot be computed in general for this model.

Rectified Hamiltonian

For general population processes, we saw that the rectified Hamiltonian follows from Eq. (4.143) where W_{s} is related to the function $\boldsymbol{U}$ appearing in the left eigenvector $\mathrm{e}^{\boldsymbol{U}}$ of the biased generator $\boldsymbol{\kappa}$ (4.40) via Eq. (4.81). For our model, the function $\boldsymbol{U}$ is related to the function $\boldsymbol{u}$ appearing in the left eigenvector $\mathrm{e}^{\boldsymbol{u}}$ of the single-process biased matrix $\tilde{\boldsymbol{\kappa}}$ through $U_{\boldsymbol{N}} = \boldsymbol{N} \cdot \boldsymbol{u}$. Indeed, one can check that the dominant left eigenvector of $\boldsymbol{\kappa}$ reads in the Dirac notation

$$\langle \boldsymbol{L}| \equiv \sum_{\boldsymbol{N}} \mathrm{e}^{\boldsymbol{N}\cdot\boldsymbol{u}} \langle \boldsymbol{N}| = \sum_{\boldsymbol{N}} \mathrm{e}^{\sum_{m=1}^{M} N_m u_m} \langle \boldsymbol{N}|, \tag{4.301}$$

with dominant eigenvalue $\Gamma = \mathcal{N}\bar{\Gamma}$, where $\bar{\Gamma}$ is the dominant eigenvalue of $\tilde{\boldsymbol{\kappa}}$. Computing the right-hand side of Eq. (4.81), it follows that W_{s} exists and is given by

$$\partial_{\boldsymbol{\mu}} W_{\mathrm{s}} = \boldsymbol{u}. \tag{4.302}$$

The rectified Hamiltonian reads then

$$H^{\mathrm{r}}_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{\mu}; \boldsymbol{\gamma}) = H_{\boldsymbol{\gamma}}(\boldsymbol{p} + \boldsymbol{u}, \boldsymbol{\mu}) - H_{\boldsymbol{\gamma}}(\boldsymbol{u}, \boldsymbol{\mu}), \tag{4.303}$$

leading to

$$H^{\mathrm{r}}_{\boldsymbol{\gamma}}(\boldsymbol{p}, \boldsymbol{\mu}; \boldsymbol{\gamma}) = \sum_{n,m\neq n} \tilde{K}_{nm}\mu_m \left[\mathrm{e}^{p_n - p_m} - 1\right], \tag{4.304}$$

which corresponds to the Hamiltonian of a non-biased process of intensive rate $\tilde{K}_{nm}(\boldsymbol{\mu}) = \tilde{K}_{nm}\mu_m$, with $\tilde{K}_{nm}$ the component of the generator $\tilde{\boldsymbol{K}}$ of the single-system driven process obtained from the Doob transform of $\tilde{\boldsymbol{\kappa}}$ using its left eigenvector $\mathrm{e}^{\boldsymbol{u}}$ [80]. This illustrates the fact that the rectification of biased Hamiltonians is equivalent to the rectification of biased generators in the linear operator formalism using the Doob transform.

4.C.2 Independent many-body Markov diffusion processes

We consider $\mathcal{N}$ independent and identical systems, each one modeled by a time-homogeneous Markov diffusion process of time-independent drift b and diffusion coefficient σ. We denote by $\nu \in \{1, 2, \ldots \mathcal{N}\}$ the ν^{th} system and by x_t^ν the stochastic process of the system ν which evolves according to the Langevin equation

$$\dot{x}_t^\nu = b(x_t^\nu) + \sigma(x_t^\nu) \circ \xi_t^\nu, \tag{4.305}$$

with ξ_t^ν a Gaussian white noise satisfying Eqs. (2.55–2.56). We are interested in the empirical occupation density

$$\rho(x,t) \equiv \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \delta(x_t^\nu - x), \tag{4.306}$$

and the empirical current

$$j(x,t) \equiv \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \delta(x_t^\nu - x) \circ \dot{x}_t^\nu, \tag{4.307}$$

playing the role of the variables $\boldsymbol{\lambda}$ and $\boldsymbol{z}$, respectively. The empirical occupation density gives the density of systems being at a state in $[x, x + \mathrm{d}x[$ at time t, and the empirical current measures the density of systems performing a displacement between x and $x + \mathrm{d}x$ within the time interval $[t, t + \mathrm{d}t[$. Both variables are related by

$$\dot{\rho}(x,t) = -\nabla j(x,t), \tag{4.308}$$

with $\nabla \equiv \partial_x$ the derivative with respect to x, and where $-\nabla$ plays the role of $\mathcal{D}$. Note that when we consider a single-diffusion process ν, the empirical observables one is interested in are the empirical current $\tilde{j}^\nu$ and the empirical occupancy $\tilde{\rho}^\nu$ defined by

$$\tilde{j}_t^\nu(x) \equiv \frac{1}{t} \int_0^t \mathrm{d}\tau \delta(x_\tau^\nu - x) \circ \dot{x}_\tau^\nu, \tag{4.309}$$

$$\tilde{\rho}_t^\nu(x) \equiv \frac{1}{t} \int_0^t d\tau \delta(x_\tau^\nu - x). \tag{4.310}$$

The observables of the single-process and of the many-body process are related by

$$\begin{aligned} \frac{1}{t} \int_0^t \mathrm{d}\tau\, j(\tau) &= \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \tilde{j}_t^\nu, \\ \frac{1}{t} \int_0^t \mathrm{d}\tau\, \rho(\tau) &= \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \tilde{\rho}_t^\nu. \end{aligned} \tag{4.311}$$

Stochastic equation for the empirical occupation density

We aim to give a coarse-grained description of the global system by deriving the stochastic equation ruling ρ. To do so, we compute the quantity: $\Delta x^\nu \equiv x^\nu_{t+\Delta t} - x^\nu_t$. Using the Langevin equation (4.305) for x^ν_t, we get:

$$\Delta x^\nu = \int_t^{t+\Delta t} \mathrm{d}\tau \left[b(x^\nu_\tau) + \sigma(x^\nu_\tau) \circ \xi^\nu_\tau \right]. \tag{4.312}$$

For an infinitesimal Δt, we have in the Stratonovich convention [222]:

$$\int_t^{t+\Delta t} \mathrm{d}\tau \, b(x^\nu_\tau) \simeq b\left(x^\nu_t + \frac{\Delta x^\nu}{2} \right) \Delta t \tag{4.313}$$

$$\int_t^{t+\Delta t} \mathrm{d}\tau \, \left[\sigma(x^\nu_\tau) \xi^\nu_\tau \right] \simeq \sigma\left(x^\nu_t + \frac{\Delta x^\nu}{2} \right) \int_t^{t+\Delta t} \mathrm{d}\tau \, \xi^\nu_\tau. \tag{4.314}$$

Expanding up to order Δt and using the identity $(\int_t^{t+\Delta t} \xi_t \mathrm{d}t)^2 = \Delta t$ when $\Delta t \to 0$ [28,125], it follows

$$\Delta x^\nu \simeq b(x^\nu_t)\Delta t + \sigma(x^\nu_t) \int_t^{t+\Delta t} \xi^\nu_\tau \mathrm{d}\tau + \frac{1}{2}\sigma(x^\nu_t)\nabla\sigma(x^\nu_t)\Delta t, \tag{4.315}$$

$$(\Delta x^\nu)^2 \simeq \sigma(x^\nu_t)^2 \Delta t. \tag{4.316}$$

We now compute $\rho(x, t+\Delta t) = \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \delta(x^\nu_{t+\Delta t} - x)$. Let φ be a test function, then

$$\int \mathrm{d}x \varphi(x) \rho(x, t+\Delta t) = \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \varphi(x^\nu_{t+\Delta t}) \tag{4.317}$$

$$= \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \varphi(x^\nu_t + \Delta x^\nu) \tag{4.318}$$

$$\simeq \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \varphi(x^\nu_t) + \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \Delta x^\nu \varphi'(x^\nu_t) + \frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \frac{1}{2} (\Delta x^\nu)^2 \varphi''(x^\nu_t), \tag{4.319}$$

where we used Taylor's formula around x^ν_t up to second order in Δx^ν in the last equation. Using Eqs. (4.315–4.316) and the fact that $\frac{1}{\mathcal{N}} \sum_{\nu=1}^{\mathcal{N}} \varphi(x^\nu_t) = \int \mathrm{d}x \varphi(x) \rho(x,t)$, Eq. (4.319) gives

$$\int \mathrm{d}x \varphi(x) \dot{\rho}(x,t) = \int \mathrm{d}x \varphi(x) \left\{ -\nabla \left[\hat{b}(x)\rho(x,t) - \frac{1}{2}\sigma(x)^2 \nabla\rho(x,t) + \sigma(x)\sqrt{\frac{\rho(x,t)}{\mathcal{N}}}\eta(x,t) \right] \right\}, \tag{4.320}$$

with $\dot{\rho}(x,t) = \lim_{\Delta t \to 0} \frac{\rho(x,t+\Delta t) - \rho(x,t)}{\Delta t}$, $\hat{b}(x) = b(x) - \frac{1}{2}\sigma(x)\nabla\sigma(x)$, and where we introduced

$$\eta(x,t) \equiv \frac{1}{\sqrt{\mathcal{N}\,\rho(x,t)}} \sum_{\nu=1}^{\mathcal{N}} \delta(x - x^\nu_t) \bar{\xi}^\nu_t, \tag{4.321}$$

with $\bar{\xi}_t^\nu \equiv \lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_t^{t+\Delta t} \mathrm{d}\tau \xi_\tau^\nu$. The stochastic process η is a Gaussian white noise in time and space [268, 269] with mean and variance

$$\langle \eta(x,t)\rangle = 0, \tag{4.322}$$
$$\langle \eta(x,t)\eta(x',t')\rangle = \delta(x-x')\delta(t-t'). \tag{4.323}$$

Since (4.320) is valid for any function φ, we obtain the stochastic equation for ρ:

$$\dot{\rho}(x,t) = -\nabla\left[\hat{b}(x)\rho(x,t) - \frac{1}{2}\sigma(x)^2\nabla\rho(x,t) + \sigma(x)\sqrt{\frac{\rho(x,t)}{\mathcal{N}}}\eta(x,t)\right]. \tag{4.324}$$

Eq. (4.324) is known as the *Dean equation*.

Derivation of the Lagrangians and Hamiltonians

In order to obtain the detailed Lagrangian $\mathscr{L}(j,\rho)$, we compute the conditional probability $P_{\delta t}(j \mid \rho)$ using Eqs. (4.308, 4.324):

$$P_{\delta t}(j\mid\rho) = \prod_x \left\langle \delta\left[j - \hat{b}\rho + \frac{1}{2}\sigma^2\nabla\rho - \sigma\sqrt{\frac{\rho}{\mathcal{N}}}\eta\right]\right\rangle_\eta, \tag{4.325}$$

The continuous product $\prod_x$ is implied to run over the states $x_\ell \equiv \ell\delta x$, with ℓ an integer and δx an infinitesimal space step, and $\langle\cdots\rangle_\eta$ is the average over the noise η:

$$\langle \mathcal{O}(\eta)\rangle_\eta \equiv \frac{1}{\mathsf{N}}\int \mathcal{O}(\eta)\mathrm{e}^{-\frac{1}{2}\delta t\delta x\eta^2}\mathrm{d}\eta, \tag{4.326}$$

with N the normalization factor, $O(\eta)$ an arbitrary function of η. We have dropped the (x,t)-dependency in the all functions for clarity. It follows

$$\begin{aligned} P_{\delta t}(j\mid\rho) &= \frac{1}{\mathsf{N}}\int \mathrm{d}\eta \prod_x \mathrm{e}^{-\frac{1}{2}\delta t\delta x\eta^2}\delta\left[j - \hat{b}\rho + \frac{1}{2}\sigma^2\nabla\rho - \sigma\sqrt{\frac{\rho}{\mathcal{N}}}\eta\right] \\ &= \frac{1}{\mathsf{N}}\int \mathrm{d}\eta \prod_x \mathrm{e}^{-\frac{1}{2}\delta t\delta x\eta^2}\frac{\sqrt{\mathcal{N}}}{\sigma\sqrt{\rho}}\delta\left[\eta - \frac{j-\hat{b}\rho+\frac{1}{2}\sigma^2\nabla\rho}{\frac{1}{\sqrt{\mathcal{N}}}\sigma\sqrt{\rho}}\right] \\ &= \frac{1}{\mathsf{N}}\mathrm{e}^{-\delta t\int_x\left[\frac{\mathcal{N}}{2\sigma^2\rho}\left(j-\hat{b}\rho+\frac{1}{2}\sigma^2\nabla\rho\right)^2+\frac{1}{2}\ln\left(\frac{\mathcal{N}}{\rho\sigma^2}\right)\right]}, \end{aligned} \tag{4.327}$$

where we used in the second equality the relation $\delta(\varphi(y)) = |\varphi'(y_0)|^{-1}\delta(y-y_0)$, for any smooth test function φ and any root y_0 of φ. In the limit of large $\mathcal{N}$, the last term in the exponential is asymptotically dominated by $\mathcal{N}$ and we obtain

$$P_{\delta t}(j\mid\rho) \underset{\mathcal{N}\to\infty}{\asymp} \mathrm{e}^{-\mathcal{N}\delta t\mathscr{L}(j,\rho)}, \tag{4.328}$$

where the Lagrangian is given by

$$\mathscr{L}(j,\rho) = \int_x \frac{1}{2\sigma^2\rho}\left(j - J^\rho\right)^2, \tag{4.329}$$

with $J^\rho = \hat{b}\rho - \frac{1}{2}\sigma^2\nabla\rho$. Computing the Legendre transform of $\mathscr{L}$ with respect to j yields the detailed Hamiltonian

$$\mathscr{H}(f,\rho) = \int_x f\left[\frac{1}{2}\sigma^2 f\rho + J^\rho\right]. \tag{4.330}$$

We are interested in the observable

$$\boldsymbol{A}_t(x) \equiv \frac{\mathcal{N}}{t}\begin{pmatrix} \int_0^t \mathrm{d}\tau j(x,\tau) \\ \int_0^t \mathrm{d}\tau \rho(x,\tau) \end{pmatrix}. \tag{4.331}$$

Using the results of Section 4.2.1, the dynamical fluctuations of $\boldsymbol{A}$ are encoded in the biased Lagrangian and Hamiltonian

$$\mathscr{L}_{\boldsymbol{\gamma}}(j,\rho) = \mathscr{L}(j,\rho) - \gamma_1\cdot j - \gamma_2\cdot\rho, \tag{4.332}$$

$$\mathscr{H}_{\boldsymbol{\gamma}}(f,\rho) = \int_x (f+\gamma_1)\left[\frac{1}{2}\sigma^2(f+\gamma_1)\rho + J^\rho\right] + \gamma_2\cdot\rho = \mathscr{H}(f+\gamma_1,\rho) + \gamma_2\rho, \tag{4.333}$$

where $\gamma_1(x)$ (resp. $\gamma_2(x)$) is conjugated to the first (resp. second) component of $\boldsymbol{A}(x)$.

SCGF and Hamilton-Jacobi equation

In order to derive the rectified Hamiltonian, we first need to translate the spectral properties of the biased generator from the linear operator formalism to the Hamiltonian formalism. Because of the independence of the $\mathcal{N}$ processes, it suffices to look at a single process. Indeed, we can relate the SCGF Γ of the global system to the SCGF $\bar{\Gamma}$ of the single process by:

$$\Gamma = \lim_{t\to\infty}\frac{1}{t}\ln\mathbb{E}_{x_\mathrm{i}}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\right] \tag{4.334}$$

$$= \lim_{t\to\infty}\frac{1}{t}\ln\mathbb{E}_{x_\mathrm{i}}\left[\mathrm{e}^{\mathcal{N}\int_0^t \mathrm{d}\tau\gamma_1\cdot j(\tau) + \mathcal{N}\int_0^t \mathrm{d}\tau\gamma_2\cdot\rho(\tau)}\right] \tag{4.335}$$

$$= \lim_{t\to\infty}\frac{1}{t}\ln\mathbb{E}_{x_\mathrm{i}}\left[\mathrm{e}^{t\sum_{\nu=1}^{\mathcal{N}}\left(\gamma_1\cdot\tilde{j}_t^\nu + \gamma_2\cdot\tilde{\rho}_t^\nu\right)}\right] \tag{4.336}$$

$$= \mathcal{N}\lim_{t\to\infty}\frac{1}{t}\ln\mathbb{E}_{x_\mathrm{i}}\left[\mathrm{e}^{t\left(\gamma_1\cdot\tilde{j}_t^\nu + \gamma_2\cdot\tilde{\rho}_t^\nu\right)}\right] \tag{4.337}$$

$$= \mathcal{N}\bar{\Gamma}, \tag{4.338}$$

where we used Eq. (4.331) in Eq. (4.335), Eq. (4.311) in Eq. (4.336) and the fact that the $\mathcal{N}$ processes are independent and identically distributed in Eq. (4.337). We know from the linear operator formalism that the single-process driven generator is obtained by taking the Doob transform of the biased generator using its dominant left eigenfunction [80]. The biased Fokker-Planck operator is given by:

$$\Lambda_{\boldsymbol{\gamma}}\varphi = (-\nabla+\gamma_1)(\hat{b}\varphi) + \frac{1}{2}(-\nabla+\gamma_1)\left[\sigma^2(-\nabla+\gamma_1)\varphi\right] + \gamma_2\varphi, \tag{4.339}$$

with φ a test function, and its adjoint operator by

$$\Lambda_{\boldsymbol{\gamma}}^\dagger\varphi = \hat{b}(\nabla+\gamma_1)\varphi + \frac{1}{2}(\nabla+\gamma_1)\left[\sigma^2(\nabla+\gamma_1)\varphi\right] + \gamma_2\varphi. \tag{4.340}$$

We also know that the SCGF $\bar{\Gamma}$ is the dominant eigenvalue of $\Lambda_{\boldsymbol{\gamma}}$. Let e^{u} be its associated dominant left eigenfunction (the positivity is ensured by the Krein-Rutman theorem):

$$\Lambda_{\boldsymbol{\gamma}}^{\dagger}\mathrm{e}^{u} = \bar{\Gamma}\mathrm{e}^{u}. \tag{4.341}$$

From Eq. (4.341) and the fact that $\rho \cdot \mathbf{1} = 0$, we have

$$\bar{\Gamma} = \int_{x} \mathrm{e}^{-u}(\Lambda_{\boldsymbol{\gamma}}^{\dagger}\mathrm{e}^{u})\rho. \tag{4.342}$$

Computing explicitly the right-hand-side of (4.342), we find

$$\bar{\Gamma} = \mathscr{H}_{\boldsymbol{\gamma}}(f = \nabla u, \rho), \tag{4.343}$$

with $\nabla = (-\nabla)^{\dagger}$. As expected, the function u appearing in the left eigenfunction of the biased Fokker-Planck operator is the solution of Hamilton-Jacobi equation and we write $u = \partial_{\rho} W_{\mathrm{s}}$.

Rectified Hamiltonian

The rectified Hamiltonian follows from Eq. (4.144):

$$\mathscr{H}^{r}(f, \rho; \boldsymbol{\gamma}) = \mathscr{H}_{\boldsymbol{\gamma}}(f + \nabla u, \rho) - \mathscr{H}_{\boldsymbol{\gamma}}(\nabla u, \rho), \tag{4.344}$$

leading after explicit computation to

$$\mathscr{H}^{r}(f, \rho; \boldsymbol{\gamma}) = \int_{x} f \left[\frac{1}{2}\sigma^{2} f\rho + J_{\boldsymbol{\gamma}}^{\mathrm{r},\rho}\right], \tag{4.345}$$

with $J_{\boldsymbol{\gamma}}^{\mathrm{r},\rho} \equiv \hat{B}_{\boldsymbol{\gamma}}\rho - \frac{1}{2}\sigma^{2}\nabla\rho$, where we introduced the rectified drift $\hat{B}_{\boldsymbol{\gamma}} \equiv \hat{b} + \sigma^{2}(\nabla u + \gamma_{1})$. Unsurprisingly, the rectified Hamiltonian corresponds to a non-biased Hamiltonian associated with the drift $\hat{B}_{\boldsymbol{\gamma}}$ of the driven process obtained from a Doob transform as seen in Eq. (3.154). This illustrates the fact that the rectification of biased Hamiltonians is equivalent to the rectification of biased generators in the linear operator formalism using the Doob transform.

4.D Phase transition in the Brownian Donkey

We saw in Section 4.2.2 that the positions of the fixed points correspond to the positions of the extrema of H_{min} defined in Eq. (4.127). Moreover, the position of the fixed point associated with the dominant trajectory in the long-time limit corresponds to the position of the absolute maxima of H_{min} as discussed in Section 4.2.3. When varying the biasing parameter $\boldsymbol{\gamma}$, the aspect of $H_{\mathrm{min}}(z)$ evolves until it undergoes a phase transition at the critical parameter $\boldsymbol{\gamma}^{c}$ for which the absolute maximum is reached at two different values of z. This is illustrated in Fig. 4.16 in which we plotted the non-biased Hamiltonian and the biased Hamiltonian with the parameter $\boldsymbol{\gamma}_{1}$ fixed at its critical value $\boldsymbol{\gamma}_{1}^{c}$ while varying the parameter γ_{2}. The non-biased Hamiltonian (dashed black curve) has one minimum and two maxima $H_{\mathrm{min}}(z_{0}^{\mathrm{NB}}) > H_{\mathrm{min}}(z_{1}^{\mathrm{NB}})$ reached at z_{0}^{NB} and z_{1}^{NB}. The biased Hamiltonian displays the same extrema: one minimum and two maxima of positions $z_{0}^{\star}$ and $z_{1}^{\star}$, with $z_{0}^{\star}$

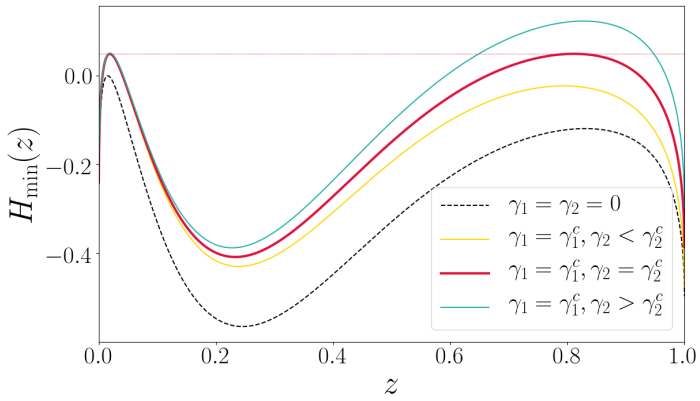

Figure 4.16 – $H_{\min}(z)$ for different values of γ_2. The dashed black curve corresponds to the non-biased Hamiltonian and the red curve corresponds to the value of the biased Hamiltonian at the critical biasing parameter $\gamma_{1,\epsilon}^{\nu,c} = 0.084$ ($\epsilon = \pm, \nu = 1, 2$) and $\gamma_2^c = 0.11$. The figures are obtained for $E = 0.8$, $N = 1000$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$.

the position of the absolute maximum: $H_{\min}(z_0^\star) > H_{\min}(z_1^\star)$. When considering values of γ_2 smaller than the critical value γ_2^c, the positions of the maxima follow the same order as for the non-biased Hamiltonian, i.e. $z_0^\star < z_1^\star$ (yellow curve). On the contrary, for values of γ_2 greater than the critical value γ_2^c, the maxima "exchange positions" and we have $z_0^\star > z_1^\star$ (blue curve). In between, at the critical value $\gamma_2 = \gamma_2^c$, the two maxima coincide, i.e. $H_{\min}(z_0^\star) = H_{\min}(z_1^\star)$ (red curve). At this value of $\boldsymbol{\gamma}$, the SCGF is not differentiable and the system undergoes a first-order phase transition.

At the phase transition, there are two fixed points associated with the dominant trajectory in the long-time limit (red trajectories in Fig. 4.17). We denote by $z_0'^\star$ (left) and $z_0^\star$ (right) their positions. The solutions $p_\pm$ (4.187) of the HJ equation (4.186) have the following stability properties:

- For $z \in [0, z_0'^\star]$:

$$\begin{cases} \text{The branch } p_+ \text{ is stable (converges to } z_0'^\star), \\ \text{The branch } p_- \text{ is unstable (exits from } z_0'^\star). \end{cases} \tag{4.346}$$

- For $z \in [z_0^\star, 1]$:

$$\begin{cases} \text{The branch } p_+ \text{ is unstable (exits from } z_0^\star), \\ \text{The branch } p_- \text{ is stable (converges to } z_0^\star). \end{cases} \tag{4.347}$$

- For $z \in [z_0'^\star, z_0^\star]$:

$$\begin{cases} \text{The branch } p_+ \text{ is unstable for } z_0'^\star \text{ and stable for } z_0^\star, \\ \text{The branch } p_- \text{ is stable for } z_0'^\star \text{ and unstable for } z_0^\star. \end{cases} \tag{4.348}$$

In order to define the rectified Hamiltonian, we need to find a globally stable solution $p_{\rm s}$ of the HJ equation. Yet, in view of Eqs. (4.346–4.348) and Fig. 4.17, such a solution does not exist. Consequently, the rectification procedure is not well-defined at the phase transition.

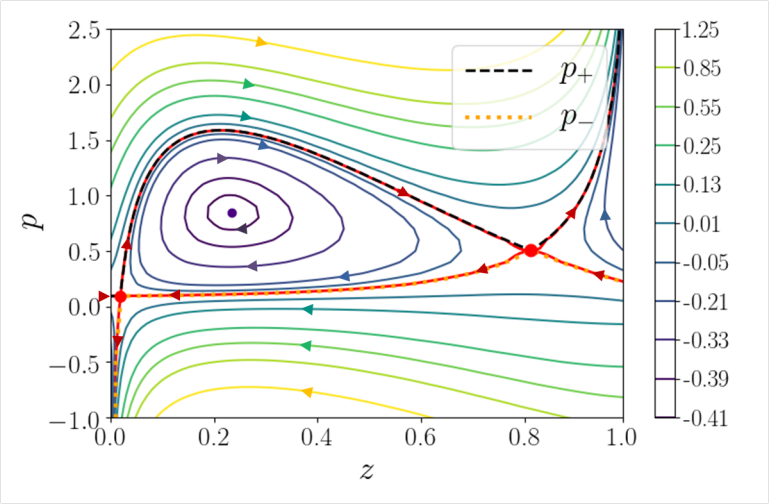

Figure 4.17 – Trajectories of the biased Hamiltonian at the phase transition. The dashed black line corresponds to the solution p_+ and the gold dotted line corresponds to the solution p_-.
The figures are obtained for $\mathcal{E} = 0.8$, $V = 2$, $E_a = 1$, $F = 1$, $\beta_1 = 1$, $\beta_2 = 2$, $\boldsymbol{\gamma}^{\nu}_{1,\epsilon} = 0.084$ $(\epsilon = \pm, \nu = 1, 2)$ and $\gamma_2 = 0.11$.

General conclusion

"Everything is theoretically impossible, until it is done."

Robert A. Heinlein

The development of non-equilibrium statistical physics in general, and the exploration of the problem of conditioning Markov processes in particular, are still in progress and many discoveries await to be made.

The latter problem seeks to find an equivalent process — the driven process — to a Markov process conditioned on a rare fluctuation of a physical observable. The generator of such a process has been derived for general time-homogeneous Markov processes in Refs. [52, 80, 81]. The main results of this manuscript are the extension of the problem of conditioning in the linear operator formalism to periodically driven Markov processes on the one hand, and the translation of this problem in the Lagrangian/Hamiltonian formalism [124] used to deal with nonlinear Markov processes on the other hand. In the first extension, we deal with time-periodic generators of jump and diffusion processes. The spectral relations of the time-homogeneous case become then initial or final value problems. In the second extension, generators become Lagrangians or Hamiltonians used to describe the dynamical large deviations in size of the system. The Lagrangian/Hamiltonian formalism allows one to reduce the number of dynamical observables when taking the large-size limit. This is useful as computing the spectral properties of large-dimension linear operators can quickly become laborious. Instead, one "simply" looks for the solutions of a Hamilton-Jacobi equation.

The conditioning, biasing and rectification of Markov processes in the linear operator formalism is rather complete as it concerns jump and overdamped diffusion processes, both in the time-homogeneous and time-periodic cases. Even so, it remains to develop the same theory for the case of underdamped diffusion processes. An underdamped diffusion process of dimension d is an overdamped diffusion process of dimension $2d$ but with a diffusion matrix $\boldsymbol{D} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\dagger$ that is not invertible, implying that the level 2.5 large deviation function is not defined. In Ref. [270], it has been shown that the latter LDF is replaced by the level 2 LDF of the empirical density. On this basis, one can question how the results of Refs. [52, 80, 81] can be adapted to the case of underdamped diffusion processes.

Concerning the conditioning, biasing and rectification in the Lagrangian/Hamiltonian formalism, there is still a lot to understand. To begin with, we hope that a proof of our conjecture on the nonlinear counterpart of the Perron-Frobenius theorem as well as for our results on the long-time limit dynamics in the presence of more than one critical manifold will be possible in future work. Afterward, it would be interesting to extend this

theory to the case of periodically driven nonlinear Markov processes on the basis of the derivation done in Section 4.4. Additionally, we touched on the fact that the rectification may fail at a first-order phase transition due to the non-existence of globally stable and unstable solutions of the Hamilton-Jacobi equation. This is equivalent to the failure of the Perron-Frobenius theorem in the linear operator formalism due to the breaking of ergodicity appearing in the continuous limit [120, 124, 271]. Further investigations are needed to understand why the rectification is not defined in the presence of critical phenomena, and what is the physical interpretation of the associated first-order phase transition.

The work provided in this thesis remains at a rather mathematical level, but we hope that it will find applications on more physical models. In particular, the results presented in this manuscript can be applied to stochastic thermodynamics by considering thermodynamic observables (heat, work, chemical current, entropy production, etc.) and transition rates related via local or instantaneous detailed balance condition to ensure thermodynamic consistency. One can wonder under what conditions the driven process models a physical system and preserves the thermodynamic consistency of the original process. An open question would be whether it is possible to show experimentally the equivalence between a system modeled by a process conditioned on always having the same value of a physical observable and a system modeled by the driven process. The latter would be obtained either by modifying the former system, or by modifying the external conditions of the same system in order to comply with the driven generator. One would then look at the distribution of the conditioning observable and check if its typical value corresponds to the imposed value.

To go further, one can try to approach the problem of conditioning from a quantum perspective. *Non-equilibrium quantum statistical mechanics* is an active field of research and a lot of effort has been put into the development of tools allowing a statistical mechanical description of quantum systems driven out of equilibrium. One approach is to make use of the *density operator formalism* and the *full counting statistics* to describes the quantum state of the system and to derive the statistical properties of physical observables [272–274]. Lately, large deviation theory has been applied to quantum systems [102, 275–278]. A fortiori, a quantum level-2.5 large deviation principle has been derived for open quantum systems [279]. Subsequently, it would be interesting to investigate how the problem of conditioning generalizes to nonequilibrium quantum systems.

I hope this manuscript will be useful to students, PhD students or anyone who wants to learn about (or continue) the work presented here. I wanted to make a pedagogical and gradual presentation of the different concepts and results of this thesis, and I hope it was the case. Lastly, I hope that the end of this manuscript will only be the beginning of other scientific adventures for other budding scientists.

Glossary

Acronyms

CGF	Cumulant Generating Function
HJ	Hamilton-Jacobi
LDF	Large Deviation Function
LDP	Large Deviation Principle
LF	Legendre-Fenchel
PF	Perron-Frobenius
SCGF	Scaled Cumulant Generating Function
TiPS	Time Periodic State

Dictionary

Critical manifold	Ensemble of compact trajectories such that at least one other trajectory converges towards it forward or backward in time.
Eigenrate	Value of the Hamiltonian at a solution of the HJ equation.
Fixed point	Stationary solution of Hamilton's equation.
Global solution	Solution of the HJ equation defined and analytic for all $\boldsymbol{z}$.
Globally stable dynamics	Reduced dynamics such that there exists a compact set C in $\boldsymbol{z}$-space such that all trajectories converge to C.
Globally stable solution	Global solution of the HJ equation for which the reduced dynamics is globally stable.
Globally unstable dynamics	Reduced dynamics such that there exists a compact set C in $\boldsymbol{z}$-space such that all trajectories exit from C.
Globally unstable solution	Global unstable solution of the HJ equation for which the reduced dynamics is globally stable.

Limit cycle	Closed critical manifold of dimension 1.
Linear operator formalism	Formalism that uses Markov generators (large deviations in time)
Reduced dynamics	Trajectories in $\boldsymbol{z}$-space solution of Hamilton's equations in the manifold $\boldsymbol{p} = \partial_{\boldsymbol{z}} W$, with W a global solution of the HJ equation.
Strange attractor	Complex geometric critical manifolds.
Transient trajectory	Trajectory that converges to a critical manifold forward (relaxing transient trajectory) or backward (fluctuating transient trajectory) in time.

Notations

We sum up in the following the mathematical definitions, notations and symbols used in the whole manuscript.

A. Mathematical symbols

$\dot{X} = \frac{\mathrm{d}X}{\mathrm{d}t}$	Time derivative of X	$\frac{\partial X}{\partial t}$	Partial derivative with respect to time
$\mathbb{1}$	Identity operator	$\mathbf{1}$	Vector of components 1 or function $\mathbf{1} : x \mapsto 1$, $\forall x$
$\equiv$	Equal by definition	$\circ$	Stratonovich stochastic integral
$\mathbb{R}$	Set of real numbers	$\mathbb{N}$	Set of natural numbers

Asymptotic equivalence $\underset{n\to\infty}{\asymp}$ $\quad \left(A_n \underset{n\to\infty}{\asymp} B_n\right) \equiv \left(\lim_{n\to\infty} \frac{A_n}{B_n} = 0\right)$

Scalar product $\cdot$

Vectors	$\boldsymbol{u} \cdot \boldsymbol{v} \equiv \sum_x u_x v_x$
Functions	$f \cdot \varphi \equiv \int \mathrm{d}x f(x)\varphi(x)$

Hadamard product $\odot$

Vectors	$(\boldsymbol{u} \odot \boldsymbol{v})_x \equiv u_x v_x$
Matrix and vector	$(\boldsymbol{M} \odot \boldsymbol{v})_{xy} \equiv M_{xy} v_y$

Adjoint operator $\dagger$ For an operator M acting on a Hilbert space $\mathbb{H}$, $M^\dagger$ is the adjoint of M and is defined by $(MX) \cdot Y \equiv X \cdot (M^\dagger Y)$, $\forall X, Y \in \mathbb{H}$. If M is a real matrix, $M^\dagger$ is the transpose of M.

Time-ordered exponentials

$\overleftarrow{\mathcal{Q}}_{\boldsymbol{M}}(t,0) \equiv \overleftarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$	Ordered exponential of $\boldsymbol{M}$/Propagator	Eq. (3.196)
$\overrightarrow{\mathcal{Q}}_{\boldsymbol{M}}(0,t) \equiv \overrightarrow{\exp} \int_0^t \boldsymbol{M}(t')\,\mathrm{d}t'$	Reverse-ordered exponential of $\boldsymbol{M}$	Eq. (3.199)

B. Parameters

Chapter 3

T	Period
n	Total number of periods
$t = 0$	Initial time
$t = nT$	Final time

Chapter 4

$\mathcal{N}$	Size-type parameter
$t = 0$	Initial time
$t = \mathcal{T}$	Final time

C. Markov jump processes (Sections 2.3 and 3.1)

Generators and probabilities

x, y	States	
$P(x,t \mid y,t')$	Transition probability from y at time t' to x at time t	
$\boldsymbol{k} = (k_{xy})$	Transition rate matrix/Markov generator	(2.5)
$\boldsymbol{\pi} = (\pi_x)$	Probability vector associated with $\boldsymbol{k}$	(2.6)
$\boldsymbol{\pi}(0)$	Initial probability vector	
$\boldsymbol{\pi}^{\mathrm{TiPS}}$	TiPS probability of the process generated by $\boldsymbol{k}$	
$\boldsymbol{j}^{\boldsymbol{\pi}} = (j^{\boldsymbol{\pi}}_{xy})$	probability current associated with $\boldsymbol{\pi}$	(2.8)
$\boldsymbol{\kappa} = (\kappa_{xy})$	Tilted (or biased) generator	(2.49)
$\boldsymbol{M}^{\boldsymbol{v}}$	Generalized Doob transform of the matrix $\boldsymbol{M}$ using the vector $\boldsymbol{v}$	(3.23)
$\boldsymbol{\mathcal{K}}^n = (\mathcal{K}^n_{xy})$	Canonical generator	(3.28)
$\boldsymbol{K} = (K_{xy})$	Driven generator	(3.41)
$\boldsymbol{\mu} = (\mu_x)$	TiPS probability of the driven process	(3.48)

Path probabilities

$z(t)$ Markov process giving the system state at time t $\quad$ $[z]$ Path or trajectory

$\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$	Path probability associated with the Markov generator $\boldsymbol{k}$	(2.13)
$\mathbb{P}_{\boldsymbol{\kappa},\boldsymbol{\pi}(0)}[z]$	Biased path probability	(2.51)
$\mathbb{P}^{\mathrm{micro}}_{\mathbf{a},\boldsymbol{\pi}(0)}[z]$	Microcanonical path probability	(2.88)
$\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\boldsymbol{\pi}(0)}[z]$	Canonical path probability	(2.89)
$\mathbb{P}_{\boldsymbol{M}^{\boldsymbol{v}},\boldsymbol{\pi}(0)}[z]$	Path probability associated with the Doob transform $\boldsymbol{M}^{\boldsymbol{v}}$	(3.25)
$\mathbb{P}_{\boldsymbol{\kappa}^{\boldsymbol{C}^n},\frac{\boldsymbol{C}^n(0)\odot\boldsymbol{\pi}(0)}{\boldsymbol{\pi}(0)\cdot\boldsymbol{C}^n(0)}}[z]$	Path probability associated with the canonical generator $\boldsymbol{\mathcal{K}}^n$	(3.35)
$\mathbb{P}_{\boldsymbol{K},\boldsymbol{\pi}(0)}[z]$	Path probability associated with the driven generator $\boldsymbol{K}$	(3.53)

Averages

$\mathbb{E}_{\boldsymbol{\pi}(0)}[\cdots]$	Path average with respect to $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$
$\mathbb{E}_{\boldsymbol{z}_0}[\cdots]$	Path average with respect to $\mathbb{P}_{\boldsymbol{k},\boldsymbol{\pi}(0)}[z]$ with initial probability $\boldsymbol{\pi}_x(0) = \delta_{x,\boldsymbol{z}_0}, \forall x$

Observables

$\boldsymbol{g} = (g_{xy})$, $\boldsymbol{h} = (h_x)$	Periodic functions	
$\boldsymbol{A}_t[z] \equiv \begin{pmatrix} \frac{1}{t}\sum_{i=0}^{M-1} g_{z_{i+1},z_i}(t_{i+1}) \\ \frac{1}{t}\int_0^t \mathrm{d}\tau h_{z(\tau)}(\tau), \end{pmatrix}$	General observable	(2.14)
$\boldsymbol{p}^n[z] = (p^n_x[z])$	Empirical occupation	(3.59)
$\boldsymbol{\omega}^n[z] = (\omega^n_{xy}[z])$	Empirical transition probability	(3.63)

Generating functions

$\boldsymbol{\gamma} = (\gamma_1 \;\; \gamma_2)$	Conjugate variable of $\boldsymbol{A}_t[z]$	
$G(t,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\right]$	Moment generating function	(2.40)
$\tilde{G}_x(t,\boldsymbol{\gamma}) = \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{t\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\delta_{x,z(t)}\right]$	Moment generating function conditioned on the final state	(2.43)
$\tilde{G}(x,t \mid y,t')$	Biased transition probability	(2.50)
$\Gamma(\boldsymbol{\gamma}) = \lim_{n\to\infty}\frac{1}{nT}\ln \mathbb{E}_{\boldsymbol{\pi}(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}}\right]$	Scaled cumulant generating function	(3.3)

Large deviation functions

$I_{2.5}(\boldsymbol{\omega},\boldsymbol{p})$	Level 2.5 large deviation function	(3.67)
$I(\boldsymbol{a})$	Large deviation function of the observable $\boldsymbol{A}$	(3.69)

Spectral elements and functions

$\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$	One-period propagator	
χ_T	Dominant eigenvalue of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$	
$\boldsymbol{r}_T$	Dominant right eigenvector of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$	(3.10)
$\boldsymbol{l}_T$	Dominant left eigenvector of $\overleftarrow{\mathcal{Q}}_{\boldsymbol{\kappa}}(T,0)$	(3.11)
$\boldsymbol{C}^n \equiv (C_x^n)$	Vector intervening in the canonical generator	(3.30)
$\boldsymbol{l} = (l_x)$	Vector intervening in the driven generator	(3.38)
$\boldsymbol{r} = (r_x)$	Vector intervening in the TiPS probability of the driven process	(3.44)

D. Markov diffusion processes (Sections 2.4 and 3.2)

Generators and probabilities

x	State	∇	Partial derivative with respect to x
W	Wiener process	$\xi_t = \frac{\mathrm{d}W}{\mathrm{d}t}$	Gaussian white noise
b	Drift	$\hat{b} = b - \frac{1}{2}\sigma\nabla\sigma$	Modified drift
σ	Diffusion coefficient	$D = \sigma^2$	Variance

ϱ	Probability density associated with b and σ	(2.64)
$\varrho(0)$	Initial probability density	
ϱ^{TiPS}	TiPS probability density of the process of drift b and diffusion σ	
J^{ϱ}	Probability current associated with ϱ	(2.66)
$\mathfrak{L}$	Fokker-Planck operator/generator	(2.67)
$\Lambda_{\boldsymbol{\gamma}}$	Tilted (or biased) generator	(2.85)
Λ^{υ}	Generalized Doob transform of Λ using the function υ	(3.122)
$b^{\upsilon}_{\boldsymbol{\gamma}}$	Drift associated with Λ^{υ}	(3.130)
$\hat{b}^{\upsilon}_{\boldsymbol{\gamma}}$	Modified drift associated with Λ^{υ}	(3.131)
$\mathbf{L}_n$	Canonical operator	(3.137)
b^c	Drift associated with $\mathbf{L}_n$	(3.138)

$\hat{b}^c$	Modified drift associated with $\mathbf{L}_n$	(3.139)
$\mathcal{L}$	Driven generator	(3.152)
$B_{\boldsymbol{\gamma}}, \hat{B}_{\boldsymbol{\gamma}}$	Drift and modified drift associated with $\mathcal{L}$	(3.154)
μ	TiPS probability of the driven process	(3.160)

Path probabilities

z_t Markov process giving the system state at time t $\quad [z]$ Path or trajectory

$\mathbb{P}_{b,\sigma,\varrho(0)}[z]$	Path probability associated with the Markov generator $\mathfrak{L}$	(2.63)
$\mathbb{P}_{\Lambda_{\boldsymbol{\gamma}},\varrho(0)}[z]$	Biased path probability	(2.87)
$\mathbb{P}^{\text{micro}}_{\mathbf{a},\varrho(0)}[z]$	Microcanonical path ensemble	
$\mathbb{P}_{b^v_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z]$	Path probability associated with the Doob transform Λ^v	(3.132)
$\mathbb{P}^{\text{cano}}_{\boldsymbol{\gamma},\varrho(0)}[z]$	Canonical path ensemble	(3.135)
$\mathbb{P}_{b^c,\sigma,\frac{C_n(0)\varrho(0)}{C_n(0)\cdot\varrho(0)}(0)}[z]$	Path probability associated with the canonical generator $\mathbf{L}_n$	(3.147)
$\mathbb{P}_{B_{\boldsymbol{\gamma}},\sigma,\varrho(0)}[z]$	Path probability associated with the driven generator $\mathcal{L}$	(3.166)

Averages

$\mathbb{E}_{\varrho(0)}[\cdots]$	Path average with respect to $\mathbb{P}_{b,\sigma,\varrho(0)}[z]$
$\mathbb{E}_{\boldsymbol{z}_0}[\cdots]$	Path average with respect to $\mathbb{P}_{b,\sigma,\varrho(0)}[z]$ with initial probability density $\varrho(x,0) = \delta(x - \boldsymbol{z}_0), \forall x$

Observables

$g(x,t), h(x,t)$	Periodic functions	
$\boldsymbol{A}_t[z] = \begin{pmatrix} \frac{1}{t}\int_0^t g(z_\tau,\tau) \circ \mathrm{d}z_\tau \\ \frac{1}{t}\int_0^t h(z_\tau,\tau)\mathrm{d}\tau \end{pmatrix}$	General observable	(3.107)
$\rho^n[z](x,\tau)$	Empirical occupation density	(3.171)
$j^n[z](x,\tau)$	Empirical current	(3.172)

Generating functions

$\boldsymbol{\gamma} = (\gamma_1 \;\; \gamma_2)$	Conjugate variable of $\boldsymbol{A}_t[z]$	
$G_{\boldsymbol{\gamma}}(t) \equiv \mathbb{E}_{\varrho(0)}\left[e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\right]$	Moment generating function	(2.81)
$\tilde{G}_{\boldsymbol{\gamma}}(x,t) \equiv \mathbb{E}_{\varrho(0)}\left[e^{t\,\boldsymbol{\gamma}\cdot\boldsymbol{A}_t[z]}\delta(z_t - x)\right]$	Moment generating function conditioned on the final state	(2.83)
$\Gamma(\boldsymbol{\gamma}) = \lim_{n\to\infty}\frac{1}{nT}\ln\mathbb{E}_{\varrho(0)}\left[\mathrm{e}^{nT\boldsymbol{\gamma}\cdot\boldsymbol{A}_{nT}}\right]$	Scaled cumulant generating function	(3.106)

Large deviation functions

$I_{2.5}(j,\rho)$	Level 2.5 large deviation function	(3.176)
$I(\boldsymbol{a})$	Large deviation function of the observable $\boldsymbol{A}$	(3.178)

Spectral elements and functions

$\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$	One-period propagator	
χ_T	Dominant eigenvalue of $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$	
r_T	Dominant right eigenfunction of $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$	(3.114)
l_T	Dominant left eigenfunction of $\overleftarrow{\mathcal{Q}}_\Lambda(T,0)$	(3.115)
C_n	Function intervening in the canonical generator	(3.142)
l	Function intervening in the driven generator	(3.149)
r	Function intervening in the TiPS probability of the driven process	(3.156)

E. Lagrangian/Hamiltonian formalism (Chapter 4)

Observables and operators

$\mathcal{D}$	Generalized differential operator	$\boldsymbol{z}$	Empirical state variable
$\boldsymbol{z}_\mathrm{i}$	Initial state	$\boldsymbol{z}_\mathrm{f}$	Final state
$\boldsymbol{\lambda}$	Empirical current variable	$\boldsymbol{f}$	Legendre conjugate variable of $\boldsymbol{\lambda}$
$\dot{\boldsymbol{z}}$	Current state variable	$\boldsymbol{p}$	Legendre conjugate variable of $\dot{\boldsymbol{z}}$

$(\boldsymbol{p}_t^*, \boldsymbol{z}_t^*)$ Solution of Hamilton's equations with initial condition $\boldsymbol{z}_\mathrm{i}$ and final condition $\boldsymbol{z}_\mathrm{f}$

$\bar{\boldsymbol{A}}_t = \begin{pmatrix} \frac{1}{t}\int_0^t \boldsymbol{\lambda}(t')\mathrm{d}t' \\ \frac{1}{t}\int_0^t \boldsymbol{z}(t')\mathrm{d}t' \end{pmatrix}$	Size-intensive observable	(4.111)
$\boldsymbol{A}_t = \mathcal{N}\bar{\boldsymbol{A}}_t$	Size-extensive observable	
$\Gamma = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \ln \mathbb{E}_{\boldsymbol{z}_\mathrm{i}}\left[\mathrm{e}^{\mathcal{T}\boldsymbol{\gamma}\cdot\boldsymbol{A}_t}\right]$	SCGF (in time)	(4.65)
$\bar{\Gamma} = \frac{\Gamma}{\mathcal{N}}$	SCGF (in time and size)	

Lagrangians

$\mathscr{L}(\boldsymbol{\lambda}, \boldsymbol{z})$	Detailed Lagrangian	(4.102)
$\mathscr{L}_{\boldsymbol{\gamma}}(\boldsymbol{\lambda}, \boldsymbol{z})$	Detailed biased Lagrangian	(4.118)
$\mathscr{L}^\mathrm{r}(\boldsymbol{\lambda}, \boldsymbol{z}; \boldsymbol{\gamma})$	Detailed rectified Lagrangian	(4.145)
$L(\dot{\boldsymbol{z}}, \boldsymbol{z})$	Standard Lagrangian	(4.105)
$L_{\boldsymbol{\gamma}}(\dot{\boldsymbol{z}}, \boldsymbol{z})$	Standard biased Lagrangian	(4.120)
$L^\mathrm{r}(\dot{\boldsymbol{z}}, \boldsymbol{z}; \boldsymbol{\gamma})$	Standard rectified Lagrangian	(4.146)

Hamiltonians

$\mathscr{H}(\boldsymbol{f}, \boldsymbol{z})$	Detailed Hamiltonian	(4.106)
$\mathscr{H}_{\boldsymbol{\gamma}}(\boldsymbol{f}, \boldsymbol{z})$	Detailed biased Hamiltonian	(4.119)
$\mathscr{H}^\mathrm{r}(\boldsymbol{f}, \boldsymbol{z}; \boldsymbol{\gamma})$	Detailed rectified Hamiltonian	(4.144)

$H(\boldsymbol{p},\boldsymbol{z})$	Standard Hamiltonian	(4.109)
$H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z})$	Standard biased Hamiltonian	(4.121)
$H^{\mathrm{r}}(\boldsymbol{p},\boldsymbol{z};\boldsymbol{\gamma})$	Standard rectified Hamiltonian	(4.143)

Probabilities

$P_{\delta t}(\boldsymbol{\lambda} \mid \boldsymbol{z})$	Conditional probability to observe $\boldsymbol{\lambda}$ given $\boldsymbol{z}$ after a time δt	
$\mathbb{P}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$	Path probability of $[z]$ given the initial state $\boldsymbol{z}_{\mathrm{i}}$	
$P_{\mathcal{T}}(\boldsymbol{z}_f \mid \boldsymbol{z}_{\mathrm{i}})$	Transition probability from $\boldsymbol{z}_{\mathrm{i}}$ to $\boldsymbol{z}_f$ after time $\mathcal{T}$	
$\mathbb{P}^{\mathrm{cano}}_{\boldsymbol{\gamma},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$	Canonical path probability given the initial state $\boldsymbol{z}_{\mathrm{i}}$	(4.154)
$\mathbb{P}^{\mathrm{r}}_{\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$	Transition probability from $\boldsymbol{z}_{\mathrm{i}}$ to $\boldsymbol{z}_f$ after time $\mathcal{T}$ in the rectified dynamics	(4.155)
$\mathbb{P}^{\mathrm{micro}}_{\boldsymbol{a},\mathcal{T}}[\boldsymbol{z} \mid \boldsymbol{z}_{\mathrm{i}}]$	Microcanonical path probability given the initial state $\boldsymbol{z}_{\mathrm{i}}$	

Spectral properties

$\boldsymbol{p}_{\min}$	Argument of the minimum (in $\boldsymbol{p}$) of the $H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z})$	(4.126)
$H_{\min}(\boldsymbol{z})$	Minimal value of $H_{\boldsymbol{\gamma}}$ for each $\boldsymbol{z}$	(4.127)
$\boldsymbol{z}^\star_\ell$	Positions of the maxima of $H_{\min}$	
$\boldsymbol{p}^\star_\ell$	Value of $\boldsymbol{p}_{\min}$ at $\boldsymbol{z}^\star_\ell$	
$(\boldsymbol{p}^\star_0,\boldsymbol{z}^\star_0)$	Dominant fixed point	
$H^\star_0(\boldsymbol{\gamma}) = \max_\ell H^\star_\ell(\boldsymbol{\gamma})$	Max-min value of $H_{\boldsymbol{\gamma}}$	
E	Eigenrate	(4.62)
$E^\star(\boldsymbol{\gamma})$	Smallest eigenrate for which the HJ equation admits global solutions	
$W_{\mathrm{s}}(\boldsymbol{z},\boldsymbol{\gamma})$	Globally stable solution of the HJ equation for $E = E^\star$	
$W_{\mathrm{u}}(\boldsymbol{z},\boldsymbol{\gamma})$	Globally unstable solution of the HJ equation for $E = E^\star$	

Fluctuation symmetries (Section 4.2.5)

θ	Current-reversal operator		$\boldsymbol{F}$	Affinity	
$\boldsymbol{F}_{\boldsymbol{\gamma}}$	Biased affinity	(4.166)	$\boldsymbol{F}^{\mathrm{r}}_{\boldsymbol{\gamma}}$	Rectified affinity	(4.172)

Population processes (Section 4.1)

X	State		
N_X	Number of particles in state X	$\boldsymbol{N}$	Population state vector of component N_X
$\boldsymbol{N}_{\mathrm{i}}$	Initial population state vector	$\boldsymbol{N}_{\mathrm{f}}$	Final population state vector

α	Transition between two state vectors
$\mathfrak{D}_{X,\alpha}$	Variation of the number of particles in state X due to the transition α
$\boldsymbol{\mathfrak{D}}_\alpha$	Vector of component $\mathfrak{D}_{X,\alpha}$
$\boldsymbol{\mathfrak{D}}$	Matrix of component $\mathfrak{D}_{X,\alpha}$
$\boldsymbol{N}(t)$	Empirical state vector at time t

$\boldsymbol{z}(t) \equiv \frac{\boldsymbol{N}(t)}{\mathcal{N}}$	Empirical density at time t	(4.4)
$[\Omega_\alpha]_t^{t+\delta t}$	Number of transitions α between times t and $t+\delta t$	
$\lambda_\alpha(t) \equiv \frac{1}{\mathcal{N}\delta t}[\Omega_\alpha]_t^{t+\delta t}$	Empirical particle current a time t due to the transition α	(4.6)
$\boldsymbol{\lambda}(t)$	Empirical particle current vector of component λ_α(t)	
$k_{\alpha,\boldsymbol{N}}$	Transition rate from $\boldsymbol{N}$ to $\boldsymbol{N}+\mathfrak{D}_\alpha$	
$k_\alpha(\boldsymbol{z})$	Intensive transition rate	(4.8)
$\boldsymbol{\kappa}$	Biased matrix	(4.40)
$\boldsymbol{K}$	Driven generator	(4.79)
$\mathrm{e}^{\boldsymbol{U}}$	Dominant left eigenvector of $\boldsymbol{\kappa}$	(4.80)

$P_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$	Transition probability from $\boldsymbol{N}_\mathrm{i}$ to $\boldsymbol{N}_\mathrm{f}$ after a time δt	(4.10)
$\mathbb{P}_t[\boldsymbol{N} \mid \boldsymbol{N}_\mathrm{i}]$	Path probability of $[\boldsymbol{N}] = (\boldsymbol{N}_\tau)_{\tau\in[0,t]}$ given the initial state $\boldsymbol{N}_\mathrm{i}$	(4.12)
$P_t(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$	Integrated transition probability after a time t	(4.14)
$\tilde{G}_{\delta t}(\boldsymbol{N}_\mathrm{f} \mid \boldsymbol{N}_\mathrm{i})$	Biased transition probability	(4.41)

Independent Markov jump processes (Section 4.C.1)

$\tilde{\boldsymbol{k}} = (\tilde{k}_{nm})$	Transition rate matrix of a single process	
n^ν	State occupied by the ν^{th} system	
$\boldsymbol{n} = (n^\nu)$	State of the global system	
$p_{\boldsymbol{n}}$	Probability that the global system is in state $\boldsymbol{n}$	
$\tilde{k}_{\boldsymbol{nm}}$	Transition rate from the global state $\boldsymbol{m}$ to the global state $\boldsymbol{n}$	(4.283)
N_n	Number of systems in state n	
$\boldsymbol{N}(\boldsymbol{n}) = (N_n(\boldsymbol{n}))$	Mesostate vector given the microstate $\boldsymbol{n}$	
$P_{\boldsymbol{N}}$	Probability that the global system is in state $\boldsymbol{N}$	(4.284)
$\boldsymbol{n}_{\boldsymbol{N}}$	Microstate compatible with the mesostate $\boldsymbol{N}$	
$\boldsymbol{k} = k_{\boldsymbol{N},\boldsymbol{N}'}$	Mesoscopic transition rate matrix	(4.286)
$\boldsymbol{D} = (D_{l,nm})$	Incidence matrix	(4.290)
$\boldsymbol{D}_{nm}$	Vector of component $D_{l,nm}$	
$\boldsymbol{\omega} = (\omega_{lm})$	Empirical transition current	(4.294)
$\boldsymbol{\mu} = (\mu_m)$	Empirical density	(4.295)
$\tilde{\boldsymbol{\kappa}}$	Biased matrix of a single process	
$\mathrm{e}^{\boldsymbol{u}}$	Dominant left eigenvector of $\tilde{\boldsymbol{\kappa}}$	
$\tilde{\boldsymbol{K}}$	Driven generator of a single process	

Independent Markov diffusion processes (Section 4.C.2)

x_t^ν	Stochastic process of the ν^{th} system	
ξ_t^ν	Gaussian white noise	
$\rho(x,t)$	Empirical occupation density	(4.306)
$j(x,t)$	Empirical current	(4.307)
$\tilde{j}_t^\nu(x)$	Empirical current for the single process ν	(4.309)
$\tilde{\rho}_t^\nu(x)$	Empirical occupancy for the single process ν	(4.310)

$\eta(x,t)$	Gaussian white noise in time and space	(4.321)

Brownian Donkey (Section 4.3.1)

$\mathcal{E}$	High state energy	V	Interaction energy
E_a	Activation energy	F	Non-conservative force
N	Number of machines in the high energy state		
U_N	Energy of the system at state N	(4.173)	
$k^{\nu}_{N+\epsilon,N}$	Transition rate from state N to state $N+\epsilon$	(4.175)	

Chemical reaction network (Section 4.3.2)

X	Chemical species	x	Concentration of the species X
$\mathfrak{K}_{\epsilon \mathsf{r}}$	kinetic constant of reaction $\epsilon \mathsf{r}$	$\mathcal{D}_{X,\epsilon \mathsf{r}}$	Variation of the number of species X due to the reaction $\epsilon \mathsf{r}$

F. Mathematical lexicon

Legendre-Fenchel transform (1.37)	$f^*(y) = \sup_x \{yx - f(x)\}$
Master equation (2.6)	$\dot{\boldsymbol{\pi}} = \boldsymbol{k\pi}$
Langevin equation (overdamped) (2.54)	$\dot{z}_t = b(z_t,t) + \sigma(z_t,t) \circ \xi_t$
Fokker-Planck equation (Stratonovich) (2.64)	$\frac{\partial \varrho(x,t)}{\partial t} = -\nabla \left[\hat{b}(x,t)\varrho(x,t) - \frac{1}{2} D(x,t) \nabla \varrho(x,t) \right]$
Generalized Doob transform (jump) (3.23)	$\boldsymbol{M^v} \equiv \mathcal{D}(\boldsymbol{v}) \boldsymbol{M}\, \mathcal{D}(\boldsymbol{v})^{-1} - \mathcal{D}(\boldsymbol{v})^{-1} \mathcal{D}(\boldsymbol{vM})$
Generalized Doob transform (diffusion) (3.122)	$M^{\upsilon} \equiv \upsilon M \upsilon^{-1} - \upsilon^{-1}(M^{\dagger}\upsilon)$
Large deviation principle (1.19)	$P(A_n = a) \underset{n\to\infty}{\asymp} \mathrm{e}^{-nI(a)}$
Euler-Lagrange equation (4.52)	$\frac{\partial L_{\boldsymbol{\gamma}}}{\partial \boldsymbol{z}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_{\boldsymbol{\gamma}}}{\partial \dot{\boldsymbol{z}}}\right) = 0$
Hamilton's equations (4.55)	$\dot{\boldsymbol{z}} = \partial_{\boldsymbol{p}} H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z}), \quad \dot{\boldsymbol{p}} = -\partial_{\boldsymbol{z}} H_{\boldsymbol{\gamma}}(\boldsymbol{p},\boldsymbol{z})$
Hamilton-Jacobi equation (4.62)	$H_{\boldsymbol{\gamma}}(\boldsymbol{p} = \partial_{\boldsymbol{z}} W, \boldsymbol{z}) = E$
Action (4.132)	$S[\boldsymbol{p},\boldsymbol{z}]_0^{\mathcal{T}} \equiv \int_0^{\mathcal{T}} [\boldsymbol{p}_t \cdot \dot{\boldsymbol{z}}_t - H_{\boldsymbol{\gamma}}(\boldsymbol{p}_t,\boldsymbol{z}_t)]\,\mathrm{d}t$
Reduced action (4.58)	$S_{\mathrm{r}}(t) \equiv \int_0^{\mathcal{T}} \boldsymbol{p}_t^* \cdot \dot{\boldsymbol{z}}_t^* \mathrm{d}t$

Résumé

En physique statistique d'équilibre, les systèmes sont étudiés dans des cadres mathématiques appelés ensembles statistiques. Les plus connus sont l'ensemble microcanonique qui traite d'un système isolé, donc d'énergie fixée; et l'ensemble canonique qui traite d'un système à l'équilibre avec un réservoir de température fixée, imposant la valeur moyenne de l'énergie. Ces deux ensembles sont équivalents dans la limite thermodynamique sous certaines conditions mathématiques. Ces notions d'ensemble et d'équivalence d'ensembles ont par la suite été généralisées dans le cadre de la physique statistique hors équilibre. En particulier, la notion de "contrainte" que l'on retrouve dans l'ensemble microcanonique a été généralisée en conditionnant des processus aléatoires sur une valeur imposée d'une observable physique, et pour lesquels on cherche un processus équivalent.

Le fil directeur de ma thèse porte sur ce problème de conditionnement qui a initialement été considéré par Doob. Ce dernier s'est intéressé un processus de Wiener conditionné à quitter l'espace des états à une certaine position. Récemment, ce problème a été formalisé et résolu pour des processus de Markov généraux homogènes en temps et conditionnés sur une valeur imposée d'une observable générale. Les observables considérées impliquent des fonctions des sauts et des états occupés, et satisfont un principe de grandes déviations dans la limite de temps long. La fonction de grande déviation associée informe non seulement sur les valeurs typiques de ces observables, mais également sur le comportement asymptotique de leurs fluctuations rares — ou grandes déviations: c'est sur ces grandes déviations que s'effectue le conditionnement. On parlera de *conditionnement microcanonique* par analogie avec l'ensemble microcanonique d'équilibre où l'énergie est fixée. Schématiquement, conditionner une observable sur une grande déviation revient à filtrer les trajectoires pour ne garder que celle pour lesquelles elle prend la valeur imposée (cf. Fig. 4.18). En général, ce processus conditionné ou microcanonique n'est pas associé à un générateur Markovien. On se pose alors la question de savoir s'il existe un processus de Markov effectif — appelé processus *drivé* — qui est équivalent au processus conditionné

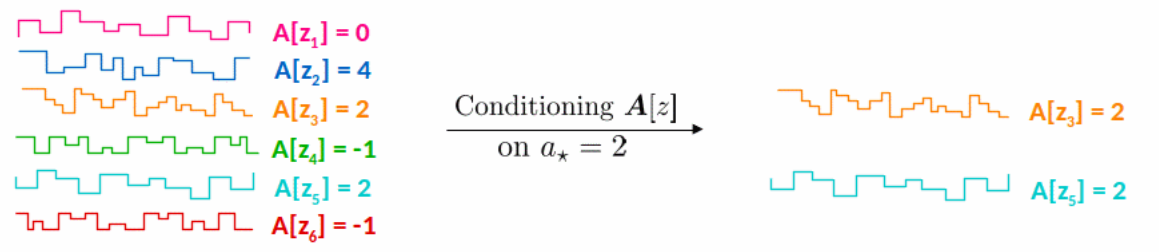

Figure 4.18 – Représentation schématique du conditionnement microcanonique.

dans la limite de temps long. La construction d'un tel processus s'appuie sur un processus auxiliaire pour lequel la contrainte sur l'observable est *transférée* vers sa variable conjuguée — de la même manière que la contrainte sur l'énergie devient une contrainte sur la température dans l'ensemble canonique d'équilibre. Ce processus dual est représenté par un générateur de Markov appelé *générateur biaisé*, mais qui n'a pas toutes les bonnes propriétés qu'un générateur Markovien doit avoir (il ne conserve pas la norme de la probabilité à tout temps). Le générateur drivé s'obtient alors en *rectifiant* le générateur biaisé via une transformation mathématique généralisant celle initialement considérée par Doob, et appelée de ce fait *transformée de Doob généralisée*. Les fonctions impliquées dans cette transformée sont telles que le processus drivé est équivalent au processus conditionné dans la limite de temps long.

L'objectif de cette thèse est de généraliser ces résultats à deux classes plus larges de processus de Markov:

1. Les processus de Markov forcés périodiquement, caractérisés par des générateurs de Markov périodiques (Chapitre 3).

2. Les processus de Markov non linéaires, impliquant des systèmes avec un grand nombre de degrés de liberté en interaction (Chapitre 4).

Dans la première généralisation, on adapte les résultats du cas homogène en temps pour construire le générateur drivé. Les générateurs indépendants du temps deviennent alors périodiques, les exponentielles de matrices deviennent des exponentielles ordonnées en temps et les problèmes spectraux deviennent des équations différentielles du premier ordre. La construction du générateur drivé se base sur l'équivalence des probabilités de chemin drivée et canonique d'une part, et des probabilités de chemin canonique et microcanonique d'autre part (sous condition de convexité de la fonction de grande déviation et pour le bon paramètre de biais). Le générateur drivé peut également s'obtenir via un problème d'optimisation sous contrainte de la fonction de grandes déviations 2.5 décrivant la probabilité des occupations empiriques et des courants de probabilité empiriques.
La seconde généralisation nécessite d'étudier à la fois une limite de temps long et une limite de grande taille. Dans cette limite, il convient de réduire le nombre de degrés de liberté du système en étudiant les valeurs typiques de deux observables empiriques, à savoir la densité empirique et le courant empirique associé. Le langage mathématique approprié est alors la théorie des grandes déviations en taille, qui s'apparente formellement à un formalisme lagrangien et hamiltonien. Ces Hamiltoniens sont en général des fonctions non-linéaires de la densité empirique, justifiant l'appellation de *processus non-linéaires*. Dans ce nouveau formalisme, la transformée de Doob généralisée menant vers le processus drivé se traduit par une transformation canonique sur les Hamiltoniens et nécessite d'étudier l'analogue non-linéaire du théorème de Perron-Frobenius. Cette étude nous a conduits à conjecturer une classification des solutions d'une équation de Hamilton-Jacobi, qui remplace les équations aux valeurs propres des générateurs dans le formalisme des opérateurs linéaires.

Avant d'exposer ces résultats, le manuscrit commence par deux chapitres introductifs portant sur les notions théoriques et outils mathématiques utilisés dans la thèse. Le chapitre 1 est une introduction à la théorie des processus aléatoires et des grandes déviations. Le chapitre 2 est une brève revue de la thermodynamique stochastique. Cette

dernière fournit le cadre permettant d'étudier la thermodynamique des systèmes hors équilibre, y compris des petits systèmes pour lesquels les fluctuations des observables physiques ne sont plus petites devant leurs valeurs moyennes. En plus de décrire l'évolution temporelle de ces moyennes, cette théorie permet d'étudier les fluctuations des observables physiques en partant d'une dynamique stochastique et en la connectant à une description thermodynamique du système au niveau d'une seule trajectoire. Pour ce faire, les systèmes sont modélisés par des processus de Markov pour lesquels l'évolution future d'une variable aléatoire à partir d'un instant présent ne dépend que de la valeur qu'elle a prise à cet instant, et non de son histoire antérieure. Un processus de Markov général est construit à partir de deux types particuliers de processus de Markov: les processus de diffusion pour lesquels la variable aléatoire évolue de manière continue (par exemple la position d'une particule Brownienne), et les processus de saut pour lesquels la variable aléatoire évolue de manière discrète en sautant d'un état vers un autre après un certain temps d'attente. Dans cette thèse, les deux formalismes sont étudiés séparément.

Le chapitre 3 aborde le problème du conditionnement des processus de Markov forcés périodiquement. La première partie de ce chapitre porte sur les processus de saut, caractérisés par des taux de transition périodiques, et la seconde partie sur les processus de diffusion, caractérisés par un "drift" et un coefficient de diffusion périodiques. Dans les deux formalismes, on s'intéresse à une observable impliquant des fonctions périodiques que l'on conditionne sur une grande déviation. On construit alors les générateurs canonique et drivé associés en utilisant les définitions des probabilités de chemin et leur équivalence. On notera que le générateur drivé est par définition la limite à temps long du générateur canonique. Dans un second temps, on montre que le générateur drivé est associé à l'occupation empirique et au courant de probabilité empirique minimisant la fonction de grande déviation 2.5 et compatibles avec la contrainte sur l'observable. Pour finir, ces résultats sur les processus forcés périodiquement sont illustrés sur un modèle simple, faisant intervenir un processus de sauts à deux états et des taux périodiques et constants par morceaux sur deux parties de la période.

Le chapitre 4 propose une théorie sur le problème de conditionnement des processus de Markov non-linéaires. Cette théorie se focalise sur les processus homogènes en temps, mais nous espérons qu'elle ouvre la voie vers une généralisation aux processus non-linéaires conduits périodiquement. La première partie de ce chapitre étudie le problème de conditionnement pour des processus de population. Il reprend les résultats connus antérieurement pour les processus de Markov usuels puis dérive leur analogue dans le formalisme lagrangien/hamiltonien en prenant la limite macroscopique. En particulier, les équations aux valeurs propres deviennent une équation de Hamilton-Jacobi, et la transformée de Doob généralisée sur le générateur biaisé devient une transformation canonique sur le Hamiltonien biaisé (cf. Fig. 4.19). Sur la base de ces résultats, une théorie générale sur la rectification des Hamiltoniens est proposée dans la deuxième partie de ce chapitre. Cette théorie est indépendante de tout modèle et s'applique à une classe particulière de Hamiltoniens — appelés *Hamiltoniens statistiques* — vérifiant un certain nombre d'hypothèses. La rectification des Hamiltoniens statistiques biaisés fait appel à deux résultats intermédiaires. Le premier résultat conjecture l'existence d'une unique solution globalement stable de l'équation de Hamilton-Jacobi pour un taux propre égal au "max-min" du Hamiltonien. Le second résultat soutient que ce taux propre correspond à la valeur du Hamiltonien le long de sa trajectoire typique dans la limite de temps long, et coïncide avec la fonction génératrice des cumulants mise à l'échelle (en temps et en taille). Une illustration de

Formalisme des opérateurs linéaires	Formalisme lagrangien/hamiltonien
Générateur de Markov	Lagrangiens et hamiltoniens
Générateur biaisé	Lagrangiens et hamiltoniens biaisés
Générateur drivé	Lagrangiens et hamiltoniens rectifiés
Théorème de Perron-Frobenius	Conjecture sur la classification des solutions de l'équation de Hamilton-Jacobi
Transformée de Doob généralisée	Transformation canonique

Figure 4.19 – Analogie entre les différents concepts de la rectification dans le formalisme des opérateurs linéaires et le formalisme lagrangien/hamiltonien.

cette théorie est faite sur des exemples spécifiques de processus de population (réactions chimiques et modèle de l'âne brownien). Enfin, des résultats préliminaires sur le conditionnement des processus non-linéaires forcés périodiquement sont présentés en fin de chapitre.

Pour conclure, les résultats principaux de cette thèse portent d'une part sur l'extension du problème de conditionnement aux processus de Markov forcés périodiquement dans le formalisme des opérateurs linéaires, et d'autre part sur la traduction de ce problème dans le formalisme lagrangien/hamiltonien. Le travail effectué dans cette seconde partie constitue un premier pas pour l'étude du conditionnement des processus non-linéaires. Ces premières investigations ont naturellement été faites pour des Hamiltoniens indépendants du temps. Même dans ce cas-là, plusieurs questions restent en suspens. La première est de fournir une preuve rigoureuse à notre conjecture. La seconde est d'analyser des systèmes plus complexes impliquant des variétés critiques autres que des points fixes, comme des cycles limites. Par la suite, il serait intéressant d'étendre ces résultats à des processus de Markov non-linéaires forcés périodiquement, impliquant des Hamiltoniens périodiques. On s'attend alors à ce que la rectification implique une des solutions de l'équation de Hamilton-Jacobi dépendante du temps. Toutefois, tout un travail reste à faire pour caractériser proprement cette solution. Enfin, le travail fourni dans cette thèse reste à un niveau plutôt abstrait, mais nous espérons qu'il servira à l'étude de systèmes réels. En particulier, les résultats présentés dans ce manuscrit peuvent être appliqués à la thermodynamique stochastique en considérant des observables thermodynamiques (chaleur, travail, courant chimique, production d'entropie, etc.) et des taux de transition satisfaisant le bilan détaillé local ou instantané. On peut alors se demander si la consistance thermodynamique est préservée après rectification.

Je tenais à faire une présentation pédagogique et progressive des différents concepts et résultats de cette thèse. J'espère avoir atteint cet objectif et que ce manuscrit sera utile aux étudiants, doctorants ou à tous ceux qui souhaitent en savoir plus (ou continuer) les travaux présentés ici.

Bibliography

[1] H. J. Kreuzer, Nonequilibrium thermodynamics and its statistical foundations. Oxford University Press, 1981. 10

[2] N. Pottier, Physique statistique hors d'équilibre. EDP Sciences, 2007. 10

[3] I. Müller, A history of thermodynamics: the doctrine of energy and entropy. Springer Science & Business Media, 2007. 10

[4] S. Carnot, Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance. Bachelier, Paris, 1824. 10

[5] J. W. Gibbs, Elementary Principles in Statistical Mechanics, developed with especial reference to the rational foundation of thermodynamics. Charles Scribner's Sons, 1902. 11

[6] L. E. Reichl, A modern course in statistical physics. Wiley-VCH Verlag GmbH, 4th ed., 2016. 11

[7] R. K. Pathria and P. D. Beale, Statistical Mechanics. Elsevier, 3rd ed., 2011. 11

[8] L. Onsager, "Reciprocal relations in irreversible processes. I.," Physical Review, vol. 37, no. 4, p. 405, 1931. 11

[9] L. Onsager, "Reciprocal relations in irreversible processes. II.," Physical Review, vol. 38, no. 12, p. 2265, 1931. 11

[10] R. Kubo, "The fluctuation-dissipation theorem," Reports on Progress in Physics, vol. 29, no. 1, p. 255, 1966. 11

[11] H. B. Callen and T. A. Welton, "Irreversibility and generalized noise," Physical Review, vol. 83, no. 1, p. 34, 1951. 11

[12] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, "Probability of second law violations in shearing steady states," Physical Review Letters, vol. 71, pp. 2401–2404, Oct 1993. 11, 30

[13] D. J. Evans and D. J. Searles, "Equilibrium microstates which generate second law violating steady states," Physical Review E, vol. 50, no. 2, p. 1645, 1994. 11, 30

[14] G. Gallavotti and E. G. D. Cohen, "Dynamical ensembles in nonequilibrium statistical mechanics," Physical Review Letters, vol. 74, no. 14, p. 2694, 1995. 11, 30

[15] J. Kurchan, "Fluctuation theorem for stochastic dynamics," Journal of Physics A: Mathematical and General, vol. 31, no. 16, p. 3719, 1998. 11, 30

[16] J. L. Lebowitz and H. Spohn, "A gallavotti–cohen-type symmetry in the large deviation functional for stochastic dynamics," Journal of Statistical Physics, vol. 95, no. 1-2, pp. 333–365, 1999. 11, 30, 32

[17] C. Jarzynski, "Nonequilibrium equality for free energy differences," Physical Review Letters, vol. 78, no. 14, p. 2690, 1997. 11

[18] C. Jarzynski, "Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach," Physical Review E, vol. 56, no. 5, p. 5018, 1997. 11

[19] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," Physical Review E, vol. 60, no. 3, p. 2721, 1999. 11

[20] G. E. Crooks, "Path-ensemble averages in systems driven far from equilibrium," Physical review E, vol. 61, no. 3, p. 2361, 2000. 11

[21] G. Hummer and A. Szabo, "Free energy reconstruction from nonequilibrium single-molecule pulling experiments," Proceedings of the National Academy of Sciences, vol. 98, no. 7, pp. 3658–3661, 2001. 11

[22] T. Hatano and S. Sasa, "Steady-state thermodynamics of Langevin systems," Physical Review Letters, vol. 86, no. 16, p. 3463, 2001. 11

[23] U. Seifert, "Stochastic thermodynamics: principles and perspectives," The European Physical Journal B, vol. 64, no. 3, pp. 423–431, 2008. 11, 26

[24] U. Seifert, "Stochastic thermodynamics, fluctuation theorems and molecular machines," Reports on progress in physics, vol. 75, no. 12, p. 126001, 2012. 11, 26, 30

[25] C. Van den Broeck, "Stochastic thermodynamics: A brief introduction," in Physics of Complex Colloids, vol. 184 of Proceedings of the International School of Physics "Enrico Fermi", pp. 155–193, IOS Press Amsterdam, 2013. 11, 26

[26] K. Sekimoto, "Kinetic characterization of heat bath and the energetics of thermal ratchet models," Journal of the physical society of Japan, vol. 66, no. 5, pp. 1234–1237, 1997. 11

[27] K. Sekimoto, "Langevin equation and thermodynamics," Progress of Theoretical Physics Supplement, vol. 130, pp. 17–27, 1998. 11

[28] K. Sekimoto, Stochastic energetics. pringer-Verlag, 2010. 11, 32, 140

[29] C. Maes and K. Netočnỳ, "Time-reversal and entropy," Journal of Statistical Physics, vol. 110, no. 1, pp. 269–310, 2003. 11

[30] C. Maes, "On the origin and the use of fluctuation relations for the entropy," Séminaire Poincaré, vol. 2, pp. 29–62, 2003. 11

[31] U. Seifert, "Entropy production along a stochastic trajectory and an integral fluctuation theorem," Physical Review Letters, vol. 95, no. 4, p. 040602, 2005. 11, 29, 30, 36

[32] S. Ciliberto, "Experiments in stochastic thermodynamics: Short history and perspectives," Physical Review X, vol. 7, no. 2, p. 021051, 2017. 11

[33] J. Bechhoefer, S. Ciliberto, S. Pigolotti, and E. Roldán, "Stochastic thermodynamics: experiment and theory," Journal of Statistical Mechanics: Theory and Experiment, vol. 2020, no. 6, p. 064001, 2020. 11

[34] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger, "Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential," Physical Review Letters, vol. 96, no. 7, p. 070603, 2006. 11, 26, 36

[35] V. Blickle and C. Bechinger, "Realization of a micrometre-sized stochastic heat engine," Nature Physics, vol. 8, no. 2, pp. 143–146, 2012. 11, 26

[36] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. Parrondo, and R. A. Rica, "Brownian Carnot engine," Nature Physics, vol. 12, no. 1, pp. 67–70, 2015. 11, 26

[37] R. M. L. Evans, "Rules for transition rates in nonequilibrium steady states," Physical Review Letters, vol. 92, p. 150601, Apr 2004. 12, 37

[38] R. M. L. Evans, "Detailed balance has a counterpart in non-equilibrium steady states," Journal of Physics A: Mathematical and General, vol. 38, no. 2, p. 293, 2004. 12, 37

[39] R. M. L. Evans, "Statistical physics of shear flow: a non-equilibrium problem," Contemporary Physics, vol. 51, no. 5, pp. 413–427, 2010. 12, 37

[40] C. Dellago, P. G. Bolhuis, and P. L. Geissler, "Transition path sampling," Advances in Chemical Physics, vol. 123, no. 1, 2002. 12, 38

[41] M. Merolle, J. P. Garrahan, and D. Chandler, "Space–time thermodynamics of the glass transition," Proceedings of the National Academy of Sciences, vol. 102, no. 31, pp. 10837–10840, 2005. 12

[42] C. Dellago, P. G. Bolhuis, and P. L. Geissler, "Transition path sampling methods," in Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1, pp. 349–391, Springer, 2006. 12, 38

[43] E. Vanden-Eijnden, "Transition path theory," in Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1, pp. 453–493, Springer, 2006. 12, 38

[44] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "Dynamical first-order phase transition in kinetically constrained models of glasses," Physical Review Letters, vol. 98, no. 19, p. 195702, 2007. 12, 32

[45] L. O. Hedges, R. L. Jack, J. P. Garrahan, and D. Chandler, "Dynamic order-disorder in atomistic models of structural glass formers," Science, vol. 323, no. 5919, pp. 1309–1313, 2009. 12

[46] C. Dellago and P. G. Bolhuis, "Transition path sampling and other advanced simulation techniques for rare events," Advanced computer simulation approaches for soft matter sciences III, pp. 167–233, 2009. 12

[47] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, "First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories," Journal of Physics A: Mathematical and Theoretical, vol. 42, no. 7, p. 075007, 2009. 12, 38

[48] F. Turci and E. Pitard, "Large deviations and heterogeneities in a driven kinetically constrained model," EPL (Europhysics Letters), vol. 94, p. 10003, mar 2011. 12

[49] D. Chandler and J. P. Garrahan, "Dynamics on the way to forming glass: Bubbles in space-time," Annual Review of Physical Chemistry, vol. 61, pp. 191–217, 2010. 12, 38

[50] T. Speck and J. P. Garrahan, "Space-time phase transitions in driven kinetically constrained lattice models," The European Physical Journal B, vol. 79, no. 1, pp. 1–6, 2011. 12

[51] C. P. Royall, F. Turci, and T. Speck, "Dynamical phase transitions and their relation to structural and thermodynamic aspects of glass physics," The Journal of Chemical Physics, vol. 153, no. 9, p. 090901, 2020. 12

[52] R. Chetrite and H. Touchette, "Nonequilibrium microcanonical and canonical ensembles and their equivalence," Physical Review Letters, vol. 111, no. 12, p. 120601, 2013. 12, 37, 39, 44, 46, 48, 51, 146

[53] A. Zeitouni and O. Dembo, Large Deviations Techniques and Applications. Springer-Verlag, New York, 1998. 12, 17

[54] A. Amann and H. Atmanspacher, "Introductory remarks on large deviation statistics," Journal of Scientific Exploration, vol. 13, no. 4, pp. 639–664, 1999. 12, 17

[55] F. den Hollander, Large deviations. Fields Institute Monographs, American Mathematical Society, 2000. 12, 17

[56] D. Ruelle, "Classical statistical mechanics of a system of particles," Helvetica Physica Acta (Switzerland), vol. 36, 1963. 12

[57] D. Ruelle, Statistical mechanics: Rigorous results. W.A. Benjamin, Amsterdam, 1969. 12

[58] R. B. Griffiths, Rigorous results and theorems, vol. 1, pp. 7–109. Academic Press, 1972. 12

[59] O. E. Lanford, "Entropy and equilibrium states in classical statistical mechanics," in Statistical Mechanics and Mathematical Problems, pp. 1–113, Springer Berlin Heidelberg, 1973. 12

[60] R. S. Ellis, Entropy, large deviations, and statistical mechanics. Springer, 1985. 12, 17, 53

[61] Y. Oono, "Large deviation and statistical physics," Progress of Theoretical Physics Supplement, vol. 99, pp. 165–205, 1989. 12

[62] D. W. Stroock and O. Zeitouni, "Microcanonical distributions, Gibbs states, and the equivalence of ensembles," in Random walks, Brownian motion, and interacting particle systems, pp. 399–424, Springer, 1991. 12

[63] J.-D. Deuschel, D. W. Stroock, and H. Zessin, "Microcanonical distributions for lattice gases," Communications in Mathematical Physics, vol. 139, no. 1, pp. 83–101, 1991. 12

[64] R. L. Dobrushin, R. Kotecký, and S. Shlosman, Wulff construction: a global shape from local interaction, vol. 104. American Mathematical Society Providence, 1992. 12

[65] H.-O. Georgii, "Large Deviations and Maximum Entropy Principle for Interacting Random Fields on $\mathbb{Z}^d$," The Annals of Probability, pp. 1845–1875, 1993. 12

[66] S. Roelly and H. Zessin, "The equivalence of equilibrium principles in statistical mechanics and some applications to large particle systems," vol. 11, pp. 384–405, 1993. 12

[67] R. H. Schonmann and S. B. Shlosman, "Complete analyticity for 2D Ising completed," Communications in Mathematical Physics, vol. 170, no. 2, pp. 453–482, 1995. 12

[68] G. Gallavotti, Statistical mechanics: A short treatise. Springer, New York, 1999. 12

[69] C. Maes, F. Redig, S. Shlosman, and A. Van Moffaert, "Percolation, path large deviations and weakly Gibbs states," Communications in Mathematical Physics, vol. 209, no. 2, pp. 517–545, 2000. 12

[70] H. Touchette, "The large deviation approach to statistical mechanics," Physics Reports, vol. 478, pp. 1–69, jul 2009. 12, 17, 21

[71] M. Paniconi and Y. Oono, "Phenomenological framework for fluctuations around steady state," Physical Review E, vol. 55, no. 1, p. 176, 1997. 12

[72] Y. Oono and M. Paniconi, "Steady state thermodynamics," Progress of Theoretical Physics Supplement, vol. 130, pp. 29–44, 1998. 12

[73] G. L. Eyink, "Dissipation and large thermodynamic fluctuations," Journal of Statistical Physics, vol. 61, no. 3, pp. 533–572, 1990. 12

[74] G. L. Eyink, "Action principle in nonequilibrium statistical dynamics," Physical Review E, vol. 54, no. 4, p. 3419, 1996. 12

[75] G. L. Eyink, "Action principle in statistical dynamics," Progress of Theoretical Physics Supplement, vol. 130, pp. 77–86, 1998. 12

[76] C. Maes and K. Netočnỳ, "Minimum entropy production principle from a dynamical fluctuation law," Journal of Mathematical Physics, vol. 48, no. 5, p. 053306, 2007. 12

[77] C. Maes and K. Netočnỳ, "Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states," EPL (Europhysics Letters), vol. 82, no. 3, p. 30003, 2008. 12

[78] C. Maes, K. Netočnỳ, and B. Shergelashvili, A Selection of Nonequilibrium Issues, pp. 247–306. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. 12

[79] H. Touchette and R. J. Harris, Large Deviation Approach to Nonequilibrium Systems, ch. 11, pp. 335–360. John Wiley & Sons, Ltd, 2013. 12

[80] R. Chetrite and H. Touchette, "Nonequilibrium Markov processes conditioned on large deviations," Annales Henri Poincaré, vol. 16, pp. 2005–2057, oct 2015. 12, 13, 17, 35, 37, 38, 44, 46, 47, 48, 50, 51, 64, 66, 67, 101, 129, 139, 142, 146

[81] R. Chetrite and H. Touchette, "Variational and optimal control representations of conditioned and driven processes," Journal of Statistical Mechanics: Theory and Experiment, vol. 2015, no. 12, p. P12001, 2015. 12, 44, 46, 50, 53, 55, 57, 64, 70, 71, 73, 129, 146

[82] J. L. Doob, "Conditional Brownian motion and the boundary limits of harmonic functions," Bulletin de la Société Mathématique de France, vol. 85, pp. 431–458, 1957. 12, 44

[83] J. Doob, Classical potential theory and its probabilistic counterpart. Springer, New York, 1984. 12, 44, 48

[84] L. C. G. Rogers and D. Williams, Diffusions, Markov processes and martingales: Volume 2, Itô calculus, vol. 2. Cambridge university press, 2000. 12

[85] F. Baudoin, "Conditioned stochastic differential equations: theory, examples and application to finance," Stochastic Processes and their Applications, vol. 100, no. 1-2, pp. 109–145, 2002. 12

[86] D. Gasbarra, T. Sottinen, and E. Valkeila, "Gaussian bridges," in Stochastic Analysis and Applications (F. E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal, and T. Zhang, eds.), Springer Berlin Heidelberg, 2007. 12

[87] T. Sottinen and A. Yazigi, "Generalized Gaussian bridges," Stochastic Processes and their Applications, vol. 124, no. 9, pp. 3084–3105, 2014. 12

[88] E. Schrödinger, Über die umkehrung der naturgesetze. Verlag der Akademie der Wissenschaften in Kommission bei Walter De Gruyter u. Co, 1931. 12

[89] E. Schrödinger, "Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique," in Annales de l'institut Henri Poincaré, vol. 2, pp. 269–310, 1932. 12

[90] B. Jamison, "The Markov processes of Schrödinger," Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, vol. 32, no. 4, pp. 323–331, 1975. 12

[91] J. C. Zambrini, "Stochastic mechanics according to E. Schrödinger," Physical review A, vol. 33, no. 3, p. 1532, 1986. 12

[92] P. Collet, S. Martínez, and J. San Martín, "Asymptotic laws for one-dimensional diffusions conditioned to nonabsorption," The Annals of Probability, pp. 1300–1314, 1995. 12

[93] R. Aebi, Schrödinger Diffusion Processes. Birkhäuser, Basel, 1996. 12

[94] J. N. Darroch and E. Seneta, "On quasi-stationary distributions in absorbing discrete-time finite Markov chains," Journal of Applied Probability, vol. 2, no. 1, pp. 88–100, 1965. 12

[95] J. N. Darroch and E. Seneta, "On quasi-stationary distributions in absorbing continuous-time finite Markov chains," Journal of Applied Probability, vol. 4, no. 1, pp. 192–196, 1967. 12

[96] S. Méléard and D. Villemonais, "Quasi-stationary distributions and population processes," Probability Surveys, vol. 9, pp. 340–410, 2012. 12

[97] P. Collet, S. Martínez, and J. S. Martín, Quasi-Stationary Distributions. Springer, New York, 1994. 12

[98] E. A. van Doorn and P. K. Pollett, "Quasi-stationary distributions for discrete-state models," European Journal of Operational Research, vol. 230, no. 1, pp. 1–14, 2013. 12

[99] V. Lecomte, C. Appert-Rolland, and F. van Wijland, "Chaotic properties of systems with Markov dynamics," Physical Review Letters, vol. 95, no. 1, p. 010601, 2005. 12, 32

[100] V. Lecomte, C. Appert-Rolland, and F. Van Wijland, "Thermodynamic formalism for systems with Markov dynamics," Journal of Statistical Physics, vol. 127, no. 1, pp. 51–106, 2007. 12, 32

[101] R. L. Jack and P. Sollich, "Large deviations and ensembles of trajectories in stochastic models," Progress of Theoretical Physics Supplement, vol. 184, pp. 304–317, 2010. 12

[102] J. P. Garrahan and I. Lesanovsky, "Thermodynamics of quantum jump trajectories," Physical Review Letters, vol. 104, no. 16, p. 160601, 2010. 12, 147

[103] V. Blickle and C. Bechinger, "Realization of a micrometre-sized stochastic heat engine," Nature Physics, vol. 8, no. 2, p. 143, 2012. 12

[104] I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. R. Parrondo, and R. A. Rica, "Brownian Carnot engine," Nature Physics, vol. 12, p. 67, oct 2016. 12

[105] S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, and A. L. Nussbaumer, "Artificial molecular machines," Chemical Reviews, vol. 115, no. 18, pp. 10081–10206, 2015. 12

[106] I. A. Martínez, É. Roldán, L. Dinis, and R. A. Rica, "Colloidal heat engines: a review," Soft Matter, vol. 13, no. 1, pp. 22–36, 2017. 12

[107] A. C. Barato and R. Chetrite, "Current fluctuations in periodically driven systems," Journal of Statistical Mechanics: Theory and Experiment, vol. 2018, p. 053207, may 2018. 12, 47

[108] G. Verley, M. Esposito, T. Willaert, and C. Van den Broeck, "The unlikely Carnot efficiency," Nature Communications, vol. 5, no. 1, pp. 1–5, 2014. 12, 17

[109] M. Doi, "Second quantization representation for classical many-particle system," Journal of Physics A: Mathematical and General, vol. 9, no. 9, p. 1465, 1976. 12, 86

[110] L. Peliti, "Path integral approach to birth-death processes on a lattice," Journal de Physique, vol. 46, no. 9, pp. 1469–1483, 1985. 12, 86

[111] R. Dickman and R. Vidigal, "Path integrals and perturbation theory for stochastic processes," Brazilian Journal of Physics, vol. 33, pp. 73–93, 2003. 12, 86

[112] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, "Large deviation approach to non equilibrium processes in stochastic lattice gases," Bulletin of the Brazilian Mathematical Society, vol. 37, no. 4, pp. 611–643, 2006. 12, 86

[113] M. Assaf and B. Meerson, "WKB theory of large deviations in stochastic populations," Journal of Physics A: Mathematical and Theoretical, vol. 50, no. 26, p. 263001, 2017. 12, 86

[114] H. J. Park and A. Traulsen, "Extinction dynamics from metastable coexistences in an evolutionary game," Physical Review E, vol. 96, no. 4, p. 042412, 2017. 12, 86

[115] R. Kraaij, A. Lazarescu, C. Maes, and M. Peletier, "Deriving generic from a generalized fluctuation symmetry," Journal of Statistical Physics, vol. 170, no. 3, pp. 492–508, 2018. 12, 86

[116] H. Vroylandt, Thermodynamique et fluctuations des petites machines. PhD thesis, Université Paris-Saclay, 2018. 12, 86

[117] C. Maes, K. Netočnỳ, and B. Wynants, "On and beyond entropy production: the case of markov jump processes," Markov Processes And Related Fields, vol. 14, pp. 445–464, 2008. 12, 86

[118] C. Monthus, "Microscopic fluctuation theory (mFT) for interacting poisson processes," Journal of Physics A: Mathematical and Theoretical, vol. 52, p. 135003, mar 2019. 12, 86

[119] H. Vroylandt, M. Esposito, and G. Verley, "Efficiency fluctuations of stochastic machines undergoing a phase transition," Physical Review Letters, vol. 124, no. 25, p. 250603, 2020. 12, 86, 87, 120

[120] H. Vroylandt and G. Verley, "Non-equivalence of dynamical ensembles and emergent non-ergodicity," Journal of Statistical Physics, vol. 174, no. 2, pp. 404–432, 2019. 12, 86, 87, 147

[121] P. Zilber, N. R. Smith, and B. Meerson, "Giant disparity and a dynamical phase transition in large deviations of the time-averaged size of stochastic populations," Physical Review E, vol. 99, no. 5, p. 052105, 2019. 12, 86

[122] J. Guioth and E. Bertin, "Large deviations and chemical potential in bulk-driven systems in contact," EPL (Europhysics Letters), vol. 123, p. 10002, aug 2018. 12, 86

[123] A. D. Wentzell, Limit theorems on large deviations for Markov stochastic processes, vol. 38 of Mathematics and Its Applications. Springer, Dordrecht, 1990. 12, 86, 87

[124] A. Lazarescu, T. Cossetto, G. Falasco, and M. Esposito, "Large deviations and dynamical phase transitions in stochastic chemical networks," The Journal of Chemical Physics, vol. 151, p. 064117, aug 2019. 12, 13, 17, 87, 89, 90, 91, 93, 129, 131, 146, 147

[125] C. W. Gardiner, Handbook of stochastic methods, vol. 3. springer Berlin, 1985. 14, 15, 16, 26, 33, 140

[126] N. G. Van Kampen, Stochastic Processes in Physics and Chemistry. Elsevier, 3rd ed., 2007. 14, 26, 33

[127] G. A. Pavliotis, Stochastic processes and applications: diffusion processes, the Fokker-Planck and La vol. 60. Springer, 2014. 14, 33

[128] R. Graham and F. Haake, "Statistical theory of instabilities in stationary nonequilibrium systems with applications to lasers and nonlinear optics," in Quantum Statistics in Optics and Solid-State Physics, Springer, 1973. 15

[129] E. Pardoux, Markov processes and applications: algorithms, networks, genome and finance, vol. 796. John Wiley & Sons, 2008. 15

[130] P. A. Gagniuc, Markov chains: from theory to implementation and experimentation. John Wiley & Sons, 2017. 15

[131] K. R. Gabriel and J. Neumann, "A Markov chain model for daily rainfall occurrence at Tel Aviv," Quarterly Journal of the Royal Meteorological Society, vol. 88, no. 375, pp. 90–95, 1962. 15

[132] R. W. Katz, "An application of chain-dependent processes to meteorology," Journal of Applied Probability, vol. 14, no. 3, pp. 598–603, 1977. 15

[133] N.-C. Lau and M. J. Nath, "The role of the "atmospheric bridge" in linking tropical pacific ENSO events to extratropical SST anomalies," Journal of Climate, vol. 9, no. 9, pp. 2036–2057, 1996. 15

[134] S. T. Rao, I. G. Zurbenko, R. Neagu, P. S. Porter, J. Y. Ku, and R. F. Henry, "Space and time scales in ambient ozone data," Bulletin of the American Meteorological Society, vol. 78, no. 10, pp. 2153–2166, 1997. 15

[135] D. S. Wilks and R. L. Wilby, "The weather generation game: a review of stochastic weather models," Progress in Physical Geography, vol. 23, no. 3, pp. 329–357, 1999. 15

[136] A. Krogh, M. Brown, I. S. Mian, K. Sjölander, and D. Haussler, "Hidden Markov models in computational biology: Applications to protein modeling," Journal of molecular biology, vol. 235, no. 5, pp. 1501–1531, 1994. 15

[137] G. Sella and A. E. Hirsh, "The application of statistical physics to evolutionary biology," Proceedings of the National Academy of Sciences, vol. 102, no. 27, pp. 9541–9546, 2005. 15

[138] D. George and J. Hawkins, "Towards a mathematical theory of cortical micro-circuits," PLoS Computational Biology, vol. 5, no. 10, p. e1000532, 2009. 15

[139] C. A. Yates, M. J. Ford, and R. L. Mort, "A multi-stage representation of cell proliferation as a Markov process," Bulletin of Mathematical Biology, vol. 79, no. 12, pp. 2905–2928, 2017. 15

[140] P. S. Kutchukian, D. Lou, and E. I. Shakhnovich, "FOG: Fragment Optimized Growth algorithm for the *de novo* generation of molecules occupying druglike chemical space," Journal of Chemical Information and Modeling, vol. 49, no. 7, pp. 1630–1642, 2009. 15

[141] D. F. Anderson and T. G. Kurtz, "Continuous time Markov chain models for chemical reaction networks," in Design and Analysis of Biomolecular Circuits, pp. 3–42, Springer, 2011. 15, 86

[142] S. Chao Du, "Correlation analysis of enzymatic reaction of a single protein molecule," The Annals of Applied Statistics, vol. 6, no. 3, p. 950, 2012. 15

[143] T. Koski, Hidden Markov models for bioinformatics, vol. 2. Springer Science & Business Media, 2001. 15

[144] W. J. Ewens and G. R. Grant, Statistical methods in bioinformatics: an introduction. Springer Science & Business Media, 2006. 15

[145] V. De Fonzo, F. Aluffi-Pentini, and V. Parisi, "Hidden Markov models in bioinformatics," Current Bioinformatics, vol. 2, no. 1, pp. 49–61, 2007. 15

[146] C. Ames, "The Markov process as a compositional model: A survey and tutorial," Leonardo, vol. 22, no. 2, pp. 175–187, 1989. 15

[147] W. Chai and B. Vercoe, "Folk music classification using hidden Markov models," in Proceedings of international conference on artificial intelligence, 2001. 15

[148] E. Pollastri and G. Simoncelli, "Classification of melodies by composer with hidden Markov models," in Proceedings first international conference on WEB delivering of music. WEDELMUSIC 2001, pp. 88–95, 2001. 15

[149] Y. Qi, J. W. Paisley, and L. Carin, "Music analysis using hidden Markov mixture models," IEEE Transactions on Signal Processing, vol. 55, no. 11, pp. 5209–5224, 2007. 15

[150] D. Herremans, S. Weisser, K. Sörensen, and D. Conklin, "Generating structured music for bagana using quality metrics based on Markov models," Expert Systems with Applications, vol. 42, no. 21, pp. 7424–7435, 2015. 15

[151] G. Pagès, H. Pham, and J. Printems, "Optimal quantization methods and applications to numerical problems in finance," in Handbook of computational and numerical methods in finance, pp. 253–297, Springer, 2004. 15

[152] K. E. Bassler, G. H. Gunaratne, and J. L. McCauley, "Markov processes, hurst exponents, and nonlinear diffusion equations: With application to finance," Physica A: Statistical Mechanics and its Applications, vol. 369, no. 2, pp. 343–353, 2006. 15

[153] N. Bäuerle and U. Rieder, Markov decision processes with applications to finance. Springer Science & Business Media, 2011. 15

[154] D. Van Ravenzwaaij, P. Cassey, and S. D. Brown, "A simple introduction to Markov Chain Monte–Carlo sampling," Psychonomic bulletin & review, vol. 25, no. 1, pp. 143–154, 2018. 15

[155] A. Golightly and D. J. Wilkinson, "Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo," Interface Focus, vol. 1, no. 6, pp. 807–820, 2011. 15

[156] A. Gupta and J. B. Rawlings, "Comparison of parameter estimation methods in stochastic chemical kinetic models: examples in systems biology," AIChE Journal, vol. 60, no. 4, pp. 1253–1268, 2014. 15

[157] J. P. Huelsenbeck, B. Larget, and M. E. Alfaro, "Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo," Molecular Biology and Evolution, vol. 21, no. 6, pp. 1123–1133, 2004. 15

[158] W. A. Link, E. Cam, J. D. Nichols, and E. G. Cooch, "Of bugs and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research," The Journal of wildlife management, pp. 277–291, 2002. 15

[159] M. Jerrum and A. Sinclair, "The Markov chain Monte Carlo method: an approach to approximate counting and integration," Approximation Algorithms for NP-hard Problems, PWS Publishing, 1996. 15

[160] C. Borgs, J. T. Chayes, A. Frieze, J. H. Kim, P. Tetali, E. Vigoda, and V. H. Vu, "Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics," in 40th Annual Symposium on Foundations of Computer Science (Cat. No. 99CB37039), pp. 218–229, 1999. 15

[161] B. A. Berg and A. Billoire, "Markov chain Monte Carlo simulations," Wiley Encyclopedia of Computer Science and Engineering, 2007. 15

[162] S. Sharma, "Markov chain Monte Carlo methods for bayesian data analysis in astronomy," Annual Review of Astronomy and Astrophysics, vol. 55, pp. 213–259, 2017. 15

[163] M. Johnson, T. L. Griffiths, and S. Goldwater, "Bayesian inference for PCFGs via Markov chain Monte Carlo," in Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference, pp. 139–146, 2007. 15

[164] A. Kolmogoroff, "Über die analytischen methoden in der wahrscheinlichkeitsrechnung," Mathematische Annalen, vol. 104, no. 1, pp. 415–458, 1931. 16

[165] I. I. Gihman and A. V. Skorohod, The theory of stochastic processes I. Springer-Verlag, 1974. 16

[166] H. Cramér, "Sur un nouveau théorème-limite de la théorie des probabilités," Actual. Sci. Ind., vol. 736, pp. 5–23, 1938. 17

[167] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," The Annals of Mathematical Statistics, vol. 23, no. 4, pp. 493–507, 1952. 17

[168] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, I," Communications on Pure and Applied Mathematics, vol. 28, no. 1, pp. 1–47, 1975. 17

[169] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time, II," Communications on Pure and Applied Mathematics, vol. 28, no. 2, pp. 279–301, 1975. 17

[170] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time—III," Communications on pure and applied Mathematics, vol. 29, no. 4, pp. 389–461, 1976. 17

[171] M. D. Donsker and S. R. S. Varadhan, "Asymptotic evaluation of certain Markov process expectations for large time. IV," Communications on Pure and Applied Mathematics, vol. 36, no. 2, pp. 183–212, 1983. 17, 21

[172] M. I. Freidlin and A. D. Wentzell, "Random perturbations of dynamical systems," Springer-Verlag, New York, 1984. 17

[173] T. R. Gingrich, G. M. Rotskoff, S. Vaikuntanathan, and P. L. Geissler, "Efficiency and large deviations in time-asymmetric stochastic heat engines," New Journal of Physics, vol. 16, no. 10, p. 102003, 2014. 17

[174] F. Ragone, J. Wouters, and F. Bouchet, "Computation of extreme heat waves in climate models using a large deviation algorithm," Proceedings of the National Academy of Sciences, vol. 115, no. 1, pp. 24–29, 2018. 17

[175] Intergovernmental Panel on Climate Change, Managing the risks of extreme events and disasters to advance climate change adaption. Cambridge University Press, 2012. 17

[176] A. AghaKouchak, D. Easterling, K. Hsu, S. Schubert, and S. Sorooshian, Extremes in a changing climate: detection, analysis and uncertainty, vol. 65. Springer Science & Business Media, 2012. 17

[177] S. C. Herring, M. P. Hoerling, T. C. Peterson, and P. A. Stott, "Explaining extreme events of 2013 from a climate perspective," Bulletin of the American Meteorological Society, vol. 95, no. 9, pp. S1 – S104, 01 Sep. 2014. 17

[178] J. Gärtner, "On large deviations from the invariant measure," Theory of Probability & Its Applications, vol. 22, no. 1, pp. 24–39, 1977. 20

[179] R. S. Ellis, "Large deviations for a general class of random vectors," The Annals of Probability, vol. 12, no. 1, pp. 1–12, 1984. 20

[180] S. R. S. Varadhan, "Asymptotic probabilities and differential equations," Communications on Pure and Applied Mathematics, vol. 19, no. 3, pp. 261–286, 1966. 21

[181] R. T. Rockafellar, Convex Analysis. Princeton Landmarks in Mathematics and Physics, Princeton University Press, 1970. 21, 23

[182] H. Touchette, "Legendre-fenchel transforms in a nutshell," URL http://www. maths. qmul. ac. uk/˜ ht/archive/lfth2. pdf, 2005. 23

[183] J. Van Tiel, Convex Analysis: An Introductory Text. John Wiley & Sons, 1984. 23

[184] A. Bossavit, A course in Convex Analysis. Warsaw: Interdisciplinary Centrum for Mathematical and Computational Modelling (ICM), Jun. 2003. 23

[185] P. T. Landsberg, Thermodynamics and statistical mechanics. Dover, 1990. 24

[186] A. Sommerfeld, Thermodynamics and statistical mechanics. Academic Press, 1964. 24

[187] W. Greiner, L. Neise, and H. Stöcker, Thermodynamics and statistical mechanics. Springer-Verlag New York, 1995. 24

[188] K. Stowe, An introduction to thermodynamics and statistical mechanics. Cambridge University Press, 2nd ed., 2007. 24

[189] C. Texier and G. Roux, Physique statistique: des processus élémentaires aux phénomènes collectifs. Dunod, 2017. 24, 83

[190] R. Balescu, Equilibrium and nonequilibrium statistical mechanics. John Wiley & Sons, 1975. 24

[191] F. Bloch and J. D. Walecka, Fundamentals of statistical mechanics: manuscript and notes of Felix Bloch. World Scientific, 2000. 24

[192] C. S. Helrich, Modern thermodynamics with statistical mechanics. Springer-Verlag Berlin Heidelberg, 2009. 24

[193] B. Linder, Thermodynamics and introductory statistical mechanics. John Wiley & Sons, 2004. 24, 25

[194] H. Touchette, R. S. Ellis, and B. Turkington, "An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles," Physica A: Statistical Mechanics and its Applications, vol. 340, no. 1-3, pp. 138–146, 2004. 25

[195] H. Touchette, "Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels," Journal of Statistical Physics, vol. 159, pp. 987–1016, feb 2015. 25, 37, 39, 52, 117

[196] J. M. Seddon and J. D. Gale, Thermodynamics and statistical mechanics. Royal Society of Chemistry, 2001. 25

[197] M. Kaufman, Principles of thermodynamics. CRC Press, 2002. 25

[198] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, "Nonequilibrium energetics of a single F1-ATPase molecule," Physical Review Letters, vol. 104, no. 19, p. 198103, 2010. 26

[199] A. Saha and R. Marathe, "Stochastic work extraction in a colloidal heat engine in the presence of colored noise," Journal of Statistical Mechanics: Theory and Experiment, vol. 2019, no. 9, p. 094012, 2019. 26

[200] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, "Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales," Physical Review Letters, vol. 89, p. 050601, Jul 2002. 26, 31

[201] G. M. Wang, J. C. Reid, D. M. Carberry, D. R. M. Williams, E. M. Sevick, and D. J. Evans, "Experimental study of the fluctuation theorem in a nonequilibrium steady state," Physical Review E, vol. 71, p. 046142, Apr 2005. 26, 31

[202] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, "Fluctuations and irreversibility: An experimental demonstration of a second-law-like theorem using a colloidal particle held in an optical trap," Physical Review Letters, vol. 92, p. 140601, Apr 2004. 26, 31

[203] A. Imparato, L. Peliti, G. Pesce, G. Rusciano, and A. Sasso, "Work and heat probability distribution of an optically driven Brownian particle: Theory and experiments," Physical Review E, vol. 76, p. 050101, Nov 2007. 26

[204] S. Schuler, T. Speck, C. Tietz, J. Wrachtrup, and U. Seifert, "Experimental test of the fluctuation theorem for a driven two-level system with time-dependent rates," Physical Review Letters, vol. 94, p. 180602, May 2005. 26, 31

[205] F. Douarche, S. Ciliberto, A. Petrosyan, and I. Rabbiosi, "An experimental test of the Jarzynski equality in a mechanical experiment," EPL (Europhysics Letters), vol. 70, no. 5, p. 593, 2005. 26

[206] J. R. Gomez-Solano, A. Petrosyan, S. Ciliberto, R. Chetrite, and K. Gawędzki, "Experimental verification of a modified fluctuation-dissipation relation for a micron-sized particle in a nonequilibrium steady state," Physical Review Letters, vol. 103, p. 040601, Jul 2009. 26

[207] S. Ciliberto and C. Laroche, "An experimental test of the Gallavotti-Cohen fluctuation theorem," Le Journal de Physique IV, vol. 8, no. PR6, pp. Pr6–215–Pr6–219, 1998. 26

[208] C. Van den Broeck and M. Esposito, "Ensemble and trajectory thermodynamics: A brief introduction," Physica A: Statistical Mechanics and its Applications, vol. 418, pp. 6–16, 2015. 28

[209] A. C. Barato and U. Seifert, "Stochastic thermodynamics with information reservoirs," Physical Review E, vol. 90, no. 4, p. 042150, 2014. 29

[210] G. Verley, Fluctuations et réponse des systèmes hors d'équillibre. PhD thesis, Université Pierre et Marie Curie, 2012. 29, 32

[211] M. Esposito, "Stochastic thermodynamics under coarse graining," Physical Review E, vol. 85, p. 041125, Apr 2012. 29

[212] G. E. Crooks, "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences," Physical Review E, vol. 60, no. 3, p. 2721, 1999. 30

[213] C. Jarzynski, "Hamiltonian derivation of a detailed fluctuation theorem," Journal of Statistical Physics, vol. 98, no. 1, pp. 77–102, 2000. 30

[214] M. Esposito and C. Van den Broeck, "Three detailed fluctuation theorems," Physical Review Letters, vol. 104, p. 090601, Mar 2010. 30

[215] D. Andrieux and P. Gaspard, "Fluctuation theorem for currents and Schnakenberg network theory," Journal of Statistical Physics, vol. 127, no. 1, pp. 107–131, 2007. 30

[216] D. J. Evans, D. J. Searles, and S. R. Williams, "On the fluctuation theorem for the dissipation function and its connection with response theory," The Journal of chemical physics, vol. 128, no. 1, p. 014504, 2008. 30

[217] D. J. Evans and D. J. Searles, "The fluctuation theorem," Advances in Physics, vol. 51, no. 7, pp. 1529–1585, 2002. 30

[218] S. Ciliberto, N. Garnier, S. Hernandez, C. Lacpatia, J.-F. Pinton, and G. R. Chavarria, "Experimental test of the Gallavotti–Cohen fluctuation theorem in turbulent flows," Physica A: Statistical Mechanics and its Applications, vol. 340, no. 1-3, pp. 240–250, 2004. 31

[219] B. Derrida and J. L. Lebowitz, "Exact large deviation function in the asymmetric exclusion process," Physical Review Letters, vol. 80, no. 2, p. 209, 1998. 32

[220] H. Risken, The Fokker-Planck equation. Springer, 2nd ed., 1989. 33, 35

[221] B. Øksendal, Stochastic differential equations. Springer, 6th ed., 2003. 34

[222] A. W. C. Lau and T. C. Lubensky, "State-dependent diffusion: Thermodynamic consistency and its path integral formulation," Physical Review E, vol. 76, no. 1, p. 011123, 2007. 34, 40, 140

[223] C. Van den Broeck and M. Esposito, "Three faces of the second law. II. Fokker-Planck formulation," Physical Review E, vol. 82, no. 1, p. 011144, 2010. 36

[224] H. Touchette, "Ensemble equivalence for general many-body systems," EPL (Europhysics Letters), vol. 96, no. 5, p. 50010, 2011. 37, 52

[225] N. Tizón-Escamilla, V. Lecomte, and E. Bertin, "Effective driven dynamics for one-dimensional conditioned Langevin processes in the weak-noise limit," vol. 2019, p. 013201, jan 2019. 44

[226] B. Derrida and T. Sadhu, "Large deviations conditioned on large deviations I: Markov chain and Langevin equation," Journal of Statistical Physics, vol. 176, no. 4, pp. 773–805, 2019. 44

[227] L. Chabane, R. Chétrite, and G. Verley, "Periodically driven jump processes conditioned on large deviations," Journal of Statistical Mechanics: Theory and Experiment, vol. 2020, no. 3, p. 033208, 2020. 45

[228] G. Verley, C. Van den Broeck, and M. Esposito, "Modulated two-level system: Exact work statistics," Physical Review E, vol. 88, no. 3, p. 032137, 2013. 47

[229] R. Chetrite and S. Gupta, "Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale," Journal of Statistical Physics, vol. 143, pp. 543–584, apr 2011. 48

[230] P. T. Nyawo and H. Touchette, "A minimal model of dynamical phase transition," EPL (Europhysics Letters), vol. 116, no. 5, p. 50009, 2017. 52, 74

[231] P. T. Nyawo and H. Touchette, "Dynamical phase transition in drifted Brownian motion," Physical Review E, vol. 98, no. 5, p. 052103, 2018. 52, 74

[232] C. Monthus, "Non-equilibrium steady states: maximization of the shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints," Journal of Statistical Mechanics: Theory and Experiment, vol. 2011, p. P03008, mar 2011. 53

[233] G. Verley, "Nonequilibrium thermodynamic potentials for continuous-time Markov chains," Physical Review E, vol. 93, no. 1, p. 012111, 2016. 53, 57

[234] L. Bertini, R. Chetrite, A. Faggionato, and D. Gabrielli, "Level 2.5 large deviations for continuous-time Markov chains with time periodic rates," Annales Henri Poincaré, vol. 19, pp. 3197–3238, Oct 2018. 54, 79

[235] M. G. Krein and M. A. Rutman, "Linear operators leaving invariant a cone in a banach space," American Mathematical Society Translations: Series 1, no. 26, 1950. Translated from Uspekhi Matematicheskikh Nauk, 1948. 64, 79

[236] N. D. Alikakos and G. Fusco, "A dynamical systems proof of the Krein–Rutman Theorem and an extension of the Perron Theorem," Proceedings of the Royal Society of Edinburgh: Section A Mathematics, vol. 117, no. 3-4, p. 209–214, 1991. 64, 79

[237] M. Kac, "On distributions of certain Wiener functionals," Transactions of the American Mathematical Society, vol. 65, no. 1, pp. 1–13, 1949. 67

[238] D. Williams, "Review: D. W. Stroock and S. R. S. Varadhan, Multidimensional diffusion processes," Bulletin (New Series) of the American Mathematical Society, vol. 2, no. 3, pp. 496–503, 1980. 67

[239] D. Revuz and M. Yor, "Continuous martingales and Brownian motion," Springer-Verlag Berlin Heidelberg, 1999. 67

[240] N. Singh and B. Wynants, "Dynamical fluctuations for periodically driven diffusions," Journal of Statistical Mechanics Theory and Experiment, no. 03, p. P03007, 2010. 70, 81

[241] R. Chetrite, "Pérégrinations sur les phénomènes aléatoires dans la nature," HDR thesis, Université de Nice-Sophia-Antipolis, France, 2018. 74

[242] L. Y. Adrianova, Introduction to Linear Systems of Differential Equation. American Mathematical Society, 1995. 76

[243] R. A. Horn and C. R. Johnson, Matrix Analysis. Cambridge University Press, 2nd ed., 2013. 78

[244] V. S. Bokharaie, Stability analysis of positive systems with applications to epidemiology. PhD thesis, National University of Ireland Maynooth, 2012. 78, 79

[245] K. Deimling, Nonlinear functional analysis. Springer-Verlag, 1985. 79

[246] O. Bai, M. Nakamura, A. Ikeda, and H. Shibasaki, "Nonlinear Markov process amplitude EEG model for nonlinear coupling interaction of spontaneous EEG," IEEE transactions on biomedical engineering, vol. 47, no. 9, pp. 1141–1146, 2000. 86

[247] K. Burrage, P. M. Burrage, A. Leier, T. Marquez-Lago, and D. V. Nicolau, "Stochastic simulation for spatial modelling of dynamic processes in a living cell," in Design and Analysis of Biomolecular Circuits, pp. 43–62, Springer, 2011. 86

[248] M. Delbrück, "Statistical fluctuations in autocatalytic reactions," The Journal of Chemical Physics, vol. 8, no. 1, pp. 120–124, 1940. 86

[249] A. F. Bartholomay, "Stochastic models for chemical reactions: II. the unimolecular rate constant," The bulletin of mathematical biophysics, vol. 21, no. 4, pp. 363–373, 1959. 86

[250] D. A. McQuarrie, "Stochastic approach to chemical kinetics," Journal of Applied Probability, vol. 4, no. 3, pp. 413–478, 1967. 86

[251] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," The journal of physical chemistry, vol. 81, no. 25, pp. 2340–2361, 1977. 86

[252] T. G. Kurtz, "The relationship between stochastic and deterministic models for chemical reactions," The Journal of Chemical Physics, vol. 57, no. 7, pp. 2976–2978, 1972. 86

[253] E. Levine and T. Hwa, "Stochastic fluctuations in metabolic pathways," Proceedings of the National Academy of Sciences, vol. 104, no. 22, pp. 9224–9229, 2007. 86

[254] S. Huang, F. Li, J. X. Zhou, and H. Qian, "Processes on the emergent landscapes of biochemical reaction networks and heterogeneous cell population dynamics: differentiation in living matters," Journal of the Royal Society Interface, vol. 14, no. 130, p. 20170097, 2017. 86

[255] J. Goutsias and G. Jenkinson, "Markovian dynamics on complex reaction networks," Physics Reports, vol. 529, no. 2, pp. 199–264, 2013. 86

[256] T. D. Frank, "Markov chains of nonlinear Markov processes and an application to a winner-takes-all model for social conformity," Journal of Physics A: Mathematical and Theoretical, vol. 41, no. 28, p. 282001, 2008. 86

[257] T. D. Frank, "Deterministic and stochastic components of nonlinear Markov models with an application to decision making during the bailout votes 2008 (USA)," The European Physical Journal B, vol. 70, no. 2, pp. 249–255, 2009. 86

[258] S. Galam and J.-D. Zucker, "From individual choice to group decision-making," Physica A: Statistical Mechanics and its Applications, vol. 287, no. 3-4, pp. 644–659, 2000. 86

[259] F. Schweitzer and J. Hołyst, "Modelling collective opinion formation by means of active Brownian particles," The European Physical Journal B - Condensed Matter and Complex Systems, vol. 15, no. 4, pp. 723–732, 2000. 86

[260] J. T. Wixted and D. Rohrer, "Analyzing the dynamics of free recall: An integrative review of the empirical literature," Psychonomic Bulletin & Review, vol. 1, no. 1, pp. 89–106, 1994. 86

[261] T. D. Frank and T. Rhodes, "Micro-dynamic associated with two-state nonlinear Markov processes: with an application to free recall," Fluctuation and Noise Letters, vol. 10, no. 01, pp. 41–58, 2011. 86

[262] B. Cleuren and C. Van den Broeck, "Ising model for a Brownian donkey," EPL (Europhysics Letters), vol. 54, no. 1, p. 1, 2001. 87, 118

[263] L. Chabane, A. Lazarescu, and G. Verley, "Conditioning Markov processes: a Lagrangian and Hamiltonian perspective," arXiv eprint 2109.06830, 2021. 87, 111

[264] M. Feinberg, "Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems," Chemical Engineering Science, vol. 42, no. 10, pp. 2229–2268, 1987. 91

[265] J.-M. Rax, Mécanique analytique: Adiabaticité, résonances, chaos. Sciences Sup, Dunod, 2020. 95, 133, 136

[266] S. H. Strogatz, Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering. Perseus Books Publishing, 1994. 96

[267] T. Herpich, T. Cossetto, G. Falasco, and M. Esposito, "Stochastic thermodynamics of all-to-all interacting many-body systems," New Journal of Physics, vol. 22, no. 6, p. 063005, 2020. 136

[268] F. Bouchet, K. Gawędzki, and C. Nardini, "Perturbative calculation of quasi-potential in non-equilibrium diffusions: a mean-field example," Journal of Statistical Physics, vol. 163, no. 5, pp. 1157–1210, 2016. 136, 141

[269] D. S. Dean, "Langevin equation for the density of a system of interacting Langevin processes," Journal of Physics A: Mathematical and General, vol. 29, no. 24, p. L613, 1996. 141

[270] L. P. Fischer, P. Pietzonka, and U. Seifert, "Large deviation function for a driven underdamped particle in a periodic potential," Physical Review E, vol. 97, no. 2, p. 022143, 2018. 146

[271] I. Oppenheim, K. E. Shuler, and G. H. Weiss, "Stochastic theory of nonlinear rate processes with multiple stationary states," Physica A: Statistical Mechanics and its Applications, vol. 88, no. 2, pp. 191–214, 1977. 147

[272] M. Esposito, U. Harbola, and S. Mukamel, "Nonequilibrium fluctuations, fluctuation theorems, and counting statistics in quantum systems," Reviews of Modern Physics, vol. 81, no. 4, p. 1665, 2009. 147

[273] W. C. Schieve and L. P. Horwitz, Quantum statistical mechanics. Cambridge University Press, 2009. 147

[274] N. N. Bogolubov and N. N. Bogolubov Jr, Introduction to quantum statistical mechanics. World Scientific Publishing Company, 2009. 147

[275] K. Netočnỳ and F. Redig, "Large deviations for quantum spin systems," Journal of Statistical Physics, vol. 117, no. 3, pp. 521–547, 2004. 147

[276] M. Žnidarič, "Exact large-deviation statistics for a nonequilibrium quantum spin chain," Physical Review Letters, vol. 112, p. 040602, Jan 2014. 147

[277] M. Žnidarič, "Large-deviation statistics of a diffusive quantum spin chain and the additivity principle," Physical Review E, vol. 89, no. 4, p. 042140, 2014. 147

[278] T. Denzler and E. Lutz, "Efficiency large deviation function of quantum heat engines," New Journal of Physics, vol. 23, no. 7, p. 075003, 2021. 147

[279] F. Carollo, R. L. Jack, and J. P. Garrahan, "Unraveling the large deviation statistics of Markovian open quantum systems," Physical Review Letters, vol. 122, no. 13, p. 130605, 2019. 147

université PARIS-SACLAY
ÉCOLE DOCTORALE
Physique en île-de-France (EDPIF)

Titre : De la rareté à la typicité : le parcours improbable d'une grande déviation

Mots clés : Grandes déviations, Processus de Markov, Physique statistique, Hors équilibre, Fluctuations

Résumé : Le problème du conditionnement de processus de Markov homogènes en temps sur une fluctuation rare a été étudié dans le cadre de la théorie des grandes déviations. Sur cette base, un nouveau processus équivalent au processus conditionné a été introduit en utilisant la transformée de Doob généralisée : il s'agit du « processus drivé ». Dans cette thèse, on ambitionne de généraliser ces résultats à une classe plus large de processus de Markov. Dans la première partie de ce manuscrit, on considère des processus de Markov conduits périodiquement, caractérisés par des générateurs périodiques. On veut conditionner ces processus sur des observables définies via des fonctions périodiques en temps. En adaptant les résultats du cas homogène en temps, on construit le processus drivé pour lequel les valeurs typiques de nos observables après un grand nombre de périodes correspondent aux valeurs utilisées pour le conditionnement. Dans le cas périodique, les générateurs indépendants du temps deviennent périodiques, les exponentielles de matrices deviennent des exponentielles ordonnées en temps et les problèmes spectraux deviennent des équations différentielles du premier ordre. Le processus drivé s'obtient soit en utilisant l'équivalence de probabilités de chemin, soit à partir d'un problème d'optimisation de fonctions de grandes déviations. Dans la deuxième partie de ce manuscrit, nous étendons ces résultats au cas général des processus de Markov non linéaires décrits par des lagrangiens et des hamiltoniens indépendants du temps. Dans ce nouveau formalisme, la transformée de Doob généralisée menant vers le processus drivé se traduit par une transformation canonique sur les hamiltoniens. Cette transformation — que l'on appellera « rectification » — nécessite d'étudier l'analogue non linéaire du théorème de Perron-Frobenius. Cette étude nous a conduits à conjecturer une classification des solutions d'une équation de Hamilton-Jacobi. Nous concluons cette partie par une ouverture sur le problème du conditionnement des processus non linéaires conduits périodiquement.

Title : From rarity to typicality: the improbable journey of a large deviation

Keywords : Large deviations, Markov processes, Statistical physics, Out of equilibrium, Fluctuations

Abstract : The problem of conditioning time-homogeneous Markov processes on a rare fluctuation has been studied within the framework of large deviation theory. On this basis, a new process equivalent to the conditioned process has been introduced using the generalized Doob transform: it is the "driven process". In this thesis, we aim to generalize these results to a larger class of Markov processes. In the first part of this manuscript, we consider periodically driven Markov processes, characterized by their time-periodic generators. We are interested in conditioning these processes on observables defined through time-periodic functions. Adapting the results of the time-homogeneous case, we derive the driven process for which the typical values of our observables after a large number of periods correspond to the values used for the conditioning. In the periodic case, time-independent generators become time-periodic, matrix exponentials become time-ordered exponentials and spectral problems become first order differential equations. The driven process can be derived either using path ensemble equivalence, or from an optimization problem on large deviation functions. In the second part of this manuscript, we extend these results to the general case of nonlinear Markov processes described by time-independent Lagrangians and Hamiltonians. In this new formalism, the generalized Doob transform leading to the driven process translates into a canonical transformation on Hamiltonians. This transformation — that we call "rectification" — requires to investigate the nonlinear counterpart of the Perron-Frobenius theorem. This investigation led us to conjecture a classification of the solutions of a Hamilton-Jacobi equation. We conclude this part by an opening on the problem of conditioning periodically driven nonlinear processes.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France