Master equations

From path integrals, to bosons, to bacteria

Markus F. Weber

Munich 2016

Revised version

This document is optimized for on-screen reading.

Master equations

From path integrals, to bosons, to bacteria

Markus F. Weber

A dissertation submitted
to the Faculty of Physics at the
Ludwig–Maximilians–Universität München
for the degree of
Doctor rerum naturalium

Munich, 8th September 2016

First referee: Prof. Dr. Erwin Frey
Second referee: Prof. Dr. Mauro Mobilia
Day of the oral examination: 27th October 2016

Zusammenfassung (Summary in German)

Die Dynamik eines komplexen physikalischen, biologischen oder chemischen Systems kann oft durch einen zeitkontinuierlichen Markov-Prozess beschrieben werden. Die grundlegenden Gleichungen dieser stochastischen Prozesse sind die Fokker-Planck- und die Mastergleichung. Beide Gleichungen nehmen an, dass die Zukunft eines Systems allein durch seinen gegenwärtigen Zustand bestimmt ist. Während die Fokker-Planck-Gleichung ein System modelliert, dessen Zustand sich mit der Zeit kontinuierlich ändert, beschreibt die Mastergleichung ein System, das zwischen seinen Zuständen springt. In dieser Arbeit studieren wir die Theorie und Anwendung dieser Gleichungen, mit einem Schwerpunkt auf der Mastergleichung.

I Mastergleichungen und die Theorie stochastischer Pfadintegrale
Nach Herleitung der Mastergleichung und ihres zeitinvertierten Gegenstücks im ersten Kapitel, diskutieren wir analytische und numerische Methoden für ihre Auswertung. Wir formulieren allgemeine Bedingungen unter denen die beiden Mastergleichungen in vier lineare partielle Differentialgleichungen transformiert werden können. Der hierfür entwickelte mathematische Formalismus verbindet verschiedene analytische Methoden, insbesondere Spektralmethoden und die sogenannte Poisson-Darstellung. Die Transformation der Mastergleichungen ermöglicht es uns, ihre Lösung durch zwei Pfadintegrale darzustellen. Beide Pfadintegrale werden in der Lösung elementarer stochastischer Prozesse veranschaulicht. Zudem diskutieren wir verschiedene Pfadintegraldarstellungen der Fokker-Planck-Gleichung.

II Evolutionäre Nullsummenspiele und getriebene, dissipative Quantensysteme
Im Anschluss an die allgemeine Theorie studieren wir die Evolution von Nullsummenspielen in endlichen Populationen. Diese Spiele werden typischerweise im Rahmen der evolutionären Spieltheorie studiert, doch ihre Mastergleichung beschreibt auch die Dynamik von nicht-interagierenden Bosonen in getriebenen, dissipativen Quantensystemen. Unsere Erklärung einer für diese Systeme vorhergesagten Kondensation, sowie der Selektion von persistenten Strategien in evolutionären Nullsummenspielen, wurde in zwei hier nachgedruckten Artikeln veröffentlicht.

III Die stochastische Dynamik expandierender Bakterienkolonien
Im abschließenden Kapitel verwenden wir die Mastergleichung zur Modellierung expandierender Bakterienkolonien von toxin-produzierenden, sensitiven und resistenten *Escherichia coli* Stämmen. In Kooperation mit der Gruppe von M. Opitz untersuchten wir, welche Faktoren das Überleben der Stämme bedingen. Unser unabhängig kalibriertes Modell reproduzierte die in Experimenten gesehenen überlebenden Stämme und ermöglichte die Vorhersage von Parametern für ihre Koexistenz. Unsere Ergebnisse wurden in einem hier nachgedruckten Artikel veröffentlicht.

Projects and contributions

The work presented in this thesis is divided into the following three chapters. Each of the chapters represents a project on which I worked during my doctoral studies. The projects are summarized on the next pages.

I Master equations and the theory of stochastic path integrals
with Erwin Frey.
The first chapter provides a comprehensive account of analytical and numerical approaches to master equations. We develop a unified mathematical framework for the study of master equations and derive new methods for their analysis. Based on the framework, we formulate general conditions under which master equations can be represented by two path integrals. The use of these path integrals is demonstrated for various elementary processes. Due to the broad scope of this chapter, its introduction and conclusion serve the whole thesis. An abridged version of the first chapter has been accepted for publication in *Reports on Progress in Physics*.

II Evolutionary zero-sum games and driven-dissipative quantum systems
with Johannes Knebel, Torben Krüger, and Erwin Frey.
The second chapter focuses on zero-sum games. These games are studied in evolutionary game theory but their master equation was also recently derived in a study on the condensation of bosons in driven-dissipative quantum systems [1]. We presented our results on these topics in the publication "Coexistence and survival in conservative Lotka-Volterra networks", *Phys. Rev. Lett.* **110**(16), 168106 (2013), to which I contributed as third author, and in the publication "Evolutionary games of condensates in coupled birth-death processes", *Nat. Commun.* **6**, 6977 (2015), to which I contributed as co-first author. The publications are reprinted in sections 2 and 3 of chapter II and will also be reprinted in the doctoral thesis of J. Knebel.

III The stochastic dynamics of bacterial range expansions
with Gabriele Poxleitner, Elke Hebisch, Erwin Frey, and Madeleine Opitz.
In the third chapter, we study range expansions of toxin-producing, sensitive, and resistant strains of *Escherichia coli*. To understand the factors determining the survival of strains, we modelled their coarse-grained dynamics using a master equation. Our independently calibrated model reproduced the surviving strains as seen in experiments and predicted experimental parameters for three-strain coexistence. The experiments were conducted by G. Poxleitner and E. Hebisch from the group of M. Opitz. Our results were published in "Chemical warfare and survival strategies in bacterial range expansions", *J. R. Soc. Interface* **11**(96), 20140172 (2014), to which I contributed as first author. The publication is reprinted in section 2 of chapter III.

Abstracts of the projects

The dynamics of a complex physical, biological, or chemical systems can often be modelled in terms of a continuous-time Markov process. The governing equations of these processes are the Fokker-Planck and the master equation. Both equations assume that the future of a system depends only on its current state, memories of its past having been wiped out by randomizing forces. Whereas the Fokker-Planck equation describes a system that evolves continuously from one state to another, the master equation models a system that performs jumps between its states.

In this thesis, we focus on master equations. We first present a comprehensive mathematical framework for the analytical and numerical analysis of master equations in chapter I. Special attention is given to their representation by path integrals. In the subsequent chapters, master equations are applied to the study of physical and biological systems. In chapter II, we study the stochastic and deterministic evolution of zero-sum games and thereby explain a condensation phenomenon expected in driven-dissipative bosonic quantum systems. Afterwards, in chapter III, we develop a coarse-grained model of microbial range expansions and use it to predict which of three strains of *Escherichia coli* survive such an expansion.

Master equations and the theory of stochastic path integrals
Mathematical analyses of master equations are often restricted to a regime of low noise, requiring ad-hoc closure schemes for the derivation of low-order moment equations, or relying on approximations such as the Kramers-Moyal or the system size expansion. In the first chapter of this thesis, we discuss methods going beyond the low-noise regime. We first derive the (forward) master equation and its backward counterpart from the Chapman-Kolmogorov equation and discuss basic analytical and numerical methods for their solution. Afterwards, we formulate general conditions under which the forward and the backward master equation can be cast into either of four partial differential equations (PDEs). Three of these PDEs are discussed in detail. The first PDE governs the time evolution of a generalized probability generating function whose basis depends on the stochastic process under consideration. Various spectral methods and a variational approach have recently been proposed for its analysis. The second PDE is novel and is obeyed by a distribution that is marginalized over an initial state. It proves useful for the computation of mean extinction times. The third PDE describes the time evolution of a "generating functional", which generalizes the so-called Poisson representation. The solutions of the PDEs are subsequently expressed in terms of two path integrals: a "forward" and a "backward" path integral. Combined with inverse transformations, one obtains two distinct path integral

representations of the conditional probability distribution solving the master equations. We exemplify both path integrals in analysing elementary chemical reactions. Moreover, we show how a well-known path integral representation of averaged observables can be recovered from them. Upon expanding the forward and the backward path integrals around stationary paths, we then discuss and extend a recent method for the computation of rare event probabilities. Besides, we also derive path integral representations for processes with continuous state spaces whose forward and backward master equations admit Kramers-Moyal expansions. A truncation of the backward expansion at the level of a diffusion approximation recovers a classic path integral representation of the (backward) Fokker-Planck equation. One can rewrite this path integral in terms of an Onsager-Machlup function and, for purely diffusive Brownian motion, it simplifies to the path integral of Wiener. Our discussion of the path integrals employs the language of probability rather than quantum (field) theory to elucidate the probabilistic structures behind coherent states, the Doi-shift, and normal-ordered observables.

Evolutionary zero-sum games and driven-dissipative quantum systems
In the second chapter, we study the stochastic and deterministic evolution of zero-sum games and their connection to driven-dissipative bosonic quantum systems. In a zero-sum game, each individual in a population of "agents" is assigned a particular strategy. The strategies of two agents are either mutually neutral, meaning that the two agents do not interact with one another, or one of their strategies dominates the other one. The agent with the inferior strategy then adopts the superior strategy when playing a game. As time progresses, the number of agents playing a particular strategy changes. Assuming that time progresses continuously, one can describe these changes by a (forward) master equation. Surprisingly, the master equation of zero-sum games was also recently derived in a study on the condensation of non-interacting bosons in driven-dissipative quantum [1]. The authors of that study observed that over time, the bosons cluster in certain condensate states while the other states in the system are being depleted. Similarly, we observed that the agents in a zero-sum game typically adopt certain persistent strategies while abandoning other strategies.

After establishing that on the leading-order time scale, the dynamics of a zero-sum game are described by the antisymmetric Lotka-Volterra equation, we could show that the selection of persistent strategies is governed by the vanishing of relative entropy production. The same principle underlies the selection of condensates in a driven-dissipative quantum system. Our insight enabled us to devise an efficient algebraic algorithm for the identification of the persistent strategies (condensates). This algorithm allowed us to study zero-sum games (quantum systems) with hundreds of strategies (states) and to explore the interplay between the condensation dynamics and the critical properties of random networks. Our results were presented in two publications [2, 3], which are included as reprints.

The stochastic dynamics of bacterial range expansions
In the third chapter, we focus on microbial range expansions. A microbial range expansion occurs after a droplet of bacterial suspension is inoculated onto an agar plate. In collaboration with the microbiology group of Madeleine Opitz, we explored the determinants of bacterial coexistence when a suspension contains three strains of *Escherichia coli*: a toxin-producing, a sensitive, and a resistant strain. These strains have been proposed as a bacterial model system for the study of non-transitive, cyclic communities [4]. In such a community, each strain dominates one other strain and is being dominated by a third strain. Thus, the hierarchy in the community resembles a "rock-paper-scissors game". By using genetic engineering to tune strain growth rates, we were in fact able to implement such a cyclic dominance between the strains as well as two other ecological scenarios.

A bacterial strain is considered a survivor of a range expansion if it managed to persist along the leading edge of an expanding colony. In order to understand the factors that decide whether a bacterial strain typically survives a range expansion, we developed a stochastic lattice-based model of range expansions based on a master equation. This master equation was evaluated numerically using the stochastic simulation algorithm of Gillespie. The key parameters of the model were identified as the initial ratios of the strains in the bacterial suspension, the growth rates of the strains, and the range and strength of the toxin produced by one of the strains (colicin E2). The values of the parameters were calibrated using independent experiments on the growth of single-strain colonies. The calibrated model successfully reproduced the surviving strains as seen in experiments and predicted parameters at which three-strain coexistence could be observed experimentally. Our analysis suggested that three-strain coexistence requires a balance between strain growth rates and either a reduced initial ratio of the toxin-producing strain, or a sufficiently short toxin range. The results of our work were presented in the publication [5], which is included as a reprint.

Contents

I Master equations and the theory of stochastic path integrals

1 Introduction

1.1 Scope of this chapter

The theory of continuous-time Markov processes is largely built on two equations: the Fokker-Planck [6–9] and the master equation [9, 10]. Both equations assume that the future of a system depends only on its current state, memories of its past having been wiped out by randomizing forces. This *Markov* assumption is sufficient to derive either of the two equations. Whereas the Fokker-Planck equation describes systems that evolve continuously from one state to another, the master equation models systems that perform jumps in state space.

Path integral representations of the master equation were first derived around 1980 [11–20], shortly after such representations had been derived for the Fokker-Planck equation [21–24]. Both approaches were heavily influenced by quantum theory, introducing such concepts as the Fock space [25] with its "bras" and "kets" [26], coherent states [27–29], and "normal-ordering" [30] into non-equilibrium theory. Some of these concepts are now well established and the original "bosonic" path integral representation has been complemented with a "fermionic" counterpart [31–36]. Nevertheless, we feel that the theory of these "stochastic" path integrals may benefit from a step back and a closer look at the probabilistic structures behind the integrals. Therefore, the objects imported from quantum theory make place for their counterparts from probability theory in this text. For example, the coherent states give way to the Poisson distribution. Moreover, we use the bras and kets as particular basis functionals and functions whose choice depends on the stochastic process at hand (a functional maps functions to numbers). Upon choosing the basis functions as Poisson distributions, one can thereby recover both a classic path integral representation of averaged observables as well as the Poisson representation of Gardiner and Chaturvedi [37, 38]. The framework presented in this chapter integrates a variety of different approaches to the master equation. Besides the Poisson representation, these approaches include a spectral method for the computation of stationary probability distributions [39], WKB approximations and other "semi-classical" methods for the computation of rare event probabilities [40–43], and a variational approach that was proposed in the context of stochastic gene expression [44]. All of these approaches can be treated

within a unified framework. Knowledge about this common framework makes it possible to systematically search for new ways of solving the master equation.

Before outlining the organization of this chapter, let us note that by focusing on the above path integral representations of master and Fokker-Planck equations, we neglect several other "stochastic" or "statistical" path integrals that have been developed. These include Edwards path integral approach to turbulence [45, 46], a path integral representation of Haken [47], path integral representations of non-Markov processes [48–63] and of polymers [64–67], and a representation of "hybrid" processes [68–70]. The dynamics of these stochastic hybrid processes are piecewise-deterministic. Moreover, we do not discuss the application of renormalization group techniques, despite their significant importance. Excellent texts exploring these techniques in the context of non-equilibrium critical phenomena [71–73] are provided by the review of Täuber, Howard, and Vollmayr-Lee [74] as well as the book by Täuber [75]. Our main interest lies in a mathematical framework unifying the different approaches from the previous paragraph and in two path integrals that are based on this framework. Both of these path integrals provide exact representations of the conditional probability distribution solving the master equation. We exemplify the use of the path integrals for elementary processes, which we choose for their pedagogic value. Most of these processes do not involve spatial degrees of freedom but the application of the presented methods to processes on spatial lattices or networks is straightforward. A process with diffusion and linear decay serves as an example of how path integrals can be evaluated perturbatively using Feynman diagrams. The particles' linear decay is treated as a perturbation to their free diffusion. The procedure readily extends to more complex processes. Moreover, we show how the two path integrals can be used for the computation of rare event probabilities. Let us emphasize that we only consider Markov processes obeying the Chapman-Kolmogorov equation and associated master equations [9]. It may be interesting to extend the discussed methods to "generalized" or "physical" master equations with memory kernels [76–79].

1.2 Organization of this chapter

The organization of this chapter is summarized in figure 1 and is as follows. In the next section 1.3, we introduce the basic concepts of the theory of continuous-time Markov processes. After discussing the roles of the forward and backward Fokker-Planck equations for processes with continuous sample paths, we turn to processes with discontinuous sample paths. The probability of finding such a "jump process" in a generic state n at time $\tau > t_0$, given that the process has been in state n_0 at time t_0, is represented by the conditional probability distribution $p(\tau, n|t_0, n_0)$. Whereas the forward master equation evolves this probability distribution forward in time, starting out at time $\tau = t_0$, the backward master equation evolves the

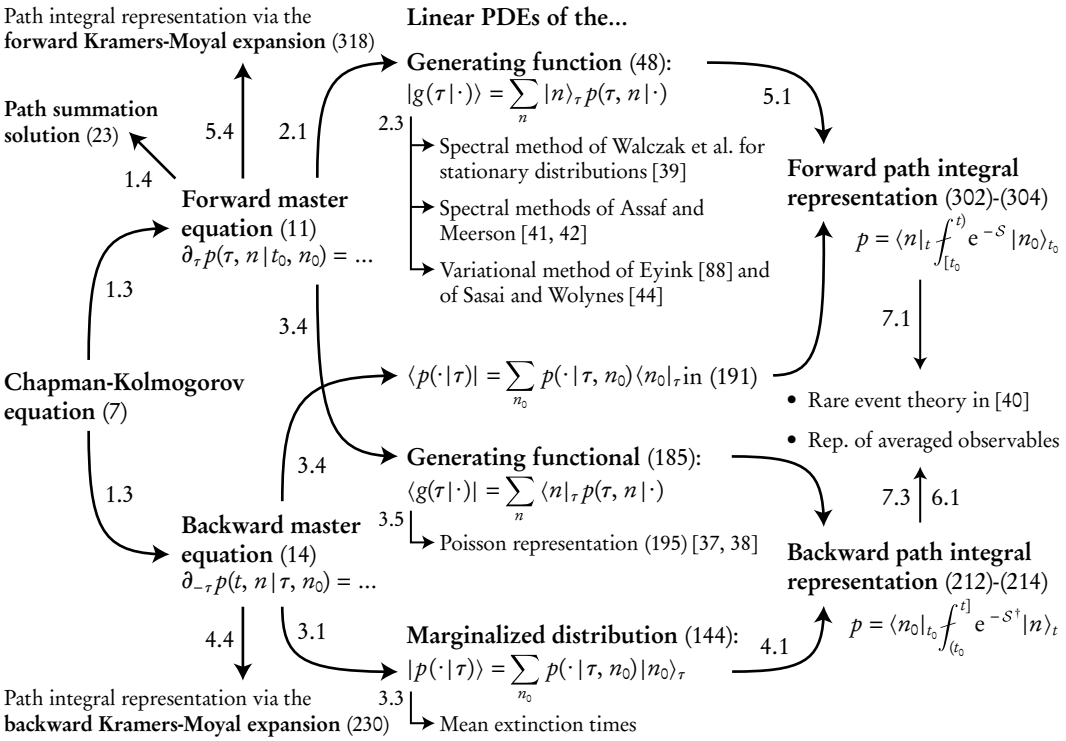


Figure 1 Roadmap and summary of the methods considered in this chapter. The arrows represent possible routes for derivations. Labelled arrows represent derivations that are explicitly treated in the respective sections. For example, the forward and backward master equations are derived from the Chapman-Kolmogorov equation in section 1.3. In this section, we also discuss the path summation representation (23) of the conditional probability distribution $p(\tau, n|t_0, n_0)$. This representation can be derived by examining the stochastic simulation algorithm (SSA) of Gillespie [80–83] or by performing a Laplace transformation of the forward master equation (11) (cf. appendix B). In sections 2 and 3, the forward and the backward master equations are cast into four linear PDEs, also called "flow equations". These equations are obeyed by a probability generating function, a probability generating functional, a marginalized distribution, and a further series expansion. The flow equations can be solved in terms of a forward and a backward path integral as shown in sections 4 and 5. Upon performing inverse transformations, the path integrals provide two distinct representations of the conditional probability distribution solving the master equations. Moreover, they can be used to represent averaged observables as explained in section 6. Besides the methods illustrated in the figure, we discuss path integral representations of processes with continuous state spaces whose master equations admit Kramers-Moyal expansions (sections 4.4 and 5.4). A truncation of the backward Kramers-Moyal expansion at the level of a diffusion approximation results in a path integral representation of the (backward) Fokker-Planck equation whose original development goes back to works of Martin, Siggia, and Rose [21], de Dominicis [22], Janssen [23, 24], and Bausch, Janssen, and Wagner [24]. The representation can be rewritten in terms of an Onsager-Machlup function [84], and it simplifies to Wiener's path integral [85, 86] for purely diffusive Brownian motion [87]. Renormalization group techniques are not considered in our text. Information on these techniques can be found in [74, 75].

distribution $p(t, n|\tau, n_0)$ backward in time, starting out at time $\tau = t$. Both master equations can be derived from the Chapman-Kolmogorov equation (cf. left side of figure 1). In section 1.4, we discuss two explicit representations of the conditional probability distribution solving the two master equations. Moreover, we comment on various numerical methods for the approximation of this distribution and for the generation of sample paths. Afterwards, section 1.5 provides a brief historical overview of contributions to the development of stochastic path integrals.

The main part of this chapter begins with section 2. We first exemplify how a generalized probability generating function can be used to determine the stationary probability distribution of an elementary chemical reaction. This example introduces the bra-ket notation used in this work. In section 2.1, we formulate conditions under which a general forward master equation can be transformed into a linear partial differential equation (PDE) obeyed by the generating function. This function is defined as the sum of the conditional probability distribution $p(\tau, n|t_0, n_0)$ over a set of basis functions $\{|n\rangle\}$, the "kets" (cf. middle column of figure 1). The explicit choice of the basis functions depends on the process being studied. We discuss different choices of the basis functions in section 2.2, first for a random walk, afterwards for chemical reactions and for processes whose particles locally exclude one another. Several methods have recently been proposed for the analysis of the PDE obeyed by the generating function. These methods include the variational method of Eyink [88] and of Sasai and Wolynes [44], the WKB approximations [89] and spectral methods of Elgart and Kamenev [40] and of Assaf and Meerson [41–43], and the spectral method of Walczak, Mugler, and Wiggins [39]. We comment on these methods in section 2.3.

In section 3.1, we formulate conditions under which a general backward master equation can be transformed into a novel, backward-time PDE obeyed by a "marginalized distribution". This object is defined as the sum of the conditional probability distribution $p(t, n|\tau, n_0)$ over a set of basis functions $\{|n_0\rangle\}$ (cf. middle column of figure 1). If the basis function $|n_0\rangle$ is chosen as a probability distribution, the marginalized distribution also constitutes a true probability distribution. Different choices of the basis function are considered in section 3.2. In section 3.3, the use of the marginalized distribution is exemplified in the calculation of mean extinction times. Afterwards, in section 3.4, we derive yet another linear PDE, which is obeyed by a "probability generating functional". This functional is defined as the sum of the conditional probability distribution $p(\tau, n|t_0, n_0)$ over a set of basis functionals $\{\langle n|\}$, the "bras". In section 3.5, we show that the way in which the generating functional "generates" probabilities generalizes the Poisson representation of Gardiner and Chaturvedi [37, 38].

Sections 4 and 5 share the goal of representing the master equations' solution by path integrals. In section 4.1, we first derive a novel *backward path integral representation* from the PDE obeyed by the marginalized distribution (cf. right

side of figure 1). Its use is exemplified in sections 4.2, 4.3, and 4.5 in which we solve several elementary processes. Although we do not discuss the application of renormalization group techniques, section 4.5 d includes a discussion of how the backward path integral representation can be evaluated in terms of a perturbation expansion. The summands of the expansion are expressed by Feynman diagrams. Besides, we derive a path integral representation for Markov processes with continuous state spaces in the "intermezzo" section 4.4. This representation is obtained by performing a Kramers-Moyal expansion of the backward master equation and it comprises a classic path integral representation [21–24] of the (backward) Fokker-Planck equation as a special case. One can rewrite the representation of the Fokker-Planck equation in terms of an Onsager-Machlup function [84] and, for purely diffusive Brownian motion [87], the representation simplifies to the path integral of Wiener [85, 86]. Moreover, we recover a Feynman-Kac like formula [90], which solves the (backward) Fokker-Planck equation in terms of an average over the paths of an Itô stochastic differential equation [91–93] (or of a Langevin equation [94]).

In section 5, we complement the backward path integral representation with a *forward path integral representation*. Its derivation in section 5.1 starts out from the PDE obeyed by the generalized generating function (cf. right side of figure 1). The forward path integral representation can, for example, be used to compute the generating function of generic linear processes as we demonstrate in section 5.2. Besides, we briefly outline how a Kramers-Moyal expansion of the (forward) master equation can be employed to derive a path integral representation for processes with continuous state spaces in section 5.4. This path integral can be expressed in terms of an average over the paths of an SDE proceeding backward in time. Its potential use remains to be explored.

Before proceeding, let us briefly point out some properties of the forward and backward path integral representations. First, the paths along which these path integrals proceed are described by real variables and all integrations are performed over the real line. Grassmann path integrals [31, 32, 34–36] for systems whose particles locally exclude one another are not considered. It is, however, explained in section 2.2 d how such systems can be treated without the need for Grassmann variables, based on a method recently proposed by van Wijland [95]. Second, transformations of the path integral variables such as the "Doi-shift" [96] are implemented on the level of the basis functions and functionals. Third, our derivations of the forward and backward path integral representations do not involve coherent states or combinatoric relations for the commutation of exponentiated operators. Last, the path integrals allow for time-dependent rate coefficients of the stochastic processes.

In section 6, we derive a path integral representation of averaged observables (cf. right side of figure 1). This representation can be derived both from the backward

and forward path integral representations (cf. section 6.1), and by representing the forward master equation in terms of the eigenvectors of creation and annihilation matrices ("coherent states"; cf. section 6.2). The duality between these two approaches resembles the duality between the wave [97] and matrix [98–100] formulations of quantum mechanics. Let us note that our resulting path integral does not involve a "second-quantized" or "normal-ordered" observable [101]. In fact, we show that this object agrees with the average of an observable over a Poisson distribution. In section 6.3, we then explain how the path integral can be evaluated perturbatively using Feynman diagrams. Such an evaluation is demonstrated for the coagulation reaction $2A \to A$ in section 6.4, restricting ourselves to the "tree level" of the diagrams.

In section 7, we review and extend a recent method of Elgart and Kamenev for the computation of rare event probabilities [40]. As explained in section 7.1, this method evaluates a probability distribution by expanding the forward path integral representation from section 5 around "stationary", or "extremal", paths. In a first step, one thereby acquires an approximation of the ordinary probability generating function. In a second step, this generating function is transformed back into the underlying probability distribution. The evaluation of this back transformation typically involves an additional saddle-point approximation. In section 7.2, we demonstrate both of the steps for the binary annihilation reaction $2A \to \emptyset$, improving an earlier approximation of the process by Elgart and Kamenev [40] by terms of sub-leading order. In section 7.3, we then extend the "stationary path method" to the backward path integral representation from section 4. The backward path integral provides direct access to a probability distribution without requiring an auxiliary saddle-point approximation. However, the leading order term of its expansion is not normalized. We demonstrate the procedure for the binary annihilation reaction in section 7.4.

Finally, section 8 closes with a summary of the different approaches discussed in this part of the thesis and outlines open challenges and promising directions for future research.

1.3 Continuous-time Markov processes and the forward and backward master equations

Our main interest lies in a special class of stochastic processes, namely in the class of continuous-time Markov processes with discontinuous sample paths. These processes are also called "jump processes". In the following, we outline the mathematical theory of jump processes and derive the central equations obeyed by them: the forward and the backward master equation. Before going into the mathematical details, let us explain when a system's time evolution can be

modelled as a continuous-time Markov process with discontinuous sample paths and what that phrase actually means.

First of all, if the evolution of a system is to be modelled as a continuous-time Markov process, it must be possible to describe the system's state by some variable n. In fact, it must be possible to do so at every point τ in time throughout an observation period $[t_0, t]$. A variable $n \in \mathbb{Z}$ could, for example, represent the position of a molecular motor along a cytoskeletal filament, or a variable $n \in \mathbb{R}_0$ the price of a stock between the opening and closing times of an exchange. The assumption of a continuous time parameter τ is rather natural and conforms to our everyday experience. Still, a discrete time parameter may sometimes be preferred, for example, to denote individual generations of an evolving population [102]. By allowing τ to take on any value between the initial time t_0 and the final time t, we can choose it to be arbitrarily close to one of those times. Below, this possibility will allow us to describe the evolution of the process in terms of a differential equation.

The (unconditional) probability of finding the system in state n at time τ is represented by the "single-time" probability distribution $p(\tau, n)$. Upon demanding that the system has visited some state n_0 at an earlier time $t_0 < \tau$, the probability of observing state n at time τ is instead encoded by the conditional probability distribution $p(\tau, n|t_0, n_0)$. If the conditional probability distribution is known, the single-time distribution can be inferred from any given initial distribution $p(t_0, n_0)$ via $p(\tau, n) = \sum_{n_0} p(\tau, n|t_0, n_0)p(t_0, n_0)$. A stochastic process is said to be *Markovian* if a distribution conditioned on multiple points in time actually depends only on the state that was realized most recently. In other words, a conditional distribution $p(t, n|\tau_k, m_k; \cdots; \tau_1, m_1; t_0, n_0)$ must agree with $p(t, n|\tau_k, m_k)$ whenever $t > \tau_k > \tau_{k-1}, \ldots, t_0$.[1] Therefore, a Markov process is fully characterized by the single-time distribution $p(t_0, n_0)$ and the conditional distribution $p(\tau, n|t_0, n_0)$. The latter function is commonly referred to as the "transition probability" for Markov processes [103].

The stochastic dynamics of a system can be modelled in terms of a Markov process if the system has no memory. Let us explain this requirement with the example of a Brownian particle suspended in a fluid [87]. Over a very short time scale, the motion of such a particle is ballistic and its velocity highly autocorrelated [104]. But as the particle collides with molecules of the fluid, that memory fades away. A significant move of the particle due to fluctuations in the isotropy of its molecular bombardment then appears to be completely uncorrelated from previous moves (provided that the observer does not look too closely [105]). Thus, on a sufficiently coarse time scale, the motion of the particle looks diffusive

1 Note that we do not distinguish between random variables and their outcomes. Moreover, we stick to the physicists' convention of ordering times in descending order. In the mathematical literature, the reverse order is more common, see e.g. [9].

and can be modelled as a Markov process. However, the validity of the Markov assumption does not extend beyond the coarse time scale.

The Brownian particle exemplifies only two of the properties that we are looking for: its position is well-defined at every time τ and its movement is effectively memoryless on the coarse time scale. But the paths of the Brownian particle are continuous, meaning that it does not spontaneously vanish and then reappear at another place. If the friction of the fluid surrounding the Brownian particle is high (over-damped motion), the probability of observing the particle at a particular place can be described by the Smoluchowski equation [106]. This equation coincides with the simple diffusion equation when the particle is not subject to an external force. In the general case, the probability of observing the particle at a particular place with a particular velocity obeys the Klein-Kramers equation [107, 108] (the book of Risken [109] provides a pedagogic introduction to these equations). From a mathematical point of view, all of these equations constitute special cases of the (forward) Fokker-Planck equation [6–9, 109]. For a single random variable $x \in \mathbb{R}$, e.g. the position of the Brownian particle, this equation has the generic form

$$\partial_\tau p(\tau, x|t_0, x_0) = -\partial_x\big[\alpha_\tau(x)p\big] + \frac{1}{2}\partial_x^2\big[\beta_\tau(x)p\big]\,. \tag{1}$$

The initial condition of this equation is given by the Dirac delta distribution (or generalized function) $p(t_0, x|t_0, x_0) = \delta(x - x_0)$. Here we used the letter x for the random variable because the letter n would suggest a discrete state space. The function α_τ is often called a drift coefficient and β_τ a diffusion coefficient (note, however, that in the context of population genetics, β_τ describes the strength of random genetic "drift" [110, 111]). For reasons addressed below, the diffusion coefficient must be non-negative at every point in time for every value of x (for a multivariate process, β_τ represents a positive-semidefinite matrix). In the mathematical community, the Fokker-Planck equation is better known as the Kolmogorov forward equation [110], honouring Kolmogorov's fundamental contributions to the theory of continuous-time Markov processes [9]. Whereas the above Fokker-Planck equation evolves the conditional probability distribution forward in time, one can also evolve this distribution backward in time, starting out from the final condition $p(t, x|t, x_0) = \delta(x - x_0)$. The corresponding equation is called the Kolmogorov backward or backward Fokker-Planck equation. It has the generic form

$$\partial_{-\tau} p(t, x|\tau, x_0) = \alpha_\tau(x_0)\partial_{x_0} p + \frac{1}{2}\beta_\tau(x_0)\partial_{x_0}^2 p\,. \tag{2}$$

The forward and backward Fokker-Planck equations provide information about the conditional probability distribution but not about the individual paths of the Brownian particle. A general theory of how partial differential equations connect to the individual sample paths of a stochastic process goes back to works of Feynman and Kac [90, 112]. Their theory allows us to write the solution of the backward Fokker-Planck equation (2) in terms of the following Kolmogorov formula, which constitutes a special case of the Feynman-Kac formula [113–115]:

$$p(t, x|\tau, x_0) = \left\langle\!\left\langle \delta\big(x - x(t)\big) \right\rangle\!\right\rangle_W . \tag{3}$$

The brackets $\langle\!\langle \cdot \rangle\!\rangle_W$ represent an average over realizations of a Wiener process W, which evolves through uncorrelated Gaussian increments $\mathrm{d}W$. The Wiener process drives the evolution of the sample path $x(s)$ from $x(\tau) = x_0$ to $x(t)$ via the Itô stochastic differential equation (SDE) [91–93]

$$\mathrm{d}x(s) = \alpha_s(x(s))\,\mathrm{d}s + \sqrt{\beta_s(x(s))}\,\mathrm{d}W(s) . \tag{4}$$

The diffusion coefficient β_s must be non-negative because $x(s)$ describes the position of a real particle. Otherwise, the sample path heads off into imaginary space (for a multivariate process, $\sqrt{\beta_s}$ may be chosen as the unique symmetric and positive-semidefinite square root of β_s [116]). Algorithms for the numerical solution of SDEs are provided in [113]. In the physical sciences, SDEs are often written as Langevin equations [94].[2] For a discussion of stochastic differential equations the reader may refer to a recent report on progress [117]. Due to the central importance of the Feynman-Kac formula, we provide a brief proof of it in appendix A. We also encounter the formula in section 4 in evaluating a path integral representation of the (backward) master equation.

After this brief detour to continuous-time Markov processes with continuous sample paths, let us return to jump processes, whose sample paths are discontinuous. A system that can be modelled as such a process are motor proteins on cytoskeletal filaments [118–120]. The uni-directional walk of a molecular motor such as myosin, kinesin, or dynein along an actin filament or a microtubule is driven by the hydrolysis of adenosine triphosphate (ATP) and is intrinsically stochastic [121]. Once a sufficient amount of energy is available, one of the two "heads" of the motor unbinds from its current binding site on the filament and moves to the next binding site. Each binding site can only be occupied by a single head. On a coarse-grained level, the state of the system at time τ is therefore characterized by the occupation of its binding sites. With only a single

2 The Langevin equation corresponding to the SDE (4) reads $\partial_s x(s) = \alpha_s(x(s)) + \sqrt{\beta_s(x(s))}\eta(s)$, with the Gaussian white noise $\eta(s)$ having zero mean and the auto-correlation function $\langle \eta(s)\eta(s') \rangle = \delta(s - s')$.

cytoskeletal filament whose binding sites are labelled as $\{0, 1, 2, \ldots\} =: \mathbb{L}$, the variable $n \equiv (n_0, n_1, n_2, \ldots) \in \{0, 1\}^{|\mathbb{L}|}$ can be used to represent the occupied and unoccupied binding sites. Here, $|\mathbb{L}|$ denotes the total number of binding sites along the filament and $n_i = 1$ signifies that the i-th binding site is occupied. Since the state space of all the binding site configurations is discrete, a change in the binding site configuration involves a "jump" in state space. Provided that the jumps are uncorrelated from one another (which needs to be verified experimentally), the dynamics of the system can be described by a continuous-time Markov process with discontinuous sample paths. Before addressing further systems for which this is the case, let us derive the fundamental equations obeyed by these processes: the forward and the backward master equation.

In his classic textbook [115], Gardiner presents a succinct derivation of both the (forward) master and the (forward) Fokker-Planck equation by distinguishing between discontinuous and continuous contributions to sample paths. In the following, we are only interested in the master equation, which governs the evolution of systems whose states change discontinuously. To prevent the occurrence of continuous changes, we assume that the state of our system is represented by a discrete variable n and that the space of all states is countable. With the state space $\mathbb{Z}$, n could, for example, represent the position of a molecular motor along a cytoskeletal filament. On the other hand, $n \in \mathbb{N}_0$ could represent the number of molecules in a chemical reaction. The minimal jump size is one in both cases. By keeping the explicit role of n unspecified, the following considerations also apply to systems harbouring different kind of molecules (e.g. $n \equiv (n^A, n^B, n^C) \in \mathbb{N}_0^3$), and to systems whose molecules perform random walks in a (discrete) space (e.g. $n \equiv \{n_i \in \mathbb{N}_0\}_{i \in \mathbb{Z}}$).

To derive the master equation, we start out by marginalizing the joint conditional distribution $p(t, n; \tau, m | t_0, n_0)$ over the state m at the intermediate time τ ($t \geq \tau \geq t_0$), resulting in

$$p(t, n | t_0, n_0) = \sum_m p(t, n; \tau, m | t_0, n_0)\,. \tag{5}$$

Whenever the range of a sum is not specified, it shall cover the whole state space of its summation variable. The above equation holds for arbitrary stochastic processes. But for a Markov process, one can employ the relation

$$p(t, n; \tau, m | t_0, n_0) = p(t, n | \tau, m; t_0, n_0) p(\tau, m | t_0, n_0) \tag{6}$$

between joint and conditional distributions to turn the equation into the Chapman-Kolmogorov equation

$$p(t, n|t_0, n_0) = \sum_m p(t, n|\tau, m)p(\tau, m|t_0, n_0)\,. \tag{7}$$

Letting $p(t|t_0)$ denote the matrix with elements $p(t, n|t_0, n_0)$, the Chapman-Kolmogorov equation can also be written as $p(t|t_0) = p(t|\tau)p(\tau|t_0)$ (semigroup property). Note that the matrix notation requires a mapping between the state space of n and n_0 and an appropriate index set $I \subset \mathbb{N}$. However, we also make use of this notation when the state space is countably infinite.

To derive the (forward) master equation from the Chapman-Kolmogorov equation (7), we define

$$Q_{\tau,\Delta t}(n, m) := \frac{p(\tau + \Delta t, n|\tau, m) - \delta_{n,m}}{\Delta t} \tag{8}$$

for all values of n and m and assume the existence and finiteness of the limits

$$Q_\tau(n, m) := \lim_{\Delta t \to 0} Q_{\tau,\Delta t}(n, m)\,. \tag{9}$$

These are the elements of the transition (rate) matrix Q_τ, which is also called the infinitesimal generator of the Markov process or is simply referred to as the Q_τ-matrix. Its off-diagonal elements $w_\tau(n, m) := Q_\tau(n, m)$ denote the rates at which probability flows from a state m to a state $n \neq m$. The "exit rates" $w_\tau(m) := -Q_\tau(m, m)$, on the other hand, describe the rates at which probability leaves state m. Both $w_\tau(n, m)$ and $w_\tau(m)$ are non-negative for all n and m. All of the processes considered here shall conserve the total probability, requiring that $\sum_n Q_\tau(n, m) = 0$ or, equivalently, $w_\tau(m) = \sum_n w_\tau(n, m)$ (with $w_\tau(m, m) := 0$). The finiteness of the exit rate $w_\tau(m)$ and the conservation of total probability imply that we consider a *stable* and *conservative* Markov process [122]. In the natural sciences, the master equation is commonly written in terms of w_τ, but most mathematicians prefer Q_τ. These matrices can be converted into one another by employing

$$Q_\tau(n, m) = w_\tau(n, m) - \delta_{n,m} w_\tau(m)\,. \tag{10}$$

We refer to both matrices as transition (rate) matrices and to their off-diagonal elements as transition rates. These rates fully specify the stochastic process.

Assuming that the limit in (9) interchanges with a sum over the state m, the (forward) master equation follows from the Chapman-Kolmogorov equation (7) as

$$\begin{aligned}\partial_\tau p(\tau, n|t_0, n_0) &= \lim_{\Delta t\to 0} \frac{p(\tau+\Delta t, n|t_0, n_0) - p(\tau, n|t_0, n_0)}{\Delta t} \\ &= \lim_{\Delta t\to 0} \sum_m \frac{p(\tau+\Delta t, n|\tau, m) - \delta_{n,m}}{\Delta t} p(\tau, m|t_0, n_0) \\ &= \sum_m Q_\tau(n, m) p(\tau, m|t_0, n_0)\,.\end{aligned} \tag{11}$$

Thus, the master equation constitutes a set of coupled, linear, first-order ordinary differential equations (ODEs). The time evolution of the distribution starts out from $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$. In matrix notation, the equation can be written as $\partial_\tau p(\tau|t_0) = Q_\tau p(\tau|t_0)$. In terms of w_τ, it assumes the intuitive gain-loss form

$$\partial_\tau p(\tau, n|t_0, n_0) = \sum_m \Big(w_\tau(n, m) - \delta_{n,m} \sum_k w_\tau(k, n) \Big) p(\tau, m|t_0, n_0) \tag{12}$$

$$= \sum_m \big[w_\tau(n, m) p(\tau, m|\cdot) - w_\tau(m, n) p(\tau, n|\cdot) \big]\,. \tag{13}$$

The dot inside the probability distribution's argument abbreviates the initial parameters t_0 and n_0, which are of secondary concern right here. That will change below in the derivation of the backward master equation. An omission of the parameters also makes it impossible to distinguish the conditional distribution $p(\tau, n|t_0, n_0)$ from the single-time distribution $p(\tau, n)$. The single-time distribution obeys the master equation as well, as can inferred directly from the relation $p(\tau, n) = \sum_{n_0} p(\tau, n|t_0, n_0) p(t_0, n_0)$ or by summing the above master equation over an initial distribution $p(t_0, n_0)$. In fact, the single-time distribution would even obey the master equation if the process was not Markovian, but without providing a complete characterization of the process [75, 79]. The master equation (11) or (12) is particularly interesting for transition rates causing an imbalance between forward and backward transitions along closed cycles of states, i.e. for rates violating Kolmogorov's criterion [123] for detailed balance [75]. Such systems are truly out of thermal equilibrium. If detailed balance is instead fulfilled, the system eventually converges to a stationary Boltzmann-Gibbs distribution with vanishing probability currents between states [75]. Whether or not detailed balance is actually fulfilled is, however, not relevant for the methods discussed in our work. Information on the existence and uniqueness of an asymptotic stationary distribution of the master equation can be found in [122].

The name "master equation" was originally coined by Nordsieck, Lamb, and Uhlenbeck [10] in their study of the Furry model of cosmic rain showers [124].

Shortly before, Feller applied an equation of the same structure to the growth of populations [125] and Delbrück to well-mixed, auto-catalytic chemical reactions [126]. Delbrück's line of research was followed by several others [127–130], most notably by McQuarrie [131–133] (see also the books [103, 115, 134]). In these articles, several elementary chemical reactions are solved by methods that also appear later in this chapter. When the particles engaging in a reaction can also diffuse in space, their density may exhibit dynamics that are not expected from observations made in well-mixed environments. Hence, reaction-diffusion master equations have been the focus of intense research and have been analysed using path integrals (see, for example, [101, 135–138] and the references in section 1.5). Master equations, and simulations algorithms based on master equations, are now being used in numerous fields of research. They are being applied in the contexts of spin dynamics [139–142], gene regulatory networks [39, 143–148], the spreading of diseases [149–152], epidermal homeostasis [153], nucleosome repositioning [154], ecological [155–168] and bacterial dynamics [5, 163, 169–171], evolutionary game theory [3, 172–182], surface growth [183], and social and economic processes [184–187]. Queuing processes are also often modelled in terms of master equations, but in this context, the equations are typically referred to as Kolmogorov equations [188]. Moreover, master equations and the SSA have helped to understand the formation of traffic jams on highways [189, 190], the walks of molecular motors along cytoskeletal filaments [118–120, 191–193], and the condensation of bosons in driven-dissipative quantum systems [1, 3, 194]. The master equation that was found to describe the coarse-grained dynamics of these bosons coincides with the master equation of the (asymmetric) inclusion process [195–198]. On the other hand, transport processes are more commonly modelled in terms of the (totally) asymmetric simple exclusion process (ASEP or TASEP) [199–202]. The ASEP describes the biased hopping of particles along a one-dimensional lattice, with each lattice site providing space for at most one particle. The ASEP and the TASEP are regarded as paradigmatic models in the field of non-equilibrium statistical mechanics, with many exact mathematical results having been established [203–214]. Some of these results were established by applying the Bethe ansatz to the master equation of the ASEP [207, 211]. The review of Chou, Mallick, and Zia provides a comprehensive account of the ASEP and of its variants [193]. The master equation of the TASEP with Langmuir kinetics was recently used to understand the length regulation of microtubules [215].

Unlike deterministic models, the master equations describing the dynamics of the above systems take into account that discrete and finite populations are prone to "demographic fluctuations". The populations of the above systems consist of genes or proteins, infected persons, bacteria or cars and they are typically small, at least compared to the number of molecules in a mole of gas. For example, the copy number of low abundance proteins in *Escherichia coli* cytosol was found to

be in the tens to hundreds [216]. Therefore, the presence or absence of a single protein is much more important than the presence or absence of an individual molecule in a mole of gas. A demographic fluctuation may even be fatal for a system, for example, when the copy number of an auto-catalytic reactant drops to zero. The master equation (11) provides a useful tool to describe such an effect.

Up to this point, we have only considered the forward master equation. But just as the (forward) Fokker-Planck equation is complemented by the backward Fokker-Planck equation, the (forward) master equation (11) is complemented by a backward master equation. It follows from the Chapman-Kolmogorov equation (7) as

$$\begin{aligned}\partial_{-\tau} p(t,n|\tau,n_0) &= \lim_{\Delta t\to 0} \frac{p(t,n|\tau-\Delta t,n_0)-p(t,n|\tau,n_0)}{\Delta t} \\ &= \lim_{\Delta t\to 0} \sum_m p(t,n|\tau,m)\frac{p(\tau,m|\tau-\Delta t,n_0)-\delta_{m,n_0}}{\Delta t} \\ &= \sum_m p(t,n|\tau,m)Q_\tau(m,n_0)\,.\end{aligned} \tag{14}$$

Here, the transition rate is obtained in the limit $\lim_{\Delta t\to 0} Q_{\tau-\Delta t,\Delta t}(m,n_0)$ (cf. (8)). In matrix notation, the backward master equation reads $\partial_{-\tau} p(t|\tau) = p(t|\tau)Q_\tau$. In terms of w_τ, it assumes the form (cf. (10))

$$\partial_{-\tau} p(t,n|\tau,n_0) = \sum_m p(t,n|\tau,m)\Big(w_\tau(m,n_0) - \delta_{m,n_0}\sum_k w_\tau(k,n_0)\Big) \tag{15}$$

$$= \sum_m \big[p(\cdot|\tau,m) - p(\cdot|\tau,n_0)\big] w_\tau(m,n_0)\,. \tag{16}$$

In this equation, the dots abbreviate the final parameters t and n. The backward master equation evolves the conditional probability distribution backward in time, starting out from the final condition $p(t,n|t,n_0) = \delta_{n,n_0}$. Just as the backward Fokker-Planck equation, the backward master equation proves useful for the computation of mean extinction and first passage times (see [115, 217] and section 3.3). Furthermore, it follows from the backward master equation (14) that the (conditional) average $\langle A\rangle(t|\tau,n_0) := \sum_n A(n)p(t,n|\tau,n_0)$ of an observable A fulfils an equation of just the same form, namely

$$\partial_{-\tau}\langle A\rangle(t|\tau,n_0) = \sum_m \langle A\rangle(t|\tau,m)Q_\tau(m,n_0)\,. \tag{17}$$

The final condition of the equation is given by $\langle A\rangle(t|t,n_0) = A(n_0)$. The validity of equation (17) is the reason why we later employ a "backward" path integral to represent the average $\langle A\rangle$ (cf. section 6).

1.4 Analytical and numerical methods for the solution of master equations

If the dynamics of a system are restricted to a finite number of states and if its transition rates are independent of time, both the forward master equation $\partial_\tau p(\tau|t_0) = Qp(\tau|t_0)$ and the backward master equation $\partial_{-\tau_0} p(t|\tau_0) = p(t|\tau_0)Q$ are solved by [218]

$$p(\tau|\tau_0) = \mathrm{e}^{Q(\tau-\tau_0)}\mathbb{1} \tag{18}$$

(recall that $p(\tau|\tau_0)$ is the matrix with elements $p(\tau, n|\tau_0, n_0)$). The Chapman-Kolmogorov equation (7) is also solved by the distribution. Although the matrix exponential inside this solution can in principle be evaluated in terms of the (convergent) Taylor expansion $\sum_{k=0}^{\infty} \frac{(\tau-\tau_0)^k}{k!} Q^k$, the actual calculation of this series is typically infeasible for non-trivial processes, both analytically and numerically (a truncation of the Taylor series may induce severe round-off errors and serves as a lower bound on the performance of algorithms in [219]). Consequently, alternative numerical algorithms have been developed to evaluate the matrix exponential. Moler and Van Loan reviewed "nineteen dubious ways" of computing the exponential in [219, 220]. Algorithms that can deal with very large state spaces are considered in [221, 222]. For time-dependent transition rates, the matrix exponential generalizes to a Magnus expansion [223, 224].

In the previous paragraph, we restricted the dynamics to a finite state space to ensure the existence of the matrix exponential in (18). Provided that the supremum $\sup_m |Q(m,m)|$ of all the exit rates is finite (uniformly bounded Q-matrix), the validity of the above solution extends to state spaces comprising a countable number of states [225]. To see that, we define the left stochastic matrix $P := \mathbb{1} + \lambda^{-1}Q$, with the parameter λ being larger than the above supremum. Writing $\mathrm{e}^{Q\Delta t} = \mathrm{e}^{-\lambda\Delta t}\mathrm{e}^{\lambda\Delta t P}$ with $\Delta t := \tau - \tau_0$, the matrix exponential can be evaluated in terms of the convergent Taylor series

$$p(\tau|\tau_0) = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-\lambda\Delta t}(\lambda\Delta t)^k}{k!} P^k \,. \tag{19}$$

Effectively, one has thereby decomposed the continuous-time Markov process with transition matrix Q into a discrete-time Markov chain with transition matrix P, subordinated to a continuous-time Poisson process with rate coefficient λ (the Poisson process acts as a "clock" with sufficiently high ticking rate λ). Such a decomposition is called a uniformization or randomization and was first proposed by Jensen [226]. The series (19) can be evaluated via numerically stable algorithms and truncation errors can be bounded [227, 228]. Nevertheless, the uniformization method requires the computation of the powers of a matrix having as many rows

and columns as the system has states. Consequently, a numerical implementation of the method is only feasible for sufficiently small state spaces. Further information on the method and on its improvements can be found in [226–231].

The mathematical study of the existence and uniqueness of solutions of the forward and backward master equations was pioneered by Feller and Doob in the 1940s [232, 233]. Feller derived an integral recurrence formula [122, 232], which essentially constitutes a single step of the "path summation representation" that we derive further below. In the following, we assume that the forward and the backward master equations have the same unique solution and we restrict our attention to processes performing only a finite number of jumps during any finite time interval. These conditions, and the conservation of total probability, are, for example, violated by processes that "explode" after a finite time.[3] More information on such processes is provided in [122, 218].

In the following, we complement the above representations of the master equations's solution with a "path summation representation". This representation can be derived by examining the steps of the stochastic simulation algorithm (SSA) of Gillespie (its "direct" version) [80, 81, 83] or by performing a Laplace transformation of the forward master equation. Here we follow the former, qualitative, approach. A formal derivation of the representation via the master equation's Laplace transform is provided in appendix B. Although the basic elements of the SSA had already been known before Gillespie's work [232–238], its popularity largely increased after Gillespie applied it to the study of chemical reactions. In chapter III, we use the SSA to evaluate a model of bacterial range expansions. As the SSA is restricted to time-independent transition rates, so is the following derivation of the path summation representation.

To derive the path summation representation, we prepare a system in state n_0 at time t_0 as illustrated in figure 2. Since the process is homogeneous in time, we may choose $t_0 = 0$. The total rate of leaving state n_0 is given by the exit rate $w(n_0) = \sum_{n_1} w(n_1, n_0)$. In order to determine how long the system actually stays in state n_0, one may draw a random waiting time τ_0 from the exponential distribution $\mathcal{W}_{n_0}(\tau_0) := w(n_0)\mathrm{e}^{-w(n_0)\tau_0}\Theta(\tau_0)$. The Heaviside step function Θ prevents the sampling of negative waiting times and is here defined as $\Theta(\tau) = 1$ for $\tau \geq 0$ and $\Theta(\tau) = 0$ for $\tau < 0$. Thus far, we only know that the system leaves n_0 but not where it ends up. It could end up in any state n_1 for which the transition rate $w(n_1, n_0)$ is positive. The probability that a particular state n_1 is realized is given by $w(n_1, n_0)/w(n_0)$. In a numerical implementation of the SSA, the state

3 Just as a population whose growth is described by the deterministic equation $\partial_\tau n = n^2$ explodes after a finite time, so does a population whose growth is described by the master equation (12) with transition rate $w(n, m) = \delta_{n,m+1} m(m-1)$ [122]. This transition rate models the elementary reaction $2\,A \to 3\,A$ as explained in section 1.5. An explosion also occurs for the rate $w(n, m) = \delta_{n,m+1} m^2$.

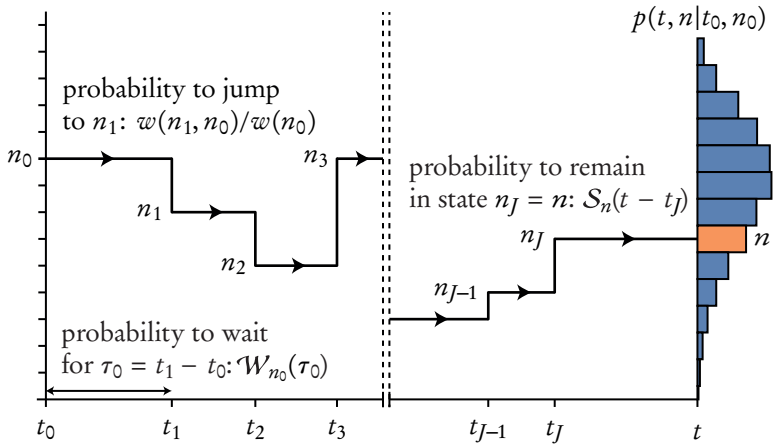


Figure 2 Illustration of the stochastic simulation algorithm (SSA) and of the path summation representation of the probability distribution $p(t, n|t_0, n_0)$. In a numerical implementation of the SSA, the system is prepared in state n_0 at time t_0. After a waiting time τ_0 that is drawn from the exponential distribution $\mathcal{W}_{n_0}(\tau_0) = w(n_0)\mathrm{e}^{-w(n_0)\tau_0}\Theta(\tau_0)$, the system transitions into a new state. An arrival state n_1 is chosen with probability $w(n_1, n_0)/w(n_0)$. The procedure is repeated until after J steps, the current time $t_J \leq t$ plus an additional waiting time exceeds t. The sample path has thus resided in state n_J at time t. This information is recorded in a histogram approximation of $p(t, n|t_0, n_0)$. The path summation representation of $p(t, n|t_0, n_0)$ requires n_J to coincide with n. The probability that the system has remained in state n over the last time interval $[t_J, t]$ is given by the survival probability $\mathcal{S}_n(t - t_J) = \mathrm{e}^{-w(n)(t-t_J)}\Theta(t - t_J)$. The total probability of the generated path, integrated over all possible waiting times, is represented by $p_\tau(\mathcal{P}_J)$ in (24). A summation of this probability over all possible sample paths results in the path summation representation (23).

n_1 is determined by drawing a second (uniformly-distributed) random number. Our goal is to derive an analytic representation of the probability $p(t, n|t_0, n_0)$ of finding the system in state n at time t. Thus, after taking $J - 1$ further steps, the sample path $\mathcal{P}_J := \{n_J \leftarrow \cdots \leftarrow n_1 \leftarrow n_0\}$ should eventually visit state $n_J = n$ at some time $t_J \leq t$. The total time $\tau_{J-1} + \cdots + \tau_0$ until the jump to state n occurs is distributed by the convolutions of the individual waiting time distributions, i.e. by $\overset{J-1}{\underset{j=0}{\star}} \mathcal{W}_{n_j}$. For example, $\tau := \tau_1 + \tau_0$ is distributed by

$$(\mathcal{W}_{n_1} \star \mathcal{W}_{n_0})(\tau) = \int_{\mathbb{R}} \mathrm{d}\tau_0\, \mathcal{W}_{n_1}(\tau - \tau_0)\mathcal{W}_{n_0}(\tau_0) \tag{20}$$

$$= w(n_1)w(n_0)\mathrm{e}^{-w(n_1)\tau} \int_0^\tau \mathrm{d}\tau_0\, \mathrm{e}^{(w(n_1)-w(n_0))\tau_0}\, \Theta(\tau) \tag{21}$$

$$= \frac{w(n_1)w(n_0)\left(\mathrm{e}^{-w(n_0)\tau} - \mathrm{e}^{-w(n_1)\tau}\right)}{w(n_1) - w(n_0)}\,\Theta(\tau)\,. \tag{22}$$

The probability that the system still resides in state $n_J = n$ at time t is determined by the "survival probability" $\mathcal{S}_n(t - t_J) = \mathrm{e}^{-w(n)(t-t_J)}\Theta(t - t_J)$. After putting all of these pieces together, we arrive at the following path summation representation of the conditional probability distribution:

$$p(t, n|t_0, n_0) = \sum_{J=0}^{\infty} \sum_{\{\mathcal{P}_J\}} p_{t-t_0}(\mathcal{P}_J) \text{ with} \tag{23}$$

$$p_\tau(\mathcal{P}_J) = \left(\mathcal{S}_n \star \left(\bigstar_{j=0}^{J-1} \frac{w(n_{j+1}, n_j)}{w(n_j)} \mathcal{W}_{n_j}\right)\right)(\tau)\,. \tag{24}$$

Here, $\sum_{\{\mathcal{P}_J\}} \coloneqq \sum_{n_1} \cdots \sum_{n_{J-1}}$ generates every path with J jumps between n_0 and $n_J = n$. The probability of such a path, integrated over all possible waiting times, is represented by $p_\tau(\mathcal{P}_J)$. By an appropriate choice of integration variables, the probability $p_\tau(\mathcal{P}_J)$ can also be written as

$$p_\tau(\mathcal{P}_J) = \left(\prod_{j=0}^{J-1} \int_0^\tau \mathrm{d}\tau_j \, \frac{w(n_{j+1}, n_j)}{w(n_j)} \mathcal{W}_{n_j}(\tau_j)\right) \mathcal{S}_n\big(\tau - (\tau_{J-1} + \ldots + \tau_0)\big)\,.$$

The survival probability $\mathcal{S}_n$ is included in these integrations. Without the integrations, the expression would represent the probability of a path with J jumps and fixed waiting times. That probability is, for example, used in the master equation formulation of stochastic thermodynamics in associating an entropy to individual paths [239]. In appendix B, we formally derive the path summation representation (23) from the Laplace transform of the forward master equation (12).

The path summation representation (23) does not only form the basis of the SSA but also of some alternative algorithms [240–245]. These algorithms either infer the path probability $p_\tau(\mathcal{P}_J)$ numerically from its Laplace transform or evaluate the convolutions in (24) analytically. The analytic expressions that arise are, however, rather cumbersome generalizations of the convolution in (22) [246, 247]. They simplify only for the most basic processes (e.g. for a random walk or for the Poisson process). Moreover, care has to be taken when the analytic expressions are evaluated numerically because they involve pairwise differences of the exit rate $w(n)$ (cf. (22)). When these exit rates differ only slightly along a path, a substantial loss of numerical significance may occur due to finite precision arithmetic. Future studies could explore how the convolutions of exponential distributions in (24) can be approximated efficiently (for example, in terms of a Gamma distribution or by analytically determining the Laplace transform of (24),

followed by a saddle-point approximation [248] of the corresponding inverse Laplace transformation). In general, both the SSA as well as its competitors suffer from the enormous number of states of non-trivial systems, as well as from the even larger number of paths connecting different states. In [242], these paths were generated using a deterministic depth-first search algorithm, combined with a filter to select the paths that arrive at the right place at the right time. In [244], a single path was first generated using the SSA and then gradually changed into new paths through a Metropolis Monte Carlo scheme. Thus far, the two methods have only been applied to relatively simple systems and their prevalence is low compared to the prevalence of the SSA. Further research is needed to explore how relevant paths can be sampled more efficiently.

The true power of the SSA lies in its generation of sample paths with the correct probability of occurrence. Thus, just a few sample paths generated with the SSA are often sufficient to infer the "typical" dynamics of a process. A look at individual paths may, for example, reveal that the dynamics of a system are dominated by some spatial pattern, e.g. by spirals [172]. Efficient variations of the above "direct" version of the SSA are, for example, described in [83, 249–251]. Algorithms for the fast simulation of biochemical networks or processes with spatial degrees of freedom are implemented in the simulation packages [252–259].

The evaluation of the average $\langle A \rangle = \sum_n A(n)p(t,n|\cdot)$ of an observable A typically requires the computation of a larger number of sample paths. However, since the occurrence probability of sample paths generated with the SSA is statistically correct, such an average typically converges comparatively fast. Furthermore, each path can be sampled independently of every other path. Therefore, the computation of paths can be distributed to individual processing units, saving real time, albeit no computation time. A distributed computation of the sample paths is most often required, but possibly not even sufficient, if one wishes to compute the full probability distribution $p(t,n|t_0,n_0)$. Vastly more sample paths are required for this purpose, especially if "rare event probabilities" in the distribution's tails are sought for. In particular, if the probability of finding a system in state n at time t is only 10^{-10}, an average of 10^{10} sample paths are needed to observe that event just once. Moreover, the probability of observing any particular state decreases with the size of a system's state space. Thus, the sampling of full distributions becomes less and less feasible as systems become larger. Various other challenges remain open as well; for example, the efficient simulation of processes evolving on multiple time scales. These processes are typically simulated using approximative techniques such as τ-leaping [83, 251, 260–269]. Another challenge is posed by the evaluation of processes with time-dependent transition rates [270–272].

For completeness, let us mention yet another numerical approach to the (forward) master equation. Since the master equation (11) constitutes a set of coupled linear first-order ODEs, it can of course be treated as such and be integrated nu-

merically. The integration is, however, only feasible if the state space is sufficiently small (or appropriately truncated) and if all transitions occur on comparable time scales (otherwise, the master equation is quite probably stiff [273]). Nevertheless, a numerical integration of the master equation may be preferable over the use of the SSA if the full probability distribution is to be computed.

Neither the matrix exponential representation $p(t|t_0) = \mathrm{e}^{\mathbb{Q}(t-t_0)}\mathbb{1}$ of the conditional probability distribution, nor its path summation representation (23) is universally applicable. Moreover, even if the requirements of these solutions are met, the size of the state space or the complexity of the transition matrix may make it infeasible to evaluate them. In the next sections, we formulate conditions under which the conditional probability distribution can be represented in terms of the "forward" path integral

$$p(t, n|t_0, n_0) = \langle n|_t \int_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}} \, |n_0\rangle_{t_0} \tag{25}$$

and in terms of the "backward" path integral

$$p(t, n|t_0, n_0) = \langle n_0|_{t_0} \int_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \, |n\rangle_t \, . \tag{26}$$

The meaning of the integral signs and of the bras $\langle n|$ and kets $|n\rangle$ will become clear over the course of our discussion. Let us only note that the integrals do not proceed along paths of the discrete variable n, but over the paths of two continuous auxiliary variables that are introduced for this purpose. The relevance of each path is weighed by the exponential factors inside the integrals.

Besides these exact representations of the conditional probability distribution solving the master equations, there exist powerful ways of approximating this distribution and the values of averaged observables. These methods include the Kramers-Moyal [108, 274] and the system-size expansion [103, 275], as well as the derivation of moment equations. Information on these methods can be found in classic text books [103, 115] and in a recent review [276]. Although moment equations encode the complete information about a stochastic process, they typically constitute an infinite hierarchy whose evaluation requires a truncation by some closure scheme [277–285]. On the other hand, the Kramers-Moyal and the system-size expansion approximate the master equation in terms of a Fokker-Planck equation. Both expansions work best if the system under consideration is "large" (more precisely, they work best if the dynamics are centred around a stable or meta-stable state at a distance $N \gg 1$ from a potentially absorbing state; the standard deviation of its surrounding distribution is then of order $\sqrt{N}$). An extension of the system-size expansion to absorbing boundaries has recently

been proposed in [286]. In sections 4.4 and 5.4, we show how Kramers-Moyal expansions of the backward and forward master equations can be used to derive path integral representations of processes with continuous state spaces. When the expansion of the backward master equation stops (or is truncated) at the level of a diffusion approximation, one recovers a classic path integral representation of the (backward) Fokkers-Planck equation [21–24].

1.5 History of stochastic path integrals

The oldest path integral, both in the theory of stochastic processes and beyond [112, 287–289], is presumably Wiener's integral for Brownian motion [85, 86]. The path integrals we consider here were devised somewhat later, namely in the 1970s and 80s: first for the Fokker-Planck (or Langevin) equation [21–24] and soon after for the master equation. For the master equation, the theoretical basis of these "stochastic" path integrals was developed by Doi [11, 12]. He first expressed the creation and annihilation of molecules in a chemical reaction by the corresponding operators for the quantum harmonic oscillator [290] (modulo a normalization factor), introducing the concept of the Fock space [25] into non-equilibrium theory. Furthermore, he employed coherent states [27–29] to express averaged observables. Similar formalisms as the one of Doi were independently developed by Zel'dovich and Ovchinnikov [13], as well as by Grassberger and Scheunert [15]. Introductions to the Fock space approach, which are in part complementary to our work, are, for example, provided in [75, 96, 291–293]. The review of Mattis and Glasser [291] provides a chronological list of contributions to the field. These contributions include Rose's renormalized kinetic theory [14], Mikhailov's path integral [16, 17, 19], which is based on Zel'dovich's and Ovchinnikov's work, and Goldenfeld's extension of the Fock space algebra to polymer crystal growth [18]. Furthermore, Peliti reviewed the Fock space approach and provided derivations of path integral representations of averaged observables and of the probability generating function [20]. Peliti also expressed the hope that future "rediscoveries" of the path integral formalism would be unnecessary in the future. However, we believe that the probabilistic structures behind path integral representations of stochastic processes have not yet been clearly exposed. As illustrated in figure 1, we show that the forward and backward master equations admit not only one but two path integral representations: the forward representation (25) and the novel backward representation (26). Although the two path integrals resemble each other, they differ conceptually. While the forward path integral representation provides a probability generating function in an intermediate step, the backward representation provides a distribution that is marginalized over an initial state. Both path integrals can be used to represent averaged observables as shown in section 6. The backward path integral, however,

will turn out to be more convenient for this purpose (upon choosing a Poisson distribution as the basis function, i.e. $|n\rangle := \frac{x^n \mathrm{e}^{-x}}{n!}$, the representation is obtained by summing (26) over an observable $A(n)$). Let us note that despite adopting some of the notation of quantum theory, our text is guided by the notion that quantum (field) theory is "totally unnecessary" for the theory of stochastic path integrals. Michael E. Fisher once stated the same about the relevance of quantum field theory for renormalization group theory [294] (while acknowledging that in order to do certain types of calculations, familiarity with quantum field theory can be very useful; the same applies to the theory of stochastic path integrals).

Thus far, path integral representations of the master equation have primarily been applied to processes whose transition rates can be decomposed additively into the rates of simple chemical reactions. Simple means that the transition rate of such a reaction is determined by combinatoric counting. Consider, for example, a reaction of the form $k\,A \to l\,A$ in which $k \in \mathbb{N}_0$ molecules of type A are replaced by $l \in \mathbb{N}_0$ molecules of the same type. Assuming that the k reactants are drawn from an urn with a total number of m molecules, the total rate of the reaction should be proportional to the falling factorial $(m)_k := m(m-1)\cdots(m-k+1)$. The global time scale of reaction is set by the rate coefficient γ_τ, which we allow to depend on time. Thus, the rate at which the chemical reaction $k\,A \to l\,A$ induces a transition from state m to state n is $w_\tau(n,m) = \gamma_\tau (m)_k \delta_{n,m-k+l}$. The Kronecker delta inside the transition rate ensures that k molecules are indeed replaced by l new ones. Note that the number of particles in the system can never become negative, provided that the initial number of particles was non-negative. Hence, the state space of n is $\mathbb{N}_0$. Insertion of the above transition rate into the forward master equation (12) results in the "chemical" master equation

$$\partial_\tau p(\tau, n|t_0, n_0) = \gamma_\tau \big[(n-l+k)_k p(\tau, n-l+k|\cdot) - (n)_k p(\tau, n|\cdot)\big]\,. \tag{27}$$

Microphysical arguments for its applicability to chemical reactions can be found in [82]. According to the chemical master equation, the mean particle number $\langle n\rangle = \sum_n n\, p(\tau, n|\cdot)$ obeys the equation $\partial_\tau \langle n\rangle = \gamma_\tau (l-k)\langle (n)_k\rangle$. For a system with a large number of particles ($n \gg k$), fluctuations can often be neglected in a first approximation, leading to the deterministic rate equation

$$\partial_\tau \bar{n} = \gamma_\tau (l-k)\bar{n}^k\,, \tag{28}$$

obeyed by a continuous particle number $\bar{n}(\tau) \in \mathbb{R}$.

Path integral representations of the chemical master equation (27) are sometimes said to be "bosonic". First, because an arbitrarily large number of molecules may in principle be present in the system (both globally and, upon extending the system to a spatial domain, locally). Second, because the path integral rep-

resentations are typically derived with the help of "creation and annihilation operators" fulfilling a "bosonic" commutation relation (see section 2.2 b). "Fermionic" path integrals, on the other hand, have been developed for systems in which the particles exclude one another. Thus, the number of particles in these systems is locally restricted to the values 0 and 1. For systems with excluding particles, the master equation's solution may be represented in terms of a path integral whose underlying creation and annihilation operators fulfil an anti-commutation relation [31–36]. However, van Wijland recently showed that the use of operators fulfilling the bosonic commutation relation is also possible [95]. These approaches are considered in section 2.2 d.

We do not intend to delve further into the historic development and applications of stochastic path integrals at this point. Doing so would require a proper introduction into renormalization group theory, which is of pivotal importance for the evaluation of the path integrals. Readers can find information on the application of renormalization group techniques in the review of Täuber, Howard, and Vollmayr-Lee [74] and in the book of Täuber [75]. Introductory texts are also provided by Cardy's (lecture) notes [96, 295]. Roughly speaking, path integral representations of the chemical master equation (27) have been used to assess how a macroscopic law of mass action changes due to fluctuations, both below [101, 135, 137, 296–307] and above the (upper) critical dimension [297, 308, 309], using either perturbative [19, 75, 135–138, 296–303, 305–307, 310–318] or non-perturbative [308, 309, 319–321] techniques. All of these articles focus on stochastic processes with spatial degrees of freedom for which alternative analytical approaches are scarce. Path integral representations of these processes, combined with renormalization group techniques, have been pivotal in understanding non-equilibrium phase transitions and they contributed significantly to the classification of these transitions in terms of universality classes [72, 75, 322]. Moreover, path integral representations of master equations have recently been employed in such diverse contexts as the study of neural networks [323–325], of ecological populations [161, 162, 167, 168, 326, 327], and of the differentiation of stem-cells [328].

1.6 Résumé

Continuous-time Markov processes with discontinuous sample paths describe a broad range of phenomena, ranging from traffic jams on highways [189] and on cytoskeletal filaments [118–120, 193] to novel forms of condensation in bosonic systems [1, 3]. In the introduction, we laid out the mathematical theory of these processes and derived the fundamental equations governing their evolution: the forward and the backward master equations. Whereas the forward master equation (11) evolves a conditional probability distribution forward in

time, the backward master equation (14) evolves the distribution backward in time. In the following main part of this chapter, we represent the conditional probability distribution solving the master equations in terms of path integrals. The framework upon which these path integrals are based unifies a broad range of approaches to the master equations, including, for example, the spectral method of Walzcak, Mugler, and Wiggins [39] and the Poisson representation of Gardiner and Chaturvedi [37, 38].

2 The probability generating function

The following two sections 2 and 3 are devoted to mapping the forward and backward master equations (11) and (14) to linear partial differential equations (PDEs). For brevity, we refer to such linear PDEs as "flow equations". In sections 4 and 5, the derived flow equations are solved in terms of path integrals.

It has been known since at least the 1940s that the (forward) master equation can be cast into a flow equation obeyed by the ordinary probability generating function

$$g(\tau; q|t_0, n_0) := \sum_{n\in\mathbb{N}_0} q^n p(\tau, n|t_0, n_0), \tag{29}$$

at least when the corresponding transition rate describes a simple chemical reaction [127–133, 329]. The generating function effectively replaces the discrete variable n by the continuous variable q. The absolute convergence of the sum in (29) is ensured (at least) for $|q| \leq 1$. The generating function "generates" probabilities in the sense that

$$p(\tau, n|t_0, n_0) = \frac{1}{n!}\partial_q^n g(\tau; q|t_0, n_0)\Big|_{q=0}. \tag{30}$$

This inverse transformation from g to p involves the application of the (real) linear functional $L_n[f] := \frac{1}{n!}\partial_q^n f(q)\big|_{q=0}$, which maps a (real-valued) function f to a (real) number. Moreover, it fulfils $L_n[f + \alpha g] = L_n[f] + \alpha L_n[g]$ for two functions f and g, and $\alpha \in \mathbb{R}$. A more convenient notation for linear functionals is introduced shortly. In the following, we generalize the probability generating function (29) and formulate conditions under which the generalized function obeys a linear PDE, i.e. a *flow equation*. But before proceeding to the general case, let us exemplify the use of a generalized probability generating function for a specific process (for brevity, we often drop the terms "probability" and "generalized" in referring to this function).

As the example, we consider the bi-directional reaction $\emptyset \rightleftharpoons A$ in which molecules of type A form at rate $\gamma \geq 0$ and degrade at per capita rate $\mu > 0$. According to the chemical master equation (27), the probability of observing

$n \in \mathbb{N}_0$ such molecules obeys the equation

$$\begin{aligned}\partial_\tau p(\tau, n|t_0, n_0) = \gamma\big[p(\tau, n-1|\cdot) - p(\tau, n|\cdot)\big] \\ + \mu\big[(n+1)p(\tau, n+1|\cdot) - np(\tau, n|\cdot)\big]\,,\end{aligned} \tag{31}$$

with initial condition $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$. This master equation respects the fact that the number of molecules cannot become negative through the reaction $A \to \emptyset$. By differentiating the probability generating function $g(\tau; q|t_0, n_0)$ in (29) with respect to the current time τ, one finds that it obeys the flow equation

$$\partial_\tau g = (q-1)(\gamma - \mu\partial_q)g\,. \tag{32}$$

Its time evolution starts out from $g(t_0; q|t_0, n_0) = q^{n_0}$. Instead of solving the flow equation right away, let us first simplify it by changing the basis function q^n of the generating function (29). As a first step, we change it to $(q+1)^n$, turning the corresponding flow equation into $\partial_\tau g = q(\gamma - \mu\partial_q)g$. As a second step, we multiply the new basis function by $\mathrm{e}^{-\frac{\gamma}{\mu}(q+1)}$ and arrive at the simplified flow equation

$$\partial_\tau g = -\mu q\partial_q g\,. \tag{33}$$

The generating function obeying this equation reads

$$g(\tau; q|\cdot) = \sum_{n\in\mathbb{N}_0} (q+1)^n \mathrm{e}^{-\frac{\gamma}{\mu}(q+1)}\, p(\tau, n|\cdot)\,. \tag{34}$$

As before, the dots inside the functions' arguments abbreviate the initial parameters t_0 and n_0.

The simplified flow equation (33) is now readily solved by separation of variables. But before doing so, let us introduce some new notation. From now on, we write the basis function as

$$|n\rangle_q := (q+1)^n \mathrm{e}^{-\frac{\gamma}{\mu}(q+1)} \tag{35}$$

and the corresponding generalized probability generating function (34) as

$$|g(\tau|\cdot)\rangle_q := \sum_{n\in\mathbb{N}_0} |n\rangle_q\, p(\tau, n|\cdot)\,. \tag{36}$$

In quantum mechanics, an object written as $|\cdot\rangle$ is called a "ket", a notation that was originally introduced by Dirac [26]. In the above two expressions, the kets simply represent ordinary functions. For brevity, we write the arguments of the

kets as subscripts and occasionally drop these subscripts altogether. In principle, the basis function could also depend on time (i.e. $|n\rangle_{\tau,q}$).

Later, in section 2.2, we introduce various basis functions for the study of different stochastic processes, including the Fourier basis function $|n\rangle_q := \mathrm{e}^{inq}$ for the solution of a random walk (with $n \in \mathbb{Z}$). Moreover, we consider a "linear algebra" approach in which $|n\rangle := \hat{e}_n$ represents the unit column vector in direction $n \in \mathbb{N}_0$ (this vector equals one at position $n+1$ and is zero everywhere else; cf. section 2.2 c). The generating function (36) corresponding to this "unit vector basis" coincides with the column vector $\boldsymbol{p}(\tau|t_0, n_0)$ of the probabilities $p(\tau, n|t_0, n_0)$. The unit vector basis will prove useful later on in recovering a path integral representation of averaged observables in section 6.2.

In our present example and in most of this this, however, the generating function (36) represents an ordinary function and obeys a linear PDE. For the basis function (35), this PDE reads

$$\partial_\tau |g\rangle = -\mu q \partial_q |g\rangle \,. \tag{37}$$

Its time evolution starts out from $|g(t_0|t_0, n_0)\rangle = |n_0\rangle$.

In order to recover the conditional probability distribution from the generating function (36), we now complement the "kets" with "bras". Such a bra is written as $\langle\cdot|$ and represents a linear functional in our present example. In particular, we define a bra $\langle m|$ for every $m \in \mathbb{N}_0$ by its following action on a test function f:

$$\langle m|f := \frac{1}{m!}\Big(\partial_q + \frac{\gamma}{\mu}\Big)^m f(q)|_{q=-1} \,. \tag{38}$$

The evaluation at $q = -1$ could also be written in integral form as $\int_{\mathbb{R}} \mathrm{d}q\, \delta(q+1)(\cdots)$. The functional $\langle m|$ is obviously linear and maps the basis function (35) to

$$\langle m|n\rangle = \delta_{m,n} \,. \tag{39}$$

Thus, the "basis functionals" in $\{\langle m|\}_{m\in\mathbb{N}_0}$ are orthogonal to the basis functions in $\{|n\rangle\}_{n\in\mathbb{N}_0}$. The orthogonality condition can be used to recover the conditional probability distribution from (36) via

$$p(\tau, n|\cdot) = \langle n|g(\tau|\cdot)\rangle \,. \tag{40}$$

Besides being orthogonal to one another, the kets (35) and bras (38) fulfil the completeness relation $\sum_n |n\rangle_q \langle n|f = f(q)$ with respect to analytic functions[4] (note that $\langle n|f$ as defined in (38) is just a real number and does not depend on q). Above, we mentioned that we later introduce alternative basis kets, including

4 As before, sums whose range is not specified cover the whole state space.

the Fourier basis function $|n\rangle_q = \mathrm{e}^{\mathrm{i}nq}$ for the solution of a random walk (with $n \in \mathbb{Z}$), and the unit column vector $|n\rangle = \hat{\boldsymbol{e}}_n$ with $n \in \mathbb{N}_0$. These kets can also be complemented to obtain orthogonal and complete bases, namely by complementing the Fourier basis function with the basis functional $\langle m|f \coloneqq \int_{-\pi}^{\pi} \frac{\mathrm{d}q}{2\pi} \mathrm{e}^{-\mathrm{i}mq} f(q)$ ($m \in \mathbb{Z}$), and by complementing the unit column vector with the unit row vector $\langle m| \coloneqq \hat{\boldsymbol{e}}_m^{\mathsf{T}}$ ($m \in \mathbb{N}_0$). For the unit vector basis $\{|n\rangle, \langle m|\}_{n,m\in\mathbb{N}_0}$, the completeness condition $\sum_n |n\rangle\langle n| = \mathbb{1}$ involves the (infinitely large) unit matrix $\mathbb{1}$.

Thanks to the new basis function (35), the simplified flow equation (37) can be easily solved by separation of variables. Making the ansatz $|g\rangle = f(\tau)h(q)$, one obtains an equation whose two sides depend either on f or on h but not on both. The equation is solved by $f(\tau) = \mathrm{e}^{-k\mu(\tau-t_0)}$ and $h(q) = q^k$, with k being a non-negative parameter. The non-negativity of k ensures the finiteness of the initial condition $|g(t_0|t_0, n_0)\rangle_q = |n_0\rangle_q$ in the limit $q \to 0$. By the completeness of the polynomial basis, the values of k can be restricted to $\mathbb{N}_0$. It proves convenient to represent also the standard polynomial basis in terms of bras and kets, namely by defining $\langle\!\langle k|f \coloneqq \frac{1}{k!}\partial_q^k f(q)|_{q=0}$ and $|k\rangle\!\rangle_q \coloneqq q^k$. These bras and kets are again orthogonal to one another in the sense of $\langle\!\langle k|l\rangle\!\rangle = \delta_{k,l}$ and they also fulfil a completeness relation ($\sum_k |k\rangle\!\rangle\langle\!\langle k|$ represents a Taylor expansion around $q = 0$ and thus acts as an identity on analytic functions). Using the auxiliary bras and kets, the solution of the flow equation (37) can be written as

$$|g(\tau|\cdot)\rangle = \sum_{k\in\mathbb{N}_0} |k\rangle\!\rangle\, \mathrm{e}^{-k\mu(\tau-t_0)} \langle\!\langle k|n_0\rangle\,. \tag{41}$$

We wrote the expansion coefficient in this solution as $\langle\!\langle k|n_0\rangle$ to respect the initial condition $|g(t_0|\cdot)\rangle = |n_0\rangle$.

The conditional probability distribution can be recovered from the generating function (41) via the inverse transformation (40) as

$$p(\tau, n|t_0, n_0) = \sum_{k\in\mathbb{N}_0} \langle n|k\rangle\!\rangle\, \mathrm{e}^{-k\mu(\tau-t_0)} \langle\!\langle k|n_0\rangle\,. \tag{42}$$

The coefficients $\langle n|k\rangle\!\rangle$ and $\langle\!\langle k|n_0\rangle$ can be computed recursively as explained in [330]. Here we are interested in the asymptotic limit $\tau \to \infty$ of the distribution (42) for which only the $k = 0$ "mode" survives. Therefore, the distribution converges to the stationary Poisson distribution

$$p(\infty, n|t_0, n_0) = \langle n|0\rangle\!\rangle\langle\!\langle 0|n_0\rangle = \frac{(\gamma/\mu)^n \mathrm{e}^{-\gamma/\mu}}{n!}\,. \tag{43}$$

The above example illustrates how the master equation can be transformed into a linear PDE obeyed by a generalized probability generating function and

how this PDE simplifies for the right basis function. The explicit choice of the basis function depends on the problem at hand. Moreover, the above example introduced the bra-ket notation used in this thesis. In section 4.2, the reaction $\emptyset \rightleftharpoons A$ will be reconsidered using a path integral representation of the probability distribution. We will then see that this process is not only solved by a Poisson distribution in the stationary limit, but actually for all times (at least, if the number of molecules in the system was initially Poisson distributed).

In the remainder of this section, as well as in section 3, we generalize the above approach and derive flow equations for the following four series expansions (with dynamic time variable $\tau \in [t_0, t]$):

$$\sum_n |n\rangle p(\tau, n|t_0, n_0)\,, \tag{44}$$

$$\sum_{n_0} p(t, n|\tau, n_0)|n_0\rangle\,, \tag{45}$$

$$\sum_n \langle n|p(\tau, n|t_0, n_0)\,, \tag{46}$$

$$\sum_{n_0} p(t, n|\tau, n_0)\langle n_0|\,. \tag{47}$$

Apparently, the series (44) coincides with the generalized probability generating function (36). In the next section 2.1, we formulate general conditions under which this function obeys a linear PDE. The remaining series (45)–(47) may not be as familiar. We call the series (45) a "marginalized distribution" (45). It will be shown in section 3 that this series does not only solve the (forward) master equation, but that it also obeys a backward-time PDE under certain conditions. The marginalized distribution proves useful in the computation of mean extinction times as we demonstrate in section 3.3. In section 3.4, we consider the "probability generating functional" (46). For a "Poisson basis function", the inverse transformation, which maps this functional to the conditional probability distribution, coincides with the Poisson representation of Gardiner and Chaturvedi [37, 38]. The potential use of the series (47) remains to be explored.

The goal of the subsequent sections 4 and 5 lies in the solution of the derived flow equations by path integrals. In section 4, we first solve the flow equations obeyed by the marginalized distribution (45) and by the generating functional (46) in terms of a "backward" path integral. Afterwards, in section 5, the flow equations obeyed by the generating function (44) and by the series expansion (47) are solved in terms of a "forward" path integral. Inverse transformations, such as (40), will then provide distinct path integral representations of the forward and backward master equations.

2.1 Flow of the generating function

We now formulate general conditions under which the forward master equation (11) can be cast into a linear PDE obeyed by the generalized probability

generating function

$$|g(\tau|t_0, n_0)\rangle_q := \sum_n |n\rangle_{\tau,q}\, p(\tau, n|t_0, n_0)\,. \tag{48}$$

The basis function $|n\rangle_{\tau,q}$ is a function of the variable q and possibly of the time variable τ. But unless one of these variables is of direct relevance, its corresponding subscript will be dropped. The explicit form of the basis function depends on the problem at hand and is chosen so that the four conditions (O), (C), (E), and (Q) below are satisfied (the conditions (O) and (C) concern the orthogonality and completeness of the basis, which we already required in the introductory example). The variable n again represents some state from a countable state space. For example, n could describe the position of a molecular motor along a cytoskeletal filament ($n \in \mathbb{Z}$), the copy number of a molecule ($n \in \mathbb{N}_0$), the local copy numbers of the molecule on a lattice ($n \equiv \{n_i \in \mathbb{N}_0\}_{i\in\mathbb{Z}}$), or the copy numbers of multiple kinds of molecules ($n \equiv (n^A, n^B, n^C) \in \mathbb{N}_0^3$). For the multivariate configurations, the basis function is typically decomposed into a product $|n_1\rangle|n_2\rangle\cdots$ of individual basis functions $|n_i\rangle$, each depending on its own variable q_i. A process with spatial degrees of freedom is considered in section 4.5 b. Besides, we also consider a system of excluding particles in section 2.2 d. There, the (local) number of particles n is restricted to the values 0 and 1.

The definition of the generating function (48) assumes the existence of a set $\{|n\rangle_\tau\}$ of basis functions for every time $\tau \in [t_0, t]$. In addition, we assume that there exists a set $\{\langle m|_\tau\}$ of linear basis functionals for every time $\tau \in [t_0, t]$. These bras shall be orthogonal to the kets in the sense that at each time point τ, they act on the kets as

$$\langle m|_\tau |n\rangle_\tau = \delta_{m,n} \tag{O}$$

(for all m and n). Here we note the possible time-dependence of the basis because the (O)rthogonality condition will only be required for equal times of the bras and kets. In addition to orthogonality, the basis shall fulfil the (C)ompleteness condition

$$\sum_n |n\rangle_\tau \langle n|_\tau f = f\,, \tag{C}$$

where f represents an appropriate test function. The completeness condition implies that the function f can be decomposed in the basis functions $|n\rangle$ with expansion coefficients $\langle n|f$. In the introduction to this section, we introduced various bases fulfilling both the orthogonality and the completeness condition. As in the introductory example, the orthogonality condition allows one to recover the conditional probability distribution via the inverse transformation

$$p(\tau, n|t_0, n_0) = \langle n|g(\tau|t_0, n_0)\rangle\,. \tag{51}$$

Before deriving the flow equation obeyed by the generating function, let us note that the (O)rthogonality condition differs slightly from the corresponding conditions used in most other texts on stochastic path integrals (see, for example, [18, 20] or the "exclusive scalar product" in [15]). Typically, the orthogonality condition includes an additional factorial $n!$ on its right hand side. The inclusion of this factorial is advantageous in that it accentuates a symmetry between the bases that we consider in sections 2.2 b and 3.2 b for the study of chemical reactions. Its inclusion would be rather unusual, however, for the Fourier basis introduced in sections 2.2 a and 3.2 a. The Fourier basis will be used to solve a simple random walk. Moreover, the factorial obscures a connection between the probability generating functional introduced in section 3.4 and the Poisson representation of Gardiner and Chaturvedi [37, 38]. We discuss this connection in section 3.5.

To derive the flow equation obeyed by the generating function $|g\rangle$, we differentiate its definition (48) with respect to the time variable τ. The resulting time derivative of the conditional probability distribution $p(\tau, n|t_0, n_0)$ can be replaced by the right-hand side of the forward master equation (11). In matrix notation, this equation reads $\partial_\tau p(\tau|t_0) = Q_\tau p(\tau|t_0)$. Eventually, one finds that

$$\partial_\tau |g\rangle = \sum_n \Big(\partial_\tau |n\rangle + \sum_m |m\rangle Q_\tau(m,n) \Big) p(\tau, n|\cdot) \,. \tag{52}$$

Our goal is to turn this expression into a partial differential equation for $|g\rangle$. For this purpose, we require two differential operators. First, we require a differential operator $\mathcal{E}_\tau(q, \partial_q)$ encoding the time evolution of the basis function. In particular, this operator should fulfil, for all values of n,

$$\mathcal{E}_\tau |n\rangle = \partial_\tau |n\rangle \,. \tag{E}$$

By the (O)rthogonality condition, one could also define this operator in a "constructive" way as

$$\mathcal{E}_\tau := \sum_n \big(\partial_\tau |n\rangle \big) \langle n| \,. \tag{54}$$

We call $\mathcal{E}_\tau$ the basis (E)volution operator. In order to arrive at a proper PDE for $|g\rangle$, $\mathcal{E}_\tau(q, \partial_q)$ should be polynomial in ∂_q (later, in section 2.2 d we also encounter a case in which it constitutes a power series with infinitely high powers of ∂_q). For now, the pre-factors of 1, ∂_q, ∂_q^2,... may be arbitrary functions of q. Later, in our derivation of a path integral in section 5, we will also require that the pre-factors can be expanded in powers of q. Note that for a multivariate configuration $n \equiv (n^A, n^B, \ldots)$, the derivative ∂_q represents individual derivatives with respect to $q \equiv (q^A, q^B, \ldots)$.

The actual dynamics of a jump process are encoded by its transition rate matrix Q_τ (see section 1.3). The off-diagonal elements of this matrix are the transition rates $w_\tau(n,m)$ from a state m to a state n, and its diagonal elements are the negatives of the exit rates $w_\tau(m) = \sum_n w_\tau(n,m)$ from a state m (with $w_\tau(m,m) = 0$; see (10)). We encode the information stored in Q_τ by a second differential operator called $\mathcal{Q}_\tau(q,\partial_q)$. This operator should fulfil, for all values of n,

$$\mathcal{Q}_\tau|n\rangle = \sum_m |m\rangle Q_\tau(m,n)\,. \tag{Q}$$

In analogy with the transition (rate) matrix Q_τ, we call $\mathcal{Q}_\tau$ the transition (rate) operator (note that we only speak of "operators" with respect to differential operators, but not with respect to matrices). Just like the basis (E)volution operator, $\mathcal{Q}_\tau(q,\partial_q)$ should be polynomial in ∂_q. In section 5, it will be assumed that $\mathcal{Q}_\tau(q,\partial_q)$ can be expanded in powers of both q and ∂_q. In a constructive approach, one could also define the operator as

$$\mathcal{Q}_\tau := \sum_{m,n} |m\rangle Q_\tau(m,n)\langle n|\,. \tag{56}$$

This constructive definition does not guarantee, however, that $\mathcal{Q}_\tau(q,\partial_q)$ has the form of a differential operator. This property is, for example, not immediately clear for the Fourier basis function $|n\rangle_q = \mathrm{e}^{inq}$ (with $n \in \mathbb{Z}$), which we complemented with the functional $\langle n|f = \int_{-\pi}^{\pi} \frac{\mathrm{d}q}{2\pi} \mathrm{e}^{-inq} f(q)$ in the introduction to this section. Most of the processes that we solve in later sections have polynomial transition rates. Suitable bases and operators for these processes are provided in the next section 2.2. It remains an open problem for the field to find such bases and operators for processes whose transition rates have different functional forms. That is, for example, the case for transition rates that saturate with the number of particles and have the form of a Hill function.

Provided that one has found a transition operator $\mathcal{Q}_\tau$ and a basis (E)volution operator $\mathcal{E}_\tau$ for a (C)omplete and (O)rthogonal basis, it follows from (52) that the generalized generating function $|g(\tau|t_0,n_0)\rangle$ obeys the flow equation[5]

$$\partial_\tau|g\rangle = (\mathcal{E}_\tau + \mathcal{Q}_\tau)|g\rangle =: \widetilde{\mathcal{Q}}_\tau|g\rangle\,. \tag{57}$$

5 In deriving the forward path integral in section 5.1, we use the finite difference approximation

$$\begin{aligned}
\partial_\tau|g(\tau|t_0,n_0)\rangle &= \lim_{\Delta t\to 0} \frac{|g(\tau+\Delta t|\cdot)\rangle - |g(\tau|\cdot)\rangle}{\Delta t} \\
&= \lim_{\Delta t\to 0}\Big(\sum_n \frac{|n\rangle_{\tau+\Delta t} - |n\rangle_\tau}{\Delta t}\big(p(\tau,n|\cdot) + \mathrm{O}(\Delta t)\big) + \sum_n |n\rangle_\tau \frac{p(\tau+\Delta t,n|\cdot) - p(\tau,n|\cdot)}{\Delta t}\Big) \\
&= \lim_{\Delta t\to 0}(\mathcal{E}_{\tau,\Delta t} + \mathcal{Q}_{\tau,\Delta t})|g(\tau|\cdot)\rangle = \lim_{\Delta t\to 0} \widetilde{\mathcal{Q}}_{\tau,\Delta t}|g(\tau|\cdot)\rangle\,,
\end{aligned}$$

Its initial condition reads $|g(t_0|t_0, n_0)\rangle = |n_0\rangle$, with the possibly time-dependent basis function $|n_0\rangle$ being evaluated at time t_0. Although time-dependent bases prove useful in section 7, we mostly work with time-independent bases in the following. The operators $\widetilde{\mathcal{Q}}_\tau$ and $\mathcal{Q}_\tau$ then agree because $\mathcal{E}_\tau$ is zero. Therefore, we refer to both $\widetilde{\mathcal{Q}}_\tau$ and $\mathcal{Q}_\tau$ as transition (rate) operators.

In our above derivation, we assumed that the ket $|n\rangle$ represents an ordinary function and that the bra $\langle n|$ represents a linear functional. To understand why we chose similar letters for the Q_τ-matrix and the $\mathcal{Q}_\tau$-operator, it is insightful to consider the unit column vectors $|n\rangle \coloneqq \hat{\boldsymbol{e}}_n$ and the unit row vectors $\langle n| \coloneqq \hat{\boldsymbol{e}}_n^\mathsf{T}$ (with $m, n \in \mathbb{N}_0$). For these vectors, the right hand side of the transition operator (56) simply constitutes a representation of the Q_τ-matrix. Hence, $\mathcal{Q}_\tau$ is not a differential operator in this case but coincides with the Q_τ-matrix. This observation does not come as a surprise because we already noted that the generating function (48) represents the vector $\boldsymbol{p}(\tau|t_0, n_0)$ of all probabilities in this case. Moreover, the corresponding flow equation (57) does not constitute a linear PDE but a vector representation of the forward master equation (11).

Following Doi, the transition operator $\widetilde{\mathcal{Q}}_\tau$ could be called a "time evolution" or "Liouville" operator [11, 12], or, following Zel'dovich and Ovchinnikov, a "Hamiltonian" [13]. The latter name is due to the formal resemblance of the flow equation (57) to the Schrödinger equation in quantum mechanics [97] (this resemblance holds for any linear PDE that is first order in ∂_τ). The name "Hamiltonian" has gained in popularity throughout the recent years, possibly because the path integrals derived later in this chapter share many formal similarities with the path integrals employed in quantum mechanics [112, 287–289]. Nevertheless, let us point out that $\widetilde{\mathcal{Q}}_\tau$ is generally not Hermitian and that the generating function $|g\rangle$ does not represent a wave function and also not a probability (unless one chooses the unit vectors $|n\rangle = \hat{\boldsymbol{e}}_n$ and $\langle n| = \hat{\boldsymbol{e}}_n^\mathsf{T}$ as basis). In our work, we stick to the name transition (rate) operator for $\widetilde{\mathcal{Q}}_\tau$ (and $\mathcal{Q}_\tau$) to emphasize its connection to the transition (rate) matrix Q_τ.

2.2 Bases for particular stochastic processes

In the previous section, we formulated four conditions under which the master equation (11) can be cast into the linear PDE (57) obeyed by the generalized

---

with the operators being defined through their following actions on the basis functions:

$$\mathcal{E}_{\tau,\Delta t}|n\rangle_\tau = \frac{|n\rangle_{\tau+\Delta t} - |n\rangle_\tau}{\Delta t} \quad \text{and} \quad \mathcal{Q}_{\tau,\Delta t}|n\rangle_\tau = \sum_m |m\rangle_\tau Q_{\tau,\Delta t}(m,n)\,.$$

The discretization scheme conforms with the derivation of the forward master equation (11).

generating function (48). In the following, we illustrate how these conditions can be met for various stochastic processes.

2.2 a Random walks

A simple process that can be solved by the method from the previous section is the one-dimensional random walk. We model this process in terms of a particle sitting at position n of the one-dimensional lattice $\mathbb{L} := \mathbb{Z}$. The particle may jump to the nearest lattice site on its right with the (possibly time-dependent) rate $r_\tau > 0$, and to the site on its left with the rate $l_\tau > 0$. Given that the particle was at position n_0 at time t_0, the probability of finding it at position n at time τ obeys the master equation

$$\begin{aligned}\partial_\tau p(\tau, n|t_0, n_0) = r_\tau\big[p(\tau, n-1|\cdot) - p(\tau, n|\cdot)\big] \\ + l_\tau\big[p(\tau, n+1|\cdot) - p(\tau, n|\cdot)\big],\end{aligned} \tag{58}$$

with initial condition $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$. One can solve this equation by solving the associated flow equation (57). But for this purpose, we first require an (O)rthogonal and (C)omplete basis, as well as a basis (E)volution operator $\mathcal{E}_\tau$ and a transition operator $\mathcal{Q}_\tau$ (condition (Q)).

An appropriate choice of the orthogonal and complete basis proves to be the time-independent Fourier basis

$$|n\rangle_q := \mathrm{e}^{\mathrm{i}nq} \text{ and } \langle n|f := \int_{-\pi}^{\pi} \frac{\mathrm{d}q}{2\pi} \mathrm{e}^{-\mathrm{i}nq} f(q) \tag{59}$$

with $n \in \mathbb{Z}$ and test function f. For this basis, the generalized generating function $|g\rangle = \sum_n \mathrm{e}^{\mathrm{i}nq} p(\tau, n|\cdot) = \langle \mathrm{e}^{\mathrm{i}nq}\rangle$ coincides with the characteristic function. Moreover, the corresponding orthogonality condition $\langle m|n\rangle = \delta_{m,n}$ agrees with a common integral representation of the Kronecker delta. The completeness of the basis can be shown with the help of a Fourier series representation of the "Dirac comb"

$$D(q) = \sum_{n\in\mathbb{Z}} \delta(q - 2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}nq}\,. \tag{60}$$

It thereby follows that

$$\sum_{n\in\mathbb{Z}} |n\rangle_q \langle n|f = \sum_{n\in\mathbb{Z}} \mathrm{e}^{\mathrm{i}nq} \int_{-\pi}^{\pi} \frac{\mathrm{d}q'}{2\pi} \mathrm{e}^{-\mathrm{i}nq'} f(q') \tag{61}$$

$$= \int_{-\pi}^{\pi} \mathrm{d}q'\, D(q - q') f(q') = f(q)\,. \tag{62}$$

Since the Fourier basis function is time-independent, the condition (E) is trivially fulfilled for the evolution operator $\mathcal{E}_\tau := 0$. The only piece still missing is the transition operator $\mathcal{Q}_\tau$. Its defining condition (Q) requires knowledge of the transition matrix Q_τ whose elements

$$Q_\tau(m,n) = r_\tau(\delta_{m,n+1} - \delta_{m,n}) + l_\tau(\delta_{m,n-1} - \delta_{m,n}) \tag{63}$$

can be inferred by comparing the master equation (58) with its general form (11). The condition (Q) therefore reads

$$\mathcal{Q}_\tau|n\rangle = r_\tau\big(|n+1\rangle - |n\rangle\big) + l_\tau\big(|n-1\rangle - |n\rangle\big)\,. \tag{64}$$

One can construct a transition operator with this property with the help of the functions $c(q) := \mathrm{e}^{\mathrm{i}q}$ and $a(q) := \mathrm{e}^{-\mathrm{i}q}$. The function c shifts the basis function $|n\rangle_q = \mathrm{e}^{\mathrm{i}nq}$ to the right via $c|n\rangle = |n+1\rangle$, and the function a shifts it to the left via $a|n\rangle = |n-1\rangle$. Thus, an operator with the property (64) can be defined as

$$\mathcal{Q}_\tau(q,\partial_q) := r_\tau\big[c(q) - 1\big] + l_\tau\big[a(q) - 1\big]\,. \tag{65}$$

This operator can also be inferred from its constructive definition (56) by making use of the Dirac comb.

$$\begin{aligned}\sum_{m,n\in\mathbb{Z}} |m\rangle_q Q_\tau(m,n)\langle n|f &= \int_{-\pi}^{\pi} \mathrm{d}q'\, D(q-q')\big[r_\tau(\mathrm{e}^{\mathrm{i}q'} - 1) + l_\tau(\mathrm{e}^{-\mathrm{i}q'} - 1)\big]f(q') \\ &= \mathcal{Q}_\tau(q,\partial_q)f(q)\,. \end{aligned} \tag{66}$$

After putting the above pieces together, one finds that the generating function $|g(\tau|t_0,n_0)\rangle_q$ obeys the flow equation

$$\partial_\tau|g\rangle = \big[r_\tau(\mathrm{e}^{\mathrm{i}q} - 1) + l_\tau(\mathrm{e}^{-\mathrm{i}q} - 1)\big]|g\rangle \tag{67}$$

for $t_0 \le \tau \le t$ with the initial value $|g(t_0|t_0,n_0)\rangle = |n_0\rangle = \mathrm{e}^{\mathrm{i}n_0 q}$. The flow equation is readily solved by

$$|g\rangle = \exp\Big((\mathrm{e}^{\mathrm{i}q} - 1)\int_{t_0}^{\tau} \mathrm{d}s\, r_s + (\mathrm{e}^{-\mathrm{i}q} - 1)\int_{t_0}^{\tau} \mathrm{d}s\, l_s\Big)|n_0\rangle\,.$$

The conditional probability distribution can be recovered from this generating function by performing the inverse Fourier transformation (51), i.e. by evaluating $p(\tau,n|t_0,n_0) = \langle n|g\rangle$. A series expansion of all of the involved exponentials and some laborious rearrangement of sums eventually result in a Skellam distribution as the solution of the process [331]. We outline the derivation of this distribution

in appendix C. The distribution's mean is $\mu = n_0 + \int_{t_0}^{\tau} \mathrm{d}s\,(r_s - l_s)$ and its variance $\sigma^2 = \int_{t_0}^{\tau} \mathrm{d}s\,(r_s + l_s)$. These moments also follow from the fact that the Skellam distribution describes the difference of two Poisson random variables; one for jumps to the right, the other for jumps to the left. The two moments can, also be obtained more easily by deriving equations for their time evolution from the master equation. Those equations do not couple for the simple random walk. Let us also note that if the two jump rates r_τ and l_τ agree, the Skellam distribution tends to a Gaussian for large times.

2.2 b Chemical reactions

As a second example, we turn to processes whose transition rates can be decomposed additively into the transition rates of simple chemical reactions. In such a reaction, $k_1, k_2, \ldots$ molecules of types $A_1, A_2, \ldots$ come together to be replaced by $l_1, l_2, \ldots$ molecules of the same types (with $k_j, l_j \in \mathbb{N}_0$). Besides reacting with each other, the molecules could also diffuse in space, which can be modelled in terms of hopping processes on a regular, d-dimensional lattice such as $\mathbb{L} := \mathbb{Z}^d$. Upon labelling particles on different lattice sites by their positions, the hopping of a molecule of type A_1 from lattice site $i \in \mathbb{L}$ to lattice site $j \in \mathbb{L}$ could be regarded as the chemical reaction $A_1^{(i)} \to A_1^{(j)}$. We consider the "chemical" master equation associated to such hopping processes in section 4.5 b.

To demonstrate the generating function approach from section 2.1, we focus on a system with only a single type of molecule A engaging in the "well-mixed" reaction $k\,A \to l\,A$ ($k, l \in \mathbb{N}_0$). Since the forward master equation (11) is linear in the transition rate $Q_\tau(n, m)$, the following considerations readily extend to networks of multiple reactions, multiple types of molecules, and processes with spatial degrees of freedom. The basis functions and functionals introduced below can, for example, be used to study branching and annihilating random walks, which are commonly modelled in terms of diffusing particles that engage in the binary annihilation reaction $2\,A \to \emptyset$ and the linear growth process $A \to (1 + m)A$ [101, 137, 319, 332, 333]. For an odd number of offspring, these walks exhibit an absorbing state phase transition falling into the universality class of directed percolation (according to perturbative calculations in one and two spatial dimensions [101, 137], according to non-perturbative calculations in up to six dimensions [320]). The decomposition of a process into elementary reactions of the form $k\,A \to l\,A$ is also possible in the contexts of growing polymer crystals [18], aggregation phenomena [296], and predator-prey ecosystems [168].

As explained in section 1.5, the transition rate $w_\tau(m, n) = \gamma_\tau \delta_{m,n-k+l}(n)_k$ of the reaction $k\,A \to l\,A$ is determined by combinatorial counting. Its proportionality to the falling factorial $(n)_k = n(n-1)\cdots(n-k+1)$ derives from picking k molecules out of a population of n molecules. The overall time scale of the

reaction is set by the rate coefficient γ_τ (this coefficient may also absorb a factorial $k!$ accounting for the indistinguishability of molecules). The above transition rate guarantees that the number of molecules (or "particles") in the system never drops below zero, provided that it was non-negative initially. Thus, the state space of n is $\mathbb{N}_0$ and we require basis functions $|n\rangle$ and basis functionals $\langle n|$ only for such values.

Before specifying the (C)omplete and (O)rthogonal basis as well as the basis (E)volution operator $\mathcal{E}_\tau$, let us first specify an appropriate transition operator $\mathcal{Q}_\tau$. Its corresponding condition (Q) depends on the transition matrix Q_τ, whose elements $Q_\tau(m,n) = \gamma_\tau(n)_k(\delta_{m,n-k+l} - \delta_{m,n})$ can be inferred from the rate $w_\tau(m,n)$ with the help of the relation (10). Consequently, the condition (Q) reads

$$\mathcal{Q}_\tau|n\rangle = \gamma_\tau(n)_k\big(|n-k+l\rangle - |n\rangle\big)\,. \tag{68}$$

This condition is met by the transition operator

$$\mathcal{Q}_\tau(c,a) \coloneqq \gamma_\tau(c^l - c^k)a^k\,, \tag{69}$$

provided that there exist a "creation" operator c and an "annihilation" operator a acting as

$$c|n\rangle = |n+1\rangle \text{ and} \tag{70}$$
$$a|n\rangle = n|n-1\rangle\,. \tag{71}$$

The second of these relations ensures that basis functions with $n < 0$ do not appear because $a|0\rangle$ vanishes. Both c and a may depend on time, just as the basis functionals and functions $\langle n|$ and $|n\rangle$ may do. It follows from (70) and (71) that the two operators fulfil the commutation relation

$$[a_\tau, c_\tau] \coloneqq a_\tau c_\tau - c_\tau a_\tau = 1 \tag{72}$$

at a fixed time τ. The commutation relation is meant with respect to functions that can be expanded in the basis functions yet to be defined. For brevity, we commonly drop the subscript τ of the creation and annihilation operators. Together with the orthogonality condition, the commutation relation implies that the operators act on the basis functionals as

$$\langle n|c = \langle n-1| \text{ and} \tag{73}$$
$$\langle n|a = (n+1)\langle n+1|\,. \tag{74}$$

In quantum mechanics, creation and annihilation operators prove useful in solving the equation of motion of a quantum particle in a quadratic potential

(i.e. in solving the Schrödinger equation of the quantum harmonic oscillator) [27, 334]. There, the ket $|n\rangle$ is interpreted as carrying n quanta of energy $\hbar\omega$ in addition to the ground state energy $\hbar\omega/2$ of the oscillator's "vacuum state" $|0\rangle$ (with Dirac constant $\hbar$ and angular frequency ω). The creation operator then adds a quantum of energy to state $|n\rangle$ via the relation $c|n\rangle = \sqrt{n+1}|n+1\rangle$, and the annihilation operator removes an energy quantum via $a|n\rangle = \sqrt{n}|n-1\rangle$ (a also destroys the vacuum state). These relations differ from the ones in (70) and (71) because in quantum mechanics, the creation and annihilation operators are defined in terms of self-adjoint position and momentum operators, forcing the former operators to be hermitian adjoints of each other (i.e. $c = a^\dagger$ and $a = c^\dagger$). Consequently, the relations corresponding to (73) and (74) read $\langle n|a^\dagger = \sqrt{n}\langle n-1|$ and $\langle n|a = \sqrt{n+1}\langle n+1|$. Nevertheless, the commutation relation (72) also holds in the quantum world in which its validity ultimately derives from the non-commutativity of the position operator Q and the momentum operator P (i.e. from $[Q,P] = \mathrm{i}\hbar$, which follows from P being the generator of spatial displacements in state space [334]). Besides, let us note that in quantum field theory, the creation operator is interpreted as adding a bosonic particle to an energy state, and the annihilation operator as removing one [335].

One may wonder whether the above interpretations can be transferred to the theory of stochastic processes in which the creation and annihilation operators need not be each other's adjoints. For example, the creation operator in (70) could be interpreted as adding a molecule to a system, and the corresponding annihilation operator (71) as removing a molecule. The commutation relation (72) may then be interpreted as that the addition of a particle to the system (one way to do it) and the removal of a particle (many ways to do it) do not commute. Let us, however, note that these interpretations only apply to the processes discussed in the present section, which can be decomposed into reactions of the form $k\,A \to l\,A$ (with possibly multiple types of particles and spatial degrees of freedom). The interpretations do not apply, for example, to the random walk of a single particle with state space $\mathbb{Z}$, which we discussed in the previous section (there, the "shift operators" with actions $c|n\rangle = |n+1\rangle$ and $a|n\rangle = |n-1\rangle$ commute).

Whether creation and annihilation operators with the properties (70) and (71) exist depends on the choice of the basis functions. For the study of chemical reactions, a useful choice proves to be the basis function

$$|n\rangle_q := (\zeta q + \tilde{q})^n \mathrm{e}^{-\tilde{x}(\zeta q+\tilde{q})}\,. \tag{75}$$

Here, $\tilde{q}(\tau)$ and $\tilde{x}(\tau)$ are arbitrary, possibly time-dependent functions and $\zeta \neq 0$ is a free parameter. This parameter only becomes relevant in section 6 in recovering a path integral representation of averaged observables. There, its value is set to $\zeta := \mathrm{i}$ but typically we choose $\zeta := 1$. For the latter choice, the basis function (75)

simplifies to

$$|n\rangle_q = (q + \tilde{q})^n \mathrm{e}^{-\tilde{x}(q+\tilde{q})} . \tag{76}$$

Alternatively, the parameter ζ could be used to rescale the variable q by a system size parameter N if such a parameter is available.

The two functions $\tilde{q}$ and $\tilde{x}$ may prove helpful in simplifying the flow equation obeyed by the generating function. For example, we chose $\tilde{q} := 1$ and $\tilde{x} := \gamma/\mu$ in the introduction to section 2 to simplify the flow equation of the reaction $\emptyset \rightleftharpoons A$ (with growth rate coefficient γ and decay rate coefficient μ). Later, in section 7, $\tilde{q}$ and $\tilde{x}$ will act as the stationary paths of a path integral with q and an auxiliary variable x being deviations from them. If both $\tilde{q}$ and $\tilde{x}$ are chosen as zero, the basis function (76) simplifies to the basis function

$$|n\rangle_q = q^n \tag{77}$$

of the ordinary probability generating function.

For the general basis function (75), creation and annihilation operators with the properties (70) and (71) can be defined as

$$c(q, \partial_q) := \zeta q + \tilde{q} \text{ and} \tag{78}$$
$$a(q, \partial_q) := \partial_{\zeta q} + \tilde{x} . \tag{79}$$

Apparently, these operators are not each other's adjoints. For the basis function $|n\rangle_q = q^n$ of the ordinary probability generating function, the operators simplify to $c = q$ and $a = \partial_q$. The corresponding transition operator (69) reads $\mathcal{Q}_\tau(q, \partial_q) = \gamma_\tau(q^l - q^k)\partial_q^k$, resulting in the flow equation

$$\partial_\tau |g\rangle = \gamma_\tau (q^l - q^k)\partial_q^k |g\rangle . \tag{80}$$

Thus far, we have not specified the basis functionals. These functionals can be defined using the annihilation operator (79). In particular, we complement the basis function (75) with the functionals

$$\langle n|f := \frac{a^n}{n!} f(q)\big|_{q=-\tilde{q}/\zeta} , \tag{81}$$

where f represents a test function. The (O)rthogonality of the basis $\{\langle m|, |n\rangle\}_{m,n\in\mathbb{N}_0}$ follows directly from the action of the annihilation operator on the basis function $|n\rangle$, and from the vanishing of this function at $q = -\tilde{q}/\zeta$, except for $n = 0$. In particular,

$$\langle m|n\rangle = \frac{a^m}{m!}|n\rangle\big|_{q=-\tilde{q}(\tau)/\zeta} = \binom{n}{m}|n-m\rangle\big|_{q=-\tilde{q}(\tau)/\zeta}\Theta_{n-m} = \delta_{m,n} , \tag{82}$$

with $\Theta_n = 1$ for $n \geq 0$ and $\Theta_n = 0$ otherwise. The (C)ompleteness of the basis is also readily established.

$$\sum_{n\in\mathbb{N}_0} |n\rangle_q\langle n|f = \sum_{n\in\mathbb{N}_0} (\zeta q + \tilde{q})^n \mathrm{e}^{-\tilde{x}(\zeta q+\tilde{q})} \frac{(\partial_{\zeta q'} + \tilde{x})^n}{n!} f(q')\big|_{q'=-\tilde{q}/\zeta} \tag{83}$$

$$= \mathrm{e}^{-\tilde{x}(\zeta q+\tilde{q})} \cdot \sum_{n\in\mathbb{N}_0} \frac{(q + \tilde{q}/\zeta)^n}{n!} \big(\partial_{q'} + \zeta\tilde{x}\big)^n f(q')\big|_{q'=-\tilde{q}/\zeta} \tag{84}$$

$$= \mathrm{e}^{-\tilde{x}(\zeta q+\tilde{q})} \cdot \sum_{n\in\mathbb{N}_0} \frac{(q + \tilde{q}/\zeta)^n}{n!} \partial_{q'}^n \big(\mathrm{e}^{\tilde{x}(\zeta q'+\tilde{q})} f(q')\big)\big|_{q'=-\tilde{q}/\zeta} \tag{85}$$

$$= \mathrm{e}^{-\tilde{x}(\zeta q+\tilde{q})} \cdot \mathrm{e}^{\tilde{x}(\zeta q+\tilde{q})} f(q) = f(q)\,. \tag{86}$$

The only piece still missing is the basis (E)volution operator $\mathcal{E}_\tau$ to encode the time-dependence of the basis function (75). Differentiation of this function with respect to time and use of the creation and annihilation operators (78) and (79) show that

$$\partial_\tau|n\rangle = n|n-1\rangle(\partial_\tau\tilde{q}) + |n\rangle\big(-(\partial_\tau\tilde{x})(\zeta q + \tilde{q}) - \tilde{x}\partial_\tau\tilde{q}\big) \tag{87}$$

$$= (\partial_\tau\tilde{q})a|n\rangle - (\partial_\tau\tilde{x})|n+1\rangle - (\partial_\tau\tilde{q})\tilde{x}|n\rangle \tag{88}$$

$$= \big((\partial_\tau\tilde{q})(a - \tilde{x}) - (\partial_\tau\tilde{x})c\big)|n\rangle\,. \tag{89}$$

Hence, the operator is given by[6]

$$\mathcal{E}_\tau(c, a) \coloneqq (\partial_\tau\tilde{q})(a - \tilde{x}) - (\partial_\tau\tilde{x})c\,. \tag{90}$$

Let us illustrate the use of the above basis for a process whose transition operator can be reduced to a mere constant by an appropriate choice of $\tilde{q}$ and $\tilde{x}$. This simplification is, however, bought by making the basis functions time-dependent. In particular, we consider the simple growth process $\emptyset \to A$ for a time-independent growth rate coefficient γ. By our previous discussions, one can readily verify that the ordinary probability generating function with basis function $|n\rangle_q = q^n$ obeys the flow equation $\partial_\tau|g\rangle = \gamma(q-1)|g\rangle$ for this process. This flow equation can be simplified by redefining the basis function of the generating function as $|n\rangle_{\tau,q} \coloneqq (q+1)^n \mathrm{e}^{-\tilde{x}(q+1)}$, with $\tilde{x}$ solving the rate equation $\partial_\tau\tilde{x} = \gamma$ of the process. Hence, the basis function depends explicitly on time through its dependence on $\tilde{x}(\tau) = \tilde{x}(t_0) + \gamma(\tau - t_0)$. Upon combining the transition

6 A forward-time discretization of the operator reads

$$\mathcal{E}_{\tau,\Delta t} \coloneqq \frac{\tilde{q}(\tau+\Delta t) - \tilde{q}(\tau)}{\Delta t}(a_\tau - \tilde{x}(\tau)) - \frac{\tilde{x}(\tau+\Delta t) - \tilde{x}(\tau)}{\Delta t} c_\tau\,.$$

operator (69) with the basis evolution operator (90), one finds that the generating function now obeys the flow equation $\partial_\tau|g\rangle = -\gamma|g\rangle$ (it follows from $\mathcal{Q}_\tau = \gamma(c_\tau - 1)$ and $\mathcal{E}_\tau = -(\partial_\tau\tilde{x})c_\tau = -\gamma c_\tau$ that $\widetilde{\mathcal{Q}} = \mathcal{Q} + \mathcal{E} = -\gamma$). The equation is readily solved by $|g(\tau|t_0, n_0)\rangle = \mathrm{e}^{-\gamma(\tau-t_0)}|n_0\rangle_{t_0}$. The conditional probability distribution can be recovered via the inverse transformation (51) as

$$p(\tau, n|t_0, n_0) = \mathrm{e}^{-\gamma(\tau-t_0)}\langle n|_\tau|n_0\rangle_{t_0}\,. \tag{91}$$

The coefficient $\langle n|_\tau|n_0\rangle_{t_0}$ can be evaluated by determining how the functional $\langle n|_\tau$ acts at time t_0. Using $a_\tau - \tilde{x}(\tau) = a_{t_0} - \tilde{x}(t_0)$ and the binomial theorem, one can rewrite the action of this functional as

$$\langle n|_\tau f = \frac{a_\tau^n}{n!} f(q)\big|_{q=-1} = \frac{1}{n!}\big(a_{t_0} + \tilde{x}(\tau) - \tilde{x}(t_0)\big)^n f(q)\big|_{q=-1} \tag{92}$$

$$= \frac{1}{n!}\sum_{m=0}^{n}\binom{n}{m} a_{t_0}^m \big(\tilde{x}(\tau) - \tilde{x}(t_0)\big)^{n-m} f(q)\big|_{q=-1} \tag{93}$$

$$= \sum_{m=0}^{n}\frac{\big(\gamma(\tau - t_0)\big)^{n-m}}{(n-m)!}\frac{a_{t_0}^m}{m!} f(q)\big|_{q=-1} = \sum_{m=0}^{n}\frac{\big(\gamma(\tau - t_0)\big)^{n-m}}{(n-m)!}\langle m|_{t_0} f\,. \tag{94}$$

By the (O)rthogonality condition, the conditional probability distribution (91) therefore evaluates to the shifted Poisson distribution

$$p(\tau, n|t_0, n_0) = \frac{\mathrm{e}^{-\gamma(\tau-t_0)}\big(\gamma(\tau - t_0)\big)^{n-n_0}}{(n-n_0)!}\Theta_{n-n_0}\,, \tag{95}$$

where $\Theta_n := 1$ for $n \geq 0$ and $\Theta_n := 0$ otherwise. The validity of this solution can be verified by solving the corresponding flow equation of the ordinary generating function.

In our above approach, we first established the form of the transition operator because the basis function (75) is not the only possible choice. Ohkubo recently proposed the use of orthogonal polynomials for this purpose [336]. In analogy to the eigenfunctions of the quantum harmonic oscillator [27, 334], one can, for example, choose the basis function as the Hermite polynomial

$$|n\rangle_q := He_n(q) = (-1)^n \mathrm{e}^{q^2/2}\partial_q^n\,\mathrm{e}^{-q^2/2}\,, \tag{96}$$

with $n \in \mathbb{N}_0$. Hermite polynomials constitute an Appell sequence, i.e. they fulfil $\partial_q He_n(q) = nHe_{n-1}(q)$ (18.9.27 in [337]). With $a := \partial_q$, this property coincides with the defining relation $a|n\rangle = n|n-1\rangle$ of the annihilation operator in (71). Furthermore, Hermite polynomials obey the recurrence relation $He_{n+1}(q) = qHe_n(q) - nHe_{n-1}(q)$ (18.9.1 in [337]). Combined with the Appell property,

$c := q - \partial_q$ therefore fulfils the defining relation $c|n\rangle = |n+1\rangle$ of a creation operator in (70). After complementing the basis function (96) with the functional

$$\langle m|f := \frac{1}{m!}\int_{-\infty}^{\infty} \mathrm{d}q\, \frac{\mathrm{e}^{-q^2/2}}{\sqrt{2\pi}} He_m(q) f(q), \tag{97}$$

the (O)rthogonality and (C)ompleteness of the basis are also established (18.3 and 18.18.6 in [337]). Thus, the chemical master equation (27) can be transformed into a flow equation obeyed by a generating function based on Hermite polynomials. To our knowledge, however, no stochastic process has thus far been solved or been approximated along these lines. Besides Hermite polynomials, Ohkubo also proposed the use of Charlier polynomials and mentioned their relation with certain birth-death processes [336].

2.2 c Intermezzo: The unit vector basis

One may wonder why we actually bother with explicit representations of the bras $\langle n|$ and kets $|n\rangle$. Often, these objects are introduced only formally as the basis of a "bosonic Fock space" [11, 12, 96, 291, 338], leaving the impression that the particles under consideration are in fact bosonic quantum particles. Although this impression takes the analogy with quantum theory too far, the analogy has helped in developing new methods for solving the master equations. In the previous sections, we showed how the forward master equation can be cast into a linear PDE obeyed by a probability generating function. Later, in section 5, this PDE is solved in terms of a path integral by which we then recover a path integral representation of averaged observables in section 6. This "analytic" derivation of the path integral is, however, not the only possible way. In the following, we sketch the mathematical basis of an alternative approach, which employs the unit vectors $\hat{\boldsymbol{e}}_n$ as the basis. The approach ultimately results in the same path integral representation of averaged observables as we show in section 6.2. The duality between the two approaches resembles the duality between the matrix mechanics formulation of quantum mechanics by Heisenberg, Born, and Jordan [98–100] and its analytic formulation in terms of the Schrödinger equation [97].

The alternative derivation of the path integral also starts out from the forward master equation $\partial_\tau p(\tau|t_0) = Qp(\tau|t_0)$. As before, $p(\tau|t_0)$ represents the matrix of the conditional probabilities $p(\tau, n|t_0, n_0)$, and Q is the transition rate matrix. For simplicity, we assume the transition rate matrix to be independent of time. Moreover, we focus on processes that can be decomposed additively into chemical reactions of the form $k\,A \to l\,A$ with time-independent rate coefficients. The elements of the transition matrix associated to this reaction read $Q(m,n) = \gamma(n)_k(\delta_{m,n-k+l} - \delta_{m,n})$. Since the particle numbers n and n_0 may assume any

values in $\mathbb{N}_0$, the probability matrix $p(\tau|t_0)$ and the transition matrix Q have infinitely many rows and columns.

In the introductory section 1.4, we formulated conditions under which the forward master equation is solved by the matrix exponential $p(\tau|t_0) = \mathrm{e}^{Q(\tau-t_0)}\mathbb{1}$ (in the sense of the expansion (19), the state space must countable and the exit rates bounded). For the above chemical reaction, those conditions are not necessarily met, and thus the matrix exponential may not exist. Nevertheless, we regard $p(\tau|t_0) = \mathrm{e}^{Q(\tau-t_0)}\mathbb{1}$ as a "formal" solution in the following. With the help of certain mathematical "tricks", this solution will be cast into a path integral in section 6. But before, let us rewrite the transition matrix in the exponential in terms of creation and annihilation *matrices*. For this purpose, we employ the (infinitely large) unit column vectors $|n\rangle := \hat{\boldsymbol{e}}_n$ and their orthogonal unit row vectors $\langle n| := \hat{\boldsymbol{e}}_n^{\mathsf{T}}$ as basis (with $n \in \mathbb{N}_0$). Individual probabilities can therefore be inferred from the above solution as

$$p(\tau, n|t_0, n_0) = \langle n|\mathrm{e}^{Q(\tau-t_0)}|n_0\rangle\,. \tag{98}$$

For multivariate or spatial processes, the basis vectors can be generalized to tensors (i.e. $|\boldsymbol{n}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \ldots$).

The orthogonality of the basis allows us to rewrite the elements of the transition rate matrix of the reaction $k\,A \to l\,A$ as

$$Q(m,n) = \langle m|\Big(\gamma\,(n)_k\big(|n-k+l\rangle - |n\rangle\big)\Big)\,. \tag{99}$$

A matrix with these elements can be written as

$$Q = \gamma(c^l - c^k)a^k\,, \tag{100}$$

with c acting as a "creation matrix" and a acting as an "annihilation matrix". In particular, we define c in terms of its sub-diagonal $(1, 1, 1, \ldots)$ and a in terms of its super-diagonal $(1, 2, 3, \ldots)$. All of their other matrix elements are set to zero. It is readily shown that these matrices fulfil $c|n\rangle = |n+1\rangle$ and $a|n\rangle = n|n-1\rangle$ with respect to the basis vectors. These relations coincide with our previous relations (70) and (71). Besides, the matrices also fulfil $\langle n|c = \langle n-1|$ and $\langle n|a = (n+1)\langle n+1|$, as well as the commutation relation $[a, c] = \mathbb{1}$. Further properties of the matrices are addressed in section 6.2. By our previous comments in section 2.1, it is not a surprise that the transition matrix (100) has the same form as the transition operator (69). Both of them are normal-ordered polynomials in c and a. This property will help us in section 6.2 in recovering a path integral representation of averaged observables.

2.2 d Processes with locally excluding particles

As noted in the introductory section 1.3, one can model the movement of molecular motors along a cytoskeletal filament in terms of a master equation [118–120]. In its simplest form, such a model describes the movement of mutually excluding motors as a hopping process on a one-dimensional lattice $\mathbb{L} \subset \mathbb{Z}$, with attachment and detachment of motors at certain boundaries. To respect the mutual exclusion of motors, their local number n_i on a lattice site $i \in \mathbb{L}$ is restricted to 0 and 1. Various other processes can be modelled in similar ways, for example, aggregation processes [31], adsorption processes [339, 340], and directed percolation [34]. In the following, we show how the generating function approach from section 2.1 and its complementary approach from section 2.2 c can be applied to such processes. To illustrate the mathematics behind these approaches while not burdening ourselves with too many indices, we demonstrate the approaches for the simple, non-spatial telegraph process [341, 342]. This process describes a system that randomly switches between two states that are called the "on" and "off" states, or, for brevity, the "1" and "0" states. The rate at which the system switches from state 1 to state 0 is denoted as μ, and the rate of the reverse transition as γ. For simplicity, we assume these rate coefficients to be time-independent. Consequently, the master equation of the telegraph process reads

$$\partial_\tau \boldsymbol{p} = \begin{pmatrix} -\mu & \gamma \\ \mu & -\gamma \end{pmatrix} \boldsymbol{p} \,, \tag{101}$$

with $\boldsymbol{p}(\tau|\cdot) = (p(\tau, 1|\cdot), p(\tau, 0|\cdot))^\mathsf{T}$ being the probability vector. Alternatively, the master equation can be written as $\partial_\tau p(\tau, n|\cdot) = \sum_{m\in\{1,0\}} Q(n, m) p(\tau, m|\cdot)$ with transition matrix elements

$$Q(n, m) = \mu(\delta_{n,0} - \delta_{n,1})\delta_{m,1} + \gamma(\delta_{n,1} - \delta_{n,0})\delta_{m,0} \,. \tag{102}$$

One can solve the above master equation in various ways, for example, by evaluating the matrix exponential in $p(\tau|\tau_0) = \mathrm{e}^{Q(\tau-\tau_0)}\mathbb{1}$ after diagonalizing the transition matrix. The resulting solution reads

$$\boldsymbol{p} = \frac{1}{\gamma+\mu} \begin{pmatrix} \gamma + \mathrm{e}^{-(\gamma+\mu)(\tau-t_0)}(\mu\delta_{1,n_0} - \gamma\delta_{0,n_0}) \\ \mu - \mathrm{e}^{-(\gamma+\mu)(\tau-t_0)}(\mu\delta_{1,n_0} - \gamma\delta_{0,n_0}) \end{pmatrix} . \tag{103}$$

In the following, the telegraph process serves as the simplest representative of processes with excluding particles and it helps in explaining how the methods from the previous sections are applied to such processes. The inclusion of spatial degrees of freedom into the procedure is straightforward.

To apply the generating function technique from section 2.1, we require orthogonal and complete basis functions $\{|0\rangle, |1\rangle\}$ and basis functionals $\{\langle 0|, \langle 1|\}$, as well as a transition operator $\mathcal{Q}(q, \partial_q)$ fulfilling condition (Q), i.e., for $n \in \{0, 1\}$,

$$\mathcal{Q}|n\rangle = \mu(|0\rangle - |1\rangle)\delta_{n,1} + \gamma(|1\rangle - |0\rangle)\delta_{n,0}\,. \tag{104}$$

An operator with this property can be constructed in (at least) two ways: using creation and annihilation operators fulfilling the "bosonic" commutation relation $[a, c] = ac - ca = 1$, or using operators fulfilling the "fermionic" anti-commutation relation $\{a, c\} := ac + ca = 1$.

Let us first outline an approach based on the commutation relation. The approach was recently proposed by van Wijland [95]. To illustrate the approach, we take the "analytic" point of view with $c(q, \partial_q)$ and $a(q, \partial_q)$ being differential operators. However, the following considerations also apply to the approach from the previous section where c and a represented creation and annihilation *matrices*. To cast a master equation such as (101) into a path integral, van Wijland employed creation and annihilation operators with the same actions as in section 2.2 b, i.e. $c|n\rangle = |n + 1\rangle$ and $a|n\rangle = n|n - 1\rangle$. Besides complying with the commutation relation $[a, c] = 1$, these relations imply that the "number operator" $\mathcal{N} := ca$ fulfils $\mathcal{N}|n\rangle = n|n\rangle$. This operator may be used to define the "Kronecker operator"

$$\delta_{\mathcal{N},m} := \int_{-\pi}^{\pi} \frac{\mathrm{d}u}{2\pi} \mathrm{e}^{\mathrm{i}u(\mathcal{N}-m)} \tag{105}$$

in terms of a series expansion of its exponential. The Kronecker operator acts on the basis functions as $\delta_{\mathcal{N},m}|n\rangle = \delta_{n,m}|n\rangle$, and thus it can be used to replace the Kronecker deltas in (104). One thereby arrives at the flow equation

$$\partial_\tau|g\rangle = \mathcal{Q}|g\rangle = \left[\mu(a - 1)\delta_{\mathcal{N},1} + \gamma(c - 1)\delta_{\mathcal{N},0}\right]|g\rangle\,. \tag{106}$$

On the downside, the flow equation now involves power series with arbitrarily high derivatives with respect to q (upon choosing the operators as $c = q$ and $a = \partial_q$). In section 5, we show how the solution of a flow equation such as (106) can be expressed in terms of a path integral. The derivation of the path integral requires that the transition operator $\mathcal{Q}(q, \partial_q)$ is normal-ordered with respect to q and ∂_q, i.e. that all the q are to the left of all the ∂_q in every summand. This order can be achieved by the repeated use of $[\partial_q, q]f(q) = f(q)$. More information on the procedure can be found in [95]. Van Wijland applied the resulting path integral to the asymmetric diffusion of excluding particles on a one-dimensional lattice. Moreover, Mobilia, Georgiev, and Täuber have employed the method in studying the stochastic Lotka-Volterra model on d-dimensional lattices [158]. The diffusion

of excluding particles on a network $\mathbb{L}$ is described by the master equation

$$\partial_\tau p(\tau, \boldsymbol{n}|\cdot) = \sum_{m\in\{0,1\}^{|\mathbb{L}|}} Q(\boldsymbol{n}, \boldsymbol{m}) p(\tau, \boldsymbol{m}|\cdot) \tag{107}$$

with the transition matrix

$$Q(\boldsymbol{n}, \boldsymbol{m}) = \sum_{\langle ij\rangle} \varepsilon_{i,j}\big(\delta_{n_i,1}\delta_{n_j,0}\delta_{\boldsymbol{n}-\hat{\boldsymbol{e}}_i+\hat{\boldsymbol{e}}_j,\boldsymbol{m}} - \delta_{n_i,0}\delta_{n_j,1}\delta_{\boldsymbol{n},\boldsymbol{m}}\big)\,. \tag{108}$$

Here, the sum proceeds over all pairs $\langle ij\rangle$ of connected nodes $i \in \mathbb{L}$ and $j \in \mathbb{L}$. The corresponding condition (Q) on the transition operator is met by

$$\mathcal{Q} = \sum_{\langle ij\rangle} \varepsilon_{i,j}(c_i a_j - 1)\delta_{\mathcal{N}_i,0}\delta_{\mathcal{N}_j,1}\,. \tag{109}$$

For the special case of a one-dimensional lattice $\mathbb{L}$ and rates $\varepsilon_{i+1,i} = D + \frac{v}{2}$ and $\varepsilon_{i,i+1} = D - \frac{v}{2}$, one obtains the diffusion operator studied by van Wijland [95].

An alternative to the above approach lies in the use of operators fulfilling the anti-commutation relation $\{a, c\} = ac + ca = 1$. This relation is clearly *not* fulfilled by $c := q$ and $a := \partial_q$, at least not if q represents an ordinary real variable. However, $\partial_q q + q\partial_q = 1$ holds true if q represents a Grassmann variable (see [343, 344] for details). Grassmann variables commute with real and complex numbers (i.e. $[q, \alpha] = 0$ for $\alpha \in \mathbb{C}$), but they anti-commute with themselves and with other Grassmann variables (i.e. $\{q, \tilde{q}\} = 0$). Grassmann variables have, for example, proven useful in calculating the correlation functions of kinetic Ising models [345], even when the system is driven far from equilibrium [346].

Instead of using Grassmann variables for the basis, let us consider a representation of the basis defined by the unit row vectors $\langle 0| = (0, 1)$ and $\langle 1| = (1, 0)$, and by the unit column vectors $|0\rangle = (0, 1)^\intercal$ and $|1\rangle = (1, 0)^\intercal$. Obviously, these vectors fulfil the orthogonality condition $\langle m|n\rangle = \delta_{m,n}$, and $\sum_n |n\rangle\langle n| = \mathbb{1}$ is the 2-by-2 unit matrix. The anti-commutation relation $\{a, c\} = 1$ is met by the creation and annihilation *matrices*

$$c := \sigma^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } a := \sigma^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{110}$$

The creation matrix acts as $c|0\rangle = |1\rangle$ and $c|1\rangle = 0$, and the annihilation matrix as $a|1\rangle = |0\rangle$ and $a|0\rangle = 0$ (zero vector). Consequently, the matrix product ca fulfils $ca|0\rangle = 0$ and $ca|1\rangle = |1\rangle$, and thus it serves the same purpose as the Kronecker delta $\delta_{n,1}$ in (104). Analogously, the operator ac serves the same purpose as $\delta_{n,0}$. The master equation (101) can therefore be written in a form resembling the flow

equation (106), namely as

$$\partial_\tau \boldsymbol{p} = \left[\mu(a - \mathbb{1})c\,a + \gamma(c - \mathbb{1})ac\right]\boldsymbol{p}\,. \tag{111}$$

In this form, the stochastic process mimics a spin-1/2 problem (especially, if the creation and annihilation matrices are written in terms of the Pauli spin matrices σ_x, σ_y, and σ_z). More information on spin-representations of master equations is provided in [71, 339, 340, 347–350]. A spin-representation has, for example, been employed in the analysis of reaction-diffusion master equations via the density matrix renormalization group [351, 352]. In their study of directed percolation of excluding particles on the one-dimensional lattice $\mathbb{Z}$, Brunel, Oerding, and van Wijland performed a Jordan-Wigner transformation of the spatially extended "spin" matrices $c_i \coloneqq \sigma_i^+$ and $a_i \coloneqq \sigma_i^-$ ($i \in \mathbb{Z}$) [34]. The transformed operators fulfill anti-commutation relations not only locally but also non-locally, and thus represent the stochastic process in terms of a "fermionic" (field) theory (the "operators" are actually tensors). While the Jordan-Wigner transformation provides an exact reformulation of the stochastic process, its applicability is largely limited to systems with one spatial dimension. Moreover, the transformation requires that the stochastic process is first rewritten in terms of a spin-1/2 chain (typically possible only for processes with a single species). Based on the (coherent) eigenstates of the resulting fermionic creation and annihilation operators, Brunel et al. then derived a path integral representation of averaged observables. The paths of these integrals proceed along the values of Grassmann variables. According to van Wijland, the path integral "proved, from a technical view, rather difficult to analyze" [95]. Further information on the Jordan-Wigner transformation, Grassmann path integrals, and alternative approaches can be found in [31–33, 35, 36, 353–355].

2.3 Methods for the analysis of the generating function's flow equation

In the following, we outline various approaches that have recently been proposed for the analysis of the generating function's flow equation.

2.3 a A spectral method for the computation of stationary distributions

In their study of a linear transcriptional regulatory cascade of genes and proteins, Walczak, Mugler, and Wiggins developed a spectral method for the computation of stationary probability distributions [39]. They described the regulatory cascade on a coarse-grained level in terms of the copy numbers of certain chemical "species" at its individual steps. By imposing a Markov approximation, the dynamics of the cascade was reduced to a succession of two-species master equations. The solution

of each master equation served as input for the next equation downstream. Every of the reduced master equations allowed for the following processes: First, each of the master equation's two species $i \in \{1,2\}$ is produced in a one-step process whose rate $\gamma_i(n_1)$ depends only on the copy number n_1 of the species coming earlier along the cascade. Second, each of the two species degrades at a constant per-capita rate μ_i. With $\boldsymbol{n} \coloneqq (n_1, n_2)^\intercal \in \mathbb{N}_0^2$, the corresponding master equation reads

$$\begin{aligned}\partial_\tau p(\tau, \boldsymbol{n}|\cdot) = {} & \gamma_1(n_1 - 1)p(\tau, \boldsymbol{n} - \hat{\boldsymbol{e}}_1|\cdot) - \gamma_1(n_1)p(\tau, \boldsymbol{n}|\cdot) \\ & + \gamma_2(n_1)\big[p(\tau, \boldsymbol{n} - \hat{\boldsymbol{e}}_2|\cdot) - p(\tau, \boldsymbol{n}|\cdot)\big] \\ & + \mu_1\big[(n_1 + 1)p(\tau, \boldsymbol{n} + \hat{\boldsymbol{e}}_1|\cdot) - n_1 p(\tau, \boldsymbol{n}|\cdot)\big] \\ & + \mu_2\big[(n_2 + 1)p(\tau, \boldsymbol{n} + \hat{\boldsymbol{e}}_2|\cdot) - n_2 p(\tau, \boldsymbol{n}|\cdot)\big]\,.\end{aligned} \tag{112}$$

Here, the unit vector $\hat{\boldsymbol{e}}_i$ points in the direction of the i-th species. The master equation can be cast into a flow equation for the generating function $|g(\tau|\cdot)\rangle = \sum_{\boldsymbol{n}} p(\tau, \boldsymbol{n}|\cdot)|\boldsymbol{n}\rangle$ by following the steps in section 2.1. For this purpose, we choose the basis function as a multivariate extension of the basis function from the introduction to section 2, i.e. as $|\boldsymbol{n}\rangle_{\boldsymbol{q}} \coloneqq |n_1\rangle_{q_1}|n_2\rangle_{q_2}$ with

$$|n_i\rangle_{q_i} \coloneqq (q_i + 1)^{n_i} \mathrm{e}^{-\frac{\bar{\gamma}_i}{\mu_i}(q_i+1)}\,. \tag{113}$$

The values of the auxiliary parameters $\bar{\gamma}_1$ and $\bar{\gamma}_2$ are only specified in a numerical implementation of the method and affect its stability. Information on how their values are chosen is provided in [39]. For our present purposes, the values of the parameters remain unspecified. By differentiating the generating function with respect to time, one finds that the generating function obeys the flow equation (with the creation and annihilation operators $c_i \coloneqq q_i + 1$ and $a_i \coloneqq \partial_{q_i} + \frac{\bar{\gamma}_i}{\mu_i}$)

$$\begin{aligned}\partial_\tau|g\rangle = {} & \sum_{\boldsymbol{n}\in\mathbb{N}_0^2} \Big(\gamma_1(n_1)\big(|\boldsymbol{n} + \hat{\boldsymbol{e}}_1\rangle - |\boldsymbol{n}\rangle\big) + \gamma_2(n_1)\big(|\boldsymbol{n} + \hat{\boldsymbol{e}}_2\rangle - |\boldsymbol{n}\rangle\big) \\ & \qquad + \mu_1 n_1\big(|\boldsymbol{n} - \hat{\boldsymbol{e}}_1\rangle - |\boldsymbol{n}\rangle\big) + \mu_2 n_2\big(|\boldsymbol{n} - \hat{\boldsymbol{e}}_2\rangle - |\boldsymbol{n}\rangle\big)\Big)p(\tau, \boldsymbol{n}|\cdot) \qquad (114) \\ = {} & \Big((c_1 - 1)\hat{\gamma}_1 + (c_2 - 1)\hat{\gamma}_2 + \mu_1(1 - c_1)a_1 + \mu_2(1 - c_2)a_2\Big)|g\rangle \qquad (115) \\ = {} & \Big(q_1\hat{\gamma}_1 + q_2\hat{\gamma}_2 - \mu_1 q_1\Big(\partial_{q_1} + \frac{\bar{\gamma}_1}{\mu_1}\Big) - \mu_2 q_2\Big(\partial_{q_2} + \frac{\bar{\gamma}_2}{\mu_2}\Big)\Big)|g\rangle \qquad (116) \\ = {} & (\mathcal{Q}_0 + \mathcal{Q}_1)|g\rangle\,. \qquad (117)\end{aligned}$$

In the last line, we introduced the transition operators

$$\mathcal{Q}_0 := -\sum_{i\in\{1,2\}} \mu_i q_i \partial_{q_i} \text{ and} \tag{118}$$

$$\mathcal{Q}_1 := -\sum_{i\in\{1,2\}} q_i(\bar{\gamma}_i - \hat{\gamma}_i)\,. \tag{119}$$

Here, the two new operators $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are defined in terms of their actions $\hat{\gamma}_i|n_1\rangle = \gamma_i(n_1)|n_1\rangle$ on the basis functions. Surprisingly, the explicit form of these operators is not needed. Thus, the spectral method even allows for non-polynomial growth rates $\gamma_i(n_1)$.

The operator $\mathcal{Q}_0$ has the same form as the transition operator of the bi-directional reaction $\emptyset \rightleftharpoons A$ from the introductory example. Without the perturbation $\mathcal{Q}_1$, the above flow equation could thus be solved by extending the previous ansatz (41) to two species. To accommodate $\mathcal{Q}_1$ as well, it proves useful to generalize that ansatz to

$$|g(\tau|\cdot)\rangle = \sum_{\boldsymbol{k}\in\mathbb{N}_0^2} |\boldsymbol{k}\rangle\!\rangle G_{\boldsymbol{k}}(\tau|\cdot)\,, \tag{120}$$

with yet to be determined expansion coefficients $G_{\boldsymbol{k}}(\tau|\cdot)$. The auxiliary ket is defined as $|\boldsymbol{k}\rangle\!\rangle_{\boldsymbol{q}} := |k_1\rangle\!\rangle_{q_1}|k_2\rangle\!\rangle_{q_2}$ with $|k_i\rangle\!\rangle_{q_i} = q_i^{k_i}$ and is orthogonal to the bra $\langle\!\langle \boldsymbol{k}| := \langle\!\langle k_1|\langle\!\langle k_2|$ with $\langle\!\langle k_i|f := \frac{1}{k_i!}\partial_{q_i}^{k_i} f(q_i)|_{q_i=0}$. These bras can be used to extract the expansion coefficient via $G_{\boldsymbol{k}}(\tau|\cdot) = \langle\!\langle \boldsymbol{k}|g(\tau|\cdot)\rangle$. Differentiation of this coefficient with respect to time and imposing stationarity eventually results in a recurrence relation for $G_{\boldsymbol{k}}$ that can be solved iteratively. The steady-state probability distribution of the stochastic process is then recovered from the generating function via the inverse transformation (51). Using the corresponding creation and annihilation operators $\bar{b}_i := q_i$ and $b_i := \partial_{q_i}$ (i.e. $\bar{b}_i|k_i\rangle\!\rangle = |k_i+1\rangle\!\rangle$ and $b_i|k_i\rangle\!\rangle = k_i|k_i-1\rangle\!\rangle$), differentiation of the expansion coefficient $G_{\boldsymbol{k}}(\tau|\cdot) = \langle\!\langle \boldsymbol{k}|g(\tau|\cdot)\rangle$ results in

$$\begin{aligned}\partial_\tau G_{\boldsymbol{k}}(\tau|\cdot) &= \langle\!\langle \boldsymbol{k}|\partial_\tau|g(\tau|\cdot)\rangle \\ &= -\sum_{i\in\{1,2\}} \mu_i k_i G_{\boldsymbol{k}} - \sum_{\boldsymbol{k}'\in\mathbb{N}_0^2} \langle\!\langle \boldsymbol{k}| \sum_{i\in\{1,2\}} \bar{b}_i(\bar{\gamma}_i - \hat{\gamma}_i)|\boldsymbol{k}'\rangle\!\rangle G_{\boldsymbol{k}'}\,.\end{aligned} \tag{121}$$

Since both $\hat{\gamma}_1$ and $\hat{\gamma}_2$ act only on the first species, we can write

$$\begin{aligned}\partial_\tau G_{\boldsymbol{k}} = &- \sum_{i\in\{1,2\}} \mu_i k_i G_{\boldsymbol{k}} - \sum_{k_1'\in\mathbb{N}_0} \langle\!\langle k_1 - 1|\bar{\gamma}_1 - \hat{\gamma}_1|k_1'\rangle\!\rangle G_{k_1',k_2} \\ &- \sum_{k_1'\in\mathbb{N}_0} \langle\!\langle k_1|\bar{\gamma}_2 - \hat{\gamma}_2|k_1'\rangle\!\rangle G_{k_1',k_2-1}\end{aligned} \tag{122}$$

with the help of $\langle\!\langle k_i|\bar{b}_i = \langle\!\langle k_i - 1|$. Defining the coefficients

$$\Gamma^{(i)}_{k_1,k_1'} := \sum_{n_1\in\mathbb{N}_0} \langle\!\langle k_1|n_1\rangle(\bar{\gamma}_i - \gamma(n_1))\langle n_1|k_1'\rangle\!\rangle\,, \tag{123}$$

insertion of the identity operator $\sum_{n_1} |n_1\rangle\langle n_1|$ into the above expression results in

$$\begin{aligned}\partial_\tau G_{k_1,k_2} = &- (\mu_1 k_1 + \mu_2 k_2) G_{k_1,k_2} \\ &- \sum_{k_1'\in\mathbb{N}_0} \left(\Gamma^{(1)}_{k_1-1,k_1'} G_{k_1',k_2} + \Gamma^{(2)}_{k_1,k_1'} G_{k_1',k_2-1}\right).\end{aligned} \tag{124}$$

Here, the operators $\hat{\gamma}_1$ and $\hat{\gamma}_2$ were replaced by the rates $\gamma(n_1)$ and $\gamma(n_2)$. The functional forms of the operators were not needed. Thus, the method is not restricted to polynomial rates. The "bra-kets" $\langle\!\langle k_1|n_1\rangle$ and $\langle n_1|k_1'\rangle\!\rangle$ can be derived recursively as explained in [39, 330].

As the next step, the flow equation (124) can be rewritten in matrix form. For that purpose, we let L denote a matrix only with ones on its lower sub-diagonal. Therefore, the matrix acts as $(LA)_{i,j} = A_{i-1,j}$ on other matrices. Furthermore, we define the diagonal matrix $D(k_2)$ via $D(k_2)_{k_1,k_1'} = (\mu_1 k_1 + \mu_2 k_2)\delta_{k_1,k_1'}$, as well as the vector $\boldsymbol{G}(k_2) := \{G_{k_1,k_2}\}_{k_1\in\mathbb{N}_0}$. In a numerical implementation, the length of G is restricted by a cut-off in the range of k_2 [39]. Using these definitions, the flow equation (124) reads

$$\partial_\tau \boldsymbol{G}(k_2) = -(D(k_2) + L\Gamma^{(1)})\boldsymbol{G}(k_2) + \Gamma^{(2)}\boldsymbol{G}(k_2 - 1)\,. \tag{125}$$

Upon imposing the stationarity condition $\partial_\tau \boldsymbol{G} = 0$, the flow equation can be written in terms of the recurrence relation

$$\boldsymbol{G}^s(k_2) = -(D(k_2) + L\Gamma^{(1)})^{-1}\Gamma^{(2)}\boldsymbol{G}^s(k_2 - 1)\,. \tag{126}$$

The initial condition of the relation is determined by

$$G^s_{k_1}(0) = \sum_{\boldsymbol{n}} \langle\!\langle k_1, 0|\boldsymbol{n}\rangle p^s(\boldsymbol{n}|\cdot) = \mathrm{e}^{-\bar{\gamma}_2/\mu_2} \sum_{\boldsymbol{n}\in\mathbb{N}_0^2} \langle\!\langle k_1|n_1\rangle p^s(\boldsymbol{n}|\cdot) \tag{127}$$

$$= \mathrm{e}^{-\bar{\gamma}_2/\mu_2} \sum_{n_1\in\mathbb{N}_0} \langle\!\langle k_1|n_1\rangle p^s(n_1|\cdot)\,. \tag{128}$$

Here, we introduced the stationary limit of the marginal distribution

$$p(\tau, n_1|t_0, \boldsymbol{n}_0) = \sum_{n_2\in\mathbb{N}_0} p(\tau, \boldsymbol{n}|t_0, \boldsymbol{n}_0)\,. \tag{129}$$

The time evolution of this distribution can be inferred by summing the master equation (112) over n_2, resulting in

$$\begin{aligned}\partial_\tau p(\tau, n_1|\cdot) = {}& \gamma_1(n_1-1)p(\tau, n_1-1|\cdot) - \gamma_1(n_1)p(\tau, n_1|\cdot)\\ &+ \mu_1\big((n_1+1)p(\tau, n_1+1|\cdot) - n_1 p(\tau, n_1|\cdot)\,.\end{aligned}$$

The steady-state of this equation is now obtained by introducing the shift operator $\mathcal{T}^+\gamma_1(n_1) = \gamma_1(n_1+1)$, so that

$$\partial_\tau p(\tau, n_1|\cdot) = (\mathcal{T}^+ - 1)\big(\mu_1 n_1 p(\tau, n_1|\cdot) - \gamma_1(n_1-1)p(\tau, n_1-1|\cdot)\big)\,.$$

The right-hand side of this expression vanishes for

$$p^s(n_1|\cdot) = \frac{p^s(n_1-1|\cdot)}{n_1}\frac{\gamma_1(n_1-1)}{\mu_1} = \frac{p^s(0|\cdot)}{n_1!}\prod_{i=0}^{n_1-1}\frac{\gamma_1(i)}{\mu_1}\,. \tag{130}$$

One can now compute $\boldsymbol{G}^s(0)$, and by iteration also $\boldsymbol{G}^s(k_2)$ for $k_2 > 0$. The stationary distribution then follows via the inverse transformation (51) as

$$p^s(\boldsymbol{n}|\cdot) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^2} \langle\boldsymbol{n}|\boldsymbol{k}\rangle\!\rangle G^s_{\boldsymbol{k}}(|\cdot)\,. \tag{131}$$

In [39], Walczak et al. computed the steady-state distribution of the transcriptional regulatory cascade by solving the above recurrence relation for $G^s_{\boldsymbol{k}}$ numerically.

In [39], Walczak et al. computed the steady-state distribution of the transcriptional regulatory cascade by solving the recurrence relation for $G^s_{\boldsymbol{k}}$ numerically. The resulting distribution was compared to distributions acquired via an iterative method and via the stochastic simulation algorithm (SSA) of Gillespie. The spectral method was found to be about 10^8 times faster than the SSA in achieving the

same accuracy. But as we already mentioned in the introductory section 1.4, the SSA generally performs poorly in the estimation of full distributions, especially in the estimation of their tails. As the spectral method has only been applied to cascades with steady-state copy numbers below $n \approx 30$ thus far, it may be challenged by a direct integration of the two-species master equation (after introducing a reasonable cut-off in the copy numbers). The integration of ~ 1000 coupled ODEs does not pose a problem for modern integrators and the integration provides the full temporal dynamics of the process (see [273] for efficient algorithms). In its current formulation, the spectral method is limited to the evaluation of steady-state distributions and to simple one-step birth-death dynamics. It would be interesting to advance the method for the application to more complex processes (possibly with spatial degrees of freedom), and to extend its scope to the temporal evolution of distributions.

2.3 b WKB approximations and related approaches

The probability distribution describing the transcriptional regulatory cascade from the previous section approaches a non-trivial stationary shape in the asymptotic time limit $\tau \to \infty$ (cf. figure 2 in [39]). Often, however, a non-trivial shape of the probability distribution persists only transiently and is said to be quasi-stationary or metastable. The lifetime and shape of such a distribution can often be approximated in terms of a WKB approximation [89]. A WKB approximation of a jump process starts out from an exponential (eikonal) ansatz for the shape of the metastable probability distribution ("real-space" approach) or for the generating function discussed in the previous sections ("momentum-space" approach). Information on the real-space approach can be found in [68, 69, 151, 175, 176, 181, 356–365]. The recent review of Assaf and Meerson provides an in-depth discussion of the applicability of the real- and momentum-space approaches [366].

In the following, we outline the momentum-space WKB approximation for a system in which particles of type A annihilate in the binary reaction $2A \to \emptyset$ with rate coefficient μ, and are replenished in the linear reaction $A \to 2A$ with rate coefficient $\gamma \gg \mu$. According to a deterministic model of the combined processes with rate equation $\partial_\tau \bar{n} = \gamma \bar{n} - 2\mu \bar{n}^2$, the particle number $\bar{n}$ converges to an asymptotic value $\bar{n}_\infty = \gamma/(2\mu) \gg 1$. However, a numerical integration of the (truncated) master equation of the stochastic process shows that the mean particle number stays close to $\bar{n}_\infty$ only for a long but finite time (cf. figure 3). Asymptotically, all particles become trapped in the "absorbing" state $n = 0$. Consequently, the conditional probability distribution converges to $p(\infty, n|t_0, n_0) = \delta_{n,0}$. Up to a pre-exponential factor, the time after which the absorbing state is reached can be readily estimated using a momentum-space WKB approximation as shown below [40]. The value of the pre-exponential factor was determined by Turner and

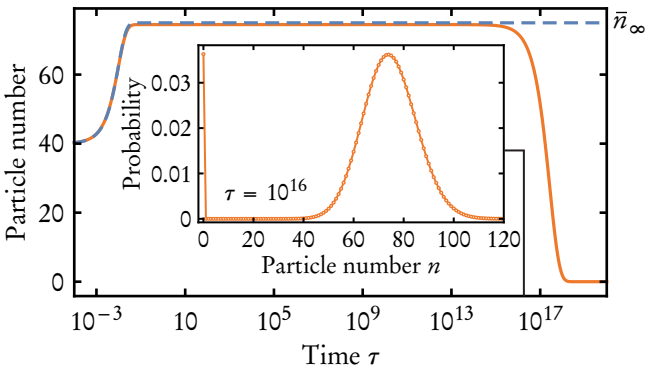


Figure 3 Comparison of the deterministic particle number $\bar{n}(\tau)$ (blue dashed line) and the mean particle number $\langle n \rangle(\tau)$ of the stochastic model (orange line) of the combined processes $2\,A \to \emptyset$ and $A \to 2\,A$. The annihilation rate coefficient was set to $\mu = 1$, the growth rate coefficient to $\gamma = 150$. The deterministic trajectory started out from $\bar{n}(0) = 40$, the numerical integration of the master equation from $p(0, n|0, 40) = \delta_{n,40}$. The deterministic trajectory converged to $\bar{n}_\infty = 75$ for very large times, whereas the mean particle number approached a quasi-stationary value close to $\bar{n}_\infty$ before converging to the absorbing state $n = 0$. The inset shows the conditional probability distribution at time $\tau = 10^{16}$. At this time, a significant share of particles was already in the absorbing state and the distribution was bimodal.

Malek-Mansour using a recurrence relation [367], by Kessler and Shnerb using a real-space WKB approximation (in terms of the decay rate) [359], and by Assaf and Meerson upon combining the generating function technique with Sturm-Liouville theory [368] (the latter method was developed in [41, 42]; see also [369]). Using this method, Assaf and Meerson also succeeded in computing the shape of the metastable distribution. Moreover, they showed how a momentum-space WKB approximation can be used to determine mean extinction times for processes with time-modulated rate coefficients [43].

Upon rescaling time as $\gamma\,\tau \to \tau$, the chemical master equation (27) of the combined processes $A \to 2\,A$ and $2\,A \to \emptyset$ translates into the flow equation

$$\partial_\tau |g\rangle = \left[(q^2 - q)\partial_q + \frac{1}{2\bar{n}_\infty}(1 - q^2)\partial_q^2\right]|g\rangle\,. \tag{132}$$

of the ordinary generating function $g(\tau; q|\cdot) = \sum_n q^n p(\tau, n|\cdot)$ (cf. sections 2.1 and 2.2 b). A WKB approximation of the flow equation can be performed by inserting the ansatz

$$g(\tau; q|\cdot) = \mathrm{e}^{S(\tau,q)} \tag{133}$$

with the "action"

$$S(\tau, q) = 2\bar{n}_\infty \sum_{k=0}^{\infty} \frac{S_k(\tau, q)}{\bar{n}_\infty^k} \tag{134}$$

into the equation, followed by a successive analysis of terms that are of the same order with respect to the power of the small parameter $1/\bar{n}_\infty$. Note that the exponential ansatz (133) often includes a minus sign in front of the action, which we neglect for convenience. The pre-factor "2" of the action (134) is also included just for convenience. Upon inserting the ansatz (133) into the flow equation (132), one obtains the equality

$$\partial_\tau S_0 + \mathrm{O}(1/\bar{n}_\infty) = H(q, \partial_q S_0) + \mathrm{O}(1/\bar{n}_\infty) \tag{135}$$

with the "Hamiltonian"

$$H(q, x) := (q^2 - q)x + (1 - q^2)x^2 \,. \tag{136}$$

Thus, at leading order of $1/\bar{n}_\infty$, we have obtained a closed equation for S_0, which has the form of a Hamilton-Jacobi equation [370]. A Hamilton-Jacobi equation can be solved by the method of characteristics [370], with the characteristic curves $q(s)$ and $x(s)$ obeying Hamilton's equations

$$\partial_s x = \frac{\partial H(q, x)}{\partial q} = (2q - 1)x - 2qx^2 \quad \text{and} \tag{137}$$

$$\partial_{-s} q = \frac{\partial H(q, x)}{\partial x} = (q^2 - q) + 2(1 - q^2)x \,. \tag{138}$$

These equations are, for example, solved by $q(s) = 1$ and $x(s)$ being a solution of $\partial_s x = x - 2x^2$. Note that this equation corresponds to a rescaled rate equation. The Hamiltonian vanishes along the characteristic curve because $H(1, x) = 0$. Further "zero-energy" lines of the Hamiltonian are given by $x = 0$ and by $x(q) = \frac{q}{1+q}$. These lines partition the phase portrait of Hamilton's equations into separate regions as shown in figure 4. The path $(q, \frac{q}{1+q})$ from the "active state" $(q, x) = (1, \frac{1}{2})$ to the "passive state" $(0, 0)$ constitutes the "optimal path to extinction" (see below) [40, 43, 371].

In addition to Hamilton's equations, the method of characteristics implies that the action S_0 obeys

$$\frac{\mathrm{d}}{\mathrm{d}s} S_0 = -\big(x\, \partial_{-s} q - H(q, x)\big) \tag{139}$$

along characteristic curves (the minus signs are only included to emphasize a similarity with an action encountered later in section 7.1). According to Elgart and Kamenev, the (negative) value of the action (134) along the optimal path

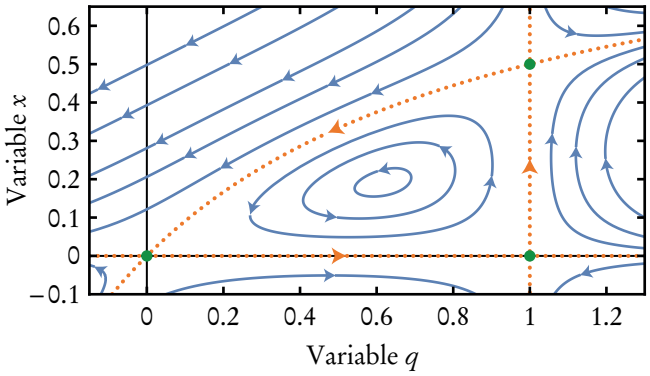


Figure 4 Phase portrait of Hamilton's equations (137) and (138). The Hamiltonian $H(q,x)$ in (136) vanishes along "zero-energy" lines (orange dotted lines). These lines connect fixed points of Hamilton's equations (green disks). Hamilton's equations reduce to the (rescaled) rate equation $\partial_s x = x - 2x^2$ along the line connecting the fixed points $(1,0)$ and $(1,1/2)$. The zero-energy line connecting the fixed point $(1,1/2)$ to $(0,0)$ denotes the "optimal path to extinction" and can be used to approximate the mean extinction time of the process.

to extinction determines the logarithm of the mean extinction time $\bar{\tau}$ from the metastable state in leading order of $\bar{n}_\infty$ [40]. As this path proceeds along a zero-energy line, the action (134) evaluates to

$$S \approx 2\bar{n}_\infty S_0 = 2\bar{n}_\infty \int_1^0 \frac{q\,\mathrm{d}q}{1+q} = -2\bar{n}_\infty(1-\ln 2)\,. \tag{140}$$

Upon returning to the original time scale via $\tau \to \gamma\,\tau$, the mean extinction time follows as

$$\bar{\tau} = A\,\gamma^{-1}\mathrm{e}^{2\bar{n}_\infty(1-\ln 2)} \tag{141}$$

with pre-exponential factor A. The value of this factor has been determined in [359, 367, 368] and reads, in terms of our rate coefficients, $A = 2\sqrt{\pi/\bar{n}_\infty}$. More information on the above procedure can be found in [40, 43]. The method has also been applied to classic epidemiological models [371] and to a variant of the Verhulst logistic model [372] (exact results on this model with added immigration have been obtained in [373] using the generating function technique).

2.3 c A variational method

The generating function's flow equation can also be analysed using a variational method as proposed by Sasai and Wolynes [44, 374]. Their approach is based on a method that Eyink had previously developed (primarily) for Fokker-Planck equations [88]. Instead of dealing with flow equation $\partial_\tau|g\rangle = \widetilde{\mathcal{Q}}_\tau|g\rangle$ of the generating

function in its differential form (or in terms of its spin matrix representation section 2.2 d), the variational method involves a functional variation of the "effective action" $\Gamma = \int_{t_0}^{t} \mathrm{d}\tau \, \langle \psi^L | (\partial_\tau - \widetilde{\mathcal{Q}}_\tau) | \psi^R \rangle$ with $|\psi^R\rangle = |g\rangle$. To perform this variation, one requires an ansatz for the objects $\langle \psi^L|$ and $|\psi^R\rangle$, being parametrized by the variables $\{\alpha_i^L\}_{i=1,\ldots,S}$ and $\{\alpha_i^R\}_{i=1,\ldots,S}$. The values of these parameters are determined by requiring that $\langle \psi^L|$ is an extremum of the action. Besides its application to networks of genetic switches [44], the variational method has been applied to signalling in enzymatic cascades in [374]. An extension of the method to multivariate processes is described in [375]. Whether or not the variational method provides useful information mainly depends on making the right ansatz for $\langle \psi^L|$ and $|\psi^R\rangle$. The method itself does not suggest their choice. It remains to be seen whether the method can be applied to processes for which only little is known about the generating function.

2.4 Résumé

In the present section, we formulated general conditions under which the forward master equation (11) can be transformed into a linear partial differential equation, a "flow equation", obeyed by the probability generating function

$$|g(\tau|t_0, n_0)\rangle = \sum_n |n\rangle \, p(\tau, n|t_0, n_0) \,. \tag{142}$$

First, the conditions (C) and (O) require that there exists a complete and orthogonal basis comprising a set of basis functions $\{|n\rangle\}$ and a set of basis functionals $\{\langle n|\}$. The basis functionals recover the conditional probability distribution via $p(\tau, n|t_0, n_0) = \langle n|g(\tau|t_0, n_0)\rangle$ and the right choice of the basis functions may help to obtain a simplified flow equation. We introduced different bases for the study of random walks, of chemical reactions, and of processes with locally excluding particles in section 2.2. Moreover, the conditions (E) and (Q) require that there exist two differential operators: a basis evolution operator $\mathcal{E}_\tau$ encoding the possible time-dependence of the basis, and a transition operator $\mathcal{Q}_\tau$ encoding the actual dynamics of the process. The generating function (142) then obeys the *flow equation*

$$\partial_\tau |g\rangle = (\mathcal{E}_\tau + \mathcal{Q}_\tau)|g\rangle = \widetilde{\mathcal{Q}}_\tau |g\rangle \,. \tag{143}$$

Various methods have recently been proposed for the study of such a flow equation and were outlined in section 2.3. These methods include the variational approach of Eyink [88] and of Sasai and Wolynes [44], WKB approximations and spectral formulations of Elgart and Kamenev and of Assaf and Meerson [40–42], and the spectral method of Walczak, Mugler, and Wiggins [39]. Thus far, most of these methods have only been applied to systems without spatial degrees of freedom

and with only one or a few types of particles. Future research is needed to overcome these limitations. Later, in section 5, we show how the solution of the flow equation (143) can be represented by a path integral. The evaluation of the path integral is demonstrated in computing the generating function of general linear processes. Moreover, we explain in section 7 how this path integral connects to a recent method of Elgart and Kamenev for the computation of rare event probabilities [40]. One can also use the path integral to derive a path integral representation of averaged observables. But a simpler route to this representation starts out from a different flow equation, a flow equation obeyed by the "marginalized distribution".

3 The marginalized distribution and the probability generating functional

In the following, we discuss two further ways of casting the forward and backward master equations into linear PDEs. The first of these flow equations is obeyed by a "marginalized distribution" and is easily derived from the backward master equation (14). The equation proves useful in the computation of mean extinction times. Moreover, it provides a most direct route to path integral representations of the conditional probability distribution and of averaged observables. To our knowledge, the flow equation of the marginalized distribution has not been considered thus far. In section 3.4, we then introduce a probability generating "functional" whose flow equation is derived from the forward master equation. The transformation mapping the functional to the conditional probability distribution is shown to generalize the Poisson representation of Gardiner and Chaturvedi [37, 38].

3.1 Flow of the marginalized distribution

The generalized probability generating function (48) was defined by summing the conditional probability distribution $p(\tau, n|t_0, n_0)$ over a set of basis functions, the kets $\{|n\rangle\}$. In the following, we consider the distribution $p(t, n|\tau, n_0)$ instead, i.e. we fix the final time t while keeping the initial time τ variable. By summing this distribution over a set of basis functions $\{|n_0\rangle_\tau\}$, one can define the series

$$|p(t, n|\tau)\rangle := \sum_{n_0} p(t, n|\tau, n_0)|n_0\rangle\,. \tag{144}$$

As before, the subscript denoting the time-dependence of the basis function is typically dropped. The variables n and n_0 again represent states from some countable state space. Assuming that the basis functions $\{|n\rangle\}$ and appropriately

chosen basis functionals $\{\langle n|\}$ form a (C)omplete and (O)rthogonal basis, the conditional probability distribution can be recovered from (144) via

$$p(t, n|\tau, n_0) = \langle n_0|p(t, n|\tau)\rangle\,. \tag{145}$$

We call the function $|p(t, n|\tau)\rangle$ a "marginalized distribution" because it proves most useful when the summation in (144) constitutes a marginalization of the conditional probability distribution $p(t, n|\tau, n_0)$ over a probability distribution $|n_0\rangle$. The marginalized distribution then represents a "single-time distribution" with respect to the random variable n in the sense of section 1.3. A basis function to which these considerations apply is the "Poisson basis function" $|n_0\rangle_x := \frac{x^{n_0}e^{-x}}{n_0!}$. We make heavy use of this basis function in the study of chemical reactions in section 3.2 b. Since the definition of the marginalized distribution in (144) does not affect the random variable n, the marginalized distribution of course solves the forward master equation $\partial_t|p(t, n|\tau)\rangle = \sum_m Q_t(n, m)|p(t, m|\tau)\rangle$ with the initial condition $|p(\tau, n|\tau)\rangle = |n_0\rangle_\tau$. In the following, we formulate conditions under which $|p\rangle$ also obeys a linear PDE evolving backward in time.

Before proceeding, let us briefly note that if the basis function of the marginalized distribution is not chosen as a probability distribution, the name marginalized "distribution" is somewhat of a misnomer; that is, for example, the case for the Fourier basis $|n_0\rangle_x := e^{in_0x}$, which we consider in section 3.2 a.

The derivation of the linear PDE obeyed by the marginalized distribution proceeds analogously to the derivation in section 2.1. But instead of employing the forward master equation, we now employ the backward master equation $\partial_{-\tau}p(t|\tau) = p(t|\tau)Q_\tau$ for this purpose (recall that $p(t|\tau)$ is the matrix with elements $p(t, n|\tau, n_0)$). Upon differentiating the marginalized distribution (144) with respect to the time parameter τ, one obtains the equation

$$\partial_{-\tau}|p\rangle = \sum_{n_0} p(\cdot|\tau, n_0)\Big(-\partial_\tau|n_0\rangle + \sum_m |m\rangle Q_\tau^\intercal(m, n_0)\Big)\,. \tag{146}$$

The rate $Q_\tau^\intercal(m, n_0) = Q_\tau(n_0, m)$ represents an element of the transposed transition matrix $Q_\tau^\intercal$.

As in section 2.1, two differential operators are required to turn the above expression into a linear PDE. First, we require a basis evolution operator $\mathcal{E}_\tau(x, \partial_x)$ fulfilling $\mathcal{E}_\tau|n_0\rangle = \partial_\tau|n_0\rangle$ for all values of n_0. The previous evolution operator (E) serves this purpose. A second differential operator $\mathcal{Q}_\tau^\dagger(x, \partial_x)$ is required to encode the information stored in the transition matrix. This operator should be a power series in ∂_x and fulfil, for all n_0,

$$\mathcal{Q}_\tau^\dagger|n_0\rangle = \sum_m |m\rangle Q_\tau^\intercal(m, n_0)\,. \tag{$\mathrm{Q}^\intercal$}$$

By the (C)ompleteness of the basis, one could also define this operator constructively as

$$\mathcal{Q}_\tau^\dagger := \sum_{m,n_0} |m\rangle Q_\tau^\intercal(m,n_0)\langle n_0| \,. \tag{148}$$

We wrote these expressions in terms of the transposed transition matrix because for the unit column vectors $|m\rangle = \hat{\boldsymbol{e}}_m$ and the unit row vectors $\langle n| = \hat{\boldsymbol{e}}_n^\intercal$, $\mathcal{Q}_\tau^\dagger$ and $Q_\tau^\intercal$ coincide (with $m, n \in \mathbb{N}_0$). As we mostly consider bases whose kets and bras represent functions and functionals, we call $\mathcal{Q}_\tau^\dagger$ the "adjoint" transition operator in the following. Often, an operator $O^\dagger(x,\partial_x)$ is said to be the adjoint of an operator $O(x,\partial_x)$ if the following relation holds with respect to two test functions f and g:

$$\int \mathrm{d}x \,[O^\dagger(x,\partial_x)f(x)]g(x) = \int \mathrm{d}x \, f(x)[O(x,\partial_x)g(x)] \,. \tag{149}$$

However, whether an operator $\mathcal{Q}_\tau$ complementing the above $\mathcal{Q}_\tau^\dagger$ actually exists will not be important in the following (except in our discussion of the Poisson representation in section 3.5).

Provided that both a basis evolution operator $\mathcal{E}_\tau$ and an adjoint transition operator $\mathcal{Q}_\tau^\dagger$ are found for a particular process, it follows from the backward-time equation (146) that the marginalized distribution $|p(t,n|\tau)\rangle$ obeys the flow equation[7]

$$\partial_{-\tau}|p\rangle = (-\mathcal{E}_\tau + \mathcal{Q}_\tau^\dagger)|p\rangle =: \widetilde{\mathcal{Q}}_\tau^\dagger|p\rangle \,. \tag{150}$$

The evolution of this equation proceeds backward in time, starting out from the final condition $|p(t,n|t)\rangle = |n\rangle$.

In section 4, we show how the flow equation (150) can be solved in terms of a "backward" path integral. Provided that the basis function $|n\rangle$ is chosen as

---

7 In the derivation of the backward path integral in section 4.1, we employ the finite difference approximation

$$\begin{aligned}
\partial_{-\tau}|p(t,n|\tau)\rangle &= \lim_{\Delta t\to 0} \frac{|p(\cdot|\tau-\Delta t)\rangle - |p(\cdot|\tau)\rangle}{\Delta t} \\
&= \lim_{\Delta t\to 0}\Big(\sum_{n_0}\big(p(\cdot|\tau,n_0) + \mathrm{O}(\Delta t)\big)\frac{|n_0\rangle_{\tau-\Delta t} - |n_0\rangle_\tau}{\Delta t} + \sum_{n_0}\frac{p(\cdot|\tau-\Delta t,n_0) - p(\cdot|\tau,n_0)}{\Delta t}|n_0\rangle_\tau\Big) \\
&= \lim_{\Delta t\to 0}(-\mathcal{E}_{\tau-\Delta t,\Delta t} + \mathcal{Q}^\dagger_{\tau-\Delta t,\Delta t})|p(\cdot|\tau)\rangle = \lim_{\Delta t\to 0}\widetilde{\mathcal{Q}}^\dagger_{\tau-\Delta t,\Delta t}|p(\cdot|\tau)\rangle \,,
\end{aligned}$$

with the operators being defined through their following actions on the basis functions:

$$\mathcal{E}_{\tau-\Delta t,\Delta t}|n_0\rangle_\tau = \frac{|n_0\rangle_\tau - |n_0\rangle_{\tau-\Delta t}}{\Delta t} \quad \text{and} \quad \mathcal{Q}^\dagger_{\tau-\Delta t,\Delta t}|n_0\rangle_\tau = \sum_m |m\rangle_\tau Q^\intercal_{\tau-\Delta t,\Delta t}(m,n_0) \,.$$

The discretization scheme conforms with the derivation of the backward master equation (14).

a probability distribution, this path integral represents a true probability distribution: the marginalized distribution. The fact that the backward path integral represents a probability distribution distinguishes it from the "forward" path integral in section 5. The forward path integral represents the probability generating function (48). Both the forward path integral and the backward path integral can be used to derive a path integral representation of averaged observables, as we show in section 6. The derivation of this representation from the backward path integral, however, is significantly easier. In fact, the path integral representation of the average of an observable A will follow directly by summing the backward path integral representation of the marginalized distribution over $A(n)$, i.e. via $\langle A\rangle = \sum_n A(n)|p(t,n|\tau)\rangle$. Note that this average also obeys the flow equation (150). In section 3.3, we demonstrate how the flow equation can be used to compute mean extinction times.

Before introducing bases for the analysis of different stochastic processes, let us briefly note that if a process is homogeneous in time, its marginalized distribution depends only on the difference $t - \tau$. The above flow equation can then be rewritten so that it evolves $|p(\tau,n|0)\rangle$ forward in time τ, starting out from the initial condition $|p(0,n|0)\rangle = |n\rangle$. In section 3.3, we make use of this property to compute mean extinction times.

3.2 Bases for particular stochastic processes

To demonstrate the application of the marginalized distribution (144), let us reconsider the random walk from section 2.2 a and the chemical reaction from section 2.2 b. The "Poisson basis function" introduced in the latter section will be employed in the computation of mean extinction times in section 3.3. Moreover, it will allow us to recover the Poisson representation of Gardiner and Chaturvedi [37, 38] in section 3.5.

3.2 a Random walks

The following solution of the random walk largely parallels the previous derivation in section 2.2 a. In particular, we again use the orthogonal and complete Fourier basis $|n\rangle_x = \mathrm{e}^{\mathrm{i}nx}$ and $\langle n|f = \int_{-\pi}^{\pi}\frac{\mathrm{d}x}{2\pi}\mathrm{e}^{-\mathrm{i}nx}f(x)$ with $n \in \mathbb{Z}$. Due to the time-independence of the basis, the (E)volution operator $\mathcal{E}_\tau$ is zero. The condition ($\mathrm{Q}^\intercal$) on the adjoint transition operator $\mathcal{Q}_\tau^\dagger$ is specified by the transition matrix (63) of the process and reads

$$\mathcal{Q}_\tau^\dagger|n_0\rangle = r_\tau(|n_0-1\rangle - |n_0\rangle) + l_\tau(|n_0+1\rangle - |n_0\rangle)\,. \tag{151}$$

As before, we employ the operators $c(x) = \mathrm{e}^{\mathrm{i}x}$ and $a(x) = \mathrm{e}^{-\mathrm{i}x}$, which act on the basis functions as $c|n\rangle = |n+1\rangle$ and $a|n\rangle = |n-1\rangle$, respectively. Therefore, the

adjoint transition operator with the above property can be defined as

$$\mathcal{Q}_\tau^\dagger(x, \partial_x) := r_\tau\big(a(x) - 1\big) + l_\tau\big(c(x) - 1\big)\,. \tag{152}$$

As this operator does not contain any derivatives, it is self-adjoint in the sense of (149). Due to a mismatch in signs, it is, however, not the adjoint of the previous operator $\mathcal{Q}_\tau$ in (65). This mismatch could be corrected by redefining the above basis function as $|n\rangle_x := \mathrm{e}^{-\mathrm{i}nx}$. Ignoring this circumstance, the flow equation of the marginalized distribution follows as

$$\partial_{-\tau}|p\rangle = \big[r_\tau(\mathrm{e}^{-\mathrm{i}x} - 1) + l_\tau(\mathrm{e}^{\mathrm{i}x} - 1)\big]|p\rangle\,, \tag{153}$$

and is solved by

$$|p\rangle = \exp\Big((\mathrm{e}^{-\mathrm{i}x} - 1)\int_\tau^t \mathrm{d}s\, r_s + (\mathrm{e}^{\mathrm{i}x} - 1)\int_\tau^t \mathrm{d}s\, l_s\Big)|n\rangle\,. \tag{154}$$

The conditional probability distribution is recovered via the inverse Fourier transformation $p(t, n|\tau, n_0) = \langle n_0|p\rangle$. Upon inserting the explicit representation of the basis, the derivation proceeds as in appendix C (with the substitutions $x \to -q$, $\tau \to t_0$ and $t \to \tau$). Eventually, one recovers a Skellam distribution as the solution of the process.

3.2 b Chemical reactions

To prepare the computation of mean extinction times in the next section as well as the derivation of the Poisson representation in section 3.5, we now reconsider processes that can be decomposed additively into chemical reactions of the form $k\,A \to l\,A$. Our later derivation of a path integral representation of averaged observables is also restricted to such processes. The state space of the number of molecules is again $\mathbb{N}_0$. Analogous to section 2.2 b, we require an (O)rthogonal and (C)omplete basis, an (E)volution operator $\mathcal{E}_\tau$ and an adjoint transition operator $\mathcal{Q}_\tau^\dagger$ (condition ($\mathrm{Q}^\intercal$)) to specify the flow equation of the marginalized distribution.

As discussed in section 2.2 b, the elements of the transition rate matrix of the reaction $k\,A \to l\,A$ with rate coefficient γ_τ are given by $Q_\tau(m, n) = \gamma_\tau(n)_k(\delta_{m,n-k+l} - \delta_{m,n})$. The condition ($\mathrm{Q}^\intercal$) on the adjoint transition operator therefore reads

$$\mathcal{Q}_\tau^\dagger|n_0\rangle = \gamma_\tau\big((n_0 - l + k)_k|n_0 - l + k\rangle - (n_0)_k|n_0\rangle\big)\,. \tag{155}$$

This condition is met by[8]

$$\mathcal{Q}_\tau^\dagger(c,a) := \gamma_\tau c^k (a^l - a^k), \tag{156}$$

provided that there exist operators c and a fulfilling

$$c|n\rangle = (n+1)|n+1\rangle \text{ and} \tag{157}$$
$$a|n\rangle = |n-1\rangle, \tag{158}$$

respectively. We again call c the creation and a the annihilation operator, even though the pre-factors in the above relations differ from the ones in (70) and (71). Similar relations also hold with respect to the basis functionals, namely $\langle n|c = n\langle n-1|$ and $\langle n|a = \langle n+1|$ (assuming the orthogonality of the basis). The operators also fulfil the commutation relation $[a,c] = 1$.

The actions of the creation and annihilation operators on the basis functions hint at how the basis functions and functionals from section 2.2 b can be adapted to meet the present requirements. In particular, the relations (157) and (158) can be fulfilled by moving the factorial from the basis functional (81) to the basis function (75) so that

$$|n\rangle_x := \frac{(\zeta x + \tilde{x})^n \mathrm{e}^{-\tilde{q}(\zeta x + \tilde{x})}}{n!} \text{ and} \tag{159}$$
$$\langle n|f := a^n f(x)\big|_{x=-\tilde{x}/\zeta}. \tag{160}$$

Let us briefly note that we changed the argument of the basis function from q to x as compared to section 2.2 b because both the generating function approach and the marginalized distribution approach thereby result in the same path integral representation of averaged observables (cf. section 6). Apart from this notational change, the basis evolution operator[9]

$$\mathcal{E}_\tau(c,a) := (\partial_\tau \tilde{x})(a - \tilde{q}) - (\partial_\tau \tilde{q})c \tag{162}$$

8 The discrete-time approximation of the operator as required in section 3.1 reads

$$\mathcal{Q}^\dagger_{\tau-\Delta t,\Delta t} = \gamma_{\tau-\Delta t} c^k_\tau (a^l_\tau - a^k_\tau)$$

Upon making an error of $\mathrm{O}(\Delta t)$, we employ the following discretization in section 7.3:

$$\mathcal{Q}^\dagger_{\tau-\Delta t,\Delta t} = \gamma_{\tau-\Delta t} c^k_{\tau-\Delta t} (a^l_{\tau-\Delta t} - a^k_{\tau-\Delta t}).$$

9 A backward-time discretization of the operator reads

$$\mathcal{E}_{\tau-\Delta t,\Delta t} = \frac{\tilde{x}(\tau) - \tilde{x}(\tau-\Delta t)}{\Delta t}(a_\tau - \tilde{q}(\tau)) - \frac{\tilde{q}(\tau) - \tilde{q}(\tau-\Delta t)}{\Delta t} c_\tau .$$

keeps the form it had in (90). The creation and annihilation operators also keep their previous forms in (78) and (79), i.e.

$$c(x, \partial_x) := \zeta x + \tilde{x} \quad \text{and} \tag{163}$$
$$a(x, \partial_x) := \partial_{\zeta x} + \tilde{q}\,. \tag{164}$$

Despite their similar appearance, the operator $\mathcal{Q}_\tau^\dagger$ in (156) is not the adjoint of the operator $\mathcal{Q}_\tau$ in (69). Nevertheless, the two operators fulfil $\mathcal{Q}_\tau(q, x) = \mathcal{Q}_\tau^\dagger(x, q)$ for scalar arguments. This relation is essentially the reason why we interchanged the letters x and q as compared to section 2.2 b. Both the generating function approach and the marginalized distribution approach thereby lead to the same path integral representation of averaged observables, as will be shown in section 6.1.

The choice of the parameters $\zeta \neq 0$, $\tilde{x}(\tau)$, and $\tilde{q}(\tau)$ in the basis function (159) depends on the problem at hand. In section 7, $\tilde{x}$ and $\tilde{q}$ will act as "stationary" or "extremal" paths, with x and an auxiliary variable q being deviations from them. For $\zeta := \tilde{q} := 1$ and $\tilde{x} := 0$, the basis function instead simplifies to the Poisson distribution

$$|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}\,. \tag{165}$$

In the following, we make heavy use of this "Poisson basis function". It will play a crucial role in the formulation of a path integral representation of averaged observables in section 6 and in recovering the Poisson representation in section 3.5.

Let us demonstrate the use of the Poisson basis function for the linear decay process $A \to \emptyset$ with rate coefficient μ_τ. For the above choice of ζ, $\tilde{x}$, and $\tilde{q}$, the creation operator (163) reads $c = x$ and the annihilation operator (164) reads $a = \partial_x + 1$. With the transition operator (156), the flow equation (150) obeyed by the marginalized distribution $|p(t, n|\tau)\rangle_x$ follows as

$$\partial_{-\tau}|p\rangle = -\mu_\tau x \partial_x |p\rangle\,. \tag{166}$$

This equation is solved by the Poisson distribution

$$|p(t, n|\tau)\rangle_x = \frac{(\alpha_{t,\tau} x)^n \mathrm{e}^{-\alpha_{t,\tau} x}}{n!} \tag{167}$$

whose mean $\alpha_{t,\tau}\, x$ decays proportionally to $\alpha_{t,\tau} := \mathrm{e}^{-\int_\tau^t \mathrm{d}s\, \mu_s}$. To interpret this solution, let us recall that the sum in the definition of the marginalized distribu-

Upon making an error of $\mathrm{O}(\Delta t)$, we employ the following discretization in section 7.3:

$$\mathcal{E}_{\tau-\Delta t,\Delta t} = \frac{\tilde{x}(\tau) - \tilde{x}(\tau - \Delta t)}{\Delta t}(a_{\tau-\Delta t} - \tilde{q}(\tau - \Delta t)) - \frac{\tilde{q}(\tau) - \tilde{q}(\tau - \Delta t)}{\Delta t} c_{\tau-\Delta t}\,. \tag{161}$$

tion (144) does not affect the particle number n. Therefore, upon relabelling the time parameters, the above solution (167) also solves the forward master equation of the process, namely

$$\partial_\tau |p(\tau, n|t_0)\rangle_x = \mu_\tau \big((n+1)|p(\tau, n+1|t_0)\rangle_x - n|p(\tau, n|t_0)\rangle_x\big)\,. \tag{168}$$

Unlike the conditional distribution $p(\tau, n|t_0, n_0)$, however, the marginalized distribution $|p(\tau, n|t_0)\rangle_x$ describes the dynamics of a population whose particle number at time t_0 is Poisson distributed with mean x. The conditional distribution is recovered from it via the inverse transformation $p(\tau, n|t_0, n_0) = \langle n_0|p(\tau, n|t_0)\rangle$ with $\langle n_0|f = (\partial_x + 1)^{n_0} f(x)|_{x=0}$. This transformation results in the Binomial distribution

$$p(\tau, n|t_0, n_0) = \binom{n_0}{n} (\alpha_{\tau,t_0})^n (1 - \alpha_{\tau,t_0})^{n_0 - n}\,. \tag{169}$$

Both the mean value $\mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\, \mu_s} n_0$ and the variance $(1 - \mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\, \mu_s})\, \mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\, \mu_s} n_0$ of this distribution decay exponentially for large times, provided that $\mu_s > 0$.

If at most two reactants and two products are involved in a reaction, the flow equation (150) of the marginalized distribution has the mathematical form of a backward Fokker-Planck equation. That is, for example, the case for the coagulation reaction $2\,A \to A$. For the Poisson basis function, the marginalized distribution $|p(t, n|\tau)\rangle_x$ of this process obeys the flow equation

$$\partial_{-\tau} |p\rangle = \alpha_\tau(x) \partial_x |p\rangle + \frac{1}{2} \beta_\tau(x) \partial_x^2 |p\rangle\,, \tag{170}$$

with the drift coefficient $\alpha_\tau(x) := -\mu_\tau x^2$ and the diffusion coefficient $\beta_\tau(x) := -2\mu_\tau x^2$. Unlike the diffusion coefficient of the "true" backward Fokker-Planck equation (2), this diffusion coefficient may be negative (e.g. for $x \in \mathbb{R}\backslash\{0\}$). The final condition $|p(t, n|t)\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$ and the Feynman-Kac formula (449) imply that the above flow equation is solved by

$$|p(t, n|\tau)\rangle_x = \left\langle\!\!\left\langle \frac{x(t)^n \mathrm{e}^{-x(t)}}{n!} \right\rangle\!\!\right\rangle_W\,. \tag{171}$$

Here, $x(s)$ obeys the Itô SDE

$$\mathrm{d}x(s) = \alpha_s(x(s))\, \mathrm{d}s + \sqrt{\beta_s(x(s))}\, \mathrm{d}W(s)\,, \tag{172}$$

which evolves $x(s)$ from $x(\tau) = x$ to $x(t)$. The symbol $\langle\!\langle \cdot \rangle\!\rangle_W$ represents an average over realizations of the Wiener process W. Since the diffusion coefficient $\beta_\tau(x) = -2\mu_\tau x^2$ of the coagulation process can take on negative values, the sample paths of the SDE may acquire imaginary components. This circumstance does

not prevent the use of (171) for the calculation of $|p(t,n|\tau)\rangle$, although it may complicate numerical evaluations. Recently, Wiese attempted the evaluation of the average in (171) via the generation of sample paths [338]. Over a short time interval $[\tau, t]$, he found a good agreement between the resulting distribution and a distribution effectively acquired via the stochastic simulation algorithm (SSA) in section 1.3. Over larger time intervals, however, the integration of the SDE encountered problems regarding its numerical convergence. Future research is needed to overcome this limitation. Moreover, it remains an open challenge to specify the boundary conditions of the PDE (170) to enable its direct numerical integration (an analogous problem is encountered for the flow equation of the generating function, cf. [368]).

3.3 Mean extinction times

One often wishes to know the mean time at which a process first hits some target in state space. Such a target could, for example, be a state in which no more particles are left in the system. If the particles only replenish through auto-catalysis, the process will then come to a halt. The mean time after which that happens is called the mean extinction time. For Markov processes with continuous sample paths, mean extinction times and, more generally, first-passage times, are commonly calculated with the help of the backward Fokker-Planck equation (2). For jump processes, one can use the backward master equation for this purpose. The calculation, however, is typically feasible only for one-step processes and involves the solution of a recurrence relation [115, 217]. In [376], Drummond et al. recently showed how the calculation can be simplified using the Poisson representation of Gardiner and Chaturvedi [37, 38]. We introduce this representation in section 3.5. In the following, we outline how mean extinction times can instead be inferred in an analogous way using the marginalized distribution from the previous sections. For the Poisson basis function $|n_0\rangle_x = \frac{x^{n_0}\mathrm{e}^{-x}}{n_0!}$, the marginalized distribution is a true (single-time) probability distribution and is more easily interpreted than the integral kernel of the Poisson representation.

We again consider a process that can be decomposed additively into chemical reactions of the form $k\,A \to l\,A$. In addition, the transition matrix $\mathcal{Q}_\tau^\dagger = \mathcal{Q}^\dagger$ of the process shall now be time-independent and the particles in the system shall be Poisson distributed with mean x at time $\tau = 0$. Consequently, the probability of finding n particles in the system at time $\tau \geq 0$ is described by the marginalized distribution $|p(\tau,n|0)\rangle_x$, provided that its basis function is chosen as $|n_0\rangle_x := \frac{x^{n_0}\mathrm{e}^{-x}}{n_0!}$ (cf. the definition (144) of the marginalized distribution). Since the Poisson basis function is independent of time and the process under consideration homogeneous in time, the marginalized distribution $|p(\tau,n|0)\rangle_x$ obeys the forward-time flow

equation (cf. (150))

$$\partial_\tau |p\rangle = \mathcal{Q}^\dagger |p\rangle \,. \tag{173}$$

We define the probability of finding the system in an "active" state with $n > 0$ particles at time $\tau \geq 0$ as

$$\alpha(\tau; x) := \sum_{n=1}^{\infty} |p(\tau, n|0)\rangle_x = 1 - |p(\tau, 0|0)\rangle_x \,. \tag{174}$$

Over time, this probability flows into the "absorbing" state $n = 0$ at rate $f(\tau, x) := -\partial_\tau \alpha(\tau; x)$. Since $f(\tau, x)\,\mathrm{d}\tau$ is the probability of becoming absorbed during the time interval $[\tau, \tau + \Delta t]$, one can define the mean extinction time as $\langle\tau\rangle_x := \int_0^\infty \mathrm{d}\tau\, \tau f(\tau, x)$. An integration by parts transforms this average into

$$\langle\tau\rangle_x = -\int_0^\infty \mathrm{d}\tau\, \tau \partial_\tau \alpha(\tau; x) = \int_0^\infty \mathrm{d}\tau\, \alpha(\tau; x) \,. \tag{175}$$

The boundary terms of the integration by parts vanished because we assume that all particles are eventually absorbed (in particular, we assume $\lim_{\tau\to\infty} \tau\alpha(\tau; x) = 0$). The definition of the probability $\alpha(\tau; x)$ in (174) implies that it fulfils the same forward-time flow equation as the marginalized distribution $|p(\tau, n|0)\rangle_x$. Since $\lim_{\tau\to\infty} \alpha(\tau; x) = 0$, the mean extinction time (175) therefore obeys

$$-\mathcal{Q}^\dagger(x, \partial_x)\langle\tau\rangle_x = \alpha(0; x) = 1 - \mathrm{e}^{-x} \,. \tag{176}$$

The last equality follows from the fact that a certain fraction of all particles, namely e^{-x}, has already been in the absorbing state initially. The above derivation readily extends to higher moments of the mean extinction time.

For the linear decay process $A \to \emptyset$ with decay rate coefficient μ, the equation (176) for the mean extinction time reads

$$\mu\partial_x\langle\tau\rangle_x = (1 - \mathrm{e}^{-x})/x \,. \tag{177}$$

This equation implies that the mean extinction time $\mu\langle\tau\rangle_x$ increases logarithmically with the particles' mean initial distance x from the absorbing state (see figure 5). The explicit solution of the equation is $\mu\langle\tau\rangle_x = \ln x + \gamma - \mathrm{Ei}(-x)$ with Euler's constant γ and the exponential integral Ei (6.2.6 in [337]; the exponential integral ensures that $\langle\tau\rangle_0 = 0$). The application of the functional $\langle n_0|f = (\partial_x + 1)^{n_0} f(x)\big|_{x=0}$ to the solution returns the mean extinction time of particles whose initial number is not Poisson distributed but that is fixed to some value $n_0 \geq 0$. The corresponding mean extinction time $\mu\langle\tau\rangle_{n_0}$ is given by the harmonic number $H_{n_0} := \sum_{i=1}^{n_0} \frac{1}{i}$.

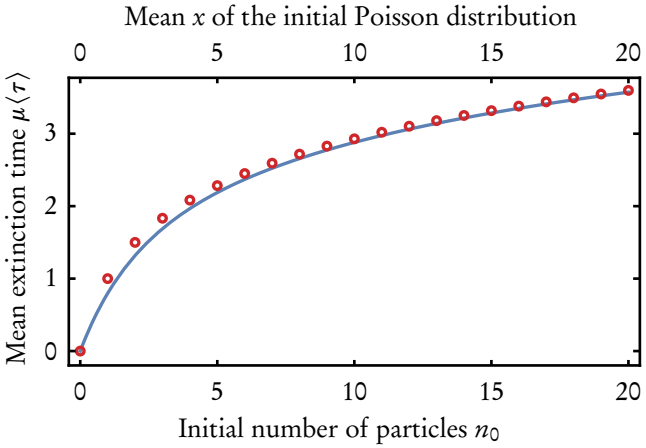


Figure 5 Mean extinction time $\mu\langle\tau\rangle$ of the linear decay process $A \to \emptyset$ with decay rate coefficient μ. The blue line represents the mean extinction time if the number of particles is initially Poisson distributed with mean $x \in \mathbb{R}_{\geq 0}$ ($\mu\langle\tau\rangle_x = \ln x + \gamma - \mathrm{Ei}(-x)$). The red circles represent the mean extinction time if the initial number of particles is set to $n_0 \in \mathbb{N}_0$ (harmonic number $\mu\langle\tau\rangle_{n_0} = H_{n_0}$).

One can also infer the above result directly from the backward master equation

$$\partial_\tau p(\tau, n|0, n_0) = \sum_{m\in\mathbb{Z}} \big(p(\tau, n|0, m) - p(\tau, n|0, n_0)\big) w(m, n_0)\,. \tag{178}$$

The probability of finding the system in an active state is now given by $\alpha(\tau, n_0) := \sum_{n>0} p(\tau, n|0, n_0)$. Choosing $n_0 > 0$, the initial probability fulfils $\alpha(0, n_0) = 1$. The rate of extinction is given by $f(\tau, x) := -\partial_\tau \alpha(\tau, n_0)$ and the mean time of extinction by $\langle\tau\rangle_{n_0} := \int_0^\infty \mathrm{d}\tau\, \tau f(\tau, n_0)$. (175) still applies in the sense of

$$\langle\tau\rangle_{n_0} = \int_0^\infty \mathrm{d}\tau\, \alpha(\tau, n_0)\,. \tag{179}$$

Moreover, the backward master equation implies

$$\partial_\tau \alpha(\tau, n_0) = \sum_{m\in\mathbb{Z}} \big(\alpha(\tau, m) - \alpha(\tau, n_0)\big) w(m, n_0) \tag{180}$$

and can be used to transform the previous equation into

$$\sum_{m\in\mathbb{Z}} \big(\langle\tau\rangle_{n_0} - \langle\tau\rangle_m\big) w(m, n_0) = 1\,. \tag{181}$$

For the linear decay process with $w(m, n_0) = \mu n_0 \delta_{m,n_0-1}$, this equality reads

$$\mu n_0 \big(\langle \tau \rangle_{n_0} - \langle \tau \rangle_{n_0-1} \big) = 1 \,. \tag{182}$$

With the initial condition $\langle \tau \rangle_0 = 0$, the solution of the equation follows as the harmonic number

$$\mu \langle \tau \rangle_{n_0} = \sum_{i=1}^{n_0} \frac{1}{i} = H_{n_0} \,. \tag{183}$$

A summation of the harmonic number over a Poisson distribution with mean x results in the previous solution.

Drummond et al. have extended the above computation to chemical reactions with at most two reactants and two products [376]. For these reactions, the flow equation (173) has the mathematical form of a (forward) Fokker-Planck equation with adjoint transition operator

$$\mathcal{Q}^\dagger(x, \partial_x) = \alpha(x)\partial_x + \frac{1}{2}\beta(x)\partial_x^2 \,. \tag{184}$$

3.4 Flow of the generating functional

Both the probability generating function (48) and the marginalized distribution (144) were defined by summing the conditional probability distribution over a set of basis functions, either at the final or at the initial time. In addition, one can define the "probability generating functional"

$$\langle g(\tau | t_0, n_0)| \coloneqq \sum_n \langle n | p(\tau, n | t_0, n_0) \tag{185}$$

by summing the conditional distribution over the set $\{\langle n|\}$ of basis functionals. The definition of the generating functional is meant with respect to test functions that can be expanded in the basis functions $\{|n\rangle\}$. As before, we assume that the basis functions and functionals constitute a (C)omplete and (O)rthogonal basis. The probability generating functional "generates" probabilities in the sense that

$$p(\tau, n | t_0, n_0) = \langle g(\tau | t_0, n_0) | n \rangle \,. \tag{186}$$

In the next section, we show how this inverse transformation reduces to the Poisson representation of Gardiner and Chaturvedi [37, 38] upon choosing the basis function $|n\rangle$ as a Poisson distribution.

The derivation of the flow equation of $\langle g|$ proceeds analogously to the derivations in sections 2.1 and 3.1. Differentiation of its definition (185) with respect to

τ and use of the forward master equation $\partial_\tau p(\tau|t_0) = Q_\tau p(\tau|t_0)$ result in

$$\partial_\tau \langle g| = \sum_n \Big(\partial_\tau \langle n| + \sum_m Q_\tau^\intercal(n,m)\langle m|\Big) p(\tau, n|\cdot)\,. \tag{187}$$

This equation can be cast into a linear PDE by using the basis (E)volution operator $\mathcal{E}_\tau$ and the adjoint transition operator $\mathcal{Q}_\tau^\dagger$ (condition ($\mathrm{Q}^\intercal$)). In particular, since the Kronecker delta $\langle n|m\rangle = \delta_{n,m}$ is independent of time, it holds that $\langle n|\partial_\tau|m\rangle = (-\partial_\tau\langle n|)|m\rangle$. Consequently, the basis evolution operator in condition (E) fulfils

$$\langle n|\mathcal{E}_\tau = -\partial_\tau \langle n| \tag{188}$$

with respect to functions that can be expanded in the basis functions. Moreover, the (O)rthogonality and (C)ompleteness of the basis imply that the adjoint transition operator in condition ($\mathrm{Q}^\intercal$) acts on basis functionals as

$$\langle n|\mathcal{Q}_\tau^\dagger = \sum_m Q_\tau^\intercal(n,m)\langle m|\,. \tag{189}$$

Both of the above expressions hold for all values of n. If the two differential operators $\mathcal{E}_\tau$ and $\mathcal{Q}_\tau^\dagger$ exist, the generating functional $\langle g(\tau|t_0, n_0)|$ obeys the functional flow equation

$$\partial_\tau \langle g| = \langle g|(-\mathcal{E}_\tau + \mathcal{Q}_\tau^\dagger) = \langle g|\widetilde{\mathcal{Q}}_\tau^\dagger\,, \tag{190}$$

with initial condition $\langle g(t_0|t_0, n_0)| = \langle n_0|$. The flow equation employs the same operator as the flow equation (150) of the marginalized distribution. Both equations can be used to derive the "backward" path integral considered in section 4.[10]

As a side note, let us remark that the flow equation $\partial_\tau|g\rangle = (\mathcal{E}_\tau + \mathcal{Q}_\tau)|g\rangle$ of the generating function in (57) also admits a functional counterpart. In particular, the series

$$\langle p(t, n|\tau)| := \sum_{n_0} p(t, n|\tau, n_0)\langle n_0| \tag{191}$$

10 The backward path integral in section 4 can be derived using the finite difference approximation

$$\begin{aligned}
\partial_\tau \langle g(\tau|t_0, n_0| &= \lim_{\Delta t\to 0} \frac{\langle g(\tau+\Delta t|\cdot)| - \langle g(\tau|\cdot)|}{\Delta t} \\
&= \lim_{\Delta t\to 0}\Big(\sum_n \frac{\langle n|_{\tau+\Delta t} - \langle n|_\tau}{\Delta t}\big[p(\tau, n|\cdot) + \mathrm{O}(\Delta t)\big] + \sum_n \langle n|_\tau \frac{p(\tau+\Delta t, n|\cdot) - p(\tau, n|\cdot)}{\Delta t}\Big) \\
&= \lim_{\Delta t\to 0} \langle g(\tau|\cdot)|(-\mathcal{E}_{\tau,\Delta t} + \mathcal{Q}_{\tau,\Delta t}^\dagger) = \lim_{\Delta t\to 0} \langle g(\tau|\cdot)|\widetilde{\mathcal{Q}}_{\tau,\Delta t}^\dagger\,,
\end{aligned}$$

with the operators being defined through their following actions on the basis functionals:

$$\langle n|_\tau \mathcal{E}_{\tau,\Delta t} = -\frac{\langle n|_{\tau+\Delta t} - \langle n|_\tau}{\Delta t} \quad \text{and} \quad \langle n|_\tau \mathcal{Q}_{\tau,\Delta t}^\dagger = \sum_m Q_{\tau,\Delta t}^\intercal(n,m)\langle m|_\tau\,.$$

obeys the flow equation

$$\partial_{-\tau}\langle p| = \sum_{n_0} p(\cdot|\tau, n_0)\Big(\partial_{-\tau}\langle n_0| + \sum_m Q_\tau(n_0, m)\langle m|\Big) \tag{192}$$

$$= \langle p|(\mathcal{E}_\tau + \mathcal{Q}_\tau) = \widetilde{\mathcal{Q}}_\tau\langle p|\,, \tag{193}$$

with final value $\langle p(t, n|t)| = \langle n|$. The corresponding inverse transformation reads $p(t, n|\tau, n_0) = \langle p(t, n|\tau)|n_0\rangle$. Both the flow equation obeyed by the generating function and the above flow equation can be used to derive the "forward" path integral representation in section 5.[11] Further uses of the series (191) remain to be explored.

3.5 The Poisson representation

Assuming that the action of the generating functional $\langle g|$ on a function f can be expressed in terms of an integral kernel (also called g) as

$$\langle g(\tau|t_0, n_0)|f = \int_{-\infty}^{\infty} \mathrm{d}x\, g(\tau; x|t_0, n_0) f(x)\,, \tag{194}$$

The discretization scheme conforms with the derivation of the forward master equation (11). The operator $\mathcal{E}_{\tau,\Delta t}$ from section 2.1 fulfils the first condition in $\mathrm{O}(\Delta t)$ because

$$\langle n|_\tau \mathcal{E}_{\tau,\Delta t}|m\rangle_\tau = \langle n|_\tau \frac{|m\rangle_{\tau+\Delta t} - |m\rangle_\tau}{\Delta t} = \frac{1}{\Delta t}\Big[\Big(\langle n|_{\tau+\Delta t} - \big(\langle n|_{\tau+\Delta t} - \langle n|_\tau\big)\Big)|m\rangle_{\tau+\Delta t} - \delta_{m,n}\Big]$$
$$= -\frac{\langle n|_{\tau+\Delta t} - \langle n|_\tau}{\Delta t}|m\rangle_{\tau+\Delta t} = -\frac{\langle n|_{\tau+\Delta t} - \langle n|_\tau}{\Delta t}|m\rangle_\tau + \mathrm{O}(\Delta t)\,.$$

11 The forward path integral in section 5 can be derived using the finite difference approximation:

$$\partial_{-\tau}\langle p(t, n|\tau)| = \lim_{\Delta t\to 0}\frac{\langle p(\cdot|\tau - \Delta t)| - \langle p(\cdot|\tau)|}{\Delta t}$$
$$= \lim_{\Delta t\to 0}\Big(\sum_{n_0}[p(\cdot|\tau, n_0) + \mathrm{O}(\Delta t)]\frac{\langle n_0|_{\tau-\Delta t} - \langle n_0|_\tau}{\Delta t} + \sum_{n_0}\frac{p(\cdot|\tau - \Delta t, n_0) - p(\cdot|\tau, n_0)}{\Delta t}\langle n_0|_\tau\Big)$$
$$= \lim_{\Delta t\to 0}\langle p(\cdot|\tau)|(\mathcal{E}_{\tau-\Delta t,\Delta t} + \mathcal{Q}_{\tau-\Delta t,\Delta t}) = \lim_{\Delta t\to 0}\langle p(\cdot|\tau)|\widetilde{\mathcal{Q}}_{\tau-\Delta t,\Delta t}\,,$$

with the operators being defined through their following actions on the basis functionals:

$$\langle n_0|_\tau \mathcal{E}_{\tau-\Delta t,\Delta t} = \frac{\langle n_0|_{\tau-\Delta t} - \langle n_0|_\tau}{\Delta t} \quad \text{and} \quad \langle n_0|_\tau \mathcal{Q}_{\tau-\Delta t,\Delta t} = \sum_m Q_{\tau-\Delta t,\Delta t}(n_0, m)\langle m|_\tau\,.$$

The discretization scheme conforms with the derivation of the backward master equation (14).

the insertion of the Poisson basis function $|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$ into the inverse transformation (186) results in

$$p(\tau, n|t_0, n_0) = \int_{-\infty}^{\infty} \mathrm{d}x\, g(\tau; x|t_0, n_0) \frac{x^n \mathrm{e}^{-x}}{n!}\,. \tag{195}$$

A representation of the probability distribution of this form is called a "Poisson representation" and was first proposed by Gardiner and Chaturvedi [37, 38]. Since the integration in (195) proceeds along the real line, the above representation is referred to as a "real" Poisson representation [377]. Although the use of a real variable may seem convenient, its use typically results in the kernel being a "generalized function", i.e. a distribution. For example, the initial condition $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$ is recovered for the integral kernel $g(t_0; x|t_0, n_0) = \delta(x)(\partial_x + 1)^{n_0}$. By the definition

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \big(\partial_x^j \delta(x)\big) f(x) := \int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x)(-\partial_x)^j f(x)$$

of distributional derivatives (1.16.12 in [337], $j \in \mathbb{N}_0$), this kernel can also be written as $[(1 - \partial_x)^{n_0} \delta(x)]$. The integral kernel $\delta(x - x_0)$ instead results for a Poisson distribution with mean x_0. In [37], Gardiner and Chaturvedi used the real Poisson representation to calculate steady-state probability distributions of various elementary reactions. Furthermore, the real Poisson representation was employed by Elderfield [378] to derive a stochastic path integral representation. Droz and McKane [379] later argued that this representation is equivalent to a path integral representation based on Doi's Fock space algebra. Our discussion of the backward and forward path integral representations in sections 4 and 5 clarifies the similarities and differences between these two approaches.

To circumvent the use of generalized functions, various alternatives to the real Poisson representation have been proposed: the "complex" and "positive" Poisson representations [115, 377] as well as the "gauge" Poisson representation [380]. The former two representations as well as the real representation are discussed in the book of Gardiner [115]. The complex Poisson representation is obtained by continuing the Poisson basis function in (195) into the complex domain and performing the integration around a closed path $\mathcal{C}$ around 0 (once in counterclockwise direction). Upon using Cauchy's differentiation formula to redefine the basis functional as

$$\langle n|f := \partial_x^n \mathrm{e}^x f(x)|_{x=0} = \oint_{\mathcal{C}} \mathrm{d}x\, \frac{n!}{2\pi \mathrm{i}} \frac{\mathrm{e}^x}{x^{n+1}} f(x)\,, \tag{196}$$

it becomes apparent that the initial condition $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$ is then recovered for the kernel $g(t_0; x|t_0, n_0) = \frac{n_0!}{2\pi \mathrm{i}} \frac{\mathrm{e}^x}{x^{n_0+1}}$. As the interpretation of this kernel is also not straightforward, we refrain from calling it a "quasi-probability distribution" [37].

Let us exemplify the flow equation obeyed by the integral kernel for the reaction $k\,A \to l\,A$ with rate coefficient γ_τ. The flow equation can be inferred from the corresponding flow equation (190) of the generating functional and the kernel's definition in (194). Since we employ the Poisson basis function, the results from section 3.2 b imply that the adjoint transition operator (156) reads

$$\mathcal{Q}_\tau^\dagger = \gamma_\tau x^k \big((\partial_x + 1)^l - (\partial_x + 1)^k\big)\,. \tag{197}$$

According to the flow equation (190) of the generating functional and the kernel's definition (194), it holds that

$$\int_{-\infty}^{\infty} \mathrm{d}x \,\Big(\partial_\tau g(\tau; x|\cdot)\Big) f(x) = \int_{-\infty}^{\infty} \mathrm{d}x \,\Big(g(\tau; x|\cdot)\mathcal{Q}_\tau^\dagger\Big) f(x)\,. \tag{198}$$

Integrations by parts therefore result in the following flow equation obeyed by the integral kernel:

$$\partial_\tau g = \mathcal{Q}_\tau(x, \partial_x) g = \gamma_\tau\big((1 - \partial_x)^l - (1 - \partial_x)^k\big) x^k g\,. \tag{199}$$

To arrive at this equation, we performed repeated integrations by parts while ignoring any potential boundary terms (cf. the definition of the adjoint operator in (149)). The importance of boundary terms is discussed in [380].

Thus far, most studies employing the Poisson representation have focused on networks of bimolecular reactions $\sum_j k_j A_j \to \sum_j l_j A_j$ with $\sum_j k_j \leq 2$ and $\sum_j l_j \leq 2$. For these networks, the flow equation (199) assumes the mathematical form of a forward Fokker-Planck equation with the derivatives being of at most second order (the corresponding diffusion coefficient may be negative). It has been attempted to map the resulting equation to an Itô SDE [37], but it should be explored whether this procedure is supported by the Feynman-Kac formula in appendix A. The numerical integration of an SDE with potentially negative diffusion coefficient was attempted for the bi-directional reaction $2\,A \leftrightharpoons \emptyset$ in [381]. The value of the integration has remained inconclusive. Recently, an exponential ansatz for the integral kernel $g(\tau; x|\cdot)$ has been considered to approximate its flow equation in the limit of weak noise [382] (cf. section 2.3 b). Moreover, a gauge Poisson representation was recently employed in a study of the coagulation reaction $2\,A \to A$ [383].

3.6 Résumé

In the previous section, we formulated general conditions under which the master equation can be transformed into a partial differential equation obeyed by the probability generating function (48). In the present section, we complemented this flow equation by a backward-time flow equation obeyed by the marginalized distribution (144) and by a functional flow equation obeyed by the probability generating functional (185). Whereas the marginalized distribution

$$|p(t,n|\tau)\rangle = \sum_{n_0} p(t,n|\tau,n_0)|n_0\rangle \tag{200}$$

was defined as the sum of the conditional probability distribution over a set of basis functions, the generating functional

$$\langle g(\tau|t_0,n_0)| = \sum_{n} \langle n|p(\tau,n|t_0,n_0) \tag{201}$$

was defined as the sum of the distribution over a set of basis functionals. In section 3.2, we introduced (O)rthogonal and (C)omplete basis functions and functionals for the study of different stochastic processes, including the Poisson basis for the study of chemical reactions ($|n\rangle_x = \frac{x^n e^{-x}}{n!}$ and $\langle m|f = (\partial_x + 1)^m f(x)\big|_{x=0}$). Provided that there also exist a basis (E)volution operator $\mathcal{E}_\tau$ and an adjoint transition operator $\mathcal{Q}_\tau^\dagger$ (condition ($\mathrm{Q}^\intercal$)), the marginalized distribution obeys the backward-time flow equation

$$\partial_{-\tau}|p\rangle = (-\mathcal{E}_\tau + \mathcal{Q}_\tau^\dagger)|p\rangle = \widetilde{\mathcal{Q}}_\tau^\dagger|p\rangle \tag{202}$$

with final condition $|p(t,n|t)\rangle = |n\rangle$. Moreover, the probability generating functional (185) then obeys the functional flow equation (190). The inverse transformation $p(\tau,n|t_0,n_0) = \langle g(\tau|t_0,n_0)|n\rangle$ of this functional generalizes the Poisson representation of Gardiner and Chaturvedi [37, 38] as shown in section 3.5. In section 3.3, we showed how the flow equation (202) obeyed by the marginalized distribution can be used to compute mean extinction times. Furthermore, the equation will prove useful in the derivation of path integral representations of the master equation and of averaged observables in sections 4 and 6. Future studies could explore whether the flow equation obeyed by the marginalized distribution can be evaluated in terms of WKB approximations or spectral methods.

4 The backward path integral representation

In the previous two sections, we showed how the forward and backward master equations can be cast into four linear PDEs for the series expansions (44)–(45). In this section, as well as in section 5, the solutions of these four equations are expressed in terms of two path integrals. Upon applying inverse transformations, the path integrals provide distinct representations of the conditional probability distribution solving the master equations. The flow equations obeyed by the generating function (44) and by the series (47) will lead us to the "forward" path integral representation

$$p(t, n|t_0, n_0) = \langle n|_t \fint_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}} \, |n_0\rangle_{t_0} \,, \tag{203}$$

and the flow equations obeyed by the marginalized distribution (45) and by the generating functional (46) to the "backward" path integral representation

$$p(t, n|t_0, n_0) = \langle n_0|_{t_0} \fint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \, |n\rangle_t \,. \tag{204}$$

The meanings of the integral signs $\fint_{[t_0}^{t)}$ and $\fint_{(t_0}^{t]}$, as well as of the exponential weights are explained below. We choose the above terms for the two representations because the forward path integral representation propagates the basis function $|n_0\rangle_{t_0}$ to time t, where it is then acted upon by the functional $\langle n|_t$. Counter-intuitively, however, this procedure requires us to solve a stochastic differential equation proceeding backward in time (cf. sections 5.2 and 5.4). Analogously, the backward path integral representation propagates the basis function $|n\rangle_t$ backward in time to t_0, where one then applies the functional $\langle n_0|_{t_0}$. This procedure requires us to solve an ordinary, forward-time Itô SDE (cf. section 4.3). The name "adjoint" path integral representation may also be used for the backward representation. As we will see below, its "action" $\mathcal{S}^\dagger$ involves the adjoint transition operator $\widetilde{\mathcal{Q}}_\tau^\dagger$.

To make the following derivations as explicit as possible, we now assume the discrete variables n and n_0 to be one-dimensional. Likewise, the variables x and q of the associated flow equations are one-dimensional real variables. The derivation below employs an exponential representation of the Dirac delta function. Although such a representation also exists for Grassmann variables [384], we do not consider that case. The extension of the following derivations to processes with multiple types of particles is straightforward and proceeds analogously to the inclusion of spatial degrees of freedom. A process with spatial degrees of freedom is considered in sections 4.5 b to 4.5 d.

After deriving the backward path integral representation in section 4.1, we exemplify how this representation can be used to solve the bi-directional reaction $\emptyset \rightleftharpoons A$ (section 4.2), the pair generating process $\emptyset \to 2A$ (section 4.5 a), and a process with diffusion and linear decay (sections 4.5 b to 4.5 d). The forward path integral representation is derived in section 5 and is exemplified in deriving the generating function of linear processes $A \to l\,A$ with $l \geq 0$. For the linear growth process $A \to 2A$, we recover a negative Binomial distribution as the solution of the master equation. Observables of the particle number are considered in section 6.

As an intermezzo, we show in sections 4.4 and 5.4 how one can derive path integral representations for jump processes with continuous state spaces (or processes whose transition rates can be extended to such spaces). The corresponding derivations are based on Kramers-Moyal expansions of the backward and forward master equations. Since the backward and forward Fokker-Planck equations constitute special cases of these expansions, we recover a classic path integral representation whose development goes back to works of Martin, Siggia, and Rose [21], de Dominicis [22], Janssen [23, 24], and Bausch, Janssen, and Wagner [24]. Moreover, we recover the Feynman-Kac formula from appendix A, an Onsager-Machlup representation [84], and Wiener's path integral for Brownian motion [85, 86].

Before starting out with the derivations of the two path integral representations (203) and (204), let us note that these representations only apply if their underlying transition operators $\widetilde{\mathcal{Q}}_\tau$ or $\widetilde{\mathcal{Q}}_\tau^\dagger$ can be written as power series in their arguments (x and ∂_x for the backward path integral and q and ∂_q for the forward path integral; the names of the variables differ because both path integrals then lead to the same path integral representation of averaged observables in section 6 without requiring a change of variable names). In addition, the power series need to be "normal-ordered", meaning that in every summand, all the x are to the left of all the ∂_x (all the q to the left of all the ∂_q). This order can be established by the repeated use of the commutation relation $[\partial_x, x]f(x) = f(x)$ (or, more directly, by invoking $(\partial_x\, x)^n f(x) = \sum_{m=0}^n \left\{ {n+1 \atop m+1} \right\} x^m \partial_x^m f(x)$ or $(x\partial_x)^n f(x) = \sum_{m=0}^n \left\{ {n \atop m} \right\} x^m \partial_x^m$, with the curly braces representing Stirling numbers of the second kind; cf. section 26.8 in [337]). The transition operator $\mathcal{Q}_\tau = \gamma_\tau (c^l - c^k) a^k$ in (69) and the adjoint transition operator $\mathcal{Q}_\tau^\dagger = \gamma_\tau c^k (a^l - a^k)$ in (156) of the chemical reaction $k\,A \to l\,A$ are already in their normal-ordered forms (with the creation and annihilation operators in (78) and (79), or (163) and (164), respectively).

4.1 Derivation

We first derive the backward path integral representation (204) because its application to actual processes is more intuitive than the application of the forward path integral representation. For this purpose, we consider the flow equation

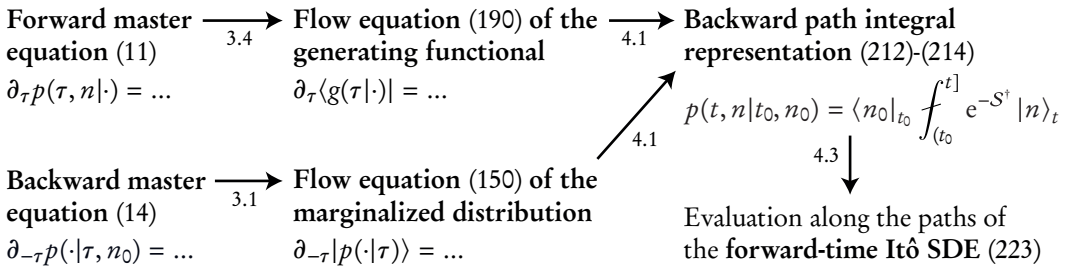


Figure 6 Outline of the derivation of the backward path integral representation and of its evaluation in terms of an average over the paths of an Itô stochastic differential equation.

$\partial_{-\tau}|p(\cdot|\tau)\rangle = \widetilde{\mathcal{Q}}_\tau^\dagger |p(\cdot|\tau)\rangle$ in (150) obeyed by the marginalized distribution. Its final condition reads $|p(t, n|t)\rangle = |n\rangle$. The conditional probability distribution is recovered from the marginalized distribution via $p(t, n|t_0, n_0) = \langle n_0|p(t, n|t_0)\rangle$. As the first step, we split the time interval $[t_0, t]$ into N pieces $t_0 \leq t_1 \leq \ldots \leq t_N := t$ of length $\Delta t := (t - t_0)/N \ll 1$. Over the time interval $[t_0, t_1]$, the flow equation is then solved by[12]

$$|p(t, n|t_0)\rangle_{x_0} = \mathcal{L}_{t_0}^\dagger(x_0, \partial_{x_0})|p(t, n|t_1)\rangle_{x_0} \tag{205}$$

with the generator $\mathcal{L}_\tau^\dagger := 1 + \widetilde{\mathcal{Q}}_\tau^\dagger \Delta t + O((\Delta t)^2)$ (see the footnote on page 58; for brevity, we now write $\widetilde{\mathcal{Q}}_\tau^\dagger$ instead of $\widetilde{\mathcal{Q}}_{\tau,\Delta t}^\dagger$ for the discrete-time operator). Alternatively, the following derivation could be performed by solving the flow equation (190) obeyed by the generating functional over the time interval $[t_{N-1}, t]$ as (cf. figure 6)

$$\langle g(t|t_0, n_0)| = \langle g(t_{N-1}|t_0, n_0)|\mathcal{L}_{t_{N-1}}^\dagger(x_{N-1}, \partial_{x_{N-1}}) \,. \tag{206}$$

Here, the generating functional $\langle g|$ acts on functions $f(x_{N-1})$, meets the initial condition $\langle g(t_0|t_0, n_0)| = \langle n_0|$, and is transformed back into the conditional probability distribution via $p(t, n|t_0, n_0) = \langle g(t_0|t_0, n_0)|n\rangle$. Both of the above approaches result in the same path integral representation. In the following, we use the equation (205) for this purpose.

As the next step of the derivation, we insert the integral representation of a Dirac delta between the generator $\mathcal{L}^\dagger$ and the marginalized distribution $|p\rangle$

12 This discretization of time conforms with the Itô prescription for forward-time SDEs and with the derivation of the backward master equation (14) (see section 4.3 and also the footnote on page 58).

in (205), turning this expression into

$$|p(t,n|t_0)\rangle_{x_0} = \mathcal{L}^\dagger_{t_0}(x_0,\partial_{x_0}) \int_{\mathbb{R}^2} \frac{\mathrm{d}x_1\,\mathrm{d}q_1}{2\pi}\, \mathrm{e}^{-\mathrm{i}q_1(x_1-x_0)} |p(t,n|t_1)\rangle_{x_1}\,. \tag{207}$$

The integrations over x_1 and q_1 are both performed along the real line from $-\infty$ to $+\infty$. If the adjoint transition operator $\widetilde{\mathcal{Q}}^\dagger_\tau$, and thus also $\mathcal{L}^\dagger_\tau$, are normal-ordered power series, we can replace ∂_{x_0} by $\mathrm{i}q_1$ in the above expression and pull $\mathcal{L}^\dagger_{t_0}$ to the right of the exponential. Making use of the final condition $|p(t,n|t)\rangle_{x_N} = |n\rangle_{t,x_N}$, the above steps can be repeated until (207) reads

$$|p(t,n|t_0)\rangle_{x_0} = \int_1^N \Big(\prod_{j=1}^N \mathrm{e}^{-\mathrm{i}q_j(x_j-x_{j-1})} \mathcal{L}^\dagger_{t_{j-1}}(x_{j-1},\mathrm{i}q_j)\Big)|n\rangle_{t,x_N}\,, \tag{208}$$

with the abbreviation

$$\int_k^l := \prod_{j=k}^l \int_{\mathbb{R}^2} \frac{\mathrm{d}x_j\,\mathrm{d}q_j}{2\pi}\,. \tag{209}$$

To proceed, we now replace $\mathcal{L}^\dagger_\tau = 1 + \widetilde{\mathcal{Q}}^\dagger_\tau \Delta t + \mathrm{O}\big((\Delta t)^2\big)$ by the exponential $\mathrm{e}^{\widetilde{\mathcal{Q}}^\dagger_\tau \Delta t}$ (for brevity, we drop the correction term in the following). Note that the exponential $\mathrm{e}^{\widetilde{\mathcal{Q}}^\dagger_\tau \Delta t}$ does not involve differential operators because those were all replaced by the new variables $\mathrm{i}q_j$. [13] Whether the exponentiation can be made mathematically rigorous is an open question and may possibly be answered positively only for a restricted class of stochastic processes. Upon performing the exponentiation, one obtains the following discrete-time path integral representation of the marginalized distribution:

$$|p(t,n|t_0)\rangle_{x_0} = \int_1^N \mathrm{e}^{-\mathcal{S}^\dagger_N}\, |n\rangle_{t,x_N} \text{ with} \tag{210}$$

$$\mathcal{S}^\dagger_N := \sum_{j=1}^N \Delta t \Big(\mathrm{i}q_j \frac{x_j - x_{j-1}}{\Delta t} - \widetilde{\mathcal{Q}}^\dagger_{t_{j-1}}(x_{j-1},\mathrm{i}q_j)\Big)\,. \tag{211}$$

The final condition $|p(t,n|t)\rangle_x = |n\rangle_{t,x}$ is trivially fulfilled because only $N = 0$ time slices fit between t and t. The object $\mathcal{S}^\dagger_N$ is called an "action". The exponential

13 The derivation of the path integral representation of averaged observables in section 6.2 involves exponentials of the form $\mathrm{e}^{c^2 a^2}$ (see also [96, 338, 385]), with c and a representing infinitely-large creation and annihilation matrices, respectively. The existence of such an exponential is not obvious.

factor $\mathrm{e}^{-\mathcal{S}_N^\dagger}$ weighs the contribution of every path $(x_1, q_1) \to (x_2, q_2) \to \ldots \to (x_N, q_N)$.

As the final step of the derivation, we take the continuous-time limit $N \to \infty$, or $\Delta t \to 0$, at least formally. The following continuous-time expressions effectively serve as abbreviations of the above discretization scheme with the identifications $x(t_0 + j\Delta t) \coloneqq x_j$ and $q(t_0 + j\Delta t) \coloneqq q_j$. Further comments on the discretization scheme are provided in section 4.4. Combined with the inverse transformation (145), we thus arrive at the following backward path integral representation of the conditional probability distribution:

$$p(t, n|t_0, n_0) = \langle n_0|_{t_0, x(t_0)} |p(t, n|t_0)\rangle_{x(t_0)} \tag{212}$$

$$\text{with } |p(t, n|t_0)\rangle_{x(t_0)} = \fint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} |n\rangle_{t, x(t)} \tag{213}$$

$$\text{and } \mathcal{S}^\dagger \coloneqq \int_{t_0}^{t} \mathrm{d}\tau \left[\mathrm{i}q\partial_\tau x - \widetilde{\mathcal{Q}}_\tau^\dagger(x, \mathrm{i}q)\right]. \tag{214}$$

Here we included $x(t_0)$ as an argument of the functional $\langle n_0|_{t_0, x(t_0)}$ to express that this functional acts on $x(\tau)$ only for $\tau = t_0$. Moreover, we defined $\fint_{(t_0}^{t]} \coloneqq \lim_{N\to\infty} \int_1^N$ to indicate that the path integral involves integrations over $x(t)$ and $q(t)$, but not over $x(t_0)$ and $q(t_0)$.

The evaluation of the above path integral representation for an explicit process involves two steps. First, the path integral (213) provides the marginalized distribution $|p(t, n|t_0)\rangle$. Second, this marginalized distribution is mapped to the conditional probability distribution $p(t, n|t_0, n_0)$ by the action of the functional $\langle n_0|$.

4.2 Simple growth and linear decay

Let us demonstrate the above two steps for the bi-directional reaction $\emptyset \rightleftharpoons A$. The stationary distribution of this process was already derived in the introduction to section 2. Both the growth rate coefficient γ and the decay rate coefficient μ shall be time-independent, but this assumption is easily relaxed. As the first step, the backward path integral (212) requires us to choose an appropriate basis. The time-independent Poisson basis with the basis function $|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$ proves to be convenient for this purpose (cf. section 3.2 b). The adjoint transition operator follows from (156), (163), and (164) as $\widetilde{\mathcal{Q}}^\dagger(x, \partial_x) = (\gamma - \mu x)\partial_x$. Insertion of this operator into the action (214) results in

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau\, \mathrm{i}q \left[\partial_\tau x - (\gamma - \mu x)\right] \tag{215}$$

The integration over the path $q(\tau)$ from $\tau = t_0$ to $\tau = t$ is performed most easily in the discrete-time approximation. The marginalized distribution thereby follows as the following product of Dirac deltas:

$$|p(t,n|t_0)\rangle_{x_0} = \Big(\prod_{j=1}^{N} \int_{\mathbb{R}} \mathrm{d}x_j\, \delta\big(x_j - x_{j-1} - (\gamma - \mu\, x_{j-1})\Delta t\big)\Big)|n\rangle_{t,x_N}\,.$$

The function $|n\rangle_{t,x_N}$ is also integrated over. Upon taking the continuous-time $\Delta t \to 0$ and performing the integration over the path $x(\tau)$, one finds that the marginalized distribution is given by the Poisson distribution

$$|p(t,n|t_0)\rangle_{x(t_0)} = |n\rangle_{t,x(t)} = \frac{x(t)^n \mathrm{e}^{-x(t)}}{n!}\,, \tag{216}$$

with $x(\tau)$ solving $\partial_\tau x = \gamma - \mu x$. This equation coincides with the rate equation of the process $\emptyset \rightleftharpoons A$. The solution $x(\tau) = \mathrm{e}^{-\mu(\tau-t_0)}(x(t_0) - \frac{\gamma}{\mu}) + \frac{\gamma}{\mu}$ of the rate equation specifies the mean and the variance of the above Poisson distribution. For $\mu > 0$, both of them converge to the asymptotic value γ/μ. In the next section and in section 4.5, we generalize these results to more general processes.

The marginalized distribution (216) does not only solve the master equation of the reaction $\emptyset \rightleftharpoons A$, but it also establishes the link between the path integral variable x and the moments of the particle number n. In particular, the mean particle number evaluates to $\langle n\rangle(t) = x(t)$, while higher order moments can determined via $\langle n^k\rangle(t) = \sum_{l=0}^{k} \{{k \atop l}\} x(t)^l$. The curly braces denote a Stirling number of the second kind (cf. section 26.8 in [337]).

The conditional probability distribution can be calculated from the marginalized distribution by applying the functional $\langle n_0|f = (\partial_{x(t_0)} + 1)^{n_0} f(x(t_0))|_{x(t_0)=0}$ to the latter (cf. (160) and (164)). In the limit of a vanishing decay rate for which $x(\tau) = x(t_0) + \gamma(\tau - t_0)$, one recovers the shifted Poisson distribution (95), i.e.

$$p(t,n|t_0,n_0) = \frac{\mathrm{e}^{-\gamma(t-t_0)}(\gamma(t-t_0))^{n-n_0}}{(n-n_0)!}\Theta_{n-n_0}\,. \tag{217}$$

Thus, the mean and variance of the conditional probability distribution grow linearly with time. If the particles decay but are not replenished ($\mu > 0$ but $\gamma = 0$), we instead recover the Binomial distribution (169).

4.3 Feynman-Kac formula for jump processes

We now show how the backward path integral representation (213) of the marginalized distribution can be expressed in terms of an average over the paths of an Itô

stochastic differential equation (SDE). The resulting expression bears similarities with the Feynman-Kac formula (3), especially when the adjoint transition operator $\widetilde{\mathcal{Q}}_\tau^\dagger(x,\partial_x)$ of the stochastic process under consideration is quadratic in ∂_x. In the general case, however, functional derivatives act on the average over paths. These derivatives can, for example, be evaluated in terms of perturbation expansions as we demonstrate in section 4.5 d. The procedure outlined below serves as a general starting point for the exact or approximate evaluation of the backward path integral (213).

In the following, we consider a stochastic process whose adjoint transition operator has the generic form

$$\widetilde{\mathcal{Q}}_\tau^\dagger(x,\partial_x) = \alpha_\tau(x)\partial_x + \frac{1}{2}\beta_\tau(x)\partial_x^2 + \mathcal{P}_\tau^\dagger(x,\partial_x)\,. \tag{218}$$

As before, we call α_τ a drift and β_τ a diffusion coefficient. The object $\mathcal{P}_\tau^\dagger$ is referred to as the (adjoint) perturbation operator, or simply as the perturbation. The perturbation operator absorbs all the terms of higher order in ∂_x and possibly also terms of lower order. Thus, the above form of $\widetilde{\mathcal{Q}}_\tau^\dagger$ is not unique and α_τ, β_τ, and $\mathcal{P}_\tau^\dagger$ should be chosen so that the evaluation of the expressions below becomes as simple as possible. If the perturbation operator $\mathcal{P}_\tau^\dagger$ is zero, those expressions simplify considerably.

As the first step, let us rewrite the backward path integral representation (213) of the marginalized distribution as

$$|p(t,n|t_0)\rangle_{x(t_0)} = \int_{\mathbb{R}^2} \frac{\mathrm{d}x_N\,\mathrm{d}q_N}{2\pi} \mathrm{e}^{-\mathrm{i}q_N x_N}\, \mathcal{Z}_{0,0}|n\rangle_{t,x_N}\,, \tag{219}$$

where $\mathcal{Z}_{Q=0,X=0}$ represents a value of the (Q,X)-generating functional

$$\mathcal{Z}_{Q,X} := \int_{(t_0}^{t)} \mathrm{e}^{\mathrm{i}q_N x_N - \mathcal{S}^\dagger + \int_{t_0}^t \mathrm{d}\tau\,[Q(\tau)x(\tau) + X(\tau)\mathrm{i}q(\tau)]}\,. \tag{220}$$

Hence, we have singled out the integrations over x_N and q_N before performing the continuous-time limit (cf. (210) and (213)). We call $\mathcal{Z}_{Q,X}$ a generating functional because a functional differentiation of $\mathcal{Z}_{Q,X}$ with respect to $Q(\tau)$ generates a factor $x(\tau)$, and a functional differentiation of $\mathcal{Z}_{Q,X}$ with respect to $X(\tau)$ generates a factor $\mathrm{i}q(\tau)$.

Given the action (214) with the adjoint transition operator (218), the above properties of the (Q,X)-generating functional can now be used to rewrite this

function as (cf. appendix D)

$$\mathcal{Z}_{Q,X} = \mathrm{e}^{\mathrm{i}q_N \frac{\delta}{\delta Q(t)} + \int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{P}_\tau^\dagger\left(\frac{\delta}{\delta Q(\tau)}, \frac{\delta}{\delta X(\tau)}\right)} \mathcal{Z}^0_{Q,X} \tag{221}$$

$$\text{with } \mathcal{Z}^0_{Q,X} := \left\langle\!\!\left\langle \mathrm{e}^{\int_{t_0}^{t} \mathrm{d}\tau Q(\tau) x(\tau)} \right\rangle\!\!\right\rangle_W . \tag{222}$$

Here, $\langle\!\langle \cdot \rangle\!\rangle_W$ represents the average over realizations of a Wiener process $W(\tau)$. This Wiener process influences the evolution of the path $x(\tau)$ through the Itô SDE

$$\mathrm{d}x(\tau) = \left[\alpha_\tau(x) + X(\tau)\right] \mathrm{d}\tau + \sqrt{\beta_\tau(x)}\, \mathrm{d}W(\tau) . \tag{223}$$

The temporal evolution of $x(\tau)$ starts out from $x(t_0)$, which is determined by the argument of the marginalized distribution (219). In section 4.5 d, we demonstrate how the marginalized distribution can be evaluated in terms of a perturbation expansion of the (Q, X)-generating functional (221) for a process of diffusing particles that are also decaying. In order to perform the perturbation expansion, let us already note that the value $x(\tau)$ of the path at time τ depends on the "source" $X(\tau')$ only for times $\tau' < \tau$, and that $\frac{\delta}{\delta Q(\tau)} \int_{t_0}^{t} \mathrm{d}\tau' Q(\tau') f(\tau') = f(\tau)$ holds for all $\tau \in (t_0, t]$. Moreover, let us note that the (Q, X)-generating functional depends on q_N and $x(t_0)$ but not on x_N. These properties follow from the derivation of the representation (221), which we outline in appendix D.

An important special case of the above representation is constituted by processes whose perturbation operator $\mathcal{P}_\tau^\dagger$ is zero. For $Q = X = 0$, the generating functional (221) then simplifies to $\mathcal{Z}_{0,0} = \mathrm{e}^{\mathrm{i}q_N x(t)}$ and the marginalized distribution (219) follows in terms of the Feynman-Kac like fomula

$$|p(t, n|t_0)\rangle_{x(t_0)} = \left\langle\!\!\left\langle |n\rangle_{t,x(t)} \right\rangle\!\!\right\rangle_W . \tag{224}$$

Here, $x(\tau)$ solves the Itô SDE

$$\mathrm{d}x(\tau) = \alpha_\tau(x)\, \mathrm{d}\tau + \sqrt{\beta_\tau(x)}\, \mathrm{d}W(\tau) \tag{225}$$

with the initial value $x(t_0)$. For the Poisson basis function $|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$, the above representation of the marginalized distribution coincides with our earlier result (171), which we encountered in the discussion of the coagulation reaction $2A \to A$. For the bi-directional reaction $\emptyset \rightleftharpoons A$ from the previous section with adjoint transition operator $\widetilde{\mathcal{Q}}^\dagger(x, \partial_x) = (\gamma - \mu x)\partial_x$, the Itô SDE (225) simplifies to the deterministic rate equation $\partial_\tau x = \gamma - \mu x$. Thus, no averaging is required to evaluate the marginalized distribution (224) and one recovers our earlier solution (216).

Our above discussion only applies to processes in well-mixed environments with a single type of particles. A generalization of the results to processes with

multiple types of particles or with spatial degrees of freedom is straightforward. A spatial process will be discussed in section 4.5. In a multivariate generalization of the above procedure, the adjoint transition operator (218) includes a vector-valued drift coefficient $\boldsymbol{\alpha}_\tau(\boldsymbol{x})$ and a diffusion matrix $\beta_\tau(\boldsymbol{x})$. Moreover, the derivation of the corresponding generating functional (221) requires that there exists a matrix $\sqrt{\beta_\tau} \coloneqq \gamma_\tau$ fulfilling $\gamma_\tau \gamma_\tau^\intercal = \beta_\tau$. If the diffusion matrix β_τ is positive-semidefinite, its positive-semidefinite and symmetric square root $\sqrt{\beta_\tau}$ can be determined via diagonalization [116].

In section 4.5, we exemplify the use of our above results for various well-mixed and spatial processes. But before, let us show how the results from the previous sections can be used to derive a path integral representation for processes with continuous state spaces.

4.4 Intermezzo: The backward Kramers-Moyal expansion

The transition rate $w_\tau(m, n)$ denotes the rate at which probability flows from a state n to a state m, or, considering an individual sample path, the rate at which particles jump from state n to state m. Thus, $\kappa_\tau(\Delta n, n) \coloneqq w_\tau(n + \Delta n, n)$ denotes the rate at which the state n is left via jumps of size Δn (with $n \in \mathbb{N}_0$ and $\Delta n \in \mathbb{Z}$). Thus far, we only considered jumps between the states of a discrete state space. But in the following, we derive a path integral representation for processes having a continuous state space.

4.4 a Processes with continuous state spaces

We consider a process whose state is characterized by a continuous variable $x \in \mathbb{R}_{\geq 0}$. The change from the letter n to the letter x is purely notational and emphasizes that the state space is now continuous (the change also highlights a formal similarity between the linear PDEs discussed so far and the ones derived below). The conditional probability distribution $p(\tau, x|t_0, x_0)$ describing the system shall be normalized as $\int_{\mathbb{R}} \mathrm{d}x\, p(\tau, x|\cdot) = 1$ and obey the master equation

$$\partial_\tau p(\tau, x|\cdot) = \int_{\mathbb{R}} \mathrm{d}\Delta x \left[\kappa_\tau(\Delta x, x - \Delta x) p(\tau, x - \Delta x|\cdot) - \kappa_\tau(\Delta x, x) p(\tau, x|\cdot)\right] \quad (226)$$

with the initial condition $p(t_0, x|t_0, x_0) = \delta(x - x_0)$. The structure of this master equation is equivalent to the structure of the master equation (12). Provided that the product $\kappa_\tau(\Delta x, x) p(\tau, x|\cdot)$ is analytic in x, one can perform a Taylor expansion of the above master equation to obtain the (forward) Kramers-Moyal expansion [6, 108, 274]

$$\partial_\tau p(\tau, x|\cdot) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \partial_x^m \left[\mathcal{M}_\tau^{(m)}(x) p(\tau, x|\cdot)\right]. \quad (227)$$

The "jump moments" $\mathcal{M}_\tau^{(m)}$ are defined as

$$\mathcal{M}_\tau^{(m)}(x) := \int_{\mathbb{R}} \mathrm{d}\Delta x \, (\Delta x)^m \kappa_\tau(\Delta x, x) \,. \tag{228}$$

By Pawula's theorem [109, 386, 387], the positivity of the conditional probability distribution requires that the Kramers-Moyal expansion either stops at its first or second summand, or that it does not stop at all. If the expansion stops at its second summand, it assumes the form of a (forward) Fokker-Planck equation. The drift coefficient of this Fokker-Planck equation is given by $\mathcal{M}_\tau^{(1)}(x)$ and its non-negative diffusion coefficient by $\mathcal{M}_\tau^{(2)}(x)$. The sample paths of the Fokker-Planck equation are continuous, however, contradicting our earlier assumption of the process making discontinuous jumps in state space. For a jump process, the Kramers-Moyal expansion cannot stop. Nevertheless, a truncation of the Kramers-Moyal expansion at the level of a Fokker-Planck equation often provides a decent approximation of a process, provided that fluctuations cause only small relative changes of its state x.

The backward analogue of the master equation (226) reads

$$\partial_{-\tau} p(t, x|\tau, x_0) = \int_{\mathbb{R}} \mathrm{d}\Delta x \, \big[p(\cdot|\tau, x_0 + \Delta x) - p(\cdot|\tau, x_0)\big] \kappa_\tau(\Delta x, x_0) \,, \tag{229}$$

with the final condition $p(t, x|t, x_0) = \delta(x - x_0)$. Given the analyticity of $p(\cdot|\tau, x_0)$ in x_0, the backward master equation can be rewritten in terms of the backward Kramers-Moyal expansion

$$\partial_{-\tau} p(\cdot|\tau, x_0) = \widetilde{\mathcal{Q}}_\tau^\dagger(x_0, \partial_{x_0}) p(\cdot|\tau, x_0) \tag{230}$$

with the adjoint transition operator

$$\widetilde{\mathcal{Q}}_\tau^\dagger(x_0, \partial_{x_0}) := \sum_{m=1}^{\infty} \frac{1}{m!} \mathcal{M}_\tau^{(m)}(x_0) \partial_{x_0}^m \,. \tag{231}$$

This operator is the adjoint of the operator in the forward Kramers-Moyal expansion (227) in the sense that $\int \mathrm{d}x \, [\widetilde{\mathcal{Q}}_\tau^\dagger(x, \partial_x) f(x)] g(x) = \int \mathrm{d}x \, f(x) [\widetilde{\mathcal{Q}}_\tau(x, \partial_x) g(x)]$ (provided that all boundary terms in the integrations by parts vanish).

4.4 b Path integral representation of the backward Kramers-Moyal expansion

The adjoint transition operator $\widetilde{\mathcal{Q}}_\tau^\dagger(x_0, \partial_{x_0})$ of the backward Kramers-Moyal expansion (231) is normal-ordered with respect to x_0 and ∂_{x_0}. Moreover, the backward Kramers-Moyal expansion (230) has the same form as the flow equation (150)

obeyed by the marginalized distribution. One can therefore follow the steps in section 4.1 to represent the backward Kramers-Moyal expansion by the path integral

$$p(t, x|t_0, x_0) = \fint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger}\, \delta(x - x(t))\big|_{x(t_0)=x_0} \tag{232}$$

with the action

$$\mathcal{S}^\dagger := \int_{t_0}^{t} \mathrm{d}\tau \left[\mathrm{i}q(\tau)\partial_\tau x(\tau) - \widetilde{\mathcal{Q}}_\tau^\dagger(x(\tau), \mathrm{i}q(\tau))\right]. \tag{233}$$

The integral sign in (232) is again defined as the continuous-time limit of (209) and traces out all paths of $x(\tau)$ and $q(\tau)$ for $\tau \in (t_0, t]$ (note that $x(\tau)$ differs from x and x_0, which are fixed parameters; for brevity, however, $x(\tau)$ is occasionally abbreviated as x below). A diagrammatic computation of multi-time correlation functions based on a path integral representation equivalent to (232) has recently been considered in [388].

Let us specify the path integral representation (232) for a model of the chemical reaction $k\,A \to l\,A$ with rate coefficient γ_τ (and $k, l \in \mathbb{N}_0$). As we assume the particle "number" x to be continuous, the model is only reasonable in an approximate sense for large values of x. In defining the transition rate of the reaction, one has to ensure the non-negativity of x and the conservation of probability. A possible choice of the transition rate is

$$\kappa_\tau(\Delta x, x) := \gamma_\tau x^k \Theta(x - k)\delta\big(\Delta x - (l - k)\big). \tag{234}$$

Here, the Heaviside step function $\Theta(x - k)$ ensures that a sufficient number of particles are present to engage in a reaction (and it prevents the loss of probability to negative values of x).

Assuming a smooth approximation of the Heaviside step function that vanishes for $x < 0$, the jump moments (228) follow as

$$\mathcal{M}_\tau^{(m)}(x) = \gamma_\tau (l - k)^m x^k \Theta(x - k). \tag{235}$$

The corresponding adjoint transition operator evaluates to

$$\widetilde{\mathcal{Q}}_\tau^\dagger(x, \mathrm{i}q) = \gamma_\tau (x\mathrm{e}^{-\mathrm{i}q})^k \big[(\mathrm{e}^{\mathrm{i}q})^l - (\mathrm{e}^{\mathrm{i}q})^k\big]\Theta(x - k). \tag{236}$$

Path integrals with such an operator have been noted in [389–391], but their potential use remains to be fully explored. Curiously, upon ignoring the step function, the operator (236) has the same structure as the transition operator (156) upon

identifying $x\mathrm{e}^{-\mathrm{i}q}$ with the creation operator c, and $\mathrm{e}^{\mathrm{i}q}$ with the annihilation operator a (Cole-Hopf transformation [389–393]). Note, however, that the stochastic processes associated to the two transition operators differ from each other (discrete vs. continuous state space; different transition rates). The connection between the two associated path integrals (213) and (232) should be further explored.

4.4 c Path integral representation of the backward Fokker-Planck equation

With the drift coefficient $\alpha_\tau(x) := \mathcal{M}^{(1)}_\tau(x)$ and the non-negative diffusion coefficient $\beta_\tau(x) := \mathcal{M}^{(2)}_\tau(x)$, a truncation of the adjoint transition operator (231) at its second summand reads

$$\widetilde{\mathcal{Q}}^\dagger_\tau(x_0, \partial_{x_0}) = \alpha_\tau(x_0)\partial_{x_0} + \frac{1}{2}\beta_\tau(x_0)\partial^2_{x_0}\,. \tag{237}$$

Thus, the corresponding Kramers-Moyal expansion (230) recovers the backward Fokker-Planck equation (2). Since the action (233) evaluates to

$$\mathcal{S}^\dagger := \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}q\big[\partial_\tau x - \alpha_\tau(x)\big] + \frac{1}{2}\beta_\tau(x)q^2\right), \tag{238}$$

the corresponding path integral (232) coincides with a classic path integral representation of the (backward) Fokker-Planck equation. The original development of this representation goes back to works of Martin, Siggia, and Rose [21], de Dominicis [22], Janssen [23, 24], and Bausch, Janssen, and Wagner [24]. The application of this path integral to stochastic processes is, for example, discussed in the book of Täuber [75].

The transition operator (237) of the backward Fokker-Planck equation has the same form as the transition operator (218) in section 4.3, but it does not involve a perturbation operator $\mathcal{P}^\dagger_\tau$. Therefore, one can follow the steps in that section to evaluate the path integral (232) in terms of an average over the paths of an Itô SDE. This procedure shows that the backward Fokker-Planck equation is solved by

$$p(t, x|t_0, x_0) = \big\langle\!\big\langle \delta(x - x(t)) \big\rangle\!\big\rangle_W\,, \tag{239}$$

with $x(\tau)$ solving the Itô SDE

$$\mathrm{d}x(\tau) = \alpha_\tau(x(\tau))\,\mathrm{d}\tau + \sqrt{\beta_\tau(x(\tau))}\,\mathrm{d}W(\tau) \tag{240}$$

with initial value $x(t_0) = x_0$. Hence, we have recovered the Feynman-Kac formula (3) (apart from a notational change in the time parameters).

4.4 d The Onsager-Machlup function

The connection between the above path integral representation of the (backward) Fokker-Planck equation and the work of Onsager and Machlup [84] becomes apparent upon the completion of a square (as in a Hubbard-Stratonovich transformation [394, 395]). For this purpose, the diffusion coefficient β_τ in the transition operator (237) must not only be non-negative but positive(-definite). One can then complete the square in the variable q as

$$\mathcal{S}_N^\dagger = \sum_{j=1}^{N} \Big((x_j - x_{j-1} - \alpha_{t_{j-1}}(x_{j-1})\Delta t)\, \mathrm{i} q_j + \frac{1}{2} \beta_{t_{j-1}}(x_{j-1}) \Delta t\, q_j^2 \Big) \tag{241}$$

$$= \sum_{j=1}^{N} \frac{\beta_{t_{j-1}}(x_{j-1})\Delta t}{2} \Big(q_j + \mathrm{i} \frac{x_j - x_{j-1} - \alpha_{t_{j-1}}(x_{j-1})\Delta t}{\beta_{t_{j-1}}(x_{j-1})\Delta t} \Big)^2 \tag{242}$$

$$+ \sum_{j=1}^{N} \frac{(x_j - x_{j-1} - \alpha_{t_{j-1}}(x_{j-1})\Delta t)^2}{2\beta_{t_{j-1}}(x_{j-1})\Delta t} \,.$$

Next, one performs the path integration over q. Returning to the discrete-time approximation (211) of the action, one obtains the following representation of the conditional probability distribution in terms of convolutions of Gaussian distributions:

$$p(t, x | t_0, x_0) = \lim_{N\to\infty} \Big(\prod_{j=0}^{N-1} \int_{\mathbb{R}} \mathrm{d}x_{j+1}\, \mathcal{G}_{\mu_j, \sigma_j^2}(x_{j+1} - x_j) \Big) \delta(x - x_N) \,. \tag{243}$$

The Dirac delta function is included in the integrations. Moreover, $\mu_j := \alpha_{t_j}(x_j)\Delta t$ acts as the mean and $\sigma_j^2 := \beta_{t_j}(x_j)\Delta t$ as the variance of the Gaussian distribution

$$\mathcal{G}_{\mu,\sigma^2}(x) := \frac{\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \,. \tag{244}$$

As before, $\Delta t = (t - t_0)/N$ denotes the intervals between $t_0 \le t_1 \le \ldots \le t_N = t$.

Identifying $x(t_0 + j\Delta t)$ with x_j, one can rewrite the representation (243) as

$$p(t, x | t_0, x_0) = \fint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger}\, \delta(x - x(t))\big|_{x(t_0)=x_0} \,. \tag{245}$$

This integral proceeds only over paths $x(\tau)$ with $\tau \in (t_0, t]$ because the q-variables have already been integrated over. Paths are weighed by the exponential factor

$\mathrm{e}^{-\mathcal{S}^\dagger}$ with the action

$$\mathcal{S}^\dagger := \lim_{N\to\infty} \sum_{j=0}^{N-1} \frac{\left[x_{j+1} - x_j - \alpha_{t_j}(x_j)\Delta t\right]^2}{2\beta_{t_j}(x_j)\Delta t} \,. \tag{246}$$

One may abbreviate the continuous-time limit of this action by the integral

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau\, \frac{\left[\partial_\tau x - \alpha_\tau(x)\right]^2}{2\beta_\tau(x)} \,. \tag{247}$$

The integrand of this action is called an Onsager-Machlup function [84, 396] and the representation (245) may, consequently, be called an Onsager-Machlup representation (or "functional" in the sense of functional integration). Note that the Onsager-Machlup representation involves the limit

$$\fint_{(t_0}^{t]} := \lim_{N\to\infty} \prod_{j=0}^{N-1} \int_{\mathbb{R}} \frac{\mathrm{d}x_{j+1}}{\sqrt{2\pi\beta_{t_j}(x_j)\Delta t}} \,. \tag{248}$$

Mathematically rigorous formulations of the Onsager-Machlup representation have been attempted in [397–399]. For the other path integrals discussed in this thesis, such attempts have not yet been made to our knowledge.

4.4 e Alternative discretization schemes

Up to this point, we have discretized time in such a way that the evaluation of the path integrals eventually proceeds via the solution of an Itô stochastic differential equation (cf. section 4.3). Alternative discretization schemes have been proposed as well and are, for example, employed in stochastic thermodynamics [239]. To illustrate these schemes, let us, for simplicity, assume that the drift coefficient $\alpha(x)$ does not depend on time and that the diffusion coefficient $D := \beta_\tau$ is constant. A general discretization scheme — called the α-scheme but here we use a κ — consists of shifting the argument x_j of the drift coefficient in the action (246) to $\bar{x}_j := \kappa\, x_{j+1} + (1-\kappa)x_j$ by writing

$$x_j = \bar{x}_j - \kappa(x_{j+1} - x_j) \tag{249}$$

(with $\kappa \in [0,1]$). Afterwards, the drift coefficient is expanded in powers of $x_{j+1} - x_j$, which is of order $\sqrt{\Delta t}$ along relevant paths. Following [400], it suffices to keep only the first two terms of the expansion of $\alpha(x_j)$ so that

$$x_{j+1} - x_j - \alpha(x_j)\Delta t \approx \left[1 + \kappa\,\alpha'(\bar{x}_j)\Delta t\right](x_{j+1} - x_j) - \alpha(\bar{x}_j)\Delta t \,. \tag{250}$$

As the next step, we rename the variables x_i to y_i and define new variables x_i via

$$x_{j+1} - x_j := \left[1 + \kappa\, \alpha'(\bar{y}_j)\Delta t\right](y_{j+1} - y_j) \tag{251}$$

(with $x_0 := y_0$). The Jacobian of this transformation vanishes everywhere except on its diagonal and sub-diagonal. Its determinant therefore contributes the following factor to the path integral

$$\prod_{j=0}^{N-1} \frac{\partial x_{j+1}}{\partial y_{j+1}} = \prod_{j=0}^{N-1} \left(1 + \kappa\, \alpha'(\bar{y}_j)\Delta t + \ldots\right) = \exp\left(\kappa \sum_{j=0}^{N-1} \Delta t\, \alpha'(\bar{y}_j)\right) + \ldots . \tag{252}$$

Consequently, the action (246) must be redefined as

$$\mathcal{S}^\dagger := \lim_{N\to\infty} \sum_{j=0}^{N-1} \left(\frac{\left[x_{j+1} - x_j - \alpha(\bar{x}_j)\Delta t\right]^2}{2D\Delta t} + \kappa\, \alpha'(\bar{x}_j)\Delta t \right).$$

One may abbreviate this limit by

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau \left(\frac{\left[\partial_\tau x - \alpha(x)\right]^2}{2D} + \kappa\, \alpha'(x) \right). \tag{253}$$

The Itô version of this action with $\kappa = 0$ has proved to be convenient in perturbation expansions of path integrals because so called "closed response loops" can be omitted right from the start (see section 4.5 in [75]). Moreover, this prescription does not require a change of variables and makes the connection to the Feynman-Kac formula in appendix A most apparent. The Stratonovich version of the action with $\kappa = \frac{1}{2}$ is, for example, employed in Seifert's review on stochastic thermodynamics [239]. For a recent, more thorough discussion of the above discretization schemes, as well as of conflicting approaches, we refer the reader to [401] (the above action is discussed in the appendices).

4.4 f Wiener's path integral

Before returning to processes with a discrete state space, let us briefly note how the Onsager-Machlup representation (245) relates to Wiener's path integral of Brownian motion [85, 86], as it is discussed in [402]. It turns out that the Onsager-Machlup representation in fact coincides with that path integral for one-dimension Brownian motion, for which the drift coefficient vanishes and the diffusion coefficient $D = \beta_\tau$ is constant. Since the convolution of two Gaussian distributions is again a Gaussian distribution, with means and variances being summed, the solution of the process follows as $p(t, x|t_0, x_0) = \mathcal{G}_{0,D(t-t_0)}(x - x_0)$.

4.5 Further exact solutions and perturbation expansions

After this detour to processes with continuous state spaces, let us return to the evaluation of the backward path integral representation (213). In the following, we show how the method introduced in section 4.3 can be applied to several elementary jump processes. We already applied a simplified version of the method in section 4.2 to solve the bi-directional reaction $\emptyset \rightleftharpoons A$. We now consider a generic reaction $k\,A \to l\,A$ with rate coefficient γ_τ. Using the Poisson basis function $|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$, the adjoint transition operator of this reaction can be written as (cf. section 3.2 b)

$$\widetilde{\mathcal{Q}}^\dagger_\tau(x,\partial_x) = \mathcal{Q}^\dagger(c,a) = \gamma_\tau c^k (a^l - a^k) \tag{254}$$

$$= \left[\gamma_\tau (l-k) x^k\right]\partial_x + \left[\gamma_\tau \big(l(l-1) - k(k-1)\big) x^k\right]\frac{\partial_x^2}{2} + \mathcal{P}^\dagger_\tau(x,\partial_x)\,. \tag{255}$$

Here we employed the creation operator $c = x$ and the annihilation operator $a = \partial_x + 1$, and we performed a series expansion with respect to ∂_x. Terms of third and higher order in ∂_x were shoved into the perturbation operator $\mathcal{P}^\dagger_\tau(x,\partial_x)$. In the following, we approximate the marginalized distribution of the reaction $k\,A \to l\,A$ by first dropping both the diffusion coefficient and the perturbation operator from (254). Afterwards, we reintroduce the diffusion coefficient and show how the marginalized distribution follows as the average of a Poisson distribution over the paths of an Itô SDE. For the pair generation process $\emptyset \to 2\,A$, this representation is exact because the perturbation operator associated to this process is zero (cf. section 4.5 a). Later, in section 4.5 d, we solve a process with a non-vanishing perturbation operator $\mathcal{P}^\dagger_\tau$.

As a first approximation of the reaction $k\,A \to l\,A$, let us drop all the terms of the adjoint transition operator (254) except for the drift coefficient $\alpha_\tau(x) = \gamma_\tau(l-k)x^k$. The SDE (223) then simplifies to the deterministic rate equation $\partial_\tau x = \gamma_\tau(l-k)x^k$ of the reaction. Its solution $x(\tau)$ acts as the mean of the marginalized distribution, which, according to (219)–(222), is again given by the Poisson distribution $|p(t,n|t_0)\rangle_{x(t_0)} = |n\rangle_{t,x(t)}$ in (216).

Going one step further, one may keep the diffusion coefficient in (254). The marginalized distribution then reads

$$|p(t,n|t_0)\rangle_{x(t_0)} = \left\langle\!\!\left\langle \frac{x(t)^n \mathrm{e}^{-x(t)}}{n!} \right\rangle\!\!\right\rangle_W \tag{256}$$

with $x(\tau)$ solving the Itô SDE

$$\mathrm{d}x = \gamma_\tau(l-k)x^k\,\mathrm{d}\tau + \sqrt{\gamma_\tau\big[l(l-1) - k(k-1)\big]x^k}\;\mathrm{d}W(\tau)\,. \tag{257}$$

Hence, the Poisson distribution in (256) is averaged over all possible sample paths of the SDE (257), whose evolution starts out from the initial value $x(t_0)$ (cf. (171)). The above expression has also recently been derived by Wiese [338], although based on the forward path integral that we will discuss in section 5. Apparently, the expression under the square root in (257) is non-negative only if $k \in \{0, 1\}$ or if $l \geq k \geq 2$. If that is not the case, $x(\tau)$ strays off into the complex domain. SDEs with imaginary noise (or the corresponding Langevin equations) have been studied in several recent articles, most notably for the binary annihilation reaction $2\,A \to \emptyset$ and for the coagulation process $2\,A \to A$ [338, 403, 404]. The numerical evaluation of such SDEs, however, often encountered severe convergence problems [338, 381]. The appearance of imaginary noise has been linked to the anti-correlation of particles in spatial systems [403], but as (257) shows, no spatial degrees of freedoms are actually required for its emergence. As Wiese pointed out, imaginary noise generally appears when, over time, the marginalized distribution (256) becomes narrower than a Poisson distribution [338].

4.5 a Pair generation

For the pair generation process $\emptyset \to 2\,A$ with growth rate coefficient γ_τ, the drift and the (squared) diffusion coefficients agree: $\alpha_\tau = \beta_\tau = 2\gamma_\tau$. As neither of them depends on $x(\tau)$, the Itô SDE is readily solved by

$$x(t) = x(t_0) + \int_{t_0}^{t} \mathrm{d}\tau 2\gamma_\tau + \int_{t_0}^{t} \mathrm{d}W(\tau)\sqrt{2\gamma_\tau}\,. \tag{258}$$

Upon introducing the (rather daunting) parameter

$$\eta_k := \Big\langle\!\Big\langle \mathrm{e}^{-\left(\int_{t_0}^{t} \mathrm{d}\tau\,\gamma_\tau + \int_{t_0}^{t} \mathrm{d}W(\tau)\sqrt{2\gamma_\tau}\right)} \left(\left(\int_{t_0}^{t} \mathrm{d}\tau\, 2\gamma_\tau \right)^{1/2} + \frac{\int_{t_0}^{t} \mathrm{d}W(\tau)\sqrt{2\gamma_\tau}}{\left(\int_{t_0}^{t} \mathrm{d}\tau\, 2\gamma_\tau \right)^{1/2}} \right)^{k} \Big\rangle\!\Big\rangle_W , \tag{259}$$

one can use the binomial theorem to write the marginalized distribution (256) as

$$|p(t, n|t_0)\rangle_{x(t_0)} = \frac{\mathrm{e}^{-\left(x(t_0) + \int_{t_0}^{t} \mathrm{d}\tau\,\gamma_\tau\right)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\int_{t_0}^{t} \mathrm{d}\tau\, 2\gamma_\tau \right)^{k/2} \eta_k\, x(t_0)^{n-k}\,. \tag{260}$$

If the rate coefficient γ is independent of time, one can show that the parameter η_k coincides with the k-th moment of a Gaussian distribution with zero mean and unit variance (i.e. $\eta_k = 0$ for odd k and $\eta_k = (k-1)!!$ for even k; note that the sum of two Gaussian random variables is again a Gaussian random variable, with its mean and variance following additively). This result is established most easily in the discrete-time approximation and by noting that the sum over independent

Gaussian random variables is again a Gaussian random variable. The mean and variance of the new variable are given as the sums of the individual means and variances. The integrals over the Gaussian increments $\Delta W_1, \Delta W_2, \ldots$ in (259) can thereby be replaced by a single integral over $W(t) = \Delta W_1 + \Delta W_2 + \ldots$ so that

$$\begin{aligned}
\eta_k &= \int_{-\infty}^{\infty} \mathrm{d}W(t)\, \frac{e^{-W(t)^2/2(t-t_0)}}{\sqrt{2\pi(t-t_0)}} \mathrm{e}^{-(\gamma(t-t_0)+\sqrt{2\gamma}W(t))} \left(\left(2\gamma(t-t_0)\right)^{1/2} + \frac{W(t)}{(t-t_0)^{1/2}} \right)^k \\
&= \int_{-\infty}^{\infty} \mathrm{d}W(t)\, \frac{e^{-(W(t)+\sqrt{2\gamma}(t-t_0))^2/2(t-t_0)}}{\sqrt{2\pi(t-t_0)}} \left(\left(2\gamma(t-t_0)\right)^{1/2} + \frac{W(t)}{(t-t_0)^{1/2}} \right)^k \qquad (261) \\
&= \int_{-\infty}^{\infty} \mathrm{d}U(t)\, \frac{e^{-U(t)^2/2}}{\sqrt{2\pi}}\, U(t)^k \,. \qquad (262)
\end{aligned}$$

The marginalized distribution therefore follows as

$$|p(t,n|t_0)\rangle_{x(t_0)} = \mathrm{e}^{-(x(t_0)+\int_{t_0}^{t} \mathrm{d}\tau\, \gamma_\tau)} \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{\left(\int_{t_0}^{t} \mathrm{d}\tau\, \gamma_\tau\right)^k}{k!} \frac{(x(t_0))^{n-2k}}{(n-2k)!} \,. \qquad (263)$$

Here, $\lfloor n/2 \rfloor$ represents the integral part of $n/2$. The marginalized distribution (263) solves the master equation of the pair generation process and is initially of Poisson shape. For $x(t_0) = 0$, the distribution effectively keeps that shape, although only on the set of all even numbers (the reaction $\emptyset \to 2A$ cannot create an odd number of particles when starting out from zero particles). Using Mathematica by Wolfram Research, we verified that the distribution also applies when γ_τ depends on time. The evolution of the distribution is shown in figure 7 for the rate coefficient $\gamma_\tau = 1/(1+\sqrt{\tau})$. We refrain from computing the conditional distribution from the marginalized distribution because the computation is unwieldy and does not shed any more light on the path integral approach.

4.5 b Diffusion on networks

The backward path integral can also be used to solve spatial processes. In fact, stochastic path integrals have been most useful in the study of such processes. If particles engaging in a chemical reaction can also diffuse in space, their density may evolve in ways that are not expected from the well-mixed, non-spatial limit. That is, for example, the case for particles that annihilate one another in the reaction $2A \to \emptyset$ while diffusing along a one-dimensional line. From the well-mixed limit, one would expect that the particle density decays asymptotically as t^{-1} with time (see section 7), but instead it decays as $t^{-1/2}$ [135, 296, 405]. The reason behind this surprising decay law is the rapid condensation of the system into a state in which isolated particles are separated by large voids. From that point on, further decay

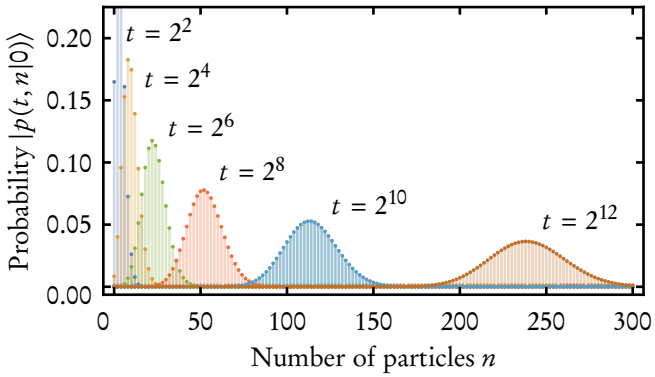


Figure 7 Evolution of the marginalized distribution $|p(t,n|t_0)\rangle$ in (263), which solves the pair generation process $\emptyset \to 2A$ (with $t_0 := 0$). The initial mean of the distribution was chosen as $x(t_0) = 0$ and the rate coefficient of the reaction as $\gamma_\tau = 1/(1+\sqrt{\tau})$.

of the particle density requires that two particles first find each other through diffusion. Therefore, the process is "diffusion-limited" and exhibits a slower decay of the particles (see e.g. [406] on the related "first-passage" problem). Path integral representations of the master equation, combined with renormalization group techniques, have been successfully applied to the computation of decay laws in spatially extended systems, both regarding the (universal) exponents and the prefactors of these laws [135–137, 296, 308, 407]. For a broader discussion of systems exhibiting a transition into an absorbing state see [73]. Before turning to the particle density and, more generally, to observables of the particle number, we now show how the backward path integral helps in computing the full probability distribution of a spatial process.

As a first step, we consider a pure diffusion process with particles hopping between neighbouring nodes of a network $\mathbb{L}$. For now, the topology of the network may be arbitrary, being either random or regular. In the next section, the network topology will be chosen as a regular lattice. In the limit of a vanishing lattice spacing (and an infinite number of lattice sites), a "field theory" will be obtained. In section 4.5 d, the particles will also be allowed to decay.

The configuration of particles on the network may be represented by the vector $\boldsymbol{n} \in \mathbb{N}_0^{|\mathbb{L}|}$, with $|\mathbb{L}|$ being the total number of network nodes. The configuration changes whenever a particle hops from some node $i \in \mathbb{L}$ to a neighbouring node $j \in \mathsf{N}_i \subset \mathbb{L}$. The probability $p(\tau, \boldsymbol{n}|t_0, \boldsymbol{n}_0)$ of finding the system in configuration $\boldsymbol{n}$ then obeys the master equation

$$\partial_\tau p(\tau, \boldsymbol{n}|\cdot) = \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} \sum_{j\in\mathsf{N}_i} \left[(n_i+1)p(\tau, \boldsymbol{n}+\hat{\boldsymbol{e}}_i-\hat{\boldsymbol{e}}_j|\cdot) - n_i p(\tau,\boldsymbol{n}|\cdot)\right]. \tag{264}$$

Here, $\hat{\boldsymbol{e}}_i$ represents a unit vector that points in direction $i \in \mathbb{L}$. Moreover, $\varepsilon_{\tau,i}$ acts as a hopping rate and may depend both on time and on the node from which a particle departs. Apparently, the above master equation has the same structure as the chemical master equation (27). It may therefore be cast into a linear PDE by extending the operators and basis functions from section 3.2 b to multiple variables. In particular, we extend the Poisson basis function to

$$|\boldsymbol{n}\rangle_{\boldsymbol{x}} := \prod_{i\in\mathbb{L}} \frac{x_i^{n_i}\mathrm{e}^{-x_i}}{n_i!} \tag{265}$$

and employ the creation and annihilation operators

$$\boldsymbol{c} := \boldsymbol{x} \text{ and } \boldsymbol{a} := \nabla + \mathbf{1}\,. \tag{266}$$

Next, we derive the adjoint transition operator from the backward analogue of the master equation (264). Using the transition rate

$$w_\tau(\boldsymbol{n},\boldsymbol{m}) = \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} \sum_{j\in\mathsf{N}_i} m_i \delta_{\boldsymbol{n}+\hat{\boldsymbol{e}}_i-\hat{\boldsymbol{e}}_j,\boldsymbol{m}}\,, \tag{267}$$

the backward master equation follows from (15) as

$$\partial_{-\tau} p(t,\boldsymbol{n}|\tau,\boldsymbol{n}_0) = \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} \sum_{j\in\mathsf{N}_i} n_{0,i}\big(p(\cdot|\tau,\boldsymbol{n}_0 - \hat{\boldsymbol{e}}_i + \hat{\boldsymbol{e}}_j) - p(\cdot|\tau,\boldsymbol{n}_0)\big)\,. \tag{268}$$

Therefore, the time evolution of the marginalized distribution

$$|p(t,\boldsymbol{n}|\tau)\rangle := \sum_{\boldsymbol{n}_0\in\mathbb{N}_0^{|\mathbb{L}|}} p(t,\boldsymbol{n}|\tau,\boldsymbol{n}_0)|\boldsymbol{n}_0\rangle \tag{269}$$

is described by the flow equation

$$\begin{aligned}\partial_{-\tau}|p(t,n|\tau)\rangle_{\boldsymbol{x}} = \sum_{\boldsymbol{n}_0\in\mathbb{N}_0^{|\mathbb{L}|}} \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} \sum_{j\in\mathsf{N}_i} p(\cdot|\tau,\boldsymbol{n}_0)\Big((n_{0,i}+1)\frac{x_i^{n_{0,i}+1}\mathrm{e}^{-x_i}}{(n_{0,i}+1)!}\frac{x_j^{n_{0,j}-1}\mathrm{e}^{-x_j}}{(n_{0,j}-1)!}\\ - n_{0,i}\frac{x_i^{n_{0,i}}\mathrm{e}^{-x_i}}{n_{0,i}!}\frac{x_j^{n_{0,j}}\mathrm{e}^{-x_j}}{n_{0,j}!}\Big)\prod_{k\neq i,j}\frac{x_k^{n_{0,k}}\mathrm{e}^{-x_k}}{n_{0,k}!} \qquad (270)\\ = \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} \sum_{j\in\mathsf{N}_i}\Big(x_i(\partial_{x_j}+1) - x_i(\partial_{x_i}+1)\Big)|p(\cdot|\tau)\rangle_{\boldsymbol{x}} \qquad (271)\end{aligned}$$

$$= \sum_{i\in\mathbb{L}} \varepsilon_{\tau,i} x_i \sum_{j\in\mathsf{N}_i} \big(\partial_{x_j} - \partial_{x_i}\big) |p(\cdot|\tau)\rangle_{\boldsymbol{x}} \tag{272}$$

$$= \sum_{i\in\mathbb{L}} \Big(\sum_{j\in\mathsf{N}_j} \big(\varepsilon_{\tau,j} x_j - \varepsilon_{\tau,i} x_i\big)\Big) \partial_{x_i} |p(\cdot|\tau)\rangle_{\boldsymbol{x}} \tag{273}$$

$$= \sum_{i\in\mathbb{L}} (\underline{\Delta}\varepsilon_{\tau,i} x_i) \partial_{x_i} |p(\cdot|\tau)\rangle_{\boldsymbol{x}} = \widetilde{\mathcal{Q}}_\tau^\dagger(\boldsymbol{x}, \nabla) |p(\cdot|\tau)\rangle_{\boldsymbol{x}} \,. \tag{274}$$

Hence, the marginalized distribution evolves via the flow equation $\partial_{-\tau}|p(t,n|\tau)\rangle = \widetilde{\mathcal{Q}}_\tau^\dagger |p(t,n|\tau)\rangle$ with the adjoint transition operator

$$\widetilde{\mathcal{Q}}_\tau^\dagger(\boldsymbol{x}, \nabla) = \sum_{i\in\mathbb{L}} (\underline{\Delta}\varepsilon_{\tau,i} x_i) \partial_{x_i} \,. \tag{275}$$

Here we introduced the discrete Laplace operator $\underline{\Delta} f_i := \sum_{j\in\mathsf{N}_i} (f_j - f_i)$, which acts both on $\varepsilon_{\tau,i}$ and on x_i.

By following the steps in section 4.3, one finds that the marginalized distribution is given by the multivariate Poisson distribution

$$|p(t, \boldsymbol{n}|t_0)\rangle_{\boldsymbol{x}(t_0)} = \prod_{i\in\mathbb{L}} \frac{x_i(t)^{n_i} \mathrm{e}^{-x_i(t)}}{n_i!} \,. \tag{276}$$

Its mean $\boldsymbol{x}(t)$ solves the discrete diffusion equation

$$\partial_\tau x_i = \sum_{j\in\mathsf{N}_i} \big(\varepsilon_{\tau,j} x_j - \varepsilon_{\tau,i} x_i\big) = \underline{\Delta}\varepsilon_{\tau,i} x_i \,, \tag{277}$$

with the initial condition $\boldsymbol{x}(t_0)$. Let us note that the marginalized distribution solves the master equation (264), but with the initial number of particles being Poisson distributed locally.

The conditional distribution $p(t, \boldsymbol{n}|t_0, \boldsymbol{n}_0)$ follows from the marginalized distribution (276) by applying the functional

$$\langle \boldsymbol{n}_0|_{\boldsymbol{x}(t_0)} f = \Big[\prod_{i\in\mathbb{L}} \big(\partial_{x_i(t_0)} + 1\big)^{n_{0,i}} \Big] f(\boldsymbol{x}(t_0)) \Big|_{\boldsymbol{x}(t_0)=0} \tag{278}$$

to it (cf. (160) with (164)). The evaluation requires a prior solution of the discrete diffusion equation (277). This equation can be written in matrix form as $\partial_\tau \boldsymbol{x} = M_\tau \boldsymbol{x}$ with $M_{\tau,ij} = (A_{ij} - |\mathsf{N}_i|\delta_{ij})\varepsilon_{\tau,j}$. Here, A represents the symmetric adjacency matrix of the network and $|\mathsf{N}_i|$ represents the number of neighbours of node $i \in \mathbb{L}$. The matrix equation can in principle be solved through a Magnus expansion. The solution has the generic form $\boldsymbol{x}(\tau) = G(\tau|t_0)\boldsymbol{x}(t_0)$ with the "propagator"

G solving $\partial_\tau G(\tau|t_0) = M_\tau G(\tau|t_0)$. We write the elements of the propagator as $G(\tau, i|t_0, j)$. Its flow starts out from $G(t_0|t_0) = \mathbb{1}$. For later purposes, let us note the time-reversal property in terms of the matrix inverse $G(t|\tau)^{-1} = G(\tau|t)$ and also the conservation law $\mathbf{1}^\intercal G(\tau|t_0) = \mathbf{1}^\intercal$.

It proves insightful to consider the master equation (264) of the multi-particle hopping process for the random walk of a single particle on the one-dimensional lattice $\mathbb{L} = \mathbb{Z}$, which we already considered in sections 2.2 a and 3.2 a (with symmetric hopping rates $\varepsilon_{\tau,i} = l_\tau = r_\tau$). The presence of only a single particle can be enforced by choosing $\boldsymbol{n}_0$ as $n_{0,k} = 1$ for one $k \in \mathbb{Z}$ and $n_{0,j} = 0$ for all $j \neq k$. By following the above steps, one eventually finds that the master equation (264) is solved by the conditional probability distribution

$$p(t, \boldsymbol{n}|t_0, \boldsymbol{n}_0) = \sum_{i\in\mathbb{Z}} \Big(\prod_{j\neq i} \delta_{n_j,0}\Big)\delta_{n_i,1} G(t, i|t_0, k)\,, \tag{279}$$

with the propagator G solving the master equation (58) of the simple random walk. Hence, the propagator evaluates to a Skellam distribution.

4.5 c Diffusion in continuous space

To make the transition to a field theory, one may specify the network as the the d-dimensional lattice $\mathbb{L} = (l\mathbb{Z})^d$ with the lattice spacing $l > 0$ going to zero. In order to take this limit, we define the variable $x(\tau, \boldsymbol{r}) := x_{\boldsymbol{r}}(\tau)$, the rescaled Laplace operator $\Delta := \underline{\Delta}/l^2$, and the rescaled diffusion coefficient $D_\tau(\boldsymbol{r}) := l^2\varepsilon_{\tau,\boldsymbol{r}}$ for $\boldsymbol{r} \in \mathbb{L}$. Assuming that these definitions can be continued to $\boldsymbol{r} \in \mathbb{R}^d$, the discrete diffusion equation (277) becomes a PDE for the "field" $x(\tau, \boldsymbol{r})$ in the limit $l \to 0$, namely $\partial_\tau x(\tau, \boldsymbol{r}) = \Delta(D_\tau(\boldsymbol{r})x(\tau, \boldsymbol{r}))$. Here, Δ represents the ordinary Laplace operator.[14] The solution of the PDE acts as the mean of the multivariate Poisson distribution (276) whose extension to $\boldsymbol{r} \in \mathbb{R}$ is, however, not quite obvious. If the diffusion coefficient is homogeneous in space, the PDE is solved by

$$x(\tau, \boldsymbol{r}) = \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{r}_0\, G(\tau, \boldsymbol{r}|t_0, \boldsymbol{r}_0) x(t_0, \boldsymbol{r}_0)\,, \tag{281}$$

14 With $\mathsf{N}_{\boldsymbol{r}}$ denoting the neighbouring lattice site of $\boldsymbol{r} \in (l\mathbb{Z})^d$, the ordinary Laplace operator follows as the limit of a finite-difference approximation, i.e. as

$$\sum_{\boldsymbol{r}'\in\mathsf{N}_{\boldsymbol{r}}} \frac{f_{\boldsymbol{r}'} - f_{\boldsymbol{r}}}{l^2} \to \sum_{i=1}^{d} \partial_{r_i}^2 f(\boldsymbol{r}) = \Delta f(\boldsymbol{r})\,. \tag{280}$$

with the Gaussian kernel

$$G(\tau, \boldsymbol{r}|\tau', \boldsymbol{r}') = \frac{\mathrm{e}^{-(\boldsymbol{r}-\boldsymbol{r}')^2/4\int_{\tau'}^{\tau} \mathrm{d}s\, D_s}}{\sqrt{4\pi \int_{\tau'}^{\tau} \mathrm{d}s\, D_s}}\,. \tag{282}$$

The action (214) can also be extended into continuous space. For that purpose, one may rescale the second path integral variable as $q(\tau, \boldsymbol{r}) \coloneqq l^{-d} q_{\boldsymbol{r}}(\tau)$ so that in the limit $l \to 0$,

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau \sum_{\boldsymbol{r}\in\mathbb{L}} \left(\mathrm{i} q_{\boldsymbol{r}} \left(\partial_\tau x_{\boldsymbol{r}} - \underline{\Delta}\varepsilon_{\tau,\boldsymbol{r}} x_{\boldsymbol{r}}\right)\right) \to \int_{t_0}^{t} \mathrm{d}\tau \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{r}\, \mathrm{i} q \left(\partial_\tau x - \Delta D_\tau x\right). \tag{283}$$

The above rescaling of x and q is not unique and depends on the problem at hand. Instead of dividing $q_{\boldsymbol{r}}$ by the volume factor l^d, this factor is sometimes employed to cast $x_{\boldsymbol{r}}$ and the particle number $n_{\boldsymbol{r}}$ into densities [135]. In the study of branching and annihilating random walks with an odd number of offspring, yet another kind of rescaling may bring the action into a Reggeon field theory [75, 408–410] like form [101]. The critical behaviour of these random walks falls into the universality class of directed percolation (DP) [101, 137]. Information on this universality class can be found in the book [73], as well as in the original articles of Janssen and Grassberger, which established the extensive scope of the DP class [411, 412].

4.5 d Diffusion and decay

Let us exemplify how one can accommodate a non-vanishing perturbation operator $\mathcal{P}_\tau^\dagger$ in the adjoint transition operator (218). As in the section before the last, we consider a system of particles that are hopping between the nodes of an arbitrary network $\mathbb{L}$. The corresponding diffusive transition operator (275) shall specify the drift coefficient $\alpha_{\tau,i}(\boldsymbol{x}) = \underline{\Delta}\varepsilon_{\tau,i} x_i$ in the multivariate extension of (218). The coefficient $\beta_{\tau,ij}(\boldsymbol{x})$ is zero. Besides allowing for diffusion, we now allow the particles to decay in the linear reaction $A \to \emptyset$. For the sake of brevity, we assume that the corresponding decay rate coefficient μ_τ is spatially homogeneous. This assumption can be relaxed but the equations that result cannot be written in matrix form and involve many indices. We treat the adjoint transition operator of the decay process as the perturbation, which reads, according to the flow equation (166),

$$\mathcal{P}_\tau^\dagger(\boldsymbol{x}, \nabla) = -\mu_\tau \sum_{i\in\mathbb{L}} x_i \partial_{x_i}\,. \tag{284}$$

The combined process could also be solved directly by treating the decay process as part of the drift coefficient $\alpha_{\tau,i}(\boldsymbol{x})$. For pedagogic reasons, however, we wish to outline a perturbative solution using Feynman diagrams.

The first step of the derivation is to solve the differential equation (223) for $\boldsymbol{x}$, which now reads $\partial_\tau x_i = \underline{\Delta}\varepsilon_{\tau,i} x_i + X_i(\tau)$. The homogeneous solution of this equation is given by $\boldsymbol{x}_h(\tau) \coloneqq G(\tau|t_0)\boldsymbol{x}(t_0)$, with G being the propagator from the end of section 4.5 b. The solution of the full equation can be written as

$$\boldsymbol{x}(\tau) = \boldsymbol{x}_h(\tau) + \int_{t_0}^{\tau} \mathrm{d}\tau' G(\tau|\tau')\boldsymbol{X}(\tau')\,. \tag{285}$$

As the next step, the evaluation of the marginalized distribution (219) requires us to compute the $(\boldsymbol{Q},\boldsymbol{X})$-generating functional (221) for $\boldsymbol{Q} = \boldsymbol{X} = \mathbb{0}$. By performing a series expansion of its leading exponential, this function can be written as

$$\mathcal{Z}_{\mathbb{0},\mathbb{0}} = \sum_{k=\mathbb{0}}^{\infty}\sum_{l=\mathbb{0}}^{k} \frac{1}{l!}\Big(\mathrm{i}\boldsymbol{q}_N \cdot \frac{\delta}{\delta \boldsymbol{Q}(t)}\Big)^l \int_{t_0}^{t} \mathrm{d}\tau_{k-l}\, \mathcal{P}^\dagger_{\tau_{k-l}} \cdots \int_{t_0}^{\tau_2} \mathrm{d}\tau_1\, \mathcal{P}^\dagger_{\tau_1}\, \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}}\big|_{\boldsymbol{Q}=\boldsymbol{X}=\mathbb{0}}\,,$$

with $\mathcal{P}^\dagger_\tau = \frac{\delta}{\delta \boldsymbol{X}(\tau)} \cdot (-\mu_\tau)\frac{\delta}{\delta \boldsymbol{Q}(\tau)}$. To evaluate $\mathcal{Z}_{\mathbb{0},\mathbb{0}}$, it helps to note that $\ln \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}} = \int_{t_0}^{t} \mathrm{d}\tau\, \boldsymbol{Q}\cdot\boldsymbol{x}$, from which it follows that for $t_0 \le \tau \le t$:

$$\frac{\delta \ln \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}}}{\delta \boldsymbol{Q}(\tau)}\bigg|_{\boldsymbol{Q}=\boldsymbol{X}=\mathbb{0}} = \boldsymbol{x}_h(\tau) \text{ and} \tag{286}$$

$$\frac{\delta^2 \ln \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}}}{\delta Q_i(\tau)\delta X_j(\tau')}\bigg|_{\boldsymbol{Q}=\boldsymbol{X}=\mathbb{0}} = G(\tau,i|\tau',j)\Theta(\tau-\tau')\,. \tag{287}$$

The Heaviside step function is used with $\Theta(\mathbb{0}) \coloneqq \mathbb{0}$ to take into account that $\boldsymbol{x}_h(\tau)$ depends on $\boldsymbol{X}(\tau')$ only for $\tau' < \tau$ (see appendix D). All other derivatives of $\ln \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}}$, as well as itself, vanish for $\boldsymbol{Q} = \boldsymbol{X} = \mathbb{0}$.

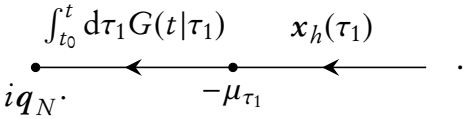
Upon inserting the perturbation (284) into (286), one observes that every summand of the expansion can be written in terms of a combination of the derivatives (286) and (287), and a terminal factor $\mathrm{i}\boldsymbol{q}_N$. For example, the summand with $k = 2$ and $l = 1$ reads

$$\mathrm{i}\boldsymbol{q}_N \cdot \frac{\delta}{\delta \boldsymbol{Q}(t)} \int_{t_0}^{t} \mathrm{d}\tau_1 \frac{\delta}{\delta \boldsymbol{X}(\tau_1)}(-\mu_{\tau_1})\frac{\delta}{\delta \boldsymbol{Q}(\tau_1)}\, \mathcal{Z}^{\mathbb{0}}_{\boldsymbol{Q},\boldsymbol{X}}\,, \tag{288}$$

with $\boldsymbol{Q} = \boldsymbol{X} = 0$ being taken after the evaluation of the functional derivatives. The evaluation of these derivatives results in

$$\mathrm{i}\boldsymbol{q}_N \cdot \int_{t_0}^{t} \mathrm{d}\tau_1 \, G(t|\tau_1)(-\mu_{\tau_1})\boldsymbol{x}_h(\tau_1)\,. \tag{289}$$

One can represent this expression graphically by a Feynman diagram according to the following rules. First, every diagram ends in a sink •←, which contributes the factor $\mathrm{i}\boldsymbol{q}_N$. The incoming line, or "leg", of the sink represents the derivative $\frac{\delta}{\delta Q(t)}$. This leg may either be left dangling, resulting in a factor $\boldsymbol{x}_h(t)$ according to (286), or it may be "contracted" with the outgoing line of a vertex ←•←. The two legs of this vertex reflect the two derivatives in the perturbation $\mathcal{P}_\tau^\dagger = \frac{\delta}{\delta X(\tau)} \cdot (-\mu_\tau)\frac{\delta}{\delta Q(\tau)}$. According to (287), the contraction results in a propagator $G(t, i|\tau, j)$ and the vertex itself contributes a factor $-\mu_\tau$. The incoming leg of the vertex may again be left dangling or it may be connected to another vertex. Therefore, each Feynman diagram is a straight line for the linear decay process. The expression (289) can therefore be represented graphically by

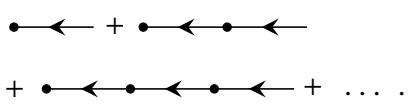


(290)

Dangling outgoing lines are not permitted because the derivative $\delta \ln \mathcal{Z}^0_{Q,X}/\delta \boldsymbol{X}(\tau)$ vanishes for $\boldsymbol{Q} = \boldsymbol{X} = 0$.

For more complex, non-linear processes, the Feynman diagrams may contain multiple kinds of vertices, each representing an individual summand of $\mathcal{P}_\tau^\dagger$. If there exist vertices with more than two legs, the diagrams may exhibit internal loops (see section 6.4). It is then usually impossible to evaluate the full perturbation series and it needs to be truncated at a certain order in the number of loops. Further information about these techniques, and about how renormalization group theory comes into play, is provided, for example, by the book of Täuber [75]. Information on a non-perturbative renormalization group technique can be found in [413].

For the simple diffusion process with decay, all the summands in the expansion above (286) are readily cast into Feynman diagrams. However, it turns out that individual summands of the expansion may be associated to multiple diagrams that are not connected to one another. Furthermore, diagrams that represent summands of lower order keep reappearing as unconnected components of summands of higher order. Hence, there appears to be redundant information involved. This redundancy is removed by a classical theorem from diagrammatic analysis. This theorem states that the logarithm $\ln \mathcal{Z}_{0,0}$ is given by the sum of only the connected diagrams [414], i.e. by

$\ln \mathcal{Z}_{0,0} =$  (291)

This sum can be evaluated with the help of $G(t|\tau)G(\tau|t_0) = G(t|t_0)$ as

$$\ln \mathcal{Z}_{0,0} = \sum_{k=0}^{\infty} \mathrm{i}\boldsymbol{q}_N \cdot \int_{t_0}^{t} \mathrm{d}\tau_k \, G(t|\tau_k)(-\mu_{\tau_k}) \tag{292}$$

$$\cdots \int_{t_0}^{\tau_2} \mathrm{d}\tau_1 \, G(\tau_2|\tau_1)(-\mu_{\tau_1})\boldsymbol{x}_h(\tau_1)$$

$$= \mathrm{i}\boldsymbol{q}_N \cdot \boldsymbol{x}_h(t)\mathrm{e}^{-\int_{t_0}^{t} \mathrm{d}\tau \, \mu_\tau} \, . \tag{293}$$

After inserting this expression into the marginalized distribution (219), one recovers the multivariate Poisson distribution (276). Its mean, however, has now acquired the pre-factor $\mathrm{e}^{-\int_{t_0}^{t} \mathrm{d}\tau \, \mu_\tau}$, reflecting the decay of the particles.

4.6 Résumé

In the present section, we introduced the novel backward path integral representation

$$|p(t,n|t_0)\rangle = \oint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \, |n\rangle_t \tag{294}$$

of the marginalized distribution (cf. section 3)

$$|p(t,n|t_0)\rangle = \sum_{n_0} p(t,n|t_0,n_0)|n_0\rangle_{t_0} \, . \tag{295}$$

When the basis function $|n\rangle_x$ is chosen as a probability distribution (e.g. $|n\rangle_x = \frac{x^n \mathrm{e}^{-x}}{n!}$ for all $n \in \mathbb{N}_0$), the backward path integral represents a true probability distribution: the marginalized distribution $|p(t,n|t_0)\rangle$. This distribution solves the forward master equation (11) for the initial condition $|p(t_0,n|t_0)\rangle = |n\rangle$ and it transforms into the conditional probability distribution as $p(t,n|\tau,n_0) = \langle n_0|p(t,n|\tau)\rangle$. In section 4.3, we showed how the backward path integral (294) can be expressed in terms of an average over the paths of an Itô stochastic differential equation. This method provided the exact solutions of various elementary stochastic processes, including the simple growth, the linear decay, and the pair generation processes. Moreover, we showed how the path integral can be evaluated perturbatively using Feynman diagrams for a process with diffusion and linear particle

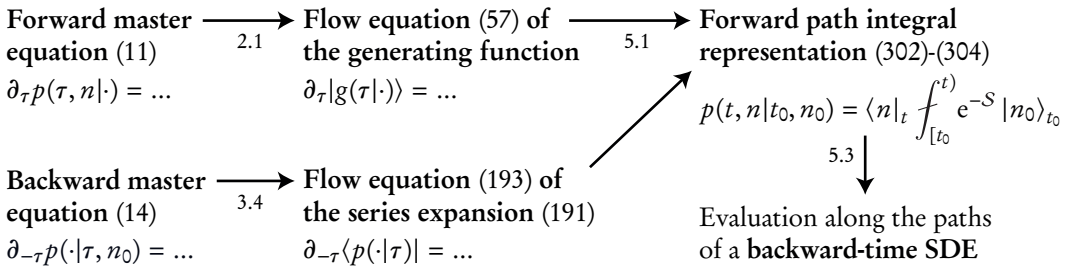


Figure 8 Outline of the derivation of the forward path integral representation and of its evaluation in terms of an average over the paths of a backward-time SDE.

decay. We hope that the backward path integral (294) will prove useful in the study of reaction-diffusion processes. Thus far, the critical behaviour of such processes could only be approached via path integral representations of averaged observables or of the generating function [74, 415]. In section 6, we show how the former representation readily follows from the backward path integral (294) upon summing the marginalized distribution over an observables $A(n)$. The corresponding representation is commonly applied in the study of diffusion-limited reactions (e.g. [74, 101, 135–137], but we here show how it can be freed of some of its quantum mechanical ballast (such as "second-quantized" or "normal-ordered" observables and coherent states). Besides, we showed in section 4.4 how one can derive a path integral representation of processes with continuous state spaces whose (backward) master equations admit a Kramers-Moyal expansion. Provided that this expansion stops at the level of a diffusion approximation, one recovers a classic path integral representation of the (backward) Fokker-Planck equation and also the Feynman-Kac formula (3). Moreover, the representation can be rewritten in terms of an Onsager-Machlup function and, for diffusive Brownian motion, it simplifies to the path integral of Wiener.

5 The forward path integral representation

Thus far, we have focused on the backward path integral (204). This integral will be used again in section 6 to derive a path integral representation of averaged observables. Moreover, we use it in section 7.4 to approximate the binary annihilation reaction $2A \to \emptyset$. In the present section, however, we shift our focus to the forward path integral (203). Its derivation proceeds analogously to the derivation of the backward solution, so we keep it brief.

5.1 Derivation

The forward path integral can be derived both from the flow equation (57) obeyed by the generating function or from the flow equation (192) obeyed by the series (191) (cf. figure 8). As it is more convenient to work with functions than with functionals, we use the flow equation of the generating function for this purpose. The derivation parallels a derivation of Elgart and Kamenev [40]. As in section 4.1, we first split the time interval $[t_0, t]$ into N pieces $t_0 \leq t_1 \leq \ldots \leq t_N := t$ of length Δt. Over a sufficiently small interval Δt, the flow equation (57) is then solved by

$$|g(t|t_0, n_0)\rangle_{q_N} = \mathcal{L}_{t_{N-1}}(q_N, \partial_{q_N})|g(t_{N-1}|t_0, n_0)\rangle_{q_N}\,, \tag{296}$$

with the generator $\mathcal{L}_\tau := 1 + \widetilde{\mathcal{Q}}_\tau \Delta t + O((\Delta t)^2)$. After inserting the integral form of a Dirac delta between $\mathcal{L}$ and $|g\rangle$, the right-hand side of the equation reads

$$\mathcal{L}_{t_{N-1}}(q_N, \partial_{q_N}) \int_{\mathbb{R}^2} \frac{\mathrm{d}q_{N-1}\,\mathrm{d}x_{N-1}}{2\pi} \mathrm{e}^{-\mathrm{i}x_{N-1}(q_{N-1}-q_N)}\, |g(t_{N-1}|\cdot)\rangle_{q_{N-1}}\,. \tag{297}$$

Assuming that the transition operator $\widetilde{\mathcal{Q}}_\tau$, and therefore also $\mathcal{L}_\tau$, are normal-ordered with respect to q and ∂_q, we may replace ∂_{q_N} by $\mathrm{i}x_{N-1}$ and interchange $\mathcal{L}_{t_{N-1}}$ with the exponential. This procedure is repeated N times before invoking the exponentiation $\mathcal{L}_\tau = \exp(\widetilde{\mathcal{Q}}_\tau \Delta t)$. Using the inverse transformation (51) and the initial condition $|g(t_0|t_0, n_0)\rangle = |n_0\rangle$, one obtains the discrete-time path integral representation

$$p(t, n|t_0, n_0) = \langle n|_{t,q_N} |g(t|t_0, n_0)\rangle_{q_N} \tag{298}$$

$$\text{with } |g(t|t_0, n_0)\rangle_{q_N} = \fint_0^{N-1} \mathrm{e}^{-\mathcal{S}_N}\, |n_0\rangle_{t_0,q_0} \text{ and} \tag{299}$$

$$\mathcal{S}_N := \sum_{j=0}^{N-1} \Delta t \left(\mathrm{i}x_j \frac{q_j - q_{j+1}}{\Delta t} - \widetilde{\mathcal{Q}}_{t_j}(q_{j+1}, \mathrm{i}x_j)\right). \tag{300}$$

Here we again employed the abbreviation (209), i.e.

$$\fint_k^l = \prod_{j=k}^{l} \int_{\mathbb{R}^2} \frac{\mathrm{d}q_j\,\mathrm{d}x_j}{2\pi}\,. \tag{301}$$

Moreover, the initial condition $p(t_0, n|t_0, n_0) = \delta_{n,n_0}$ is again trivially fulfilled for $N = 0$.

Upon taking the continuous-time limit $N \to \infty$, so that $\Delta t \to 0$, the forward path integral representation of the master equation now follows as

$$p(t, n|t_0, n_0) = \langle n|_{t,q(t)} |g(t|t_0, n_0)\rangle_{q(t)} \tag{302}$$

$$\text{with } |g(t|t_0, n_0)\rangle_{q(t)} = \fint_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}} \, |n_0\rangle_{t_0,q(t_0)} \tag{303}$$

$$\text{and } \mathcal{S} := \int_{t_0}^{t} \mathrm{d}\tau \left[\mathrm{i}x\partial_{-\tau} q - \widetilde{\mathcal{Q}}_\tau(q, \mathrm{i}x)\right]. \tag{304}$$

The limit $\fint_{[t_0}^{t)} := \lim_{N\to\infty} \int_0^{N-1}$ now involves integrations over $x(t_0)$ and $q(t_0)$, but not over $x(t)$ and $q(t)$.

5.2 Linear processes

The forward path integral (303) can, for example, be used to derive the generating function of the generic linear process $A \to l\,A$ with rate coefficient μ_τ (and $l \in \mathbb{N}_0$). For that purpose, we choose the basis function as $|n\rangle_{\tau,q} = q^n$ so that $|g\rangle$ coincides with the ordinary generating function. Since the basis function does not depend on time, (69) alone specifies the transition operator

$$\widetilde{\mathcal{Q}}_\tau(q, \partial_q) = \mu_\tau(q^l - q)\partial_q \tag{305}$$

of the flow equation $\partial_\tau |g(\tau|\cdot)\rangle = \widetilde{\mathcal{Q}}_\tau(q, \partial_q)|g(\tau|\cdot)\rangle$. Consequently, the action

$$\mathcal{S} = \int_{t_0}^{t} \mathrm{d}\tau \, \mathrm{i}x\big[\partial_{-\tau} q - \mu_\tau(q^l - q)\big] \tag{306}$$

is linear in $\mathrm{i}x$. To evaluate the path integral (303), it helps to reconsider the discrete-time approximation

$$\mathcal{S}_N = \sum_{j=0}^{N-1} \mathrm{i}x_j \Big(q_j - (q_{j+1} + \mu_{t_j}(q_{j+1}^l - q_{j+1})\Delta t)\Big). \tag{307}$$

of the action. Upon integrating over all the q_j-variables and taking the limit $\Delta t \to 0$, one obtains the generating function

$$|g(t|t_0, n_0)\rangle_{q(t)} = q(t_0)^{n_0} \tag{308}$$

with $q(\tau)$ solving $\partial_{-\tau} q = \mu_\tau (q^l - q)$. The unique real solution of this equation with final value $q(t)$ reads

$$q(\tau) = \frac{q(t)}{\left[q(t)^{l-1} + \mathrm{e}^{(l-1)\int_\tau^t \mathrm{d}s\, \mu_s} (1 - q(t)^{l-1})\right]^{1/(l-1)}} . \tag{309}$$

For the linear growth, or Yule-Furry [124, 416], process $A \to 2A$ ($l = 2$), the inverse transformation (302) casts the generating function into

$$p(t, n|t_0, n_0) = \left(\frac{1}{n!} \partial^n_{q(t)} |g(t|t_0, n_0)\rangle_{q(t)} \right)\Big|_{q(t)=0} \tag{310}$$

$$= \binom{n-1}{n-n_0} \left(\mathrm{e}^{-\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau}\right)^{n_0} \left(1 - \mathrm{e}^{-\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau}\right)^{n-n_0} \tag{311}$$

for $n > 0$ and into δ_{0,n_0} for $n = 0$. The above distribution has the form of a negative Binomial distribution with probability of success $\mathrm{e}^{-\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau}$, number of failures $n - n_0$, and number of successes n_0 [417]. The mean value $\mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau} n_0$ of the marginalized distribution (311) grows exponentially with time as long as $\mu_\tau > 0$, and so does its variance $(\mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau} - 1)\, \mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau} n_0$.

For the linear decay process $A \to \emptyset$ ($l = 0$), the inverse transformation in (310) instead recovers the Binomial distribution (169).

5.3 Feynman-Kac formula for jump processes

The above procedure can be generalized to more complex processes, just as we did in section 4.3 for the backward solution. Allowing for multiple types of particles or for spatial degrees of freedom on a lattice $\mathbb{L}$, the procedure applies to every process whose transition operator can be written as

$$\widetilde{\mathcal{Q}}_\tau(\boldsymbol{q}, \nabla_q) = \boldsymbol{\alpha}_\tau(\boldsymbol{q}) \cdot \nabla + \frac{1}{2} \sum_{k,l} \beta_{\tau,kl}(\boldsymbol{q}) \partial_{q_k} \partial_{q_l} + \mathcal{P}_\tau(\boldsymbol{q}, \nabla) . \tag{312}$$

The drift coefficient $\boldsymbol{\alpha}_\tau$ and the perturbation operator $\mathcal{P}_\tau$ may be arbitrary functions. For the diffusion coefficient β_τ, we require a matrix $\sqrt{\beta_\tau} \coloneqq \gamma_\tau$ fulfilling $\gamma_\tau \gamma_\tau^\intercal = \beta_\tau$. As shown in appendix E, the forward path integral (303) can then be rewritten as

$$|g(t|t_0, \boldsymbol{n}_0)\rangle_{\boldsymbol{q}(t)} = \int_{\mathbb{R}^{2|\mathbb{L}|}} \frac{\mathrm{d}\boldsymbol{q}_0\, \mathrm{d}\boldsymbol{x}_0}{(2\pi)^{|\mathbb{L}|}} \mathrm{e}^{-\mathrm{i}\boldsymbol{x}_0 \cdot \boldsymbol{q}_0} \mathcal{Z}_{0,0} |\boldsymbol{n}_0\rangle_{t_0, \boldsymbol{q}_0} \tag{313}$$

with the $(\boldsymbol{X}, \boldsymbol{Q})$-generating functional

$$\mathcal{Z}_{X,Q}(\boldsymbol{x}_0, \boldsymbol{q}(t)) = \mathrm{e}^{\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{P}_\tau\left(\frac{\delta}{\delta X}, \frac{\delta}{\delta Q}\right) + \frac{\delta}{\delta X(t_0)} \cdot \mathrm{i}\boldsymbol{x}_0}\, \mathcal{Z}^0_{X,Q} \tag{314}$$

$$\text{with } \mathcal{Z}^0_{X,Q}(\boldsymbol{q}(t)) = \left\langle\!\!\left\langle \mathrm{e}^{\int_{t_0}^{t} \mathrm{d}\tau \boldsymbol{X}\cdot\boldsymbol{q}} \right\rangle\!\!\right\rangle_{\boldsymbol{W}} . \tag{315}$$

The average proceeds over realizations of the Wiener process $\boldsymbol{W}(\tau)$. This process governs the evolution of $\boldsymbol{q}(\tau)$ through the "backward-time" Itô SDE

$$-\mathrm{d}\boldsymbol{q}(\tau) = (\boldsymbol{\alpha}_\tau(\boldsymbol{q}) + \boldsymbol{Q}(\tau))\,\mathrm{d}\tau + \sqrt{\boldsymbol{\beta}_\tau(\boldsymbol{q})}\,\mathrm{d}\boldsymbol{W}(\tau) . \tag{316}$$

with final value $\boldsymbol{q}(t)$. If the perturbation operator $\mathcal{P}_\tau$ does not vanish, a series expansion of the exponential in (314) can again be organized in terms of Feynman diagrams.

5.4 Intermezzo: The forward Kramers-Moyal expansion

The above procedure can be generalized to processes whose transition operator is of the form

$$\widetilde{\mathcal{Q}}_\tau(q, \partial_q) = \alpha_\tau(q)\partial_q + \frac{1}{2}\beta_\tau(q)\partial_q^2 + \mathcal{P}_\tau(q, \partial_q) . \tag{317}$$

The derivation proceeds analogously to section 4.3 and appendix D. The probability distribution is thereby expressed as an average over the paths of a stochastic differential equation proceeding backward in time. The merit of such a representation remains to be explored. In the following, we briefly outline how the procedure is applied to processes with continuous sample paths.

In section 4.4, we explained how the forward master equation (226) of a process with a continuous state space can be written in terms of the (forward) Kramers-Moyal expansion

$$\partial_\tau p(\tau, q|\cdot) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \partial_q^m \left[\mathcal{M}_\tau^{(m)}(q) p(\tau, q|\cdot)\right] \tag{318}$$

with initial condition $p(\tau, q|t_0, q_0) = \delta(q - q_0)$. Here we changed the letter from x to q to keep in line with the notation used in sections 5.1 and 5.2. If the Kramers-Moyal expansion stops with its second summand, it coincides with the (forward) Fokker-Planck equation

$$\partial_\tau p(\tau, q|\cdot) = -\partial_q\left[\mathcal{M}_\tau^{(1)}(q) p\right] + \frac{1}{2}\partial_q^2\left[\mathcal{M}_\tau^{(2)}(q) p\right] . \tag{319}$$

To apply the procedure from the previous section to this equation, it needs to be brought into the form $\partial_\tau p(\tau, q|\cdot) = \widetilde{\mathcal{Q}}_\tau(q, \partial_q) p(\tau, q|\cdot)$ with a normal-ordered transition operator $\widetilde{\mathcal{Q}}_\tau(q, \partial_q)$. It is readily established that this operator can be expressed in the form of (317) with the coefficients

$$\alpha_\tau(q) := -\mathcal{M}_\tau^{(1)} + \partial_q \mathcal{M}_\tau^{(2)}, \tag{320}$$

$$\beta_\tau(q) := \mathcal{M}_\tau^{(2)}, \text{ and} \tag{321}$$

$$\mathcal{P}_\tau(q) := -\partial_q \mathcal{M}_\tau^{(1)} + \frac{1}{2}\partial_q^2 \mathcal{M}_\tau^{(2)}. \tag{322}$$

The Fokker-Planck equation now has the same form as the flow equation obeyed by the generating function in section 5.1. Therefore, we can follow the steps in that section to represent the solution of the Fokker-Planck equation by the path integral

$$p(t, q|t_0, q_0) = \int_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}}\, \delta(q_0 - q(t_0))\big|_{q(t)=q} \tag{323}$$

$$\text{with } \mathcal{S} := \int_{t_0}^{t} \mathrm{d}\tau \left[\mathrm{i}x\partial_{-\tau} q - \widetilde{\mathcal{Q}}_\tau(q, \mathrm{i}x)\right]. \tag{324}$$

The evaluation of this path integral proceeds analogously to the derivation in section 5.2 and appendix D. In particular, one can rewrite the probability distribution as

$$p(t, q|t_0, q_0) = \left\langle\!\!\left\langle \mathrm{e}^{\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{P}_\tau(q(\tau))} \delta(q_0 - q(t_0)) \right\rangle\!\!\right\rangle_W, \tag{325}$$

with $q(\tau)$ solving the backward-time SDE

$$-\,\mathrm{d}q(\tau) = \alpha_\tau(q(\tau))\,\mathrm{d}\tau + \sqrt{\beta_\tau(q(\tau))}\,\mathrm{d}W(\tau). \tag{326}$$

The time evolution of this SDE starts out from the final value $q(t) = q$. In a discrete-time approximation, the SDE reads

$$q_j - q_{j+1} = \alpha_{t_j}(q_{j+1})\Delta t + \sqrt{\beta_{t_j}(q_{j+1})}\,\Delta W_j. \tag{327}$$

The increments ΔW_j are Gaussian distributed with mean 0 and variance Δt. Let us note that the two path integral representations (232) and (323) of the conditional probability distribution belong to an infinite class of representations [47, 109, 418] (see also section 4.4 e). We focus on the two representations that follow from the backward and forward Kramers-Moyal expansions via the step-by-step derivations in sections 4.1 and 5.1, respectively.

To exemplify the validity of the path integral (323) and of the representation (325), we consider a process with linear drift and no diffusion, i.e. a process with jump moments $\mathcal{M}_\tau^{(1)} = \mu q$ and $\mathcal{M}_\tau^{(2)} = 0$. For this process, the representation (325) evaluates to

$$p(t, q|t_0, q_0) = \mathrm{e}^{-\mu(t-t_0)}\delta(q_0 - q\,\mathrm{e}^{-\mu(t-t_0)}) = \delta(q - q_0\,\mathrm{e}^{\mu(t-t_0)})\,. \tag{328}$$

Thus, the leading exponential in (325) converted the argument of the Dirac delta from the solution of a final value problem to the solution of an initial value problem. It is, of course, no surprise that the probability distribution is given by a Dirac delta function because the process is purely deterministic.

Another simple process that can be solved with the help of (325) is the pure diffusion process with $\mathcal{M}_\tau^{(1)} = 0$ and $\mathcal{M}_\tau^{(2)} = D$. The derivation is performed most easily in the discrete-time approximation and results in the Wiener path integral

$$p(t, q|t_0, q_0) = \lim_{N\to\infty}\Big(\prod_{j=1}^{N}\int_{\mathbb{R}} \mathrm{d}\tilde{q}_{j-1}\Big)\Big(\prod_{j=1}^{N}\mathcal{G}_{0,D\Delta t}(\tilde{q}_{j-1} - \tilde{q}_j)\Big)\delta(\tilde{q}_0 - q_0)\,, \tag{329}$$

with $\tilde{q}_N := q$ and Gaussian distribution $\mathcal{G}_{\mu,\sigma^2}$ (cf. (244)). An evaluation of the convolutions of Gaussian distributions shows that the process is solved by $\mathcal{G}_{0,D(t-t_0)}(q - q_0)$. Further uses of the path integral representation (323) of the (forward) Fokker-Planck equation remain to be explored.

5.5 Résumé

Here we derived the forward path integral representation

$$|g(t|t_0, n_0)\rangle = \fint_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}}\,|n_0\rangle_{t_0} \tag{330}$$

of the probability generating function $|g(t|t_0, n_0)\rangle = \sum_n |n\rangle_t\, p(t, n|t_0, n_0)$. The conditional probability distribution is recovered from the generating function via the inverse transformation $p(t, n|t_0, n_0) = \langle n|g(t|t_0, n_0)\rangle$. The path integral representation of the generating function has, for example, been employed to compute rare event probabilities [40] by a method that we discuss in section 7. Most often, however, the representation has only served as an intermediate step in deriving a path integral representation of averaged observables. Such a representation is considered in the next section. In section 5.2, we showed how the forward path integral (330) can be evaluated along the paths of a differential equation proceeding backward in time. We thereby obtained the generating function of generic linear processes. Besides, we derived a novel path integral representation of processes

with continuous state spaces in section 5.4, based on the (forward) Kramers-Moyal expansion. The potential use of this representation remains to be explored.

6 Path integral representation of averaged observables

The backward and forward path integral representations of the conditional probability distribution provide a full characterization of a Markov process. Yet, an intuitive understanding of how a process evolves is often attained more easily by looking at the mean particle number $\langle n \rangle$ and at its variance $\langle (n - \langle n \rangle)^2 \rangle$. Although both of these averages can in principle be inferred from a given probability distribution, it often proves convenient to bypass the computation of the distribution and to focus directly on the observables. Thus, we now show how one can derive a path integral representation of the average $\langle A \rangle$ of an observable $A(n)$. The path integral representation applies to all processes that can be decomposed additively into reactions of the form $k\,A \to l\,A$ in a well-mixed, non-spatial environment. The extension of the path integral to multiple types of interacting particles and to processes with spatial degrees of freedom is straightforward. In the derivation, we assume that the number of particles in the system is initially Poisson distributed with mean $x(t_0)$. This assumption is common in the study of reaction-diffusion master equations and will allow us to focus on the marginalized distribution from section 3 instead of on the conditional probability distribution.

6.1 Derivation

Assuming that the number of particles is initially Poisson distributed with mean $x(t_0)$, the probability of finding n particles at time t is

$$p(t,n) = \sum_{n_0} p(t,n|t_0,n_0)p(t_0,n_0)\,. \tag{331}$$

Here we employ the initial (single-time) distribution

$$p(t_0,n_0) = \frac{x(t_0)^{n_0}\mathrm{e}^{-x(t_0)}}{n_0!}\,. \tag{332}$$

The average value of an observable $A(n)$ at time t therefore evaluates to

$$\langle A \rangle_{x(t_0)} = \sum_{n=0}^{\infty} A(n)p(t,n)\,. \tag{333}$$

Here we emphasize that the average depends on the mean of the initial Poisson distribution. As the single-time distribution coincides with the marginalized distribution $|p(t,n|t_0)\rangle$ in (144) for the Poisson basis function $|n_0\rangle_{x(t_0)} = \frac{x(t_0)^{n_0}\mathrm{e}^{-x(t_0)}}{n_0!}$, the above average can equivalently be written as

$$\langle A\rangle_{x(t_0)} = \sum_{n=0}^{\infty} A(n)|p(t,n|t_0)\rangle_{x(t_0)}\,. \tag{334}$$

The path integral representation of $\langle A\rangle_{x(t_0)}$ then follows directly from the backward path integral representation (213) of the marginalized distribution as

$$\langle A\rangle_{x(t_0)} = \int_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger}\, \langle\!\langle A\rangle\!\rangle_{x(t)} \tag{335}$$

$$\text{with } \langle\!\langle A\rangle\!\rangle_x := \sum_{n=0}^{\infty} \frac{x^n \mathrm{e}^{-x}}{n!} A(n)\,. \tag{336}$$

With the adjoint transition operator $\mathcal{Q}_\tau^\dagger(c,a) = \gamma_\tau c^k(a^l - a^k)$ of the reaction $k\,A \to l\,A$ (cf. section 3.2 b), the action $\mathcal{S}^\dagger$ in (214) reads

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau \left[\mathrm{i}q\partial_\tau x - \mathcal{Q}_\tau^\dagger(x, \mathrm{i}q + 1)\right]. \tag{337}$$

The "+1" in the transition operator is called the "Doi-shift" [96]. In our above derivation, this shift followed from choosing a Poisson distribution as the basis function. The unshifted version of the path integral can be derived by choosing the basis function of the marginalized distribution (144) as $|n_0\rangle_{x(t_0)} = \frac{x(t_0)^{n_0}}{n_0!}$, turning the average (334) into

$$\langle A\rangle_{x(t_0)} = \mathrm{e}^{-x(t_0)} \sum_{n=0}^{\infty} A(n)|p(t,n|t_0)\rangle_{x(t_0)}\,. \tag{338}$$

Upon rewriting the marginalized distribution in terms of the backward path integral (213), the action (337) acquires the addition summand $x(t_0) - x(t)$ and now involves the unshifted transition operator $\mathcal{Q}_\tau^\dagger(x, \mathrm{i}q)$.

The average over a Poisson distribution in (336) establishes the link between the particle number n and the path integral variable x. For the simplest observable $A(n) := n$, i.e. for the particle number itself, it holds that $\langle\!\langle n\rangle\!\rangle_x = x$. This relation generalizes to factorial moments of order $k \in \mathbb{N}$ for which $\langle\!\langle (n)_k\rangle\!\rangle_x = x^k$ (recall that $(n)_k := n(n-1)\cdots(n-k+1)$). The computation of factorial moments may serve as an intermediate step in obtaining ordinary moments of the particle

number. For this purpose, one may use the relation $n^k = \sum_{l=0}^{k} \{{k \atop l}\} (n)_l$, where the curly braces represent a Stirling number of the second kind (cf. section 26.8 in [337]). An extension of the path integral representation (335) to multi-time averages of the form

$$\sum_{n_2,n_1=0}^{\infty} A(n_2,n_1) p(\tau_2,n_2;\tau_1,n_1|t_0,n_0) \tag{339}$$

remains open (with $\tau_2 > \tau_1 > t_0$). The results of Elderfield [378] may prove helpful for this purpose.

In a slightly rewritten form, the Doi-shifted path integral (335) was, for example, employed by Lee in his study of the diffusion-controlled annihilation reaction $k\,A \to \emptyset$ with $k \geq 2$ [135, 385]. He found that below the critical dimension $d_c = 2/(k-1)$, the particle density asymptotically decays as $n \sim A_k t^{-d/2}$ with a universal amplitude A_k. At the critical dimension, the particle density instead obeys $n \sim A_k (\ln t/t)^{1/(k-1)}$. Neglecting the diffusion of particles, the above action (337) of the reaction $k\,A \to \emptyset$ reads

$$\mathcal{S}^\dagger = \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}q\partial_\tau x + \gamma_\tau x^k \sum_{j=1}^{k} \binom{k}{j} (\mathrm{i}q)^j \right). \tag{340}$$

Besides using different names for integration variables ($\mathrm{i}q \to \overline{\psi}$ and $x \to \psi$), the action employed by Lee involves an additional boundary term. This term can be introduced by rewriting the path integral representation (335) as

$$\langle A \rangle_{\lambda_0} = \oint_{[t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \langle\!\langle A \rangle\!\rangle_{x(t)} . \tag{341}$$

This path integral involves integrations over the variables $q(t_0)$ and $x(t_0)$, and the mean of the initial Poisson distribution is denoted by λ_0. Consequently, one needs to add the boundary term $\mathrm{i}q(t_0)(x(t_0) - \lambda_0)$ to the action (340) to equate $x(t_0)$ with λ_0. The factor $\mathrm{i}q(t_0)x(t_0)$ of the new term is, however, often dropped eventually [135, 385]. The convention of Lee is also commonly used, for example, by Täuber [75]. The unshifted version of the path integral with transition operator $\mathcal{Q}_\tau^\dagger(x,\mathrm{i}q)$ is recovered via $\mathrm{i}q + 1 \to \mathrm{i}q$.

For completeness, let us note that the path integral representation (335) can also be derived from the forward path integral (299), provided that the observable $A(n)$ is analytic in n. It then suffices to consider the factorial moment $A(n) := (n)_k$. To perform the derivation, one may choose the basis function of the generating function (48) as $|n\rangle := (\mathrm{i}q + 1)^n$ (insert $\zeta := \mathrm{i}$, $\tilde{q} := 1$ and $\tilde{x} := 0$ into (75)). As

the first step, the forward path integral (303) is summed over an initial Poisson distribution with mean $x(t_0)$. The average of the factorial moment is then obtained via $\langle (n)_k \rangle_{x(t_0)} = \partial^k_{\mathrm{i}q_N} \langle g(t|t_0; x(t_0)) \rangle_{q_N} |_{q_N=0}$. To recover the action (337), one may note that the operator $\mathcal{Q}_\tau(c, a)$ in (69) and the operator $\mathcal{Q}^\dagger_\tau(c, a)$ in (156) fulfil $\mathcal{Q}_\tau(\mathrm{i}q + 1, x) = \mathcal{Q}^\dagger_\tau(x, \mathrm{i}q + 1)$ for scalar arguments. Thus, both the marginalized distribution approach and the generating function approach result in the same path integral representation of averaged observables. A detailed derivation of the path integral (335) from the generating function is, for example, included in the article of Dickman and Vidigal [415] (see their equations (106) and (108)).

6.2 Intermezzo: Alternative derivation based on coherent states

We noted in section 2.2 c that the path integral representation (335) of the average

$$\langle A \rangle_{x(t_0)} = \sum_{n_0, n=0}^{\infty} A(n)\, p(t, n|t_0, n_0) \frac{x(t_0)^{n_0} \mathrm{e}^{-x(t_0)}}{n_0!} \tag{342}$$

can be derived without first casting the master equation into a linear PDE [96, 135, 291, 338, 385]. The following section outlines this derivation for the process $k\,A \to l\,A$ with time-independent rate coefficient γ. Moreover, we assume $A(n)$ to be analytic in n.

The alternative derivation of the path integral representation (335) starts out from the exponential solution of the master equation, i.e. from (cf. (98))

$$p(t, n|t_0, n_0) = \langle n| \mathrm{e}^{Q(t-t_0)} |n_0\rangle \,. \tag{343}$$

The bras are chosen as the unit row vectors $\langle n| = \hat{\boldsymbol{e}}_n^{\mathsf{T}}$ and the kets as the unit column vectors $|n\rangle = \hat{\boldsymbol{e}}_n$. As before, the transition matrix of the reaction $k\,A \to l\,A$ with rate coefficient γ reads (cf. (100))

$$Q(c, a) = \gamma(c^l - c^k)a^k \,. \tag{344}$$

The creation matrix fulfils $c|n\rangle = |n + 1\rangle$ and $\langle n|c = \langle n - 1|$, and the annihilation matrix $a|n\rangle = n|n - 1\rangle$ and $\langle n|a = (n + 1)\langle n + 1|$. The basis column vectors can be generated incrementally via $|n\rangle = c^n|0\rangle$, the basis row vectors via $\langle n| = \langle 0|\frac{a^n}{n!}$.

Since the bra $\langle n|$ is a left eigenvector of the number matrix $\mathcal{N} := ca$ with eigenvalue n, one may write $A(n)\langle n| = \langle n|A(\mathcal{N})$ for an analytic observable A. After inserting the exponential solution (343) into the averaged observable (342), an evaluation of the sums therefore results in

$$\langle A \rangle_{x_0} = \langle 0| \mathrm{e}^a A(ca) \mathrm{e}^{Q(c,a)(t-t_0)} \mathrm{e}^{x_0(c-\mathbb{1})} |0\rangle \,. \tag{345}$$

Here we employed the (infinitely-large) unit matrix $\mathbb{1}$. Moreover, we changed the variable $x(t_0)$ to x_0 in anticipation of a discrete-time approximation.

Following the lecture notes of Cardy [96], we now perform the Doi-shift by shifting the first exponential in the above expression to the right. To do so, we require certain relations, which are all based on the commutation relation $[a, c] = \mathbb{1}$. First, it follows from this relation that $[a, c^n] = nc^{n-1}\mathbb{1}$ holds for all $n \in \mathbb{N}_0$, and more generally that $[a, [a \cdots [a, c^n]]] = (n)_j c^{n-j}\mathbb{1}$ holds for nested commutators with $j \leq n$ annihilation matrices. Nested commutators of higher order vanish. Using the Hadamard lemma [419], it follows for $z \in \mathbb{C}$ that

$$\mathrm{e}^{za}c^n = (\mathrm{e}^{za}c^n\mathrm{e}^{-za})\mathrm{e}^{za} = \big(c^n + [a, c^n]z + \frac{1}{2}[a, [a, c^n]]z^2 + \ldots\big)\mathrm{e}^{za} \tag{346}$$

$$= \big(c^n + nc^{n-1}z\,\mathbb{1} + \binom{n}{2}c^{n-2}z^2\,\mathbb{1} + \ldots\big)\mathrm{e}^{za} = (c + z\mathbb{1})^n\mathrm{e}^{za}\,. \tag{347}$$

This expression generalizes to shift operations for analytic functions, i.e.

$$\mathrm{e}^{za}f(c) = f(c + z\mathbb{1})\mathrm{e}^{za} \text{ and} \tag{348}$$

$$f(a)\mathrm{e}^{zc} = \mathrm{e}^{zc}f(a + z\mathbb{1})\,. \tag{349}$$

Shifting of the first exponential in (345) to the right therefore results in

$$\langle A\rangle_{x_0} = \langle 0|A\big((c + \mathbb{1})a\big)\mathrm{e}^{Q(c+\mathbb{1},a)(t-t_0)}\mathrm{e}^{x_0 c}|0\rangle\,. \tag{350}$$

Next, we split the time interval $[t_0, t]$ into N pieces of length $\Delta t := (t - t_0)/N$. The Trotter formula [420] can then be used to write

$$\langle A\rangle_{x_0} = \lim_{N\to\infty}\langle 0|A\big((c + \mathbb{1})a\big)\big[1 + Q(c + \mathbb{1}, a)\Delta t\big]^N\mathrm{e}^{x_0 c}|0\rangle. \tag{351}$$

This expression is now rewritten by inserting N of the identity matrices [20]

$$\mathbb{1} = \sum_{m=0}^{\infty}|m\rangle\langle m| = \sum_{m,n=0}^{\infty}\frac{1}{m!}\Big(\int_{\mathbb{R}}\mathrm{d}x\,(\partial_x^n\,x^m)\,\delta(x)\Big)|m\rangle\langle n| \tag{352}$$

$$= \sum_{m,n=0}^{\infty}\frac{1}{m!}\Big(\int_{\mathbb{R}}\mathrm{d}x\,x^m\,(-\partial_x)^n\int_{\mathbb{R}}\frac{\mathrm{d}q}{2\pi}\,\mathrm{e}^{-\mathrm{i}qx}\Big)c^m|0\rangle\langle 0|\frac{a^n}{n!} \tag{353}$$

$$= \int_{\mathbb{R}^2}\frac{\mathrm{d}x\,\mathrm{d}q}{2\pi}\mathrm{e}^{-(\mathrm{i}q-1)x}\,\mathrm{e}^{-x}\mathrm{e}^{xc}|0\rangle\langle 0|\mathrm{e}^{\mathrm{i}qa} = \int_{\mathbb{R}^2}\frac{\mathrm{d}x\,\mathrm{d}q}{2\pi}\mathrm{e}^{-(\mathrm{i}q-1)x}|x\rangle\!\rangle\langle\!\langle -\mathrm{i}q|\,. \tag{354}$$

Note that potential boundary terms were neglected in the integrations by parts (such integrations were not required in the analytic derivations in sections 4 and 5).

Moreover, we introduced the right eigenvector $|z\rangle\!\rangle := \mathrm{e}^{z(c-\mathbb{1})}|0\rangle$ of the annihilation matrix ($a|z\rangle\!\rangle = z|z\rangle\!\rangle$ with $z \in \mathbb{C}$), and the left eigenvector $\langle\!\langle z^\star| := \langle 0|\mathrm{e}^{za}$ of the creation matrix ($\langle\!\langle z^\star|c = z\langle\!\langle z^\star|$). These vectors are commonly referred to as "coherent states". The eigenvector conditions can be easily verified by rewriting the vectors as

$$|z\rangle\!\rangle = \sum_{m=0}^{\infty} \frac{z^m \mathrm{e}^{-z}}{m!}|m\rangle \text{ and } \langle\!\langle z^\star| = \sum_{n=0}^{\infty} z^n \langle n|\,. \tag{355}$$

Insertion of the identity matrices into (351) results in

$$\begin{aligned}\langle A\rangle_{x_0} = \lim_{N\to\infty} \int_1^N & \langle 0|A((c+\mathbb{1})a)\mathrm{e}^{x_N c}|0\rangle \\ & \cdot \prod_{j=1}^{N} \langle 0|\mathrm{e}^{-\mathrm{i}q_j(x_j-a)}\big[1+Q(c+\mathbb{1},a)\Delta t\big]\mathrm{e}^{x_{j-1}c}|0\rangle\,.\end{aligned} \tag{356}$$

This expression can be simplified with the help of the shift operations in (348) and (349). Since the Q-matrix (344) is a normal-ordered polynomial in c and a (i.e. all the c are to the left of all the a), and both $\langle 0|c$ and $a|0\rangle$ vanish, the above expression evaluates to

$$\begin{aligned}\langle A\rangle_{x_0} = \lim_{N\to\infty} \int_1^N & \langle 0|A((c+\mathbb{1})(a+x_N\mathbb{1}))|0\rangle \\ & \cdot \prod_{j=1}^{N} \mathrm{e}^{-\mathrm{i}q_j(x_j-x_{j-1})}\big[1+Q(\mathrm{i}q_j+1,x_{j-1})\Delta t\big]\,.\end{aligned} \tag{357}$$

The factor $\langle 0|A((c+\mathbb{1})(a+x_N\mathbb{1}))|0\rangle$ could be evaluated by normal-ordering the observable with respect to c and a before employing $\langle 0|c = 0$ and $a|0\rangle = 0$ again. The resulting object is sometimes called a "normal-ordered observable" [338]. In the following, we show that this object agrees with the average over a Poisson distribution in (336). The proof of this assertion is based on the observation that $f_j(x) := \langle 0|((c+\mathbb{1})(a+x\mathbb{1}))^j|0\rangle$ fulfils the following defining relation of Touchard polynomials (see [421, 422] for information on these polynomials):

$$\begin{aligned} f_{j+1}(x) &= \langle 0|\big([(c+\mathbb{1})(a+x\mathbb{1})\big]^{j+1}|0\rangle && (358)\\ &= \langle 0|(c+\mathbb{1})\big[(a+x\mathbb{1})(c+\mathbb{1})\big]^j(a+x\mathbb{1})|0\rangle \\ &= \langle 0|\big[(c+\mathbb{1})(a+x\mathbb{1})+\mathbb{1}\big]^j x\mathbb{1}\,|0\rangle && (359)\\ &= x\sum_{i=0}^{j}\binom{j}{i}\langle 0|\big[(c+\mathbb{1})(a+x\mathbb{1})\big]^i|0\rangle = x\sum_{i=0}^{j}\binom{j}{i}f_i(x)\,. && (360)\end{aligned}$$

Here we used $[a, c] = \mathbb{1}$, $a|0\rangle = 0$ and $\langle 0|c = 0$. The recurrence relation starts out from $f_0(x) = 1$. Since the j-th Touchard polynomial $f_j(x)$ agrees with the j-th moment of a Poisson distribution with mean x, i.e. with $\langle\!\langle n^j \rangle\!\rangle_x$ as defined in (336), our above assertion holds true.

The path integral representation (335) of averaged observables is recovered as the continuous-time limit of (357) (upon rewriting $1 + Q\Delta t$ as an exponential). the action (337) is recovered because the transition matrix in (344) and the adjoint transition operator in (156) fulfil $Q(\mathrm{i}q + 1, x) = \mathcal{Q}^\dagger(x, \mathrm{i}q + 1)$ for scalar arguments.

6.3 Perturbation expansions

The path integral representation (335) of factorial moments can be rewritten in terms of a (Q, X)-generating functional analogously to section 4.3. The resulting expression may serve as the starting point for a perturbative [20, 75] or a non-perturbative analysis [413] of the path integral. Here we focus on the perturbative approach. To derive the representation, we assume that the adjoint transition operator can be split into drift, diffusion, and perturbation parts as in (218), i.e.

$$\mathcal{Q}_\tau^\dagger(x, \partial_x + 1) = \alpha_\tau(x)\partial_x + \beta_\tau(x)\frac{\partial_x^2}{2} + \mathcal{P}_\tau^\dagger(x, \partial_x)\,. \tag{361}$$

By following the steps in appendix D (with N being replaced by $N + 1$), the path integral representation of a factorial moment can be written as

$$\langle (n)_j \rangle_{x(t_0)} = \frac{\delta^j}{\delta Q(t)^j} \mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau\, \mathcal{P}_\tau^\dagger\left(\frac{\delta}{\delta Q(\tau)}, \frac{\delta}{\delta X(\tau)}\right)} \mathcal{Z}_{Q,X}^0 \Big|_{Q=X=0}\,, \tag{362}$$

where $\mathcal{Z}_{Q,X}^0 = \left\langle\!\!\left\langle \mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau Q(\tau)x(\tau)} \right\rangle\!\!\right\rangle_W$ represents the (Q, X)-generating functional (222). Moreover, $x(\tau)$ solves the Itô SDE

$$\mathrm{d}x(\tau) = \big[\alpha_\tau(x) + X(\tau)\big]\,\mathrm{d}\tau + \sqrt{\beta_\tau(x)}\,\mathrm{d}W(\tau) \tag{363}$$

with initial condition $x(t_0)$. If both the diffusion and perturbation parts of the transition operator (361) vanish, the factorial moment at time t obeys the statistics of a Poisson distribution, i.e. $\langle (n)_j \rangle = x(t)^j$. The representation (362) can be evaluated perturbatively by expanding its exponential as [20]

$$\mathrm{e}^{\int_{t_0}^t \mathrm{d}\tau\, \mathcal{P}_\tau^\dagger} = \sum_{m=0}^\infty \int_{t_0}^t \mathrm{d}\tau_m\, \mathcal{P}_{\tau_m}^\dagger \cdots \int_{t_0}^{\tau_2} \mathrm{d}\tau_1\, \mathcal{P}_{\tau_1}^\dagger\,. \tag{364}$$

Unlike the expansion (286), this expansion involves only a single sum.

6.4 Coagulation

Let us exemplify the perturbation expansion of (362) for particles that diffuse and coagulate in the reaction $2A \to A$. This reaction exhibits the same asymptotic particle decay as the binary annihilation reaction $2A \to \emptyset$ and thus belongs to the same universality class [296, 423]. Only a pre-factor differs between the perturbation operators of the two reactions (see below). A full renormalization group analysis of general annihilation reactions $kA \to lA$ with $l < k$ was performed by Lee [135, 385] (see also [74]). The procedure of Lee differs slightly from the one presented below, but it involves the same Feynman diagrams. We restrict ourselves to the "tree level" of the diagrams. The coagulation reaction $2A \to A$ and the annihilation reaction $2A \to \emptyset$ eventually cause all but possibly one particle to vanish from a system. The transition into an absorbing steady state can be prevented; for example, by allowing for the creation of particles through the reaction $\emptyset \to A$. The fluctuations around the ensuing non-trivial steady state have been explored with the help of path integrals in [300, 310].

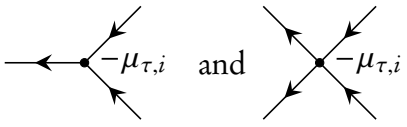
In the following, we consider particles that diffuse on an arbitrary network $\mathbb{L}$ and that coagulate in the reaction $2A \to A$. Hence, all results from section 4.5 d apply here as well, except that the transition operator of the coagulation reaction now acts as a perturbation. With the local coagulation rate coefficient $\mu_{\tau,i}$ at lattice site $i \in \mathbb{L}$, this perturbation reads (cf. (156))

$$\mathcal{P}_\tau^\dagger(\boldsymbol{x}, \mathrm{i}\boldsymbol{q}) = \sum_{i\in\mathbb{L}} \mu_{\tau,i} x_i^2 \left[(\mathrm{i}q_i + 1) - (\mathrm{i}q_i + 1)^2\right] \tag{365}$$

$$= \sum_{i\in\mathbb{L}} \left[(-\mu_{\tau,i}) x_i^2 \mathrm{i}q_i + (-\mu_{\tau,i}) x_i^2 (\mathrm{i}q_i)^2\right]. \tag{366}$$

For the binary annihilation reaction $2A \to \emptyset$, the first rate coefficient $\mu_{\tau,i}$ in this expression differs by an additional pre-factor 2. Note that we allow the rate coefficient to depend both on time and on the node $i \in \mathbb{L}$.

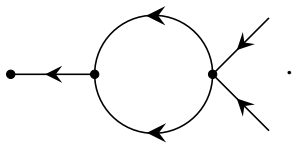
As explained in section 4.5 d, summands of the expansion of (362) can be represented by Feynman diagrams. For the mean local particle number $\langle n_i \rangle$, each diagram is composed of a sink with one incoming leg, and possibly of the vertices



(367)

According to the functional derivative (287), contracted lines, either between two vertices or between a vertex and the sink, are associated to the propagator $G(\tau|\tau')$. Dangling incoming lines, on the other hand, introduce a factor $\boldsymbol{x}_h(\tau)$, representing the homogeneous solution of $\partial_\tau x_i = \underline{\Delta}\varepsilon_{\tau,i} x_i + X_i(\tau)$ (cf. (285) and (286)).

In general, diagrams constructed from the above building blocks exhibit internal loops. The simplest connected diagram with such a loop and with a single sink is given by



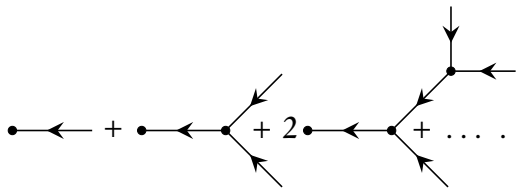
(368)

This loop is part of the $m = 2$ summand of (364). Its mathematical expression is obtained by tracing the diagram from right to left and reads

$$\int_{t_0}^{t} \mathrm{d}\tau_2 \sum_{j\in\mathbb{L}} G(t,i|\tau_2,j)(-\mu_{\tau_2,j}) \int_{t_0}^{\tau_2} \mathrm{d}\tau_1 \sum_{k\in\mathbb{L}} 2\big[G(\tau_2,j|\tau_1,k)\big]^2(-\mu_{\tau_1,k})\big[x_{h,k}(\tau_1)\big]^2 . \tag{369}$$

The combinatorial factor 2 stems from the two possible ways of connecting the two vertices (either outgoing leg can connect to either incoming leg). Note that the sink, which corresponds to the final derivative $\delta/\delta Q(t)$, is not associated to an additional pre-factor (unlike in section 4.5 d).

In the following, we focus on the contribution of "tree diagrams" to the mean particle number. These diagrams do not exhibit internal loops. Thus, upon removing all the diagrams containing loops from the expansion of (362), one can define the "tree-level average"

$\bar{n}_i(t) :=$ 

(370)

The corresponding mathematical expression reads

$$\begin{aligned}\bar{n}_i(t) = x_{h,i}(t) &+ \int_{t_0}^{t} \mathrm{d}\tau_1 \sum_{j\in\mathbb{L}} G(t,i|\tau_1,j)(-\mu_{\tau_1,j})\big[x_{h,j}(\tau_1)\big]^2 \\ &+ 2\int_{t_0}^{t} \mathrm{d}\tau_2 \sum_{j\in\mathbb{L}} G(t,i|\tau_2,j)(-\mu_{\tau_2,j})x_{h,j}(\tau_2) \\ &\cdot \int_{t_0}^{\tau_2} \mathrm{d}\tau_1 \sum_{k\in\mathbb{L}} G(\tau_2,j|\tau_1,k)(-\mu_{\tau_1,k})\big[x_{h,k}(\tau_1)\big]^2 + \dots .\end{aligned} \tag{371}$$

The inclusion of diagrams with loops would correct $\bar{n}_i$ to the true mean $\langle n_i \rangle$. For the treatment of loops, see for example [75, 135, 295, 385].

The tree-level average $\bar{n}_i$ defined above fulfils the deterministic rate equation of the coagulation process, i.e. it fulfils

$$\partial_\tau \bar{n}_i = \underline{\Delta} \varepsilon_{\tau,i} \bar{n}_i - \mu_{\tau,i} \bar{n}_i^2 \,. \tag{372}$$

If both the rates ε_τ and μ_τ, and also the mean $x(t_0)$ of the initial Poisson distribution are spatially homogeneous, the well-mixed analogue of (372) can be derived by direct resummation of (371). In the general case, the validity of the rate equation (372) can be established by exploiting a self-similarity of $\bar{n}_i$. For that purpose, we assume that $\bar{n}_i(\tau)$ is known for $\tau < t$ and we want to extend its validity up to time t. To do so, we remove all the sinks from the diagrams in (370) and connect their now dangling outgoing legs with the incoming leg of another three-vertex at time τ. The second incoming leg of this vertex is contracted with every other tree diagram and its outgoing leg with a sink at time t. The corresponding expression reads $\sum_{j\in\mathbb{L}} G(t,i|\tau,j)(-\mu_{\tau,j})\bar{n}_j(\tau)^2$ and it contributes to $\bar{n}_i(t)$ for every time τ. To respect also the initial condition $\bar{n}_i(t_0) = x_i(t_0)$, we introduce the first diagram of (370) by hand so that in total

$$\bar{n}_i(t) = x_{h,i}(t) + \int_{t_0}^{t} \mathrm{d}\tau \sum_{j\in\mathbb{L}} G(t,i|\tau,j)(-\mu_{\tau,j})\bar{n}_j(\tau)^2 \,. \tag{373}$$

Differentiation of $\bar{n}_i(t)$ with respect to t confirms the validity of (372).

6.5 Résumé

Path integral representations of averaged observables have proved useful in a variety of contexts. Their use has deepened our knowledge about the critical behaviour of diffusion-limited annihilation and coagulation reactions [74, 75, 135, 295, 296, 304, 306, 307, 318, 403], of branching and annihilating random walks and percolation processes [74, 75, 101, 137, 138, 303, 305, 424, 425], and of elementary multi-species reactions [136, 297–299, 301, 302, 311–313, 315, 426, 427]. In the present section, we derived a path integral representation of averaged observables for processes that can be decomposed additively into reactions of the form $k\,A \to l\,A$. Provided that the number of particles n in the system is initially Poisson distributed with mean $x(t_0)$, the average of an observable $A(n)$ can be represented by the path integral

$$\langle A \rangle_{x(t_0)} = \int_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \langle\!\langle A \rangle\!\rangle_{x(t)} \tag{374}$$

$$\text{with } \langle\!\langle A \rangle\!\rangle_x = \sum_{n=0}^{\infty} \frac{x^n \mathrm{e}^{-x}}{n!} A(n) \,. \tag{375}$$

We derived this representation in section 6.1 by summing the backward path integral representation (213) of the marginalized distribution over the observable $A(n)$ (using the Poisson basis function $|n\rangle = \frac{x^n e^{-x}}{n!}$). The generalization of the above path integral to reactions with multiple types of particles and spatial degrees of freedom is straightforward. The above path integral representation was found to be equivalent to Doi-shifted path integrals used in the literature [75, 101, 135–138]. Its unshifted version is obtained for a redefined basis function. Unlike path integral representations encountered in the literature, our representation (374) does not involve a so-called "normal-ordered observable". By using a defining relation of Touchard polynomials, we could show that this object agrees with the average (375) of $A(n)$ over a Poisson distribution (cf. section 6.2).

As shown in section 6.3, the path integral representation (374) can be rewritten in terms of a perturbation expansion. We demonstrated the evaluation of this expansion in section 6.4 for diffusing particles that coagulate according to the reaction $2A \to A$. In doing so, we restricted ourselves to the tree-level of the Feynman diagrams associated to the perturbation expansion. Information on perturbative renormalization group techniques for the treatment of diagrams with loops can be found in [74, 75]. Recently, non-perturbative renormalization group techniques have been developed for the evaluation of stochastic path integrals [319, 413]. These techniques have proved particularly useful in studying branching and annihilating random walks [319, 320] and annihilation processes [308, 309].

7 Stationary paths

In the previous sections, we outlined how path integrals can be expressed in terms of averages over the paths of stochastic differential equations (SDEs). Corrections to those paths were treated in terms of perturbation expansions. In the following, we formulate an alternative method in which the variables of the path integrals act as deviations from "stationary", or "extremal", paths. The basic equations of the method have the form of Hamilton's equations from classical mechanics. Their application to stochastic path integrals goes back at least to the work of Mikhailov [16].

More recently, Elgart and Kamenev extended the method for the study of rare event probabilities [40] and for the classification of phase transitions in reaction-diffusion models with a single type of particles [322]. These studies are effectively based on the forward path integral representation (299) of the generating function. After reviewing how this approach can be used for the approximation of probability distributions, we extend it to the backward path integral representation (210) of the marginalized distribution. Whereas the generating function approach requires an auxiliary saddle-point approximation to extract probabilities from the

generating function, the backward approach provides direct access to probabilities. A proper normalization of the resulting probability distribution is, however, only attained beyond leading order. The generating function technique respects the normalization of the distribution even at leading order, but this normalization may be violated by the subsequent saddle-point approximation.

The methods discussed in the following all apply to the chemical master equation (27) and employ the basis functions that we introduced in sections 2.2 b and 3.2 b. Moreover, we assume the number of particles to be initially Poisson distributed with mean $x(t_0)$. This assumption proves to be convenient in the analysis of explicit stochastic processes but it can be easily relaxed.

7.1 Forward path integral approach

Our goal lies in the approximation of the marginalized probability distribution

$$|p(t,n|t_0)\rangle_{x(t_0)} = \sum_{n_0=0}^{\infty} p(t,n|t_0,n_0)\frac{x(t_0)^{n_0}\mathrm{e}^{-x(t_0)}}{n_0!}\,. \tag{376}$$

For this purpose, we employ the forward path integral (299) to formulate an alternative representation of the ordinary probability generating function

$$|g(t|t_0,n_0)\rangle_{q(t)} := \sum_{n=0}^{\infty} q(t)^n p(t,n|t_0,n_0)\,. \tag{377}$$

As the first step, we rewrite the argument $q(t)$ of this function in terms of a deviation $\Delta q(t)$ from a "stationary path" $\tilde{q}(t)$, i.e. $q(t) = \tilde{q}(t) + \Delta q(t)$. The path $\tilde{q}(t)$ and an auxiliary path $\tilde{x}(\tau)$ are chosen so that the action of the resulting path integral is free of terms that are linear in the path integral variables $\Delta q(\tau)$ and $\Delta x(\tau)$ (with $\tau \in [t_0, t]$). Thus, the approach bears similarities with the stationary phase approximation of oscillatory integrals [428]. In the next section, we apply the method to the binary annihilation process $2\,A \to \emptyset$.

In order to implement the above steps, we define the basis function

$$|n\rangle_{\tau,\Delta q(\tau)} := (\zeta\Delta q(\tau) + \tilde{q}(\tau))^n \mathrm{e}^{-\tilde{x}(\tau)(\zeta\Delta q(\tau)+\tilde{q}(\tau))} \tag{378}$$

for yet to be specified paths $\tilde{q}(\tau)$ and $\tilde{x}(\tau)$, and a free parameter ζ. The basis function has the same form as the basis function (75) but with its second argument being written as $\Delta q(\tau)$. Provided that the path $\tilde{q}(\tau)$ fulfils the final condition $\tilde{q}(t) = q(t)$, one can trivially rewrite the generating function (377) in terms of the

above basis function as

$$|g\rangle_{q(t)} = \mathrm{e}^{\tilde{x}(t)\tilde{q}(t)}\Big(\sum_{n=0}^{\infty}|n\rangle_{t,\Delta q(t)}\,p(t,n|\cdot)\Big)\Big|_{\Delta q(t)=0}\,. \tag{379}$$

The term in brackets has the form of the generalized generating function (48), and thus it can be rewritten in terms of the forward path integral (303) by following the steps in section 5.1. Using $\Delta q(\tau)$ and $\Delta x(\tau)$ as labels for the path integral variables, one arrives at the representation

$$|g\rangle_{q(t)} = \mathrm{e}^{\tilde{x}(t)\tilde{q}(t)}\Big(\fint_{[t_0}^{t)}\mathrm{e}^{-\mathcal{S}}\,|n_0\rangle_{t_0,\Delta q(t_0)}\Big)\Big|_{\Delta q(t)=0}\,. \tag{380}$$

In analogy with (301), the integral sign is defined as the continuous-time limit

$$\fint_{[t_0}^{t)} = \lim_{N\to\infty}\prod_{j=0}^{N-1}\int_{\mathbb{R}^2}\frac{\mathrm{d}\Delta x_j\,\mathrm{d}\Delta q_j}{2\pi}\,. \tag{381}$$

The action (304) inside the generating function (380) reads

$$\mathcal{S} = \int_{t_0}^{t}\mathrm{d}\tau\left[\mathrm{i}\Delta x\partial_{-\tau}\Delta q - \widetilde{\mathcal{Q}}_\tau(\Delta q,\mathrm{i}\Delta x)\right]. \tag{382}$$

As emphasized above, our interest lies in a system whose initial number of particles is Poisson distributed with mean $x(t_0)$. Thus, instead of dealing with the ordinary generating function (377), it proves convenient to work in terms of the generating function

$$|g(t|t_0;x(t_0))\rangle_{q(t)} \coloneqq \sum_{n_0=0}^{\infty}|g(t|t_0,n_0)\rangle_{q(t)}\frac{x(t_0)^{n_0}\mathrm{e}^{-x(t_0)}}{n_0!} \tag{383}$$

$$= \mathrm{e}^{\tilde{x}(t)\tilde{q}(t)-x(t_0)}\fint_{[t_0}^{t)}\mathrm{e}^{-\mathcal{S}}\Big|_{\Delta q(t)=0}\,. \tag{384}$$

To arrive at the second line, we made use of the path integral representation (380) while requiring that the path $\tilde{x}(\tau)$ meets the initial condition $\tilde{x}(t_0) = x(t_0)$. Note that the marginalized distribution (376) is recovered from the redefined generating function via

$$|p(t,n|t_0)\rangle_{x(t_0)} = \frac{1}{n!}\partial^n_{q(t)}|g(t|t_0;x(t_0))\rangle_{q(t)}\Big|_{q(t)=0}\,. \tag{385}$$

Thus far, we have not yet specified the paths $\tilde{q}(\tau)$ and $\tilde{x}(\tau)$, besides requiring that they fulfil the boundary conditions $\tilde{q}(t) = q(t)$ and $\tilde{x}(t_0) = x(t_0)$. We specify these paths in such a way that the action (382) becomes free of terms that are linear in the deviations Δq and Δx. For this purpose, let us recall the definitions of the creation operator $c = \zeta \Delta q + \tilde{q}$ and of the annihilation operator $a = \partial_{\zeta \Delta q} + \tilde{x}$ from section section 2.2 b, as well as the definition of the transition operator

$$\widetilde{\mathcal{Q}}_\tau(\Delta q, \partial_{\Delta q}) = \mathcal{Q}_\tau(c, a) + \mathcal{E}_\tau(c, a)\,. \tag{386}$$

The basis evolution operator is specified in (90) as

$$\mathcal{E}_\tau(c, a) = (\partial_\tau \tilde{q})(a - \tilde{x}) - (\partial_\tau \tilde{x})c\,. \tag{387}$$

Upon performing a Taylor expansion of the operator $\widetilde{\mathcal{Q}}_\tau(\Delta q, \mathrm{i}\Delta x)$ in the action (382) with respect to the deviations Δq and Δx, one observes that terms that are linear in the deviations vanish from the action if the paths $\tilde{x}(\tau)$ and $\tilde{q}(\tau)$ fulfil[15]

$$\partial_\tau \tilde{x} = \frac{\partial \mathcal{Q}_\tau(\tilde{q}, \tilde{x})}{\partial \tilde{q}} \text{ with } \tilde{x}(t_0) = x(t_0) \text{ and} \tag{388}$$

$$\partial_{-\tau} \tilde{q} = \frac{\partial \mathcal{Q}_\tau(\tilde{q}, \tilde{x})}{\partial \tilde{x}} \text{ with } \tilde{q}(t) = q(t)\,. \tag{389}$$

These equations resemble Hamilton's equations from classical mechanics. Just as in classical mechanics, $\mathcal{Q}_\tau$ is conserved along solutions of the equations if it does not depend on time itself. This property follows from $\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{Q}_\tau(\tilde{q}, \tilde{x}) = \partial_\tau \mathcal{Q}_\tau(\tilde{q}, \tilde{x})$, with $\frac{\mathrm{d}}{\mathrm{d}\tau}$ being the total time derivative. Since the ordinary generating function obeys the flow equation $\partial_\tau |g\rangle_q = \mathcal{Q}_\tau(q, \partial_q)|g\rangle_q$, the conservation of total probability requires that $\mathcal{Q}_\tau(1, \partial_q) = 0$. This condition is, for example, fulfilled by the transition operator $\mathcal{Q}_\tau(q, \partial_q) = \gamma_\tau(q^l - q^k)\partial_q^k$ of the generic reaction $k\,A \to l\,A$ (cf. (69)). For the final value $q(t) = 1$, Hamilton's equations (388) and (389) are solved by $\tilde{q}(\tau) = 1$ with $\tilde{x}(\tau)$ solving the rate equation $\partial_\tau \tilde{x} = \gamma_\tau(l - k)\tilde{x}^k$ of the process. Elgart and Kamenev analysed the topology of $\mathcal{Q}_\tau(\tilde{q}, \tilde{x}) = 0$ lines

15 If Hamilton's equations are fulfilled, it holds that

$$\begin{aligned}\widetilde{\mathcal{Q}}_\tau(\Delta q, \mathrm{i}\Delta x) &= \big(\mathcal{Q}_\tau(\tilde{q}, \tilde{x}) + \frac{\partial \mathcal{Q}_\tau}{\partial \tilde{q}}\zeta \Delta q + \frac{\partial \mathcal{Q}_\tau}{\partial \tilde{x}}\zeta^{-1}\mathrm{i}\Delta x + \Delta \mathcal{Q}_\tau\big) \\ &\quad + \big((\partial_\tau \tilde{q})\zeta^{-1}\mathrm{i}\Delta x - (\partial_\tau \tilde{x})(\zeta \Delta q + \tilde{q})\big) \\ &= -\partial_\tau(\tilde{x}\tilde{q}) - \big(\tilde{x}\partial_{-\tau}\tilde{q} - \mathcal{Q}_\tau(\tilde{q}, \tilde{x})\big) + \Delta \mathcal{Q}_\tau\,.\end{aligned}$$

Here, $\Delta \mathcal{Q}_\tau$ represents all terms of the Taylor expansion of $\mathcal{Q}_\tau$ that are of second or higher order in the deviations.

of reaction-diffusion models with a single type of particles to classify the phase transitions of these models [322] (with label p instead of $\tilde{q}$, and q instead of $\tilde{x}$).

Provided that Hamilton's equations (388) and (389) are fulfilled, the action (382) evaluates to

$$\mathcal{S} = \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}\Delta x \partial_{-\tau} \Delta q - \widetilde{\mathcal{Q}}_\tau(\Delta q, \mathrm{i}\Delta x) \right) \tag{390}$$

$$= \tilde{x}(t)\tilde{q}(t) - \tilde{x}(t_0)\tilde{q}(t_0) + \int_{t_0}^{t} \mathrm{d}\tau \left(\tilde{x}\partial_{-\tau}\tilde{q} - \mathcal{Q}_\tau(\tilde{q}, \tilde{x}) \right) \tag{391}$$

$$+ \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}\Delta x \partial_{-\tau} \Delta q - \Delta\mathcal{Q}_\tau \right)$$

$$= \tilde{x}(t)\tilde{q}(t) - x(t_0) + \tilde{\mathcal{S}} + \Delta\mathcal{S}\,, \tag{392}$$

with the definitions

$$\tilde{\mathcal{S}} := x(t_0)(1 - \tilde{q}(t_0)) + \int_{t_0}^{t} \mathrm{d}\tau \left[\tilde{x}\partial_{-\tau}\tilde{q} - \mathcal{Q}_\tau(\tilde{q}, \tilde{x}) \right] \tag{393}$$

$$\text{and } \Delta\mathcal{S} := \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}\Delta x \partial_{-\tau} \Delta q - \Delta\mathcal{Q}_\tau \right). \tag{394}$$

The transition operator $\Delta\mathcal{Q}_\tau$ absorbs all terms of the Taylor expansion of $\mathcal{Q}_\tau$ that are of second or higher order in the deviations. Combined with (384), the above action results in the following path integral representation of the generating function:

$$|g(t|t_0; x(t_0))\rangle_{q(t)} = \mathrm{e}^{-\tilde{\mathcal{S}}} \int_{[t_0}^{t)} \mathrm{e}^{-\Delta\mathcal{S}}\big|_{\Delta q(t)=0}\,. \tag{395}$$

Although this path integral representation may seem daunting, our primary interest lies only in its leading-order approximation $|g\rangle \approx \mathrm{e}^{-\tilde{\mathcal{S}}}$. This approximation is exact if $\Delta\mathcal{Q}_\tau$ vanishes. That is, for example, the case for the simple growth process $\emptyset \to A$. For the binary annihilation reaction $2A \to \emptyset$, the leading-order approximation was evaluated by Elgart and Kamenev up to a pre-exponential factor [40]. The pre-exponential factor was later approximated by Assaf and Meerson for large times [41]. In the next section, we evaluate the leading-order approximation of the binary annihilation reaction and evaluate its pre-exponential factor for arbitrary times.

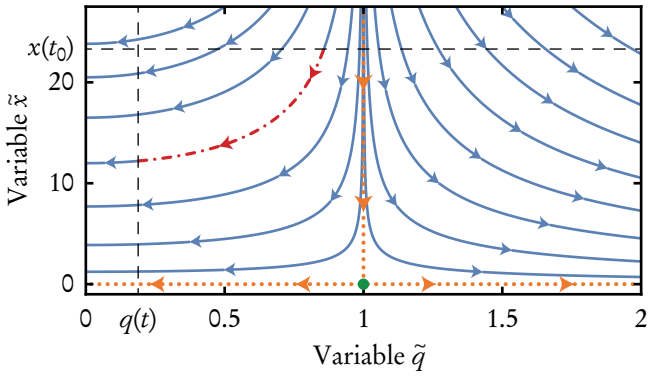


Figure 9 Phase portrait of Hamilton's equations (396) and (397) for the rate coefficient $\mu_\tau = 1$. The transition operator $\mathcal{Q}_\tau(\tilde{q}, \tilde{x}) = \mu_\tau(1 - \tilde{q}^2)\tilde{x}^2$ of the binary annihilation process vanishes for $\tilde{x} = 0$ and $\tilde{q} = 1$ (orange dotted lines). The green disk represents the fixed point $(\tilde{q}, \tilde{x}) = (1, 0)$ of Hamilton's equations. The red dash-dotted line exemplifies a particular solution of Hamilton's equations for a given time interval (t_0, t), a given initial value $x(t_0)$ and a given final value $q(t)$ (black dashed lines).

7.2 Binary annihilation

Let us demonstrate the use of the representation (395) for the binary annihilation reaction $2A \to \emptyset$ with rate coefficient μ_τ. First, we evaluate the generating function in leading order. Afterwards, the generating function is cast into a probability distribution using the inverse transformation (51). The derivatives involved in the inverse transformation are expressed by Cauchy's differentiation formula, which is evaluated in terms of a saddle-point approximation.

The transition operator of the binary annihilation reaction follows from (69) as $\mathcal{Q}_\tau(c_\tau, a_\tau) = \mu_\tau(1 - c_\tau^2)a_\tau^2$ with the annihilation rate coefficient μ_τ. Therefore, Hamilton's equations (388) and (389) read

$$\partial_\tau \tilde{x} = -2\mu_\tau \tilde{q}\tilde{x}^2 \text{ with } \tilde{x}(t_0) = x(t_0) \text{ and} \tag{396}$$

$$\partial_{-\tau} \tilde{q} = 2\mu_\tau(1 - \tilde{q}^2)\tilde{x} \text{ with } \tilde{q}(t) = q(t)\,. \tag{397}$$

A phase portrait of these equations is shown in figure 9. Equation (397) is solved by $\tilde{q} = 1$, for which the previous equation simplifies to the rate equation $\partial_\tau \tilde{x} = -2\mu_\tau \tilde{x}^2$ of the process. This rate equation is solved by

$$\bar{x}(t) := \tilde{x}(t) = \frac{x(t_0)}{1 + x(t_0)/\bar{x}_\infty(t)} \tag{398}$$

with the asymptotic limit $\bar{x}_\infty(t) := (2\int_{t_0}^{t} \mathrm{d}\tau\, \mu_\tau)^{-1} \to 0$. Since Hamilton's equations conserve $(\tilde{q}^2 - 1)\tilde{x}^2$, one can rewrite the equation (397) as

$$\partial_\tau \tilde{q} = -2\mu_\tau x(t_0)\sqrt{1-\tilde{q}(t_0)^2}\sqrt{1-\tilde{q}^2}\,. \tag{399}$$

Here we assume $\tilde{q}(t) = q(t) < 1$ but the derivation can also be performed for $q(t) > 1$ (these inequalities are preserved along the flow). The above equation only allows for an implicit solution that provides $\tilde{q}(t_0)$ for a given $q(t)$, namely[16]

$$\arccos \tilde{q}(t_0) + \frac{x(t_0)}{\bar{x}_\infty(t)}\sqrt{1-\tilde{q}(t_0)^2} = \arccos q(t)\,. \tag{400}$$

The first term of this equation was neglected in previous studies [40, 41] (for large times t, $\tilde{q}(t_0) \approx 1$). Using the conservation of $(\tilde{q}^2 - 1)\tilde{x}^2$ once again, the action in the leading-order approximation $|g\rangle_{q(t)} = \mathrm{e}^{-\tilde{\mathcal{S}}(q(t))}$ can be written as

$$\tilde{\mathcal{S}}(q(t)) = \big[1-\tilde{q}(t_0)\big]x(t_0) + \frac{\big[1-\tilde{q}(t_0)^2\big]x(t_0)^2}{2\bar{x}_\infty(t)}\,. \tag{401}$$

We have thus fully specified the generating function in terms of its argument $q(t)$ and the mean $x(t_0)$ of the initial Poisson distribution. The leading-order approximation $|g\rangle_{q(t)} = \mathrm{e}^{-\tilde{\mathcal{S}}(q(t))}$ respects the normalization of the underlying probability distribution because

$$\sum_{n=0}^{\infty} p(t,n|t_0,n_0) = |g\rangle_{q(t)=1} = 1\,. \tag{402}$$

Here we used that for $q(t) = 1$, Hamilton's equation (397) is solved by $\tilde{q}(\tau) = 1$, implying that $\tilde{\mathcal{S}}(q(t) = 1) = 0$.

The probability distribution follows from the generating function via the inverse transformation $|p(t,n|t_0)\rangle_{x(t_0)} = \frac{1}{n!}\partial^n_{q(t)}|g\rangle_{q(t)}\big|_{q(t)=0}$. This transformation

16 For $q(t) > 1$, the rate equation reads

$$\partial_\tau \tilde{q} = 2\mu_\tau x(t_0)\sqrt{\tilde{q}(t_0)^2-1}\sqrt{\tilde{q}^2-1}\,.$$

It preserves $q(\tau) > 1$ for all times τ and it is solved by

$$\operatorname{arcosh} \tilde{q}(t_0) + \frac{x(t_0)}{\bar{x}_\infty(t)}\sqrt{\tilde{q}(t_0)^2-1} = \operatorname{arcosh} q(t)\,.$$

Since $\operatorname{arcosh} y = -\mathrm{i}\arccos y$ holds for $y > 0$ and $\sqrt{y} = -\mathrm{i}\sqrt{-y}$ for $y > 0$, the solution agrees with (400).

is trivial for $n = 0$, for which one obtains the probability of observing an empty system. For other values of n, the derivatives can be expressed by Cauchy's differentiation formula so that in leading order

$$|p(t,n|t_0)\rangle_{x(t_0)} = \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}} \frac{\mathrm{d}q(t)}{q(t)} \mathrm{e}^{-\tilde{S}(q(t))-n\ln q(t)} . \tag{403}$$

Here, $\mathcal{C}$ represents a closed path around zero in the complex domain and is integrated over once in counter-clockwise direction.

The contour integral in (403) can be evaluated in a saddle-point approximation:

$$|p(t,n|t_0)\rangle_{x(t_0)} \approx \frac{q_s^{-n}\mathrm{e}^{-\tilde{S}(q_s)}}{\sqrt{2\pi(n-q_s^2\tilde{S}''(q_s))}} . \tag{404}$$

The saddle-point q_s is found by solving the equation $n/q_s = -\tilde{S}'(q_s)$. To obtain a closed expression for q_s, we first differentiate both sides of (400) with respect to $q(t)$ and find

$$-\frac{1}{\sqrt{1-q(t)^2}} = -\frac{1}{\sqrt{1-\tilde{q}(t_0)^2}}\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)} - \frac{x(t_0)}{\bar{x}_\infty(t)}\frac{\tilde{q}(t_0)}{\sqrt{1-\tilde{q}(t_0)^2}}\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)} \tag{405}$$

$$= -\frac{1}{\sqrt{1-\tilde{q}(t_0)^2}}\Big(1+\frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}\Big)\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)} . \tag{406}$$

The saddle-point condition $n/q_s = -\tilde{S}'(q_s)$ therefore becomes

$$\frac{n}{q_s} = \Big(x(t_0)\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)} + \frac{\tilde{q}(t_0)x(t_0)^2}{\bar{x}_\infty(t)}\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)}\Big)\Big|_{q(t)=q_s} \tag{407}$$

$$= x(t_0)\Big(1+\frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}\Big)\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)}\Big|_{q(t)=q_s} = x(t_0)\frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q_s^2}} . \tag{408}$$

A closed equation for q_s is obtained by combining this equation with the implicit solution (400). The resulting equation is solved by $q_s = 1$ for $n = \bar{x}(t)$, i.e. if n lies on the trajectory of the rate equation. For other values of n, the equation has to be solved numerically. Once q_s has been obtained, $\tilde{q}(t_0)$ can be inferred from (408).

One piece is still missing for the numeric evaluation of the probability distribution (404), namely its denominator. It evaluates to[17]

$$n - q_s^2 \tilde{\mathcal{S}}''(q_s) = \bar{x}_\infty(t) \frac{q_s^2 \alpha - n/\bar{x}_\infty(t)}{q_s^2 - 1} \tag{409}$$

$$\text{with } \alpha := 1 - \Big(1 + \frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}\Big)^{-1} . \tag{410}$$

The pre-exponential factor of the distribution (404) computed by Assaf and Meerson [41] is recovered for large times for which $\bar{x}_\infty(t) \to 0$ and thus $\alpha \to 1$. The above expressions hold both for $q_s < 1$ and $q_s > 1$. Care has to be taken in evaluating the limit $q_s \to 1$. A rather lengthy calculation employing L'Hôpital's rule shows that the left-hand side of (409) evaluates to[18]

$$\frac{2}{3}\bar{x}(t)\Big(1 + \frac{1}{2}\Big[1 + \frac{x(t_0)}{\bar{x}_\infty(t)}\Big]^{-3}\Big) . \tag{411}$$

17 The derivation proceeds as

$$\begin{aligned} n - q_s^2 \tilde{\mathcal{S}}''(q_s) &= n + q_s^2 \frac{\mathrm{d}}{\mathrm{d}q(t)} \Big(x(t_0) \frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q(t)^2}} \Big)\Big|_{q(t)=q_s} \\ &= n + \frac{q_s^2}{1-q_s^2} \Big(q_s x(t_0) \frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q_s^2}} - \tilde{q}(t_0) x(t_0) \frac{\sqrt{1-q_s^2}}{\sqrt{1-\tilde{q}(t_0)^2}} \frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q(t)}\Big|_{q(t)=q_s} \Big) \\ &= n + \frac{q_s^2}{1-q_s^2} \Big(n - \frac{x(t_0)\tilde{q}(t_0)}{1 + \frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}} \Big) = n - \frac{q_s^2}{q_s^2 - 1} \Big(n - \bar{x}_\infty(t)\alpha \Big) = \bar{x}_\infty(t) \frac{q_s^2 \alpha - n/\bar{x}_\infty(t)}{q_s^2 - 1} . \end{aligned}$$

18 With the help of the saddle-point condition (408), equation (409) can be rewritten as

$$\lim_{q_s \to 1} (n - q_s^2 \tilde{\mathcal{S}}''(q_s)) = -\bar{x}_\infty(t) \lim_{q_s \to 1} \frac{q_s \sqrt{1-q_s^2}\, \alpha - \frac{x(t_0)}{\bar{x}_\infty(t)} \sqrt{1-\tilde{q}(t_0)^2}}{(1-q_s^2)^{3/2}/q_s} = (\star) .$$

To apply L'Hôpital's rule, both the numerator and the denominator must be differentiated with respect to q_s. Differentiation of the numerator results in

$$\frac{1-2q_s^2}{\sqrt{1-q_s^2}} \alpha + q_s \sqrt{1-q_s^2} \Big(1 + \frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}\Big)^{-2} \frac{x(t_0)}{\bar{x}_\infty(t)} \frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q_s} + \frac{x(t_0)}{\bar{x}_\infty(t)} \frac{\tilde{q}(t_0)}{\sqrt{1-\tilde{q}(t_0)^2}} \frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q_s}$$

and of the denominator in

$$-\sqrt{1-q_s^2}(2q_s^2 + 1)/q_s^2 .$$

The pre-factor of this expression matches the normalization constant that Elgart and Kamenev inserted by hand [40].

After putting all of the above pieces together, the probability distribution (404) provides a decent approximation of the binary annihilation process. Figure 10 compares the distribution with a distribution that was obtained through a numerical integration of the master equation. For very large times, the quality of the approximation deteriorates. In particular, the approximation does not capture the final state of the process in which it is equally likely to find a single surviving particle or none at all. Limiting cases of the approximation (404) are provided in [40, 41].

Although the above evaluation of the saddle-point approximation is rather elaborate, it is still feasible. It becomes infeasible if one goes beyond the leading-order term of the generating function. In the section after the next, we show how higher order terms can be included in a dual approach, which is based on the backward path integral (213).

By combining the two, one finds that

$$
\begin{aligned}
(\star) &= \bar{x}_\infty(t) \lim_{q_s\to 1}\Big(\frac{q_s^2(1-2q_s^2)}{(1-q_s^2)(2q_s^2+1)}\alpha + \frac{q_s^3}{2q_s^2+1}\Big(1+\frac{x(t_0)\tilde{q}(t_0)}{\bar{x}_\infty(t)}\Big)^{-3}\frac{x(t_0)}{\bar{x}_\infty(t)}\frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q_s^2}} \\
&\qquad + \frac{q_s^2}{(1-q_s^2)(2q_s^2+1)}\alpha\Big) \\
&= \bar{x}_\infty(t)\Big(\lim_{q_s\to 1}\frac{2q_s^2}{2q_s^2+1}\alpha + \frac{1}{3}\Big(1+\frac{x(t_0)}{\bar{x}_\infty(t)}\Big)^{-3}\frac{x(t_0)}{\bar{x}_\infty(t)}\lim_{q_s\to 1}\frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q_s^2}}\Big) \\
&= \frac{2}{3}\frac{x(t_0)}{1+x(t_0)/\bar{x}_\infty(t)}\Big(1+\frac{1}{2}\Big(1+\frac{x(t_0)}{\bar{x}_\infty(t)}\Big)^{-3}\Big) \\
&= \frac{2}{3}\bar{x}(t)\Big(1+\frac{1}{2}\Big(1+\frac{x(t_0)}{\bar{x}_\infty(t)}\Big)^{-3}\Big).
\end{aligned}
$$

In the third step, we used L'Hopital's rule again:

$$
\lim_{q_s\to 1}\frac{\sqrt{1-\tilde{q}(t_0)^2}}{\sqrt{1-q_s^2}} = \lim_{q_s\to 1}\frac{\tilde{q}(t_0)}{q_s}\frac{\sqrt{1-q_s^2}}{\sqrt{1-\tilde{q}(t_0)^2}}\frac{\mathrm{d}\tilde{q}(t_0)}{\mathrm{d}q_s} = \Big(1+\frac{x(t_0)}{\bar{x}_\infty(t)}\Big)^{-1}.
$$

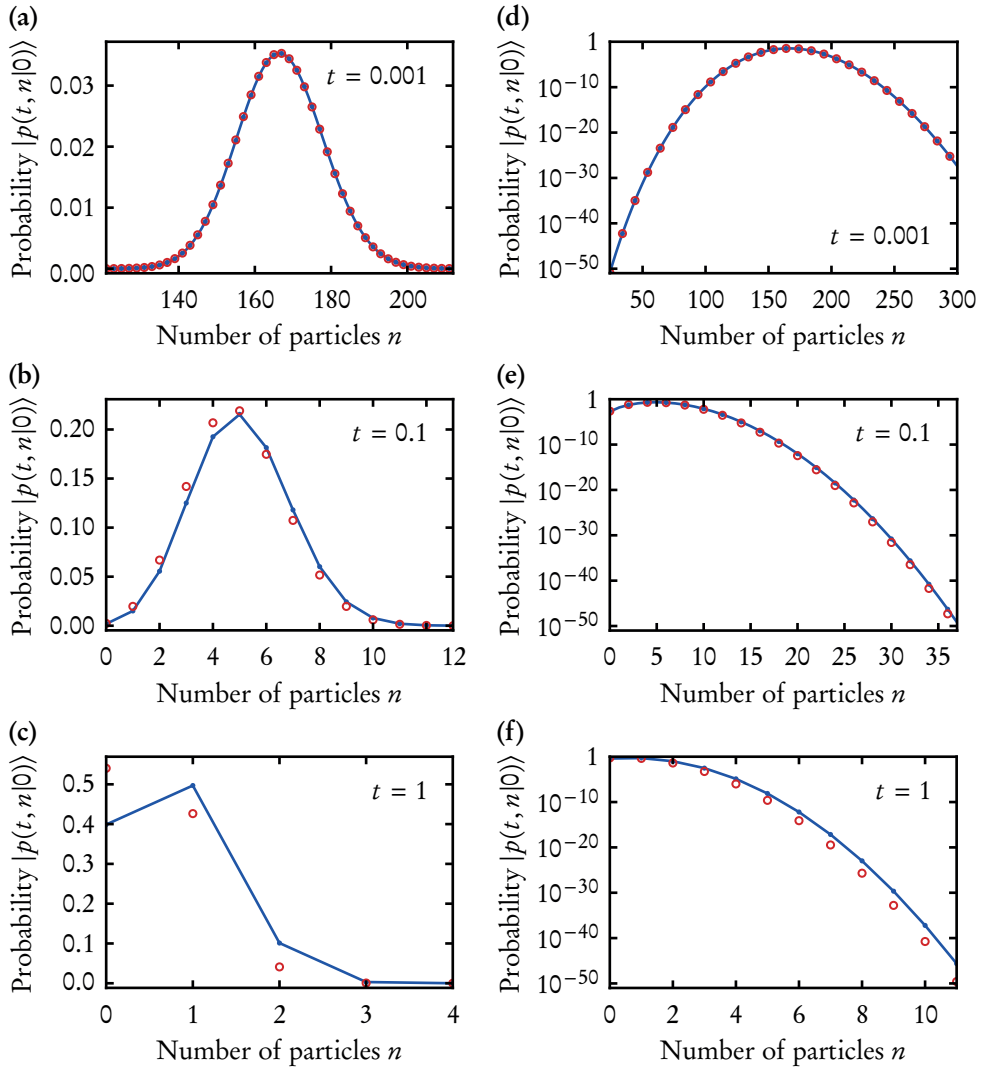


Figure 10 Solution of the binary annihilation reaction $2A \to \emptyset$ via the numerical integration of the master equation (blue lines) and approximation of the process using the stationary path method in section 7.2 (red circles). The figures in the left column show the marginalized probability distribution $|p(t,n|0)\rangle$ at times (*a*) $t = 0.001$, (*b*) $t = 0.1$, and (*c*) $t = 1$ on a linear scale. The figures in the right column show the distribution at times (d) $t = 0.001$, (e) $t = 0.1$, and (f) $t = 1$ on a logarithmic scale. The rate coefficient of the process was set to $\mu_\tau = 1$. For the numerical integration, the master equation was truncated at $n = 450$. The marginalized distribution was initially of Poisson shape with mean $x(0) = 250$. Its leading-order approximation is given in (404).

7.3 Backward path integral approach

The above method can be readily extended to a novel path integral representation of the marginalized distribution

$$|p(t,n|t_0)\rangle_{x(t_0)} = \sum_{n_0=0}^{\infty} p(t,n|t_0,n_0)\frac{x(t_0)^{n_0}\mathrm{e}^{-x(t_0)}}{n_0!}\,. \tag{412}$$

Unlike in the previous section, no saddle-point approximation is required to evaluate $|p(t,n|t_0)\rangle_{x(t_0)}$. The resulting path integral can also be evaluated beyond leading order, at least numerically. In the next section, we outline such an evaluation for the binary annihilation reaction $2\,A \to \emptyset$. On the downside, the leading order of the "backward" approach does not respect the normalization of $|p(t,n|t_0)\rangle_{x(t_0)}$. The normalization of the distribution has to be implemented by hand or by evaluating higher order terms.

To represent the marginalized distribution (412) by a path integral, we require that for a given particle number n and for a given initial mean $x(t_0)$, there exist functions $\tilde{x}(\tau)$ and $\tilde{q}(\tau)$ fulfilling

$$\partial_\tau \tilde{x} = \frac{\partial \mathcal{Q}_\tau^\dagger(\tilde{x},\tilde{q})}{\partial \tilde{q}} \text{ with } \tilde{x}(t_0) = x(t_0) \text{ and} \tag{413}$$

$$\partial_{-\tau} \tilde{q} = \frac{\partial \mathcal{Q}_\tau^\dagger(\tilde{x},\tilde{q})}{\partial \tilde{x}} \text{ with } \tilde{q}(t) = \frac{n}{\tilde{x}(t)}\,. \tag{414}$$

Since the transition operators $\mathcal{Q}_\tau^\dagger$ in (156) and $\mathcal{Q}_\tau$ in (69) are connected via $\mathcal{Q}_\tau^\dagger(\tilde{x},\tilde{q}) = \mathcal{Q}_\tau(\tilde{q},\tilde{x})$, the above equations differ from the previous equations (388) and (389) only in the final condition on $\tilde{q}$. As shown below, the marginalized distribution (412) can then be expressed as

$$|p(t,n|t_0)\rangle_{x(t_0)} = \frac{\tilde{x}(t)^n \mathrm{e}^{-\tilde{\mathcal{S}}^\dagger}}{n!} \fint_{(t_0}^{t]} \mathrm{e}^{-\Delta\mathcal{S}^\dagger}\big|_{\Delta x(t_0)=0} \tag{415}$$

$$\text{with } \tilde{\mathcal{S}}^\dagger := x(t_0) + \int_{t_0}^{t} \mathrm{d}\tau \left(\tilde{q}\partial_\tau \tilde{x} - \mathcal{Q}_\tau^\dagger(\tilde{x},\tilde{q})\right) \tag{416}$$

$$\text{and } \Delta\mathcal{S}^\dagger := n\Big(\frac{\zeta\Delta x(t)}{\tilde{x}(t)} - \ln\Big[1 + \frac{\zeta\Delta x(t)}{\tilde{x}(t)}\Big]\Big) + \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}\Delta q\,\partial_\tau \Delta x - \Delta\mathcal{Q}_\tau^\dagger\right). \tag{417}$$

The variables Δx and Δq again represent deviations from the stationary paths $\tilde{x}$ and $\tilde{q}$. The transition operator $\Delta\mathcal{Q}_\tau^\dagger$ encompasses all the terms of an expansion of $\mathcal{Q}_\tau^\dagger(\zeta\Delta x + \tilde{x}, \zeta^{-1}\mathrm{i}\Delta q + \tilde{q})$ that are of second or higher order in the deviations. A

Taylor expansion of the action shows that it is free of terms that are linear in Δx and Δq.

The derivation of the above representation proceeds analogously to section 7.1. It begins by using $\tilde{x}(t_0) = x(t_0)$ and the basis function $|n\rangle_{\tau,\Delta x} = \frac{1}{n!}(\zeta\Delta x + \tilde{x})^n \mathrm{e}^{-\tilde{q}(\zeta\Delta x+\tilde{x})}$ from (159) to rewrite the right-hand side of the marginalized distribution (412) as

$$\mathrm{e}^{(\tilde{q}(t_0)-1)\tilde{x}(t_0)} \sum_{n_0=0}^{\infty} p(t,n|t_0,n_0)|n_0\rangle_{t_0,\Delta x(t_0)}\big|_{\Delta x(t_0)=0}\,. \tag{418}$$

The sum in this expression can be represented by the backward path integral (210), turning the expression into

$$\mathrm{e}^{(\tilde{q}(t_0)-1)\tilde{x}(t_0)} \int_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger}\,|n\rangle_{t,\Delta x(t)}\big|_{\Delta x(t_0)=0}\,. \tag{419}$$

Recalling the definition of $\mathcal{E}_\tau$ in (162), the transition operator

$$\widetilde{\mathcal{Q}}^\dagger_\tau(\Delta x, \mathrm{i}\Delta q) = \mathcal{Q}^\dagger_\tau(\zeta\Delta x + \tilde{x}, \zeta^{-1}\mathrm{i}\Delta q + \tilde{q}) - \mathcal{E}_\tau(\zeta\Delta x + \tilde{x}, \zeta^{-1}\mathrm{i}\Delta q + \tilde{q}) \tag{420}$$

can now be expanded in the deviations. We thereby obtain the action[19]

$$\begin{aligned}\mathcal{S}^\dagger &= \int_{t_0}^{t} \mathrm{d}\tau\,\big(\mathrm{i}\Delta q\,\partial_\tau\Delta x - \widetilde{\mathcal{Q}}^\dagger_\tau(\Delta x, \mathrm{i}\Delta q)\big) \qquad (421)\\ &= \tilde{q}(t_0)\tilde{x}(t_0) - \tilde{q}(t)\tilde{x}(t) + \int_{t_0}^{t} \mathrm{d}\tau\,\big(\tilde{q}\,\partial_\tau\tilde{x} - \mathcal{Q}^\dagger_\tau(\tilde{x},\tilde{q})\big) + \int_{t_0}^{t} \mathrm{d}\tau\,\big(\mathrm{i}\Delta q\,\partial_\tau\Delta x - \Delta\mathcal{Q}^\dagger_\tau\big)\,.\end{aligned}$$

The path integral representation (415) follows upon inserting this action into (419) and employing the final condition $\tilde{q}(t) = n/\tilde{x}(t)$.

7.4 Binary annihilation

To complement the approximation of the binary annihilation reaction $2\,A \to \emptyset$ with rate coefficient μ_τ in section 7.2, we now perform an approximation of the process using the backward path integral representation (415) of the marginalized

19 The proof employs

$$\begin{aligned}\widetilde{\mathcal{Q}}^\dagger_\tau(\Delta x, \mathrm{i}\Delta q) &= \mathcal{Q}^\dagger_\tau(\tilde{x},\tilde{q}) + \frac{\partial\mathcal{Q}^\dagger_\tau}{\partial\tilde{x}}\zeta\Delta x + \frac{\partial\mathcal{Q}^\dagger_\tau}{\partial\tilde{q}}\zeta^{-1}\mathrm{i}\Delta q + \Delta\mathcal{Q}^\dagger_\tau - (\partial_\tau\tilde{x})\zeta^{-1}\mathrm{i}\Delta q + (\partial_\tau\tilde{q})(\zeta\Delta x + \tilde{x})\\ &= \partial_\tau(\tilde{x}\tilde{q}) - \big(\tilde{q}\,\partial_\tau\tilde{x} - \mathcal{Q}^\dagger_\tau(\tilde{x},\tilde{q})\big) + \Delta\mathcal{Q}^\dagger_\tau\,.\end{aligned}$$

distribution. We first perform the approximation in leading order, which amounts to the evaluation of the pre-factor of (415). For the simple growth process $\emptyset \to A$ and for the linear decay process $A \to \emptyset$, the pre-factor provides exact solutions.

7.4 a Leading order

The evaluation of the leading-order approximation

$$|p(t,n|t_0)\rangle_{x(t_0)} \approx \frac{\tilde{x}(t)^n \mathrm{e}^{-\tilde{\mathcal{S}}^\dagger}}{n!} \tag{422}$$

proceeds very much like the derivation in section 7.2. Using the adjoint transition operator $\mathcal{Q}_\tau^\dagger(c_\tau, a_\tau) = \mu_\tau c_\tau^2(1-a_\tau^2)$ from (156), one finds that Hamilton's equations

$$\partial_\tau \tilde{x} = -2\mu_\tau \tilde{q}\tilde{x}^2 \text{ with } \tilde{x}(t_0) = x(t_0) \text{ and} \tag{423}$$

$$\partial_{-\tau}\tilde{q} = 2\mu_\tau(1-\tilde{q}^2)\tilde{x} \text{ with } \tilde{q}(t) = n/\tilde{x}(t) \tag{424}$$

agree with (396) and (397), apart from the final condition on $\tilde{q}$. The conservation of $(\tilde{q}^2-1)\tilde{x}^2$ and the asymptotic deterministic limit $\bar{x}_\infty(t) = (2\int_{t_0}^t \mathrm{d}\tau\, \mu_\tau)^{-1}$ can be used to rewrite the action (416) as

$$\tilde{\mathcal{S}}^\dagger = x(t_0) + \tilde{q}(t)\tilde{x}(t) - \tilde{q}(t_0)\tilde{x}(t_0) + \int_{t_0}^t \mathrm{d}\tau \left(\tilde{x}\partial_{-\tau}\tilde{q} - \mathcal{Q}_\tau^\dagger(\tilde{x},\tilde{q})\right) \tag{425}$$

$$= n + (1-\tilde{q}(t_0))\tilde{x}(t_0) + \frac{(1-\tilde{q}(t_0)^2)x(t_0)^2}{2\bar{x}_\infty(t)} \tag{426}$$

$$= n + (1-\tilde{q}(t_0))\tilde{x}(t_0) + \frac{(1-\tilde{q}(t)^2)\tilde{x}(t)^2}{2\bar{x}_\infty(t)} \tag{427}$$

$$= n + \Big(1 - \sqrt{1 - \frac{\tilde{x}(t)^2 - n^2}{x(t_0)^2}}\Big)x(t_0) + \frac{\tilde{x}(t)^2 - n^2}{2\bar{x}_\infty(t)}\,. \tag{428}$$

In addition, the conservation law can be used to infer the flow equation

$$\partial_\tau \tilde{q} = -2\mu_\tau \sqrt{1-\tilde{q}^2}\sqrt{\tilde{x}(t)^2 - n^2}\,. \tag{429}$$

Here we assume $\tilde{q}(t) < 1$ but the derivation can also be performed for $\tilde{q}(t) > 1$. The equation is implicitly solved by

$$\arccos\sqrt{1 - \frac{\tilde{x}(t)^2 - n^2}{x(t_0)^2}} + \frac{\sqrt{\tilde{x}(t)^2 - n^2}}{\bar{\tilde{x}}_\infty(t)} = \arccos\frac{n}{\tilde{x}(t)}\,. \tag{430}$$

For a given initial mean $x(t_0)$, this equation can be solved numerically for $\tilde{x}(t)$, which is then inserted into (422). Upon normalizing the distribution by hand, it provides a reasonable approximation of the process. The quality of the approximation comes close to the quality of the approximation discussed in section 7.2.

7.4 b Beyond leading order

Instead of normalizing the function (422) by hand, it can be normalized by evaluating the path integral (415). In the following, we perform this evaluation, but only after restricting the action $\Delta\mathcal{S}^\dagger$ in (417) to terms that are of second order in the deviations Δx and Δq, i.e. to

$$\Delta\mathcal{S}^\dagger \approx -\frac{(\Delta x(t))^2}{2}\frac{n}{\tilde{x}(t)^2} + \int_{t_0}^{t} \mathrm{d}\tau \left(\mathrm{i}\Delta q\,\partial_\tau \Delta x - \Delta\mathcal{Q}_\tau^\dagger\right). \tag{431}$$

The corresponding transition operator

$$\Delta\mathcal{Q}_\tau^\dagger \approx \Delta x\mathrm{i}\Delta q\,\frac{\partial^2\mathcal{Q}_\tau^\dagger}{\partial\tilde{x}\partial\tilde{q}} + \frac{(\zeta^{-1}\mathrm{i}\Delta q)^2}{2}\frac{\partial^2\mathcal{Q}_\tau^\dagger}{\partial\tilde{q}^2} + \frac{(\zeta\Delta x)^2}{2}\frac{\partial^2\mathcal{Q}_\tau^\dagger}{\partial\tilde{x}^2} \tag{432}$$

$$= \mathrm{i}\Delta q\Delta x\,\alpha_\tau - \frac{(\Delta q)^2}{2}\beta_\tau + \frac{(\Delta x)^2}{2}c_\tau \tag{433}$$

is obtained as explained in section 7.3 (with $\zeta := \mathrm{i}$). Its coefficients read

$$\alpha_\tau := -4\mu_\tau\tilde{x}\tilde{q}\,, \quad \beta_\tau := 2\mu_\tau\tilde{x}^2\,, \quad c_\tau := 2\mu_\tau(\tilde{q}^2 - 1)\,. \tag{434}$$

The path integral can now be evaluated in one of the following two ways.

On the one hand, the transition operator $\Delta\mathcal{Q}_\tau^\dagger$ has the same form as the one in section 4.3 and thus can be treated in the same way. Following appendix D, the path integral (415) can be expressed in terms of the following average over a Wiener process W:

$$\int_{(t_0}^{t]} \mathrm{e}^{-\Delta\mathcal{S}^\dagger}\big|_{\Delta x(t_0)=0} = \left\langle\!\!\left\langle \exp\left(\frac{(\Delta x(t))^2}{2}\frac{n}{\tilde{x}(t)^2} + \int_{t_0}^{t}\mathrm{d}\tau\,\frac{(\Delta x)^2}{2}c_\tau\right)\right\rangle\!\!\right\rangle_W. \tag{435}$$

Here, $\Delta x(\tau)$ solves the Itô SDE

$$\mathrm{d}\Delta x = \alpha_\tau\Delta x\,\mathrm{d}\tau + \sqrt{\beta_\tau}\,\mathrm{d}W(\tau), \tag{436}$$

with initial value $\Delta x(t_0) = 0$. Unfortunately, the computation of a sufficient number of sample paths to approximate the average (435) was found to be rather slow.

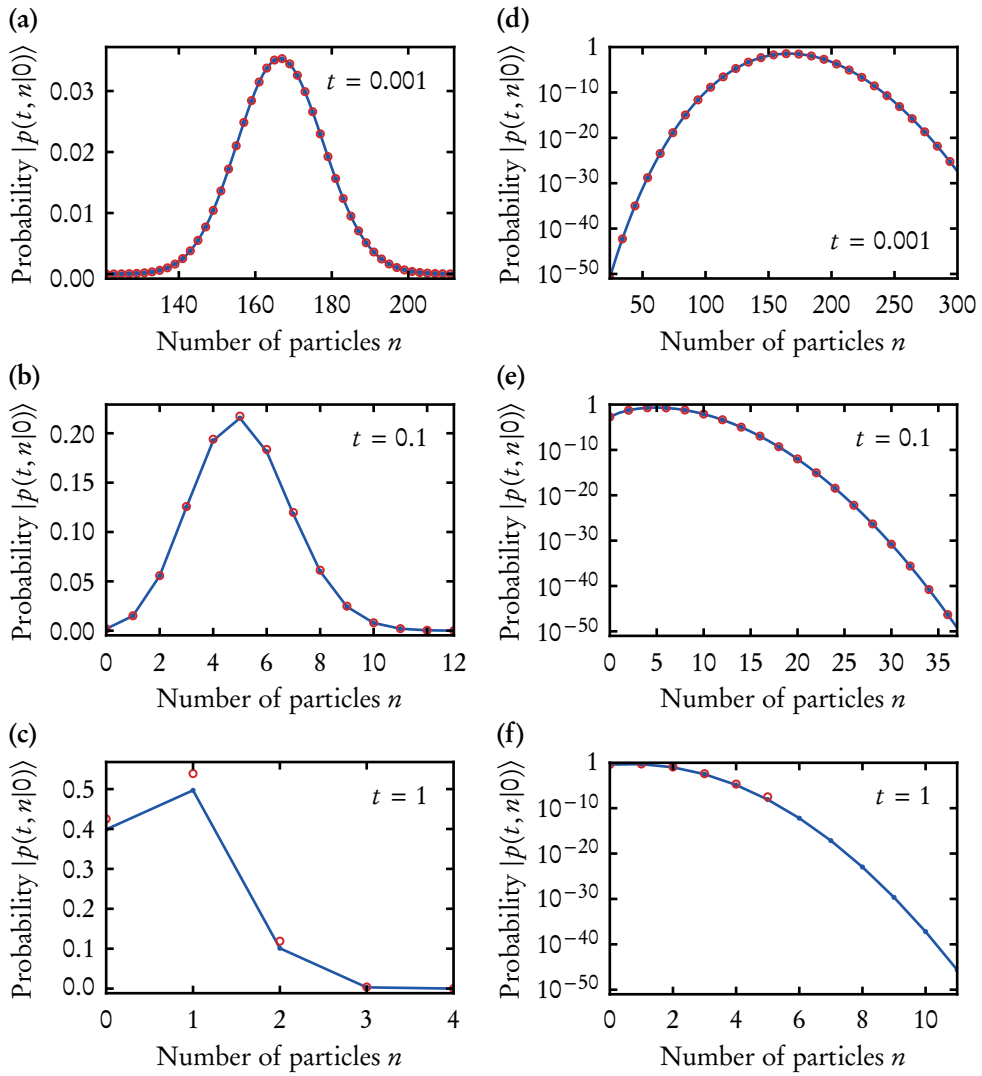


Figure 11 Solution of the binary annihilation reaction $2A \to \emptyset$ via the numerical integration of the master equation (blue lines) and approximation of the process using the stationary path method in section 7.4 (red circles). The figures in the left column show the marginalized probability distribution $|p(t,n|0)\rangle$ at times (*a*) $t = 0.001$, (*b*) $t = 0.1$, and (*c*) $t = 1$ on a linear scale. The figures in the right column show the distribution at times (d) $t = 0.001$, (e) $t = 0.1$, and (f) $t = 1$ on a logarithmic scale. The rate coefficient of the process was set to $\mu_\tau = 1$. For the numerical integration, the master equation was truncated at $n = 450$. The marginalized distribution was initially of Poisson shape with mean $x(0) = 250$. Its leading-order approximation (422) was corrected by the factor $1/\sqrt{\det A}$ (cf. section 7.4 b). The numerical solution of the fixed point equation (430) failed for large times (missing circles in (f)).

As an alternative to the above way, one can discretize the action $\Delta\mathcal{S}^\dagger$ and cast it into the quadratic form $\Delta\mathcal{S}^\dagger = \frac{1}{2}\xi^\top A\xi$ with $\xi := \{\Delta q_j, \Delta x_j\}_{j=1,\dots,N}$. The matrix A is symmetric and tridiagonal. It is specified by its diagonal

$$(\beta_0\Delta t, -c_1\Delta t, \dots, -c_{N-1}\Delta t, \beta_{N-1}\Delta t, -n/\tilde{x}(t)^2)$$

and by its super- and sub-diagonal

$$(\mathrm{i}, -\mathrm{i} - \mathrm{i}\alpha_1\Delta t, \mathrm{i}, \dots, -\mathrm{i} - \mathrm{i}\alpha_{N-1}\Delta t, \mathrm{i})\,.$$

If A is also positive-definite, the path integral (415) evaluates to $1/\sqrt{\det A}$. Curiously, we found that this factor even matches the average (435) when A is not positive-definite. Therefore, we used the factor $1/\sqrt{\det A}$ to correct the approximation (422) and evaluated the resulting expression for various times t. A comparison with distributions obtained via a numeric integration of the master equation is provided in figure 11. Apparently, the quality of the approximation is even better than the quality of the approximation discussed in section 7.2, at least for early times. However, upon approaching the asymptotic limit of the process, it becomes increasingly difficult to numerically solve (430) for a fixed point $\tilde{x}(t)$.

7.5 Résumé

The solution of a master equation can be approximated by expanding its forward or backward path integral representations around stationary paths of their respective actions. Such an expansion proves particularly useful in the approximation of "rare-event" probabilities in a distribution's tail. Algorithms such as the stochastic simulation algorithm (SSA) of Gillespie typically perform poorly for this purpose. Whereas the expansion of the forward path integral representation provides the ordinary probability generating function (377) as an intermediate step, the expansion of the backward path integral representation provides the marginalized distribution (412). In both cases, the stationary paths obey differential equations resembling Hamilton's equations from classical mechanics. In sections 7.2 and 7.4, we demonstrated the two approaches in approximating the binary annihilation reaction $2\,A \to \emptyset$. Our approximation based on the forward path integral amends an earlier computation of Elgart and Kamenev [40]. The computation requires a saddle-point approximation of Cauchy's differentiation formula to extract probabilities from the generating function. The advantage of the approach lies in the fact that its leading order term respects the normalization of the underlying probability distribution. The backward approach does not require the auxiliary saddle-point approximation but its leading order term is not normalized. In the approximation of the binary annihilation reaction, we demonstrated how the expansion of the backward path integral can be evaluated beyond leading order.

Future studies are needed to explore whether the two approaches are helpful in analysing processes with multiple types of particles and spatial degrees of freedom. The efficient evaluation of contributions beyond the leading order approximations also remains a challenge.

8 Summary and outlook

On sufficiently coarse time and length scales, many complex systems appear to evolve through finite jumps. Such a jump may be the production of an mRNA or protein in a gene regulatory network [143, 144, 147], the step of a molecular motor along a cytoskeletal filament [118–120], or the flipping of a spin [139–142]. The stochastic evolution of such processes is commonly modelled in terms of a master equation [9, 10]. This equation provides a generic description of a system's stochastic evolution under the following three conditions: First, time proceeds continuously. Second, the system's future is fully determined by the system's present state. Third, changes of the system's state proceed via discontinuous jumps.

In this chapter, we reviewed the mathematical theory of master equations and discussed analytical and numerical methods for their solution. Special attention was paid to methods that apply even when stochastic fluctuations are strong and to the representation of master equations by path integrals. In the following, we provide brief summaries of the discussed methods, which all have their own merits and limitations. Information on complementary approximation methods can be found in the text books [103, 115] and in the review [276].

The stochastic simulation algorithm (section 1.4)
The SSA [80–83] and its variations [83, 249–251] come closest to being all-purpose tools in solving the (forward) master equation numerically. The SSA enables the computation of sample paths with the correct probability of occurrence. Since these paths are statistically independent of one another, their computation can be easily distributed to individual processing units. Besides providing insight into a system's "typical" dynamics, the sample paths can be used to compute a histogram approximation of the master equation's solution or to approximate observables. In principle, the SSA can be applied to arbitrarily complex systems with multiple types of particles and possibly spatial degrees of freedom. Consequently, the algorithm is commonly used in biological studies [143, 144, 147] (see chapter III for its application to microbial range expansions). The SSA has been implemented in various large-scale simulation packages [252–259]. A fast but only approximate alternative to the SSA for the simulation of processes evolving on multiple time scales is τ-leaping [83, 251, 260–269]. Algorithms for the simulation of processes with time-dependent transition rates are discussed in [270, 271]. Let us note,

however, that for small systems, a direct numerical integration of the master equation may be more efficient than the use of the SSA.

Alternative path summation algorithms (section 1.4)
Given enough time, the SSA could be used to generate every possible sample path of a process (at least if the state space is finite). If every path is recorded just once, the corresponding histogram approximation of the conditional probability distribution $p(\tau, n|\tau_0, n_0)$ would agree with the "path summation representation" (23). Various alternatives to the SSA have been proposed for the numerical evaluation of this representation, mostly based on its Laplace transform [240–245]. Thus far, these algorithms have remained restricted to rather simple systems. The analytical evaluation of the sums and convolutions that are involved proves to be a challenge.

Exponentiation and uniformization (section 1.4)
If the state space of a system is finite, the (forward) master equation $\partial_\tau p(\tau|t_0) = Qp(\tau|t_0)$ is solved by the matrix exponential $p(\tau|\tau_0) = \mathrm{e}^{Q(\tau-\tau_0)}\mathbb{1}$. Here, Q represents the transition matrix of a process and $p(\tau|t_0)$ the matrix of the conditional probabilities $p(\tau, n|\tau_0, n_0)$. If the state space of the system is countably infinite but the exit rates from all states are bounded (i.e. $\sup_m |Q(m,m)| < \infty$), the conditional probability distribution can be represented by the "uniformization" formula (19). As the evaluation of this formula involves the computation of an infinite sum of matrix powers, its use is restricted to small systems and to systems whose transition matrices exhibit special symmetries. The same applies to the evaluation of the above matrix exponential note, however, the projection techniques proposed in [221, 222].

Flow equations (sections 2 and 3)
For the chemical master equation (27) of the reaction $k\,A \to l\,A$ with rate coefficient γ_τ, it is readily shown that the probability generating function

$$|g(\tau|t_0, n_0)\rangle = \sum_n |n\rangle\, p(\tau, n|t_0, n_0) \tag{437}$$

with basis function $|n\rangle_q = q^n$ obeys the linear partial differential equation, or "flow equation"

$$\partial_\tau |g\rangle = \gamma_\tau (q^l - q^k)\partial_q^k |g\rangle\,. \tag{438}$$

Analogously, the marginalized distribution

$$|p(t, n|\tau)\rangle = \sum_{n_0} p(t, n|\tau, n_0)|n_0\rangle \tag{439}$$

with Poisson basis function $|n_0\rangle_x = \frac{x^{n_0}\mathrm{e}^{-x}}{n_0!}$ obeys the backward-time flow equation

$$\partial_{-\tau}|p\rangle = \gamma_\tau x^k\left[(\partial_x+1)^l - (\partial_x+1)^k\right]|p\rangle\,. \tag{440}$$

After one has solved one of these equations, the conditional probability distribution can be recovered via the inverse transformations $p(\tau,n|t_0,n_0) = \langle n|g(\tau|t_0,n_0)\rangle$ or $p(t,n|\tau,n_0) = \langle n_0|p(t,n|\tau)\rangle$. In sections 2 and 3, we generalized the above approaches and formulated conditions under which the forward and backward master equations can be transformed into flow equations. Besides the flow equations obeyed by the generating function and the marginalized distribution, we also derived a flow equation obeyed by the generating functional

$$\langle g(\tau|t_0,n_0)| = \sum_n \langle n|p(\tau,n|t_0,n_0)\,. \tag{441}$$

For the Poisson basis function $|n\rangle_x = \frac{x^n\mathrm{e}^{-x}}{n!}$, the inverse transformation $p(\tau,n|t_0,n_0) = \langle g(\tau|t_0,n_0)|n\rangle$ of this functional recovers the Poisson representation of Gardiner and Chaturvedi [37, 38]. The Poisson representation can, for example, be used for the computation of mean extinction times [376], but we found the marginalized distribution to be more convenient for this purpose (cf. section 3.3).

Thus far, most methods that have been proposed for the analysis of the generating function's flow equation have only been applied to systems without spatial degrees of freedom. These methods include a variational approach [44, 88], spectral formulations and WKB approximations [39, 41–43, 330, 368, 369, 372, 429]. We outlined some of these methods in section 2.3. "Real-space" WKB approximations, which employ an exponential ansatz for the probability distribution rather than for the generating function, were not discussed in this thesis (information on these approximations can be found in [68, 69, 151, 175, 176, 181, 356–362, 364, 365]). WKB approximations often prove helpful in computing a mean extinction time or a (quasi)stationary probability distribution. Future studies could explore whether the flow equation obeyed by the marginalized distribution can be evaluated in terms of a WKB approximation or using spectral methods. Moreover, further research is needed to investigate whether the above methods may be helpful in studying processes with spatial degrees of freedom and multiple types of particles. For that purpose, it will be crucial to specify satisfactory boundary conditions for the flow equations obeyed by the generating function and the marginalized distribution (see [368] for a discussion of "lacking" boundary conditions in a study of the branching-annihilation reaction $A \to 2\,A$ and $2\,A \to \emptyset$).

Recently, we have been made aware of novel approaches to the generating function's flow equation based on so-called duality relations [430–434]. These approaches are not yet covered here.

Stochastic path integrals (sections 4–6)
Path integral representations of the master equation have proved invaluable tools in gaining analytical and numerical insight into the behaviour of stochastic processes, particularly in the vicinity of phase transitions [19, 74, 75, 101, 135–138, 167, 295–321, 323, 325–327, 403, 424–427]. A classical example of such a phase transition is the transition between an active and an absorbing state of a system [73]. In section 6, we showed that for a process whose initial number of particles is Poisson distributed with mean $x(t_0)$, the average of an observable $A(n)$ at time t can be represented by the path integral

$$\langle A\rangle_{x(t_0)} = \oint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} \langle\!\langle A\rangle\!\rangle_{x(t)} \tag{442}$$

$$\text{with } \langle\!\langle A\rangle\!\rangle_x := \sum_{n=0}^{\infty} \frac{x^n \mathrm{e}^{-x}}{n!} A(n)\,. \tag{443}$$

The above path integral can be derived both from the novel *backward path integral representation*

$$p(t,n|t_0,n_0) = \langle n_0|_{t_0} \oint_{(t_0}^{t]} \mathrm{e}^{-\mathcal{S}^\dagger} |n\rangle_t\,, \tag{444}$$

of the conditional probability distribution (cf. section 4), and from the *forward path integral representation*

$$p(t,n|t_0,n_0) = \langle n|_t \oint_{[t_0}^{t)} \mathrm{e}^{-\mathcal{S}} |n_0\rangle_{t_0} \tag{445}$$

(cf. section 5). The meanings of the integral signs and of the actions $\mathcal{S}$ and $\mathcal{S}^\dagger$ were explained in the respective sections. We derived both of the above representations of the conditional probability distribution from the flow equations discussed in the previous paragraph. An extension of the path integrals to systems with multiple types of particles or spatial degrees of freedom is straightforward. We demonstrated the use of the integrals in solving various elementary processes, including the pair generation process and a process with linear decay of diffusing particles. Although we did not discuss the application of renormalization group techniques, we showed how the above path integrals can be evaluated in terms of perturbation expansions using Feynman diagrams. Information on perturbative renormalization group techniques can be found in [74, 75], information on non-perturbative techniques in [413]. Besides the above path integrals, we showed how one can derive path integral representations for processes with continuous state spaces. These path integrals were based on Kramers-Moyal expansions of the respective backward and forward master equations. Upon truncating the expansion of the backward master

equation at the level of a diffusion approximation, we recovered a classic path integral representation of the (backward) Fokker-Planck equation [21–24]. We hope that our exposition of the path integrals helps in developing new methods for the analysis of stochastic processes. Future studies may focus directly on the backward or forward path integral representations (444) and (445) of the conditional probability distribution, or use the representation (442) to explore how (arbitrarily high) moments of the particle number behave in the vicinity of phase transitions.

Stationary path approximations (section 7)
Elgart and Kamenev recently showed how the forward path integral representation (445) can be approximated by expanding its action around "stationary paths" [40]. The paths obey Hamilton's equations of the form

$$\partial_\tau \tilde{x} = \frac{\partial \mathcal{Q}_\tau}{\partial \tilde{q}} \text{ and } \partial_{-\tau} \tilde{q} = \frac{\partial \mathcal{Q}_\tau}{\partial \tilde{x}}\,, \tag{446}$$

with the transition operator $\mathcal{Q}_\tau$ acting as the "Hamiltonian". In section 7, we reviewed this "stationary path method" and showed how it can be extended to the backward path integral representation (444). We found that this backward approach does not require an auxiliary saddle-point approximation if the number of particles is initially Poisson distributed, but that a proper normalization of the probability distribution is only attained beyond leading order. Future work is needed to apply the method to systems with spatial degrees of freedom and multiple types of particles. The latter point also applies to the classification of phase transitions based on phase-space trajectories of the equations (446) as has been proposed in [322].

We hope that our discussion of the above methods inspires future research on master equations and that it helps researchers who are new to the field of stochastic processes to become acquainted with the theory of "stochastic" path integrals.

Acknowledgement

For critical reading and for providing valuable feedback to this chapter, I would like to thank Michael Assaf, Léonie Canet, Alexander Dobrinevski, Nigel Goldenfeld, Peter Grassberger, Baruch Meerson, Ralf Metzler, Mauro Mobilia, Harold P. de Vladar, Herbert Wagner, and Royce Zia. Moreover, I would like to thank Johannes Knebel, Isabella Krämer, and Cornelius Weig for helpful discussions during the preparation of the text.

A Proof of the Feynman-Kac formula in section 1.3

Here we provide a brief proof of a special case of the Feynman-Kac, or Kolmogorov, formula (see section 4.3.5 in [115]). In particular, we assume that for $\tau \in [t_0, t]$, a function $u(\tau, x)$ obeys the linear PDE

$$\partial_{-\tau} u(\tau, x) = \alpha_\tau(x)\partial_x u + \frac{1}{2}\beta_\tau(x)\partial_x^2 u \tag{447}$$

$$\text{with } \textit{final} \text{ value } u(t, x) = G(x)\,. \tag{448}$$

The function α_τ is called a drift coefficient and β_τ a diffusion coefficient. According to the Feynman-Kac formula, the above PDE is solved by

$$u(\tau, x) = \left\langle\!\left\langle G(X(t)) \right\rangle\!\right\rangle_W \tag{449}$$

with $\langle\!\langle \cdot \rangle\!\rangle_W$ representing an average over realizations of the Wiener process W. The function $X(s)$ with $s \in [\tau, t]$ and $\tau \geq t_0$ represents a solution of the Itô stochastic differential equation (SDE)

$$\mathrm{d}X(s) = \alpha_s(X(s))\,\mathrm{d}s + \sqrt{\beta_s(X(s))}\,\mathrm{d}W(s) \tag{450}$$

$$\text{with } \textit{initial} \text{ value } X(\tau) = x\,. \tag{451}$$

If the initial value x of the SDE is chosen as a real number, the drift coefficient α_s is a real function, and the diffusion coefficient β_s as a real non-negative function, the "sample path" $X(s)$ assumes only real values along its temporal evolution. In a multivariate extension of the Feynman-Kac formula, one requires a matrix $\sqrt{\beta_s} \coloneqq \gamma_s$ fulfilling $\gamma_s \gamma_s^\intercal = \beta_s$. If β_s is symmetric and positive-semidefinite, one may choose γ_s as its unique symmetric and positive-semidefinite square root [116].

The solution (3) of the backward Fokker-Planck equation (2) constitutes a special case of the above formula. There, the independent variable is x_0 instead of x, and $u(\tau, x_0)$ is the conditional probability distribution $p(t, x|\tau, x_0)$ (with x being an arbitrary parameter). The final value of the distribution is $G(x_0) = \delta(x - x_0)$. The Feynman-Kac formula (449) then implies that $\left\langle\!\left\langle \delta(x - X(t)) \right\rangle\!\right\rangle_W$ solves the backward Fokker-Planck equation (2). Note that in the main text, we use a small letter to denote the sample path.

To prove the Feynman-Kac formula (449), we assume that $X(s)$ solves the SDE (450). By Itô's Lemma and the PDE (447), it then holds that

$$\mathrm{d}u(s,X(s)) = \frac{\partial u}{\partial s}\,\mathrm{d}s + \frac{\partial u}{\partial X(s)}\,\mathrm{d}X(s) + \frac{1}{2}\frac{\partial^2 u}{\partial X(s)^2}\,\mathrm{d}X(s)^2 \tag{452}$$

$$= \Big(\frac{\partial u}{\partial s} + \alpha_s(X(s))\frac{\partial u}{\partial X(s)} + \frac{1}{2}\beta_s(X(s))\frac{\partial^2 u}{\partial X(s)^2}\Big)\,\mathrm{d}s + \frac{\partial u}{\partial X(s)}\sqrt{\beta_s(X(s))}\,\mathrm{d}W(s)$$

$$= \frac{\partial u(s,X(s))}{\partial X(s)}\sqrt{\beta_s(X(s))}\,\mathrm{d}W(s)\,. \tag{453}$$

As the next step, we integrate this differential from $s = \tau$ to $s = t$ and average the result over realizations of the Wiener process W. Since $\langle\!\langle \mathrm{d}W \rangle\!\rangle_W = 0$, it follows that $\langle\!\langle u(\tau, X(\tau)) \rangle\!\rangle_W = \langle\!\langle u(t, X(t)) \rangle\!\rangle_W$. This expression coincides with the Feynman-Kac formula (449) because the initial condition (451) implies that $u(\tau, X(\tau)) = u(\tau, x)$ does not depend on the Wiener process, and because the final condition (448) implies that $u(t, X(t)) = G(X(t))$.

B Proof of the path summation representation in section 1.4

In the following, we prove that the path summation representation (23) solves the forward master equation (12) if the transition rate $w(n,m)$ is independent of time. Hence, the process is homogeneous in time and we may choose $t_0 := 0$. After defining $d(n,m) := \delta_{n,m} w(m) = \delta_{n,m} \sum_k w(k,m)$, we first rewrite the master equation as

$$\partial_\tau \boldsymbol{p}(\tau|0,n_0) = (w-d)\boldsymbol{p}(\tau|\cdot) \tag{454}$$

with the probability vector $\boldsymbol{p}(\tau|t_0,n_0)$. Note that the matrix notation assumes some mapping between the state space of n and an index set $I \subset \mathbb{N}$. Following [241], the Laplace transform

$$\mathcal{L}f(s) := \int_0^\infty \mathrm{d}\tau\, \mathrm{e}^{-s\tau} f(\tau) \tag{455}$$

of the above master equation is given by

$$s\mathcal{L}\boldsymbol{p}(s|\cdot) - \hat{\boldsymbol{e}}_{n_0} = (w-d)\mathcal{L}\boldsymbol{p}(s|\cdot)\,. \tag{456}$$

Here, $\boldsymbol{p}(0|0,n_0) = \hat{\boldsymbol{e}}_{n_0}$ represents a unit vector pointing in direction n_0. Note that s is a scalar but that w and d are matrices. Making use of the unit matrix $\mathbb{1}$, the

above expression can be solved for $\mathcal{L}\boldsymbol{p}(s|\cdot)$ and be rewritten as

$$\mathcal{L}\boldsymbol{p}(s|\cdot) = \left[\mathbb{1} - (s\mathbb{1} + d)^{-1}w\right]^{-1}(s\mathbb{1} + d)^{-1}\hat{\boldsymbol{e}}_{n_0}\,. \tag{457}$$

The value of the real part of s is determined by the parameter ε in the inverse Laplace transformation

$$f(t) = \frac{1}{2\pi\mathrm{i}}\int_{\varepsilon-\mathrm{i}\infty}^{\varepsilon+\mathrm{i}\infty} \mathrm{d}s\, \mathrm{e}^{st}\mathcal{L}f(s)\,. \tag{458}$$

We assume that ε can be chosen so large that the first factor in the solution (457) can be rewritten as a geometric series, i.e. as

$$\mathcal{L}\boldsymbol{p}(s|\cdot) = \sum_{J=0}^{\infty}\left[(s\mathbb{1} + d)^{-1}w\right]^{J}(s\mathbb{1} + d)^{-1}\hat{\boldsymbol{e}}_{n_0}\,. \tag{459}$$

This expression may be simplified by noting that

$$\left[(s\mathbb{1} + d)^{-1}w\right]_{n,m} = \frac{w(n,m)}{s + w(n)}\,. \tag{460}$$

By defining $\sum_{\{\mathcal{P}_J\}} := \sum_{n_1}\cdots\sum_{n_{J-1}}$, (459) becomes

$$\mathcal{L}p(s,n|\cdot) = \sum_{J=0}^{\infty}\sum_{\{\mathcal{P}_J\}}\left(\prod_{j=0}^{J-1} w(n_{j+1},n_j)\right)\prod_{j=0}^{J}\frac{1}{s + w(n_j)}\Big|_{n_J=n}\,. \tag{461}$$

Since an exponential function $f(\tau) := \mathrm{e}^{-\alpha\tau}\Theta(\tau)$ transforms as $\mathcal{L}f(s) = (s + \alpha)^{-1}$ and since the Laplace transform converts convolutions into products, we thus find that (461) agrees with the Laplace transform of the path summation representation (23).

C Solution of the random walk (sections 2.2 a and 3.2 a)

We here provide the conditional probability distribution $p(\tau, n|t_0, n_0)$ solving the random walk from sections 2.2 a and 3.2 a. In the first of these sections, we showed that this distribution is obtained by applying the functional $\langle n|f = \int_{-\pi}^{\pi}\frac{\mathrm{d}q}{2\pi}\mathrm{e}^{-\mathrm{i}nq}f(q)$ to the generating function $|g(\tau|t_0, n_0)\rangle$, which we derived as

$$\exp\Big((\mathrm{e}^{\mathrm{i}q} - 1)\int_{t_0}^{\tau}\mathrm{d}s\, r_s + (\mathrm{e}^{-\mathrm{i}q} - 1)\int_{t_0}^{\tau}\mathrm{d}s\, l_s + \mathrm{i}n_0 q\Big)\,.$$

Series expansions show that this expression can be rewritten as $\mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\,(r_s+l_s)}$ times

$$\sum_{k=0}^{\infty} \sum_{m=0}^{k} \frac{\left(\int_{t_0}^{\tau} \mathrm{d}s\, r_s\right)^m}{m!} \frac{\left(\int_{t_0}^{\tau} \mathrm{d}s\, l_s\right)^{k-m}}{(k-m)!} \mathrm{e}^{\mathrm{i}(2m-k+n_0)q} .$$

We ignore the constant pre-factor $\mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\,(r_s+l_s)}$ for now and apply the functional $\langle n|$ to this expression. After carefully noting the restrictions imposed by Kronecker deltas, the resulting expression reads

$$\sum_{k=0}^{\infty} \frac{\left(\int_{t_0}^{\tau} \mathrm{d}s\, r_s \cdot \int_{t_0}^{\tau} \mathrm{d}s\, l_s\right)^k}{k!(k+|n-n_0|)!} . \tag{462}$$

This sum can be expressed by a modified Bessel function of the first kind (10.25.2 in [337]), resulting in

$$\left(\frac{\int_{t_0}^{\tau} \mathrm{d}s\, r_s}{\int_{t_0}^{\tau} \mathrm{d}s\, l_s}\right)^{\frac{n-n_0}{2}} I_{n-n_0}\left(2\left(\int_{t_0}^{\tau} \mathrm{d}s\, r_s \int_{t_0}^{\tau} \mathrm{d}s\, l_s\right)^{\frac{1}{2}}\right) .$$

Multiplied with $\mathrm{e}^{-\int_{t_0}^{\tau} \mathrm{d}s\,(r_s+l_s)}$, this expression corresponds to a Skellam distribution with mean $\mu = n_0 + \int_{t_0}^{\tau} \mathrm{d}s\,(r_s - l_s)$ and variance $\sigma^2 = \int_{t_0}^{\tau} \mathrm{d}s\,(r_s + l_s)$.

D Evaluation of the backward path integral representation in section 4.3

In the following, we fill out the missing steps in section 4.3 and rewrite the backward path integral representation (213) of the marginalized distribution in terms of the (Q, X)-generating functional (221). Upon comparing the discrete-time approximation of the marginalized distribution (210) with its representation in (219), one observes that the (Q, X)-generating functional should read

$$\mathcal{Z}_{Q,X,N} := \int_1^{N-1} \mathrm{e}^{\mathrm{i}q_N x_N - \mathcal{S}_N^{\dagger}} \mathrm{e}^{Z_N} \tag{463}$$

$$\text{with } Z_N := \sum_{j=1}^{N} \Delta t\, Q_j x_{j-1} + \sum_{j=1}^{N-1} \Delta t\, X_{j-1} \mathrm{i}q_j . \tag{464}$$

A differentiation of (463) with respect to $\Delta t\, Q_j$ generates a factor x_{j-1}, a differentiation with respect to $\Delta t\, X_{j-1}$ a factor $\mathrm{i}q_j$. Upon recalling the definition of the

action $\mathcal{S}_N^\dagger$ in (211), the first exponential in (463) can be rewritten via

$$\begin{aligned} \mathrm{i}q_N x_N - \mathcal{S}_N^\dagger = & -\sum_{j=1}^{N-1} \mathrm{i}q_j \big(x_j - \big[x_{j-1} + \alpha_{t_{j-1}}(x_{j-1})\Delta t\big]\big) \\ & -\sum_{j=1}^{N-1} \frac{q_j^2}{2} \beta_{t_{j-1}}(x_{j-1})\Delta t + \sum_{j=1}^{N-1} \Delta t\, \mathcal{P}^\dagger_{t_{j-1}}(x_{j-1}, \mathrm{i}q_j) + \mathrm{i}q_N x_{N-1}\,. \end{aligned} \tag{465}$$

The equality holds up to corrections of $\mathrm{O}(\Delta t)$ (because of the sums, all remaining terms are of $\mathrm{O}(1)$). Note that the right hand side of (465) is independent of x_N. One can linearize the terms that are quadratic in q_j by the completion of a square. In particular, we write

$$\exp\Big(-\sum_{j=1}^{N-1} \frac{q_j^2}{2} \beta_{t_{j-1}}(x_{j-1})\Delta t\Big) = \Big\langle\!\Big\langle \exp\Big(\sum_{j=1}^{N-1} \mathrm{i}q_j \sqrt{\beta_{t_{j-1}}(x_{j-1})}\Delta W_j\Big)\Big\rangle\!\Big\rangle_W \tag{466}$$

with the average being defined as

$$\langle\!\langle\ \cdot\ \rangle\!\rangle_W := \Big(\prod_{j=1}^{N-1} \int_{\mathbb{R}} \mathrm{d}\Delta W_j\, \mathcal{G}_{0,\Delta t}(\Delta W_j)\Big)(\,\cdot\,)\,. \tag{467}$$

The average employs the Gaussian distribution

$$\mathcal{G}_{0,\Delta t}(\Delta W_j) = \frac{\mathrm{e}^{-(\Delta W_j)^2/(2\Delta t)}}{\sqrt{2\pi\Delta t}} \tag{468}$$

with zero mean and variance Δt (cf. (244)). The sequence $\Delta W_1, \ldots, \Delta W_{N-1}$ can be interpreted as the steps of a discretized Wiener process. To proceed with the derivation, we now move the perturbation operator $\mathcal{P}^\dagger$ to the front of the (Q, X)-generating functional (463) by writing

$$\mathrm{e}^{\sum_{j=1}^{N-1} \Delta t\, \mathcal{P}^\dagger_{t_{j-1}}(x_{j-1}, \mathrm{i}q_j)} \mathrm{e}^{Z_N} = \mathrm{e}^{\sum_{j=1}^{N-1} \Delta t\, \mathcal{P}^\dagger_{t_{j-1}}\left(\frac{1}{\Delta t}\frac{\partial}{\partial Q_j}, \frac{1}{\Delta t}\frac{\partial}{\partial X_{j-1}}\right)} \mathrm{e}^{Z_N}\,. \tag{469}$$

Upon combining all of the above steps, one finds that

$$\mathcal{Z}_{Q,X,N} = \mathrm{e}^{\mathrm{i}q_N \frac{1}{\Delta t}\frac{\partial}{\partial Q_N} + \sum_{j=1}^{N-1} \Delta t\, \mathcal{P}^\dagger_{t_{j-1}}\left(\frac{1}{\Delta t}\frac{\partial}{\partial Q_j}, \frac{1}{\Delta t}\frac{\partial}{\partial X_{j-1}}\right)} \mathcal{Z}^0_{Q,X,N} \tag{470}$$

with the definition

$$\mathcal{Z}_N^0 := \Big\langle\!\Big\langle \Big(\prod_{j=1}^{N-1} \int_{\mathbb{R}} \mathrm{d}x_j\, \delta\Big(x_j - x_{j-1} - \Big\{[\alpha_{t_{j-1}}(x_{j-1}) + X_{j-1}]\Delta t + \sqrt{\beta_{t_{j-1}}(x_{j-1})}\,\Delta W_j\Big\}\Big)\Big) \cdot \mathrm{e}^{\sum_{j=1}^{N} \Delta t\, Q_j x_{j-1} + \mathrm{O}(\Delta t)} \Big\rangle\!\Big\rangle_W . \tag{471}$$

The sequence of Dirac delta functions implies that x_j depends on X_i only for $i < j$. This property has to be kept in mind when the functional derivatives in (470) are evaluated in continuous-time. In the continuous-time limit, i.e. for $N \to \infty$ and $\Delta t \to 0$, one recovers the marginalized distribution (219) with generating functional (221) and Itô SDE (223).

E Evaluation of the forward path integral representation in section 5.3

The forward path integral (303) can be rewritten in terms of the $(\boldsymbol{X}, \boldsymbol{Q})$-generating functional (314) following the same steps as above. First, we compare (313) with the discrete-time representation (299). The comparison shows that the $(\boldsymbol{X}, \boldsymbol{Q})$-generating functional should be defined as

$$\mathcal{Z}_{X,Q,N} := \oint_1^{N-1} \mathrm{e}^{\mathrm{i}\boldsymbol{x}_0 \cdot \boldsymbol{q}_0} \mathrm{e}^{-\mathcal{S}_N} \mathrm{e}^{Z_N} . \tag{472}$$

$$\text{with } Z_N := \sum_{j=0}^{N-1} \Delta t\, \boldsymbol{X}_j \cdot \boldsymbol{q}_{j+1} + \sum_{j=1}^{N-1} \Delta t\, \boldsymbol{Q}_{j+1} \cdot \mathrm{i}\boldsymbol{x}_j . \tag{473}$$

Up to corrections of $\mathrm{O}(\Delta t)$, the action (300) can be rewritten as

$$\begin{aligned} \mathcal{S}_N = {} & \mathrm{i}\boldsymbol{x}_0 \cdot \boldsymbol{q}_0 + \sum_{j=1}^{N-1} \mathrm{i}\boldsymbol{x}_j \cdot \Big(\boldsymbol{q}_j - \big(\boldsymbol{q}_{j+1} + \boldsymbol{\alpha}_{t_j}(\boldsymbol{q}_{j+1})\Delta t\big)\Big) \\ & + \sum_{j=1}^{N-1} \frac{1}{2} \boldsymbol{x}_j^{\intercal} \beta_{t_j}(\boldsymbol{q}_{j+1}) \boldsymbol{x}_j \Delta t - \sum_{j=1}^{N-1} \Delta t\, \mathcal{P}_{t_j}(\boldsymbol{q}_{j+1}, \mathrm{i}\boldsymbol{x}_j) - \mathrm{i}\boldsymbol{x}_0 \cdot \boldsymbol{q}_1 . \end{aligned} \tag{474}$$

By making use of $\sqrt{\beta_\tau}\sqrt{\beta_\tau}^{\intercal} = \beta_\tau$ and

$$-\frac{1}{2}\boldsymbol{x}^{\intercal}\beta\boldsymbol{x}\Delta t - \frac{1}{2\Delta t}\big(\Delta \boldsymbol{W} - \mathrm{i}\sqrt{\beta}^{\intercal}\boldsymbol{x}\Delta t\big)^2 = -\frac{(\Delta \boldsymbol{W})^2}{2\Delta t} + \mathrm{i}\boldsymbol{x} \cdot \sqrt{\beta}\Delta \boldsymbol{W} , \tag{475}$$

the quadratic term can be linearized via

$$\exp\Big(-\sum_{j=1}^{N-1}\frac{1}{2}\boldsymbol{x}_j^{\intercal}\beta_{t_j}(\boldsymbol{q}_{j+1})\boldsymbol{x}_j\Delta t\Big)=\Big\langle\!\Big\langle\exp\Big(\sum_{j=1}^{N-1}\mathrm{i}\boldsymbol{x}_j\cdot\sqrt{\beta_{t_j}(\boldsymbol{q}_{j+1})}\Delta\boldsymbol{W}_j\Big)\Big\rangle\!\Big\rangle_W. \quad (476)$$

Here we introduced the following average over Gaussian distributions with variance Δt:

$$\langle\!\langle\,\cdot\,\rangle\!\rangle_W:=\Big(\prod_{j=1}^{N-1}\int_{\mathbb{R}^{|\mathbb{L}|}}\mathrm{d}\Delta\boldsymbol{W}_j\,\mathcal{G}_{0,\Delta t}(\Delta\boldsymbol{W}_j)\Big)(\,\cdot\,). \quad (477)$$

As the next step, the perturbation operator $\mathcal{P}$ is moved to the front of (472) by writing

$$\mathrm{e}^{\sum_{j=1}^{N-1}\Delta t\,\mathcal{P}_{t_j}(\boldsymbol{q}_{j+1},\mathrm{i}\boldsymbol{x}_j)}\mathrm{e}^{Z_N}=\mathrm{e}^{\sum_{j=1}^{N-1}\Delta t\,\mathcal{P}_{t_j}\left(\frac{1}{\Delta t}\frac{\partial}{\partial\boldsymbol{X}_j},\frac{1}{\Delta t}\frac{\partial}{\partial\boldsymbol{Q}_{j+1}}\right)}\mathrm{e}^{Z_N}. \quad (478)$$

After combining all of the above steps, one obtains

$$\mathcal{Z}_N=\mathrm{e}^{\sum_{j=1}^{N-1}\Delta t\,\mathcal{P}_{t_j}\left(\frac{1}{\Delta t}\frac{\partial}{\partial\boldsymbol{X}_j},\frac{1}{\Delta t}\frac{\partial}{\partial\boldsymbol{Q}_{j+1}}\right)+\frac{1}{\Delta t}\frac{\partial}{\partial\boldsymbol{X}_0}\cdot\mathrm{i}\boldsymbol{x}_0}\mathcal{Z}_N^0 \quad (479)$$

with the definition

$$\mathcal{Z}_N^0=\Big\langle\!\Big\langle\Big(\prod_{j=1}^{N-1}\int_{\mathbb{R}^{|\mathbb{L}|}}\mathrm{d}\boldsymbol{q}_j\,\delta\Big(\boldsymbol{q}_j-\boldsymbol{q}_{j+1}-\big((\boldsymbol{\alpha}_{t_j}(\boldsymbol{q}_{j+1})+\boldsymbol{Q}_{j+1})\Delta t+\sqrt{\beta_{t_j}(\boldsymbol{q}_{j+1})}\,\Delta\boldsymbol{W}_j\big)\Big)\Big)$$
$$\cdot\,\mathrm{e}^{\sum_{j=0}^{N-1}\Delta t\,\boldsymbol{X}_j\cdot\boldsymbol{q}_{j+1}+\mathrm{O}(\Delta t)}\Big\rangle\!\Big\rangle_W. \quad (480)$$

In the continuous-time limit $\Delta t\to 0$, one obtains the representation (313)–(315). Note that the above derivation implies that q_j depends on Q_i only for $i>j$.

II Evolutionary zero-sum games and driven-dissipative quantum systems

1 Introduction

Evolutionary game theory (EGT) explores how the popularity of certain "strategies" changes in a population [177, 435–437]. Such a strategy may be to keep silent or blame your accomplice after a robbery (prisoner's dilemma [438]), to swerve your car or hold it steady while driving on the wrong side of the road (game of chicken [439]), or to play rock in a game of rock-paper-scissors [177, 436]. The rock-paper-scissors game constitutes the simplest game with a cyclic, non-transitive dominance between its strategies. A biological system in which such a cyclic dominance can be implemented is explored in chapter III, where we study range expansions of three *Escherichia coli* strains. But for our present purpose, another property of the rock-paper-scissors game is more relevant — namely, that it constitutes a zero-sum game [440].

The following sections introduce the mathematical theory of zero-sum games. The defining properties of these games can be understood from the viewpoint of classical game theory (CGT) by considering the interaction of two "agents". We then turn towards EGT, in which games are played within a whole population of agents. The success and failure of the agents' strategies feed back into the popularity of the strategies. The frequency at which the strategies are used in the population can be modelled both on a deterministic and a stochastic level. The stochastic model is based on a master equation. Surprisingly, just the same master equation was recently derived also in another context, namely in a study on the condensation of bosons in driven-dissipative quantum systems [1]. This condensation phenomenon, and the general stability of zero-sum games, were the focus of our publications "Coexistence and survival in conservative Lotka-Volterra networks" [2] and "Evolutionary games of condensates in coupled birth-death processes" [3], which are reprinted in sections 2 and 3. The following text serves as an introduction to these publications.

1.1 Zero-sum games

To set the stage for our discussion of zero-sum games, we consider a game between two agents. Both of the agents may choose one strategy out of the S different choices $\{E_1, \ldots, E_S\}$. In the following, we assume that agent 1 has opted for

(a)

Agent 1 \ Agent 2	E_1	E_2	E_3
E_1	0	A_{12}	$-A_{31}$
E_2	$-A_{12}$	0	A_{23}
E_3	A_{31}	$-A_{23}$	0

(b)

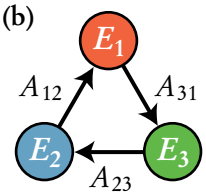


Figure 1 Two representations of the payoff matrix A of the rock-paper-scissors game (with strategies $E_1 :=$ rock, $E_2 :=$ scissors, and $E_3 :=$ paper). **(a)** A matrix element A_{ij} represents the payoff an agent 1 with strategy E_i receives or looses in a game against an agent 2 with strategy E_j. The payoff matrix is antisymmetric because the gain of one agent equals the loss of the other agent (zero-sum game). **(b)** Representation of the payoff matrix as a directed graph. A labelled arrow from strategy E_j to E_i (coloured disks) denotes the payoff A_{ij} the agent with strategy E_j provides to the agent with strategy E_i during a game.

strategy E_i and that agent 2 has opted for strategy E_j (with $i, j \in \{1, \ldots, S\}$). The performance of a strategy shall not depend on the agent using it. Games with this property are called symmetric games. The outcome of a game between the two agents is now decided by assigning the "payoff" A_{ij} to agent 1 and the payoff A_{ji} to agent 2. Both of these payoffs can be identified with elements of the same payoff matrix A because the game is symmetric. Provided the two payoffs balance one another irrespective of the strategies that are used (i.e. $A_{ji} = -A_{ij}$ for all $i, j \in \{1, \ldots, S\}$), the game is called a zero-sum game. In other words, a game is a zero-sum game if its payoff matrix is antisymmetric (i.e. $A^\intercal = -A$). Figure 1 shows the antisymmetric payoff matrix of the rock-paper-scissors game.

1.2 Game theory and the deterministic time evolution of zero-sum games

Classical game theory (CGT) was developed to explain the decision-making of rational (egocentric) agents in parlour games [441], assuming that what is learned here extends to the decision-making of rational people in economics [440]. A typical question posed in CGT is: how can an agent maximize the payoff it accumulates over multiple rounds of a game? After each round, the agent may decide on which strategy it uses in the next round. On the mathematical level, this decision-making process is typically encoded in a pre-defined rule. The rule should prepare the agent in the best possible way against any opponent. Its details depend both on the game that is played and on the memory span of the agents. A large number of publications have been devoted to iterated two-player games, including Axelrod and Hamilton's classic article on the iterated prisoner's dilemma [442] and a recent article by Press and Dyson on "zero-determinant" rules [443].

Our interest lies in evolutionary game theory (EGT), which differs from CGT in two central points: First, EGT typically considers a population of agents. Second, it assumes that the payoff received by agents with a particular strategy determines the reproductive success of theses agents, i.e. their "fitness". The motivation for the second assumption stems from the application of game theory to biological systems in which the strategies represent species. The idea of relating payoff to fitness was first formulated by Maynard Smith and Price [444].

One can implement the idea that payoff relates to fitness in a mathematical model of how the popularity of the strategies changes over time. This model can be formulated both in discrete and in continuous time. The discrete-time formulation results in a set of difference equations and is, for example, appropriate when a population changes only seasonally. The continuous-time formulation instead results in a set of first-order ordinary differential equations and admits the population to change at any time. Both of these cases were first discussed by Taylor and Jonker [445]. Our interest lies in the continuous-time evolution. To derive the corresponding equations, we consider pairwise games in a population of $N \in \mathbb{N}_0$ agents. The variable $n_i \in \mathbb{N}_0$ shall denote the number of agents having adopted strategy E_i at a certain time τ, and $x_i := n_i/N$ shall be their fraction of the whole population (with $N = \sum_i n_i$). We also refer to x_i as a "concentration".

By our previous discussion of two-player games, an agent with strategy E_i should expect the payoff $p_i(\boldsymbol{x}) \approx \sum_j A_{ij} x_j$ when playing against a randomly chosen opponent. The central hypothesis of Taylor and Jonker's model is that this payoff coincides with the effective rate at which the agents with strategy E_i reproduce. In other words, one assumes that the payoff $p_i(\boldsymbol{x})$ coincides with the fitness $f_i(\boldsymbol{x})$ in the rate equation $\partial_\tau n_i = n_i f_i(\boldsymbol{x})$. After defining $\bar{f}(\boldsymbol{x}) := \sum_j x_j f_j(\boldsymbol{x})$ as the average fitness of the agents, this rate equation is readily converted into a differential equation for the concentration $x_i = n_i/N$. The resulting equation

$$\partial_\tau x_i = x_i(f_i(\boldsymbol{x}) - \bar{f}(\boldsymbol{x})) \tag{1}$$

is called a *replicator equation* [436, 446] and holds for all $i \in \{1, \dots, S\}$. The total concentration $\sum_i x_i = 1$ is conserved along its temporal evolution. The replicator equation provides a deterministic approximation of how the popularity of strategy E_i changes over time. The discreteness and finiteness of the population of agents are neglected by the equation and the concentration x_i is treated as a continuous variable. The validity and use of such a deterministic approximation regarding zero-sum games is discussed below and in our publication [3].

For zero-sum games, the antisymmetry of the payoff matrix A implies the vanishing of the average fitness $\bar{f}(\boldsymbol{x}) = \boldsymbol{x}^\intercal A \boldsymbol{x}$ because $\boldsymbol{x}^\intercal A \boldsymbol{x} = -\boldsymbol{x}^\intercal A \boldsymbol{x} = 0$. Therefore, for all $i \in \{1, \dots, S\}$, the replicator equation (1) simplifies to the *antisymmet-*

ric Lotka-Volterra equation (ALVE) [447–449]

$$\partial_\tau x_i = x_i(A\boldsymbol{x})_i \,. \tag{2}$$

According to the ALVE, any change in the concentration x_i due to games being played between agents with strategies E_i and E_j is accompanied by an equal but negative change in the concentration x_j. In other words, any term $x_i A_{ij} x_j$ in $\partial_\tau x_i$ is accompanied by a term $x_j A_{ji} x_i = -x_i A_{ij} x_j$ in $\partial_\tau x_j$. In our above discussion, we assumed that changes in the concentrations are caused only indirectly by the birth and death of agents. But for zero-sum games, the balancing of gain and loss admits also the following interpretation: a positive payoff A_{ij} can be interpreted as the rate at which an agent with strategy E_j adopts strategy E_i. The ways in which the agents change their strategies can be visualized in terms of a directed graph. Figure 1**(b)** shows such a graph for the rock-paper-scissors game.

A natural extension of the rock-paper-scissors game to five strategies is depicted in figure 2**(a)** [450]. This game served as an example in our publication "Coexistence and survival in conservative Lotka-Volterra networks" [2], which I wrote together with Johannes Knebel, Torben Krüger, and Erwin Frey. A reprint of the publication is included in section 2 of this chapter. In the publication, we studied the ALVE in a biological context and formulated criteria under which a zero-sum game admits the coexistence of its strategies (its species; the ALVE is referred to as a conservative Lotka-Volterra model in the publication). For the five-strategy game in figure 2**(a)**, the coexistence of all strategies is only possible if the rate A_{12} is chosen as $A_{12} = 5$. This rate makes the kernel of the antisymmetric rate matrix A three-dimensional, implying three conservation laws besides the conservation of the total concentration ($\sum_i x_i = 1$). The conservation laws constrain all strategy concentrations to neutrally-stable periodic orbits at a finite distance from zero, guaranteeing the coexistence of all strategies. For $A_{12} \neq 5$, the kernel of A is only one-dimensional, leading to the extinction of strategy E_2 for $A_{12} > 5$ and of strategy E_1 for $A_{12} < 5$. The agents abandon these strategies exponentially fast as shown in figure 2**(b)**. The time until extinction was found to obey a power law in the distance from the critical rate $A_{12} = 5$ [2].

The extinction of individual strategies is typical for the evolution of a zero-sum game. In our above publication [2], we asked how the strategies that do not go extinct and thus persist in the population can be identified for a given rate matrix A. This question was answered in our later publication "Evolutionary games of condensates in coupled birth-death processes" [3] (see section 3 for a reprint). Here, a "condensate" denotes a persistent strategy and the selection of these strategies is interpreted as a "condensation" phenomenon. The reason for this terminology derives from a recent study of Vorberg et al. on the condensation of non-interacting bosons in driven-dissipative quantum systems [1]. To understand

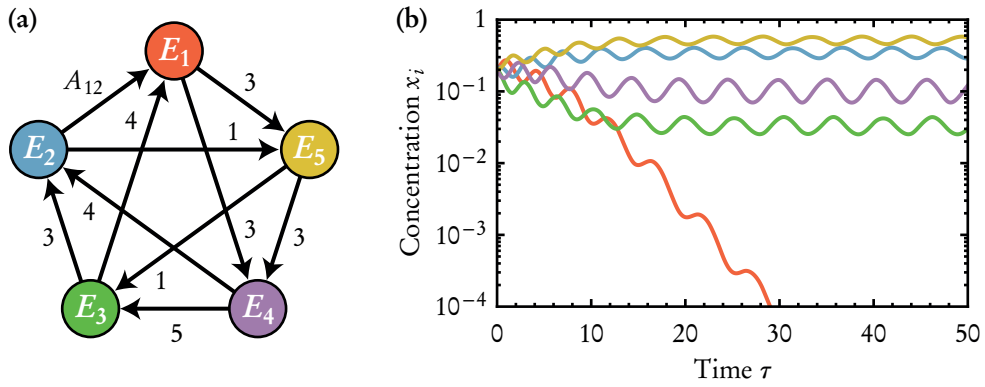


Figure 2 Deterministic evolution of a cyclic zero-sum game with five strategies. **(a)** Representation of the rate matrix A of the game as a directed graph (cf. figure 1). Every strategy E_j (coloured disk) is dominated by the two strategies E_{j+1} and E_{j+2} with the corresponding rates $A_{j+1,j}$ and $A_{j+2,j}$ (indices modulo 5). As shown in our publication [2], the choice $A_{12} = 5$ results in the coexistence of all strategies, $A_{12} > 5$ in the extinction of strategy E_2, and $A_{12} < 5$ in the extinction of strategy E_1 (a different convention for the direction of arrows is used in the publication). **(b)** Deterministic trajectories obtained from a numerical integration of the ALVE (2) for the rate $A_{12} = 4$. The colors correspond to the strategies in **(a)**. The agents abandon strategy E_1 exponentially fast. The concentrations (fractions) of the remaining strategies converge to a neutrally-stable periodic orbit.

why the selection of persistent strategies in a zero-sum game is mathematically equivalent to the selection of condensates in a driven-dissipative quantum system, we introduce the stochastic basis of the ALVE (2) in the following.

1.3 The stochastic time evolution of zero-sum games

Both the replicator equation (1) and the ALVE (2) neglect fluctuations that are caused by the discreteness and finiteness of the population of agents. Provided that the agents have no memory, these aspects can be incorporated in a mathematical model based on the (forward) master equation

$$\partial_\tau p(\tau, \boldsymbol{n}|t_0, \boldsymbol{n}_0) = \sum_{\boldsymbol{m}} \big(w_\tau(\boldsymbol{n}, \boldsymbol{m}) p(\tau, \boldsymbol{m}|\cdot) - w_\tau(\boldsymbol{m}, \boldsymbol{n}) p(\tau, \boldsymbol{n}|\cdot) \big) \,. \tag{3}$$

The dots "$\cdot$" in this equation abbreviate the initial parameters $(t_0, \boldsymbol{n}_0)$. The temporal evolution of the conditional probability distribution starts out from $p(t_0, \boldsymbol{n}|t_0, \boldsymbol{n}_0) = \delta_{\boldsymbol{n},\boldsymbol{n}_0}$, with the multivariate Kronecker delta being defined as $\delta_{\boldsymbol{m},\boldsymbol{n}} = \prod_{i=1}^{S} \delta_{m_i,n_i}$. As explained in section 1.3 of chapter I, a master equation describes the evolution of a continuous-time Markov process with discontinuous

sample paths. Here, these paths proceed along realizations of the discrete variable $\boldsymbol{n} = (n_1, \ldots, n_S)^\intercal \in \mathbb{N}_0^S$, denoting the number of agents with strategies $E_1, \ldots, E_S$.

The master equation of zero-sum games follows from equation (3) by specifying its transition rate $w_\tau(\boldsymbol{m}, \boldsymbol{n})$. Since n_i agents with strategy E_i can potentially play against n_j agents with strategy E_j, the rate at which games between these agents occur is proportional to the product $n_i n_j$. Consequently, the rate at which agents with strategy E_j adopt strategy E_i is given by $\Gamma_{i\leftarrow j}(n_i, n_j) = R_{ij} n_i n_j$, with $R_{ij} \geq 0$ (for $i, j \in \{1, \ldots, S\}$ and $R_{ii} = 0$). Note that both R_{ij} and R_{ji} may be positive for given indices i and j. The relation between the rate matrix R and the antisymmetric rate matrix A is discussed shortly. After complementing $\Gamma_{i\leftarrow j}$ with the correct reaction kinetics in the transition rate

$$w_\tau(\boldsymbol{m}, \boldsymbol{n}) = \sum_{i,j=1}^{S} \Gamma_{i\leftarrow j}(n_i, n_j) \delta_{\boldsymbol{m}, \boldsymbol{n} - \hat{\boldsymbol{e}}_j + \hat{\boldsymbol{e}}_i}\,, \tag{4}$$

equation (3) simplifies to the master equation of zero-sum games

$$\partial_\tau p(\tau, \boldsymbol{n}|\cdot) = \sum_{i,j=1}^{S} \left(\Gamma_{i\leftarrow j}(n_i - 1, n_j + 1) p(\tau, \boldsymbol{n} + \hat{\boldsymbol{e}}_j - \hat{\boldsymbol{e}}_i|\cdot) - \Gamma_{i\leftarrow j}(n_i, n_j) p(\tau, \boldsymbol{n}|\cdot) \right). \tag{5}$$

Sample paths of the master equation can be generated using the stochastic simulation algorithm (SSA) of Gillespie [80, 81] (cf. section 1.4 of chapter I). The total number of agents $N = \sum_i n_i$ is conserved along these paths.

Before proceeding, let us introduce a small generalization of the master equation. In particular, let us allow agents to spontaneously switch their strategy from E_j to E_i with per-capita rate $R_{ij} S_{ij}$ ($S_{ij} \geq 0$). The combined transition rate reads

$$\Gamma_{i\leftarrow j}(n_i, n_j) = R_{ij}(S_{ij} + n_i) n_j\,. \tag{6}$$

1.4 Condensation in driven-dissipative quantum systems

The reason for our generalization of the transition rate $\Gamma_{i\leftarrow j}$ is that for $S_{ij} = 1$, the master equation (5) coincides with the master equation that Vorberg et al. derived in their study on the condensation of non-interacting bosons in driven-dissipative quantum systems [1]. These quantum systems are weakly coupled to a reservoir for dissipation and are driven by an external time-periodic force (Floquet system). In this context, the configuration $\boldsymbol{n} = (n_1, \ldots, n_S)^\intercal \in \mathbb{N}_0^S$ denotes the numbers of bosons occupying the (pseudo-)energy states $E_1, \ldots, E_S$. The rate $\Gamma_{i\leftarrow j} = R_{ij}(1 + n_i) n_j$ describes the rate at which bosons transition from the state E_j to the state E_i. These transitions are effectively classical because of certain assumptions that were made in the derivation of the master equation (Born,

Markov, and rotating wave approximation). The bosonic quantum statistics is only apparent from the proportionality of $\Gamma_{i\leftarrow j}$ to the (future) occupation $1 + n_i$ of the arrival state. The rate coefficients R_{ij} can be inferred from microscopic parameters of the quantum system as explained in the supplemental material of [1].

Depending on the values of the rate coefficients R_{ij}, the master equation (5) predicts the formation of two groups of states. One of these groups comprises states that are depleted exponentially fast. In the other group, the boson concentrations of states approach metastable levels at a finite distance from zero. Consequently, each of these states carries a finite fraction of all bosons. Their boson concentrations may either remain effectively constant, or they may oscillate due to the exchange of bosons with other metastable states. Figure 3 demonstrates both possibilities for a system with seven states. The duration over which the metastable states persist scales linearly with the total number of bosons $N = \sum_i n_i$ [2].

Vorberg et al. interpreted the splitting into depleted states and states with a finite fraction of all bosons as a condensation phenomenon [1]. Given a rate matrix R, they asked how one can identify the states with a positive boson occupation. We answered this question and our previous question regarding the selection of persistent strategies in zero-sum games in our publication "Evolutionary games of condensates in coupled birth-death processes" [3].

1.5 Evolutionary games of condensates

In the following, the main results of our publication [3] are briefly summarized. The publication is reprinted in section 3 of this chapter. Author contributions are listed on page 8 of the publication (page 173 of this thesis).

In the first part of our publication, we showed that the metastable states (i.e. the condensates/the persistent strategies) of the master equation (5) with rate (6) agree with the neutrally-stable states of the ALVE (2) with rate matrix $A = R - R^\intercal$. We established this result by performing a Kramers-Moyal expansion of the master equation and by identifying the time scales on which the individual terms of the expansion act. The Kramers-Moyal expansion was introduced in section 4.4 of chapter I in this thesis. Our analysis showed that on the leading-order time scale, the dynamics of the master equation are described by the deterministic ALVE (2) with the antisymmetric rate matrix $A = R - R^\intercal$. Figure 3 provides a comparison between stochastic trajectories of the master equation and deterministic trajectories of the ALVE for a system with seven states (strategies).

Demographic fluctuations act only on sub-leading time scales and are suppressed by at least a factor $1/\sqrt{N}$ as compared to the terms of the ALVE (see the Supplementary Note 1 of [3], which is reprinted on pages 179–183). The fluctuations cause slow amplitude changes of the deterministic trajectories (cf. figure 3). The linear contribution $R_{ij}S_{ij}n_j$ to the rate (6) also becomes relevant on

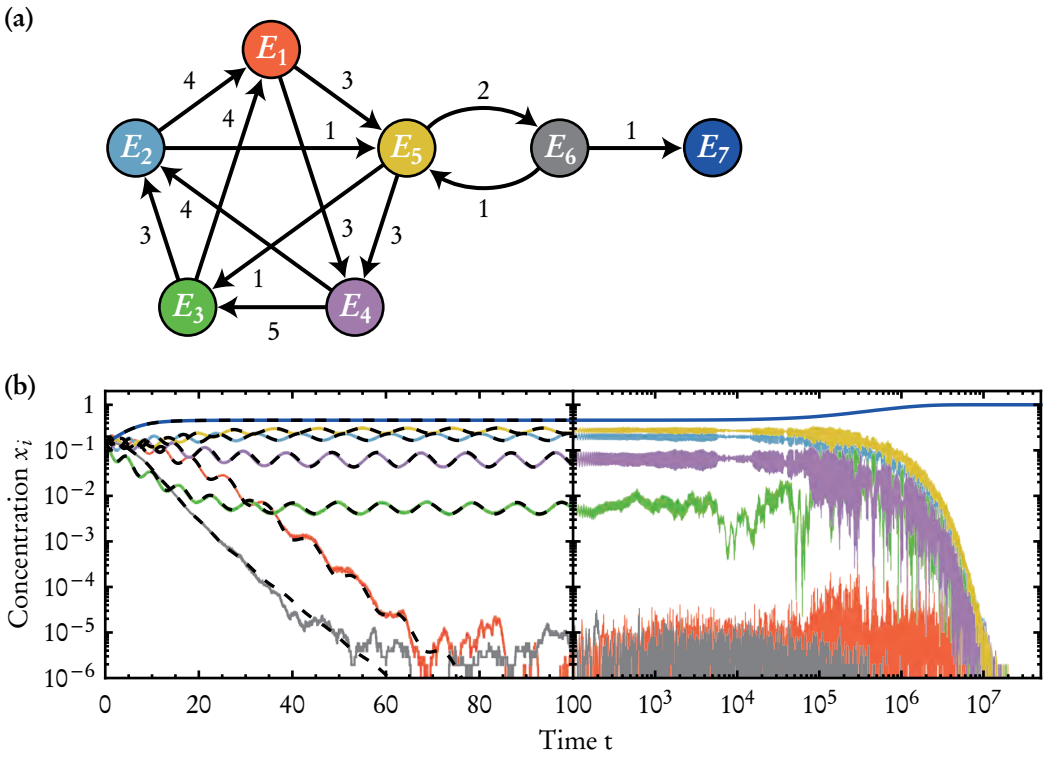


Figure 3 Stochastic and deterministic evolution of a system (zero-sum game) with seven states (strategies). **(a)** A labelled arrow represents the rate R_{ij} at which bosons transition from an initial state E_j to an arrival state E_i (the rate at which agents with strategy E_j adopt strategy E_i. **(b)** Coloured trajectories show the stochastic temporal evolution of the strategy concentrations $x_i = n_i/N$. The total number of bosons was chosen as $N = \sum_i n_i = 10^6$. The trajectories were obtained from a Gillespie simulation of the master equation (5). The corresponding rate matrix R is represented by the graph in **(a)** (with $S_{ij} = 1$). The scale of the horizontal axis changes from linear to logarithmic at time $t = 100$. The dashed black lines in the left half of the figure represent trajectories obtained from a numerical integration of the ALVE (2) with rate matrix $A = R - R^\intercal$. The ALVE approximates the stochastic dynamics on the leading-order time scale. In [3], we showed how the metastable states (condensates/persistent strategies) can be identified from the condensate vector $\boldsymbol{c}$ of the antisymmetric matrix A (here, $\boldsymbol{c} = (0, 4, 1, 1, 7, 0, 8)^\intercal$). Over longer time scales, fluctuations on sub-leading time scales cause amplitude changes and eventually a transition into the absorbing state E_7 (the "absorbing state" of the system in the sense of non-equilibrium theory is given by the vector $\boldsymbol{n} = (0, 0, 0, 0, 0, 0, N)$).

sub-leading time scales. A positive value of this contribution may affect whether or not the corresponding initial state E_j is absorbing.

After establishing that the condensates correspond to the neutrally-stable states of the ALVE (2), we developed an algebraic method to identify these states. The identification employed a theorem from linear programming theory [451] by which there exists a "condensate vector" $\boldsymbol{c}$ with the following properties: its entries c_i are positive for $i \in I \subseteq \{1, \ldots, S\}$ and zero for $i \in \overline{I} := \{1, \ldots, S\} \setminus I$, whereas $(A\boldsymbol{c})_i$ is zero for $i \in I$ and negative for $i \in \overline{I}$. Although multiple condensate vectors may exist, the index set I is unique. After proving that the relative entropy

$$D(\boldsymbol{c}||\boldsymbol{x}) = \sum_{i \in I} c_i \log(c_i/x_i) \tag{7}$$

of the "condensate vector" $\boldsymbol{c}$ from the state concentration $\boldsymbol{x}$ is a Lyapunov function, we were able to show that the states E_i with $i \in I$ become the condensates and that the states with $i \in \overline{I}$ are depleted. The depletion occurs exponentially fast. For systems whose rates R_{ij} (or A_{ij}) are sampled from continuous probability distributions, the condensate vector is unique (up to normalization) and $|(A\boldsymbol{c})_i|$ denotes the rate at which a state E_i with $i \in \overline{I}$ is depleted.

The above algebraic relations fulfilled by the condensate vector $\boldsymbol{c}$ can be rephrased in terms of the linear inequalities

$$A\boldsymbol{c} \leq \mathbf{0} \quad \text{and} \quad \boldsymbol{c} - A\boldsymbol{c} > \mathbf{0}, \tag{8}$$

which are both understood component-wise. Given an antisymmetric matrix A, one can solve the inequalities numerically via linear programming. We implemented an efficient algorithm for this purpose and applied the algorithm to large networks of states. The networks were characterized by their total number number of states S and by their connectivity C. The connectivity denotes the probability that two states E_i and E_j are connected by a positive effective rate $A_{ij} = R_{ij} - R_{ji}$ or A_{ji}. Effective rates were sampled from a Gaussian distribution with zero mean and unit variance. Our algorithm succeeded in the identification of condensates for fully connected systems with up to $S = 2000$ states and for sparsely connected systems with up to $S = 500$ states. In studying the average number of condensates supported by a system per the number of its states, we found that this number exhibits a minimum along the power law $C \sim 1/S^{\gamma}$ with $\gamma = 0.998 \pm 0.008$ (s.e.m.). Our results showed that the number of condensates supported by a system depends on an interplay between the condensation dynamics and the critical properties of random matrices. Finally, we explained how one can design systems that condense into a particular set of states and formulated conditions under which a system condenses into a "rock-paper-scissors game of condensates".

2 Publication

Coexistence and survival in conservative Lotka-Volterra networks

by

J. Knebel,[1] T. Krüger,[2] M. F. Weber,[1] and E. Frey[1]

[1]Department of Physics, Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany,

[2]Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstraße 38, 80333 München, Germany

reprinted on pages 156–160

with permission from

Phys. Rev. Lett. **110(16), 168106 (2013),**

DOI: 10.1103/PhysRevLett.110.168106.

Supplemental Material reproduced on pages 161–163.

Coexistence and Survival in Conservative Lotka-Volterra Networks

Johannes Knebel,[1] Torben Krüger,[2] Markus F. Weber,[1] and Erwin Frey[1,*]

[1]*Department of Physics, Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Ludwig-Maximilians-Universität München, Theresienstrasse 37, 80333 München, Germany*

[2]*Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstrasse 38, 80333 München, Germany*

(Received 14 September 2012; published 19 April 2013)

Analyzing coexistence and survival scenarios of Lotka-Volterra (LV) networks in which the total biomass is conserved is of vital importance for the characterization of long-term dynamics of ecological communities. Here, we introduce a classification scheme for coexistence scenarios in these conservative LV models and quantify the extinction process by employing the Pfaffian of the network's interaction matrix. We illustrate our findings on global stability properties for general systems of four and five species and find a generalized scaling law for the extinction time.

DOI: 10.1103/PhysRevLett.110.168106 PACS numbers: 87.23.Cc, 02.50.Ey, 05.40.−a, 87.10.Mn

Understanding the stability of ecological networks is of pivotal importance in theoretical biology [1,2]. Coexistence and extinction of species depend on many factors such as inter- and intraspecies interactions [3,4], population size [5–9], and mobility of individuals [10–16]. An intriguing question is how the stability of ecosystems depends on the interaction network between species. Is it the topology of the network (whose links may arise through predation, competition over common resources, or mutual cooperation) that sets the level of biodiversity? And how important is the strength of a single interaction link? Stable coexistence can, for example, be observed for natural populations in nonhierarchical networks that are comprised of species that interact in a competitive and predator-prey-like manner [17,18]. By understanding the interplay between the structure of the interaction network and the strengths of its links, it is possible to reveal mechanisms that underlie this stability.

A paradigm in addressing these ecologically important questions from a theoretical perspective are Lotka-Volterra (LV) models [19,20] in which the total biomass of species is conserved. These conservative LV systems [12,21,22] originate in the well-mixed limit from agent-based formulations of reaction-diffusion systems, where individuals of S different species $A_1, A_2, \ldots, A_S$ compete directly with each other following the simplified reaction scheme [23]: $A_i + A_j \to A_i + A_i$. Species A_i beats species A_j with rate k_{ij} and immediately replaces an individual of species A_j with an own offspring. Species A_j is thus degraded at the same rate such that the interaction matrix $G_S = \{k_{ij}\}_{i,j}$ is *skew-symmetric*. The interaction network can be visualized by a graph; see Fig. 1. Neglecting demographic fluctuations [24], the deterministic dynamics for the species' *concentration* vector $\mathbf{x} = (x_1, \ldots, x_S)^T$ is given by the rate equations (REs)

$$\partial_t x_i = x_i (G_S \mathbf{x})_i, \quad \text{for all } i = 1, \ldots, S. \tag{1}$$

This conservative LV model has been investigated as a prototype to understand principles of biodiversity from a theoretical point of view [8,25]. While these systems are also of central importance to many other fields of science (e.g., plasma physics [26], evolutionary game theory [27,28], and chemical kinetics [29]), no general scheme to classify coexistence, survival, and extinction of species has been established so far. It is frequently assumed that the topology of the interaction network alone determines coexistence of species [30,31], i.e., that such systems can be regarded as Boolean networks [32]. Recent investigations of specific topologies indicate, however, that knowledge about the network topology may not suffice to conclude whether all species coexist or if some of them go extinct [33–35]. These questions on global stability properties have been previously addressed successfully for various particular LV systems [27,36] and for hierarchical networks [37,38].

In this Letter, we present a general classification of coexistence scenarios in conservative LV networks with an arbitrary number of species. We elucidate the consequences of the interplay between the network structure and

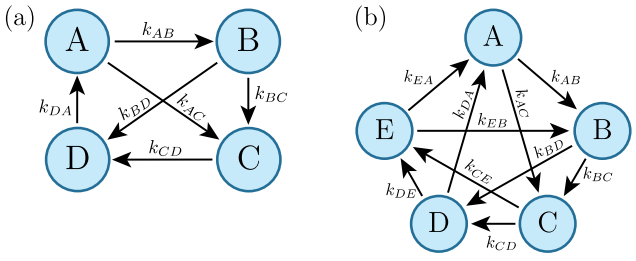


FIG. 1 (color online). Two interaction topologies specifying the conservative LV systems. (a) The general cyclic four species systems (4SS). (b) The general cyclic five species system (5SS) as a natural extension of the rock-papers-scissors configuration (RPS) [44].

the strengths of its interaction links on global stability. By analyzing conserved quantities, we find conditions on the reaction rates that yield coexistence of all species. In our mathematical framework this amounts to the characterization of positive kernel elements of the interaction matrix: By employing the algebraic concept of the Pfaffian of a skew-symmetric matrix, we are able to generalize previous approaches [34,39] and to quantify the extinction process when no conserved quantities exist. We illustrate our general results for coexistence and survival scenarios of four and five species systems (4SS and 5SS); cf. Fig. 1. Moreover, we demonstrate the implications of our findings for the stability of stochastic systems: We show how the extinction time diverges with the distance to the critical rate at which coexistence of all species is observed.

First, we discuss some general results for the REs (1) before the specific interaction topologies in Fig. 1 are analyzed. In order to characterize the stability of the generic LV system, we study conserved quantities. We elaborate on the form of conserved quantities, under which conditions they exist at all, and how many conserved quantities there are for a given interaction network. Since the interaction matrix G_S is skew-symmetric, the REs (1) conserve the sum over all species' concentrations $\tau_0 = x_1 + \ldots x_S$, independent of the interaction scheme. Hence, the dynamics can be normalized onto the $(S-1)$-dimensional simplex where all concentrations are nonnegative and add up to 1. The vertices of the simplex correspond to the extinction of all but one species, its edges reflect the extinction of all but two species, and so on. Further conserved quantities have previously been derived as $\tau = x_1^{p_1} \ldots x_S^{p_S}$ [20,27,39]. Interestingly, these conserved quantities can be obtained from solutions of the linear problem $G_S\mathbf{p} = \mathbf{0}$ because $\dot{\tau} = -\tau\langle G_S\mathbf{p}, \mathbf{x}\rangle$, with $\mathbf{p} = (p_1, \ldots, p_S)^T$. One infers that τ is conserved if the exponent vector $\mathbf{p}$ is an eigenvector corresponding to eigenvalue 0 [39], or in other words, if $\mathbf{p}$ lies in the kernel of the matrix G_S.

Coexistence means that all concentrations stay away from the boundary of the simplex by a finite distance for all times. Since the species' concentrations are bounded to the interval [0,1], one concludes from the structure of the conserved quantity τ that *all* S species coexist if the kernel of the interaction matrix is positive; i.e., one finds an element $\mathbf{p}$ in the kernel of G_S with positive entries $p_i > 0$ for all i. Hence, to reveal coexistence scenarios in the conservative LV model, one has to characterize the kernel of the interaction matrix G_S and identify its positive elements. Note that this conclusion goes beyond stating that a positive kernel element corresponds to a stationary point in the inside of the simplex; see REs (1).

The existence of conserved quantities constrains the dynamics to a submanifold of the simplex whose dimension D_c is determined as follows. The rank of a skew-symmetric matrix is always even, because its nonzero eigenvalues are purely imaginary, conjugate pairs. The rank-nullity theorem [40] then implies that the dimension of the kernel of G_S is odd whenever S is odd, and even whenever S is even. Each linearly independent kernel element gives rise to an independent conserved quantity τ which constrains the degrees of freedom of the trajectory. Together with τ_0, one finds that the dynamics in the case of nonstationary motion is constrained to a deformed sphere of dimension $D_c = S - 1 - \dim\mathrm{Ker}(G_S)$ for a positive kernel; see the Supplemental Material [41] for mathematical details. Thus, coexistence in high-dimensional systems is generically observed on nonperiodic trajectories ($D_c > 1$); see Movie M1 in Supplemental Material [41]. Only if the reaction rates are fine-tuned to a positive and maximal kernel of dimension $S-2$ is the dynamics restricted to periodic orbits ($D_c = 1$); see Fig. 2(a) and Movie M2 in Supplemental Material [41]. In particular, for $S = 3$ or 4, a positive kernel immediately implies coexistence on periodic orbits. This follows from the fact that with three species, the kernel is always one-dimensional. For the general 4SS, the dimension of the kernel of the interaction matrix is either 0 or 2. A two-dimensional, positive kernel yields coexistence on periodic orbits; see Fig. 2(a). If $\dim\mathrm{Ker}(G_S) = 0$, i.e., if the kernel is trivial, one observes extinction of species as detailed below.

Next, we focus on the mapping between the reaction rates in G_S and its kernel elements in order to find the stationary points. To this end, we apply the concepts of the Pfaffian and of the adjugate matrix [40,42]. The Pfaffian is a simpler form of the determinant tailored to skew-symmetric matrices with the property that its square equals the value of the determinant. In contrast to the nonnegative determinant of skew-symmetric matrices, the Pfaffian carries a sign which will turn out to be crucial

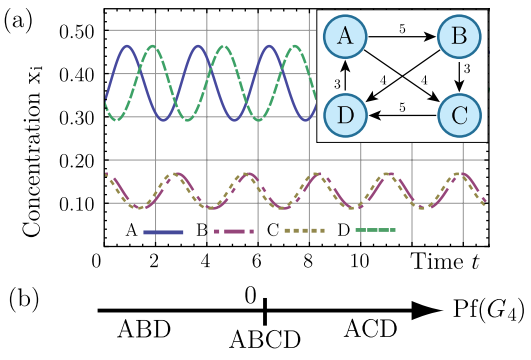


FIG. 2 (color online). Coexistence and survival in the general cyclic 4SS are controlled by the Pfaffian of the interaction matrix. (a) For $\mathrm{Pf}(G_4) = 0$, one obtains coexistence of all species on periodic orbits. (b) Deterministic survival diagram: for $\mathrm{Pf}(G_4) < 0$, species A, B, and D survive in a stable RPS configuration, whereas A, C, and D survive for $\mathrm{Pf}(G_4) > 0$.

for our purposes. For a skew-symmetric matrix, the Pfaffian can be computed recursively as

$$\mathrm{Pf}(G_S) = \sum_{i=2}^{S} (-1)^i k_{1i} \mathrm{Pf}(G_{\hat{1}\hat{i}}), \tag{2}$$

with $G_{\hat{1}\hat{i}}$ being the matrix where both the first and ith column and row have been removed from the matrix G_S. The Pfaffian of a 2×2 skew-symmetric matrix $G_2 = \{k_{AB}\}$, is given by $\mathrm{Pf}(G_2) = k_{AB}$. For the interaction matrix corresponding to the LV network in Fig. 1(a),

$$G_4 = \begin{pmatrix} 0 & k_{AB} & k_{AC} & -k_{DA} \\ -k_{AB} & 0 & k_{BC} & k_{BD} \\ -k_{AC} & -k_{BC} & 0 & k_{CD} \\ k_{DA} & -k_{BD} & -k_{CD} & 0 \end{pmatrix},$$

the Pfaffian is $\mathrm{Pf}(G_4) = k_{AB}k_{CD} - k_{AC}k_{BD} - k_{DA}k_{BC}$.

The Pfaffian always vanishes for odd S as opposed to systems with an even number of species [42]. In the latter case, the Pfaffian is zero only if a constraint on the reaction rates is fulfilled. If the Pfaffian vanishes, one finds more kernel elements than just the null vector and, thus, conserved quantities of form τ exist. In the following, we distinguish between even and odd S.

For an even number of species and a two-dimensional kernel, positive kernel elements can be identified via the adjugate matrix R_S which is a generalized inverse of the interaction matrix such that $G_S R_S = -\mathrm{Pf}(G_S)\mathbb{I}_S$, with $\mathbb{I}_S$ being the identity matrix [42]. The adjugate matrix can be computed as $(R_S)_{ij} = (-1)^\sigma \mathrm{Pf}(G_{\hat{i}\hat{j}})$, where $(-1)^\sigma$ denotes the sign of the permutation $\sigma = (ij1 \ldots \hat{i} \ldots \hat{j} \ldots S)$, and the columns of R_S give two independent kernel elements of G_S.

As an example, consider again the general cyclic 4SS depicted in Fig. 1(a). By setting all reaction rates equal to each other (e.g., to 1), the Pfaffian does not vanish and, therefore, not all species can coexist. Only when the rates are chosen such that $\mathrm{Pf}(G_4) = 0$, do we obtain two independent kernel elements of G_4: From its adjugate matrix, R_4, we identify $\mathbf{p}_1 = (k_{CD}, 0, k_{DA}, k_{AC})^T$ and $\mathbf{p}_2 = (k_{BD}, k_{DA}, 0, k_{AB})^T$. We infer the two conserved quantities $\tau_1 = x_A^{k_{CD}} x_C^{k_{DA}} x_D^{k_{AC}}$ and $\tau_2 = x_A^{k_{BD}} x_B^{k_{DA}} x_D^{k_{AB}}$, and conclude that the kernel is positive and coexistence occurs on periodic orbits; see Fig. 2(a). Hence, classifying LV networks in terms of their topology is incomplete; the strengths of the interaction links are crucial in general.

In general, if the Pfaffian for a system with even S is nonzero, i.e., when only the null vector lies in the kernel, coexistence of all species is not possible. Still one can quantify the extinction process by generalizing an approach of Durney *et al.* [34] for a system with $S = 4$ to systems composed of an arbitrary even number of species. We define the function $\rho = x_1^{q_1} \ldots x_S^{q_S}$ in the same way as the conserved quantity τ, but this time choosing the exponent vector $\mathbf{q}_S = -R_S\mathbf{1}$ with $\mathbf{1} = (1, \ldots, 1)^T$. It is straightforward to show that this function evolves exponentially in time:

$$\rho(t) = \rho(0) e^{-\mathrm{Pf}(G_S)t}, \tag{3}$$

generalizing previous investigations [24,33–35]. It is quite remarkable that ρ quantifies the global collective dynamics of systems with an arbitrary interaction topology and even S. Depending on the sign of the Pfaffian, ρ grows or decays exponentially fast with the Pfaffian of the interaction matrix as rate. Since the system's dynamics is driven towards the boundary of the simplex, one can conclude on the extinction of some species. This feature of ρ is reminiscent of a Lyapunov function; note also that ρ becomes a conserved quantity τ if the Pfaffian is zero. An interesting question for future investigations is to ask whether further quantities exist that characterize the dynamics of conservative LV networks.

For the general 4SS shown in Fig. 1(a), we find $\mathbf{q}_4 = (-k_{CD} + k_{BD} - k_{BC}, k_{CD} + k_{DA} + k_{AC}, -k_{BD} - k_{DA} - k_{AB}, k_{BC} - k_{AC} + k_{AB})^T$. The fact that $(\mathbf{q}_4)_2$ is always positive suggests that species B goes extinct for a positive Pfaffian, and that the converse holds true for $(\mathbf{q}_4)_3$ and species C for a negative Pfaffian. In both cases, the system tends to a stable rock-paper-scissors (RPS) configuration. In summary, we derive the survival diagram shown in Fig. 2(b). Interestingly, A and D always survive in this topology although D can be easily tuned to be the weakest species. We emphasize that this result depends on the sign of the Pfaffian and cannot be obtained from applying the concept of the determinant. Again, since the Pfaffian of the interaction matrix characterizes the dynamics of this 4SS, its topology alone does not determine the long-time dynamics. These findings unify previous results for other 4SS [24,33,34], and show that rules like "survival of the strongest" or "survival of the weakest" [25,43] cannot be formulated in general.

For an odd number of species, the kernel of G_S is always nontrivial. In general, if $\dim\ker G_S = 1$, we determine the independent kernel element via the adjugate vector [42], $\mathbf{r}_S = (\mathrm{Pf}(G_{\hat{1}}), -\mathrm{Pf}(G_{\hat{2}}), \ldots, \mathrm{Pf}(G_{\hat{S}}))^T$, which enables us to investigate the influence of the reaction rates on the survival scenarios. For $S = 3$, only the well-studied RPS topology [8,27] leads to a positive adjugate vector $\mathbf{r}_3$. In other words, coexistence of all three species depends only on the topology of the network. This behavior is unique to $S = 3$ and changes dramatically for systems with more than three species.

We illustrate the importance of the reaction rates for a system of five interacting species; see Fig. 1(b). This interaction topology where each species dominates two species and is outperformed by the two remaining species, recently gained attention as a natural extension of the RPS game [30,44,45]. For specificity, we investigate the dependence of the survival scenarios on the rate k_{AB} with which

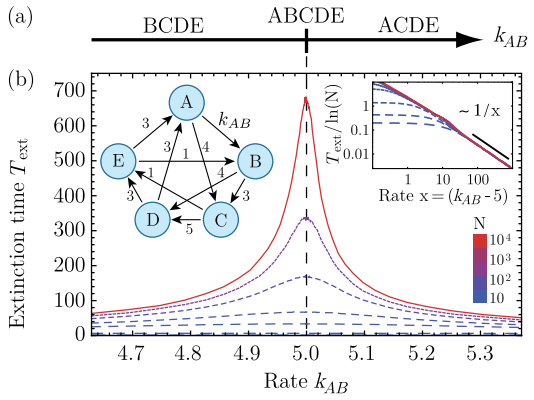


FIG. 3 (color online). Stability of the cyclic 5SS. (a) For the interaction scheme [left inset of (b)], one obtains coexistence of all species for the critical rate $k_{AB} = 5$. (b) Stability of the stochastic system, reflected by the extinction time T_{ext}, peaks at the critical rate, which becomes more pronounced as $N \to \infty$. We find a scaling law for T_{ext} in the distance to the critical rate (right inset). Initial conditions were chosen as $\mathbf{x}(0) = 1/5 \times \mathbf{1}$. Larger line gap corresponds to smaller N.

species A beats species B and chose the other rates [see Fig. 3(b), left inset] such that either five or four species survive depending on the value of k_{AB}; see Fig. 3(a). The kernel of the interaction matrix depends on k_{AB} and is characterized by the adjugate vector $\mathbf{r}_5 = (0, 0, 3k_{AB} - 15, 5 - k_{AB}, 5k_{AB} - 25)^T$. For $k_{AB} \neq 5$, the kernel is one-dimensional and nonpositive, and four species survive. In contrast, for $k_{AB} = 5$, $\mathbf{r}_5$ equals the null vector which in turn means that the kernel becomes three-dimensional [42]. Since we have ensured that the kernel is also positive, we obtain coexistence of all five species on periodic orbits ($D_c = 1$).

Finally, we discuss the implications of our findings by asking how demographic noise affects the stability of stochastic LV systems. We analyze ecological LV systems with a finite number N of interacting individuals in the eye of the knowledge gained from the deterministic analysis. It has been shown that due to demographic fluctuations the system ultimately reaches an absorbing state that is characterized by the extinction of all but one species [46–49]. Moreover, the scaling behavior of the mean extinction time with the system size N characterizes the stability of the interaction network [14,48].

As an example, we continue the discussion of the 5SS from Fig. 3(b), left inset. We have carried out extensive computer simulations employing the Gillespie algorithm [50] to measure the time T_{ext} until the first species has become extinct for different system sizes N and different reaction rates k_{AB}. The results are displayed in Fig. 3(b) and highlight the significance of the deterministic drift underlying the stochastic dynamics. We observe a peak in the extinction time as the reaction rate k_{AB} approaches the critical value $k_{\text{cr}} = 5$ for which we obtain coexistence of all species in the deterministic case. The divergence of the extinction time for $k_{AB} \to k_{\text{cr}}$ becomes more pronounced for larger system sizes as the system reaches the deterministic limit for $N \to \infty$.

A scaling analysis reveals how the extinction time peaks in the vicinity of the coexistence scenario. Near the critical rate, the extinction time scales linearly with the system size leading to neutrally stable interaction networks [8,24,51]. At larger distance from the critical rate, the deterministic driving force to the absorbing boundary becomes more dominant than the demographic fluctuations; see Fig. 3(b), right inset. The interplay between the stochastic and deterministic forces is reflected by the scaling law

$$T_{\text{ext}} \propto \begin{cases} N & \text{for } k_{AB} = k_{\text{cr}}, \\ \frac{\ln N}{|k_{AB} - k_{\text{cr}}|} & \text{for } k_{AB} \neq k_{\text{cr}}, \end{cases} \tag{4}$$

which extends the linear scaling $T_{\text{ext}} \propto N$ of neutral coexistence. We observe a power-law dependence in the distance of the reaction rates to the critical rate and logarithmic scaling with N for attracting boundaries [8,52].

The observed scaling law (4) for $k_{AB} \neq k_{\text{cr}}$ can be attributed to the exponentially fast extinction of species $x_i = x_i(0) \exp(-\alpha_i t)$; see Eq. (1). The extinction rate α_i is computed via the temporal average over the trajectory $\langle \mathbf{x} \rangle$ as $\alpha_i = -(G_S \langle \mathbf{x} \rangle)_i$, which becomes linear in the distance to the critical rate $|k_{AB} - k_{\text{cr}}|$ for large times. The logarithmic dependence on N follows by defining that a species with concentration x_i less than $1/N$ has become extinct. With this scaling behavior at hand, we are able to compare the ecological stability of different interaction networks based on our analysis of the REs (1).

In this Letter, we investigated global stability properties of conservative LV networks. By employing the Pfaffian of the interaction matrix, we revealed the relation between the reaction rates and the conditions for coexistence, and exemplified the implications for the stability of ecological networks with finite populations. We expect that our results will also stimulate further progress for the investigation of extinction scenarios. Beyond analyzing whether an ecosystem is stable or unstable, it would be highly interesting to actually predict *which* of its species ultimately survive for a general conservative LV system. This would, for example, allow us to predict the eventual outcome of an unstable version of the five species system shown in Fig. 1(b), and to formulate the conditions under which 3- or 4-species cycles are attained. First insights into these extinction dynamics will be outlined in a future publication [53]. We believe that a full characterization of general conservative LV dynamics is possible.

We would like to thank Steffen Rulands, Lukas Darnstädt, Falk Töppel, and Michal Oszmaniec for helpful discussions. This project was supported by the Deutsche Forschungsgemeinschaft in the framework of the SFB TR 12 "Symmetry and Universality in Mesoscopic Systems",

and the German Excellence Initiative via the program "Nanosystems Initiative Munich" (NIM). J. K. acknowledges funding by the Studienstiftung des Deutschen Volkes.

*frey@lmu.de

[1] R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, NJ, 1973).
[2] J. M. Montoya, S. L. Pimm, and R. V. Solé, Nature (London) **442**, 259 (2006).
[3] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).
[4] J. Mathiesen, N. Mitarai, K. Sneppen, and A. Trusina, Phys. Rev. Lett. **107**, 188101 (2011).
[5] C. Taylor, D. Fudenberg, A. Sasaki, and M. Nowak, Bull. Math. Biol. **66**, 1621 (2004).
[6] A. J. McKane and T. J. Newman, Phys. Rev. Lett. **94**, 218102 (2005).
[7] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. Lett. **95**, 238701 (2005).
[8] T. Reichenbach, M. Mobilia, and E. Frey, Phys. Rev. E **74**, 051907 (2006).
[9] A. Melbinger, J. Cremer, and E. Frey, Phys. Rev. Lett. **105**, 178101 (2010).
[10] M. Nowak and R. May, Nature (London) **359**, 826 (1992).
[11] R. Durrett and S. Levin, Theor. Popul. Biol. **46**, 363 (1994).
[12] L. Frachebourg, P. L. Krapivsky, and E. Ben-Naim, Phys. Rev. E **54**, 6186 (1996).
[13] M. Mobilia, I. T. Georgiev, and U. C. Täuber, J. Stat. Phys. **128**, 447 (2007).
[14] T. Reichenbach, M. Mobilia, and E. Frey, Nature (London) **448**, 1046 (2007).
[15] R. Abta, M. Schiffer, A. Ben-Ishay, and N. M. Schnerb, Theor. Popul. Biol. **74**, 273 (2008).
[16] S. Rulands, T. Reichenbach, and E. Frey, J. Stat. Phys. 2011, L01003 (2011).
[17] L. W. Buss and J. B. C. Jackson, Am. Nat. **113**, 223 (1979).
[18] B. Sinervo and C. Lively, Nature (London) **380**, 240 (1996).
[19] A. Lotka, J. Am. Chem. Soc. **42**, 1595 (1920).
[20] V. Volterra, *Leçons sur la Théorie Mathématique de la Lutte pour la Vie* (Gauthier-Villars, Paris, 1931), 1st ed.
[21] N. S. Goel, S. C. Maitra, and E. W. Montroll, Rev. Mod. Phys. **43**, 231 (1971).
[22] Y. Itoh, Proc. Jpn. Acad. **47**, 854 (1971).
[23] K. I. Tainaka, Phys. Rev. Lett. **63**, 2688 (1989).
[24] A. Dobrinevski and E. Frey, Phys. Rev. E **85**, 051903 (2012).
[25] M. Frean and E. R. Abraham, Proc. R. Soc. B **268**, 1323 (2001).
[26] V. Zakharov, S. Musher, and A. Rubenchik, JETP Lett. **19**, 151 (1974); S. Manakov, Sov. Phys. JETP **40**, 269 (1975).
[27] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, England, 1998), 1st ed.
[28] M. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006).
[29] N. G. Van Kampen, *Stochastic Process in Physics and Chemistry* (Elsevier, Amsterdam, 2007), 3rd ed.
[30] R. Laird and B. S. Schamp, J. Theor. Biol. **256**, 90 (2009).
[31] Y. Li, L. Dong, and G. Yang, Physica (Amsterdam) **391A**, 125 (2011).
[32] K. Klemm and S. Bornholdt, Phys. Rev. E **72**, 055101 (2005).
[33] S. O. Case, C. H. Durney, M. Pleimling, and R. K. P. Zia, Europhys. Lett. **92**, 58 003 (2010).
[34] C. H. Durney, S. O. Case, M. Pleimling, and R. K. P. Zia, Phys. Rev. E **83**, 051108 (2011).
[35] C. H. Durney, S. O. Case, M. Pleimling, and R. K. P. Zia, J. Stat. Phys. 2012, P06014 (2012).
[36] B. S. Goh, Am. Nat. **111**, 135 (1977); R. Redheffer, J. Diff. Equ. **52**, 245 (1984); M. Zeeman, Proc. Am. Math. Soc. **123**, 87 (1995); Y. Takeuchi, *Global Dynamical Properties of Lotka-Volterra Systems* (World Scientific, Singapore, 1996).
[37] E. Chauvet, J. Paullet, J. Previte, and Z. Walls, Mathematics magazine **75**, 243 (2002).
[38] T. Cheon, Phys. Rev. Lett. **90**, 258105 (2003).
[39] R. K. P. Zia, arXiv:1101.0018.
[40] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Philadelphia, 2000), 3rd ed.
[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.168106 for mathematical details and movies.
[42] C. E. Cullis, *Matrices and Determinoids* (Cambridge University Press, Cambridge, England, 1913), 1st ed., Vols. I–II.
[43] M. Berr, T. Reichenbach, M. Schottenloher, and E. Frey, Phys. Rev. Lett. **102**, 048102 (2009).
[44] The Big Bang Theory, "*The Lizard-Spock Expansion*" Season 2, Episode 8 (2008).
[45] K. Hawick, *Cycles, Diversity and Competition in Rock-Paper-Scissors-Lizard-Spock Spatial Game Agent Simulations, Tech. Rep.* (Massey University, Auckland, 2011).
[46] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge University Press, Cambridge, England, 2010), 1st ed.
[47] L. Frachebourg and P. L. Krapivsky, J. Phys. A **31**, L287 (1998).
[48] T. Antal and I. Scheuring, Bull. Math. Biol. **68**, 1923 (2006).
[49] M. Parker and A. Kamenev, Phys. Rev. E **80**, 021129 (2009).
[50] D. T. Gillespie, J. Comp. Physiol. **22**, 403 (1976).
[51] J. Cremer, T. Reichenbach, and E. Frey, New J. Phys. **11**, 093029 (2009).
[52] E. Frey, Physica (Amsterdam) **389A**, 4265 (2010).
[53] J. Knebel, T. Krüger, M. F. Weber, and E. Frey (in preparation).

Supplemental Material
Coexistence and Survival in Conservative Lotka-Volterra Networks

Johannes Knebel[1], Torben Krüger[2], Markus F. Weber[1], and Erwin Frey[1]

[1]*Arnold Sommerfeld Center for Theoretical Physics (ASC) and Center for NanoScience (CeNS), Department of Physics, Ludwig-Maximilians-Universität München, Theresienstrasse 37, 80333 München, Germany*

[2]*Department of Mathematics, Ludwig-Maximilians-Universität München, Theresienstrasse 38, 80333 München, Germany*

Characterization of coexistence scenarios of conservative LV networks

In this supplemental material, we extend the characterization of the qualitative behavior of coexistence scenarios for the conservative LV model as defined in the main text,

$$\partial_t x_i = x_i \cdot (G_S \mathbf{x})_i \ , \tag{1}$$

where G_S is a skew-symmetric matrix. We show that trajectories lie on odd dimensional, deformed spheres. In case of a positive and maximal kernel of the interaction matrix this behavior translates to periodic orbits.

In the main text of this letter, we define coexistence of all species for the deterministic case if the species' concentrations retain a finite distance to the absorbing boundary for all times and a given set of initial conditions. Furthermore, we show that a positive kernel of the interaction matrix implies this coexistence. We call the kernel of the interaction matrix positive if there exists an element $\mathbf{p}$ in the kernel of G_S with $p_i > 0$ for all i.

First, we show that the motion of the LV system defined in Eq. (1) is either stationary or restricted to a $(S-1-n)$-dimensional manifold M, where n denotes the dimension of the kernel of G_S. Let us assume a LV system characterized by an interaction matrix G_S with a positive, n-dimensional kernel. Note that n has the same parity as S since the rank of a skew-symmetric matrix is always even. In a sufficiently small neighborhood of a positive element $\mathbf{p}$ of the kernel of the interaction matrix, the kernel is still positive and contains n

linearly independent vectors $\mathbf{p}^{(1)}, \ldots, \mathbf{p}^{(n)}$. We can assume these vectors to be normalized such that $\sum_{i=1}^{S} p_i^{(l)} = 1$ holds true for $l = 1, \ldots, n$. They give rise to n constants of motion of the form

$$\tau_l = x_1^{p_1^{(l)}} \ldots x_S^{p_S^{(l)}} \ ,$$

as shown in the main text. In addition, the trivially conserved quantity

$$\tau_0 = \sum_{i=1}^{S} x_i = 1 \ ,$$

always exists. In order to prove linear independence of these conserved quantities, we compute:

$$0 = \sum_{l=1}^{n} c_l \cdot \mathrm{d} \log \tau_l + c_0 \cdot \mathrm{d}\tau_0 = \sum_{i=1}^{S} \left(\sum_{l=1}^{n} c_l p_i^{(l)} + c_0 x_i \right) \frac{\mathrm{d}x_i}{x_i} \ ,$$

with arbitrary real constants $c_0, c_1, \ldots, c_n$. In case of $\mathbf{x} \notin \ker G_S$, the latter equation holds true only if $c_l = 0$ for all $l = 0, \ldots, n$. This result shows the linear independence of $\mathrm{d}\tau_0, \mathrm{d}\tau_1, \ldots, \mathrm{d}\tau_n$. For $\mathbf{x} \in \ker G_S$, the motion of the system is stationary as can be seen from Eq. (1). Hence, we conclude that the motion is either stationary or restricted to a $(S-1-n)$-dimensional manifold M and the dimension of M is always odd.

In the following, we elucidate that this manifold is diffeomorphic to a sphere of dimension $(S-1-n)$. We introduce the coordinates

$$u_1 = \log x_1, \ldots, u_S = \log x_S \ ,$$

and note that the manifold M can be characterized in these new coordinates by the intersection $U \cap \{f(\mathbf{u}) = 1\}$. The latter U denotes the set

$$U = \left\{ \mathbf{u} \in \mathbb{R}^S : \sum_i p_i^{(l)} u_i = \alpha_l; \ l = 1, \ldots, n \right\} \ ,$$

which is an $(S-n)$-dimensional affine subspace given by the intersection of the hypersurfaces defined by the conserved quantities $\tau_1, \ldots, \tau_n$. The constants $\alpha_1, \ldots, \alpha_S < 0$ are determined by the initial conditions. The function

$$f : \mathbf{u} \mapsto \sum_i e^{u_i}$$

corresponds to the sum over the species' concentrations x_i and is a strictly convex function. Provided that $\min_{\mathbf{u}\in U} f(\mathbf{u}) < 1$, this property implies that the set $\{\mathbf{u} \in U : f(\mathbf{u}) \leq 1\}$ is a strictly convex bounded subset of U with open interior and smooth boundary which is, therefore, diffeomorphic to a sphere. If on the other hand $\min_{\mathbf{u}\in U} f(\mathbf{u}) = 1$, then U consists of only one point and the motion is stationary.

As a consequence, the motion along solutions of Eq. (1) is either stationary or takes place on an odd dimensional, deformed sphere in case of a positive kernel. From the observation that M cannot contain an element of the kernel of G_S for non-stationary motion, it follows via the equations of motion (1) that $\partial_t x_i \neq 0$ and together with the compactness of M, we conclude that $|\partial_t x_i| \geq const > 0$ for all i and all times. In other words, the dynamics on these odd-dimensional spheres does not come to rest.

In summary, if the kernel is positive, quasi-periodic and non-periodic trajectories are typically observed as can be seen from the corresponding Fourier spectrum; see also Movie M1 of the Supplemental Material (SM). If the dimension of the kernel is maximal, that is if $\dim\ker G_S = S - 2$, the non-stationary trajectories are restricted to deformed circles and they never come to rest. Therefore, the motion occurs on periodic orbits; see Movie M2 of the SM.

3 Publication

Evolutionary games of condensates in coupled birth-death processes

by

J. Knebel,*,[1] M. F. Weber,*,[1] T. Krüger,*,[2] and E. Frey[1]

*equal contribution, [1]Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany, [2]IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

reprinted on pages 166–174

from

Nat. Commun. **6, 6977 (2015),**

DOI: 10.1038/ncomms7977.

Supplementary Information reproduced on pages 175–188.

nature COMMUNICATIONS

ARTICLE

Received 21 Nov 2014 | Accepted 20 Mar 2015 | Published 24 Apr 2015

DOI: 10.1038/ncomms7977 OPEN

Evolutionary games of condensates in coupled birth-death processes

Johannes Knebel[1,*], Markus F. Weber[1,*], Torben Krüger[2,*] & Erwin Frey[1]

Condensation phenomena arise through a collective behaviour of particles. They are observed in both classical and quantum systems, ranging from the formation of traffic jams in mass transport models to the macroscopic occupation of the energetic ground state in ultra-cold bosonic gases (Bose–Einstein condensation). Recently, it has been shown that a driven and dissipative system of bosons may form multiple condensates. Which states become the condensates has, however, remained elusive thus far. The dynamics of this condensation are described by coupled birth–death processes, which also occur in evolutionary game theory. Here we apply concepts from evolutionary game theory to explain the formation of multiple condensates in such driven-dissipative bosonic systems. We show that the vanishing of relative entropy production determines their selection. The condensation proceeds exponentially fast, but the system never comes to rest. Instead, the occupation numbers of condensates may oscillate, as we demonstrate for a rock–paper–scissors game of condensates.

[1] Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany. [2] IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to E.F. (email: frey@lmu.de).

Condensation phenomena occur in a broad range of contexts in both classical and quantum systems. Networks such as the World Wide Web or the citation network perpetually grow by the addition of nodes or links and they evolve by rewiring. Over time, a finite fraction of the links of a network may be attached to particular nodes. These nodes become hubs and thereby dominate the dynamics of the whole network; they become condensate nodes[1–3]. Condensation also occurs in models for the jamming of traffic[4–7] and in related mass transport models in which particles hop between sites on a lattice[3,8,9]. A condensate forms when a finite fraction of all particles aggregates into a cluster that dominates the total particle flow. Bose–Einstein condensation, on the other hand, is a quintessentially quantum mechanical phenomenon. When an equilibrated, dilute gas of bosonic particles is cooled to a temperature near absolute zero, a finite fraction of bosons may condense into the energetic ground state[10–12]. Long-range phase coherence builds up and quantum physics becomes manifest on the macroscopic scale[13,14].

In both the classical and the quantum mechanical context, condensation occurs when one or multiple states become macroscopically occupied (they become condensates), whereas the other states become depleted[15,16]. However, the physical origins of condensation in the above examples differ from each other. Why and how condensation arises in a particular system remains a topic of general interest and vivid research.

Here we study condensation in two systems from different fields of research: incoherently driven-dissipative systems of non-interacting bosons and evolutionary games of competing agents. As we show below, the physical principle of vanishing entropy production governs the formation of condensates in both of these systems. The entities that constitute the respective system shall be called particles. They may be quantum or classical particles (bosons or agents). The dynamics of these particles eventually lead to the condensation into particular states (quantum states or strategies). Before describing the above two systems, we now introduce the mathematical framework of our study.

On an abstract level, we consider a system of S (non-degenerate) states $E_i, i = 1, \ldots, S$, each of which is occupied by $N_i \geq 0$ indistinguishable particles, see Fig. 1a. The configuration of the system at time t is fully characterized by the occupation numbers $\mathbf{N} = (N_1, N_2, \ldots, N_S)$. This configuration changes continuously in time due to the transition of particles between states. The total number of particles in this coupled birth–death process is conserved ($N = \sum_i N_i$). We are interested in the probability $P(\mathbf{N}, t)$ of finding the system in configuration $\mathbf{N}$ at time t. The temporal evolution of the probability distribution $P(\mathbf{N}, t)$ is governed by the classical master equation[17,18]:

$$\partial_t P(\mathbf{N},t) = \sum_{\substack{i,j=1 \\ j\neq i}}^{S} \big(\Gamma_{i\leftarrow j}(N_i-1,N_j+1)P(\mathbf{N}-\mathbf{e}_i+\mathbf{e}_j,t) - \Gamma_{i\leftarrow j}(N_i,N_j)P(\mathbf{N},t)\big), \quad (1)$$

where $\mathbf{e}_i \in \mathbb{Z}^S$ denotes the unit vector in direction i (equal to 1 at index i, otherwise 0). The rate for the transition of particles from state E_j to E_i depends linearly on the number of particles in the departure and in the arrival state:

$$\Gamma_{i\leftarrow j} = r_{ij}(N_i + s_{ij})N_j \ , \quad (2)$$

with rate constant $r_{ij} \geq 0$ and constant $s_{ij} \geq 0$.

Condensation in this framework is understood as the macroscopic occupation of one or multiple states[15,16]: We consider a state E_i as a condensate when the long-time average of the number of particles in this state scales linearly with the system size ($\langle N_i \rangle_t \sim \mathcal{O}(N)$ for large t). Hence, a condensate harbours a finite fraction of the total number of particles for large systems ($N \gg 1$). We refer to a state as depleted when its average occupation number scales less than linearly with the system size. Therefore, the fraction of particles in a depleted state vanishes in the limit of large systems.

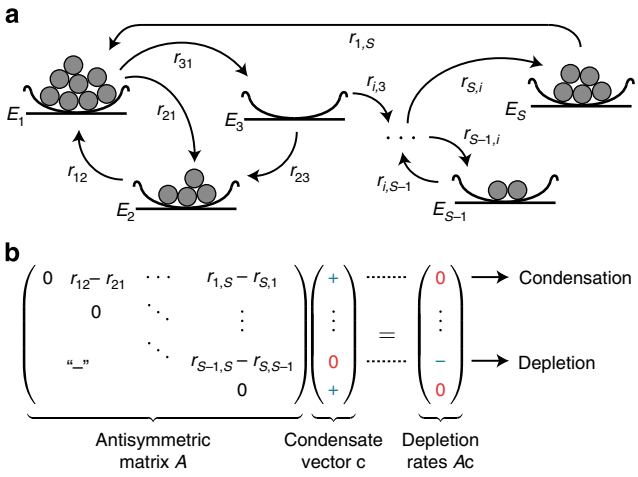


Figure 1 | Condensation into multiple states due to particle transitions between states and mathematics of condensate selection. (**a**) With respect to condensation in an incoherently driven-dissipative quantum system, each bowl represents a state E_i that is occupied by N_i non-interacting bosons (filled circles). If indicated by an arrow, bosons may undergo transitions from state E_j to state E_i at a rate $\Gamma_{i\leftarrow j} = r_{ij}(N_i+1)N_j$, with rate constant r_{ij}. In the language of evolutionary game theory, the figure depicts the interaction of N_i agents (filled circles) playing strategies E_i (bowls) for $i = 1,\ldots,S$. An agent playing strategy E_j adopts strategy E_i at a rate $\Gamma_{i\leftarrow j} = r_{ij}N_iN_j$. The above rate of bosonic condensate selection is recovered if agents may also spontaneously mutate from E_j to E_i at a rate r_{ij}. (**b**) A condensate vector c for an antisymmetric matrix A has two properties: its entries are positive for indices for which Ac is zero, and they are zero for indices for which Ac is negative ("-" signifies the antisymmetry of matrix A). Temporal evolution of the relative entropy of the condensate vector to the state concentrations under the ALVE (3) relates positive entries of the condensate vector to condensates, and its zero entries to depleted states. Generically, positive entries of c represent the asymptotic temporal average of oscillating condensate concentrations according to the ALVE (3), and negative entries of Ac represent depletion rates.

Depending on the values of the rate constants r_{ij}, numerical simulation of equations (1) with rates (2) reveals that all states, multiple states or only one state become condensates when the particle density N/S is large enough to detect condensation[19]. Thus far, various questions about condensation have remained elusive for the coupled birth–death process defined by equation (1): Which of the states become condensates? How does this selection of condensates proceed? Is it possible to construct systems that condense into a specific set of condensates?

In the following, we answer these questions by illuminating the physical principle that governs the formation of multiple condensates on the leading-order timescale. We show that the vanishing of relative entropy production determines the selection of condensates (see equations (3) and (4) below). We elaborate how condensate selection is determined by the rate constants r_{ij}. The condensation proceeds exponentially fast into a dynamic, metastable steady state within which the occupation numbers of condensates may oscillate. By applying our general results to systems with many states, we show that the interplay between critical properties of such networks of states[20] and dynamically stable network motifs[21] determines the selection of condensates. The results of our analysis apply to any system whose dynamics

are described by the coupled birth–death processes (1) with rates (2). Before proceeding to the mathematical and numerical analysis of condensation in these processes, we now give a brief overview of such systems.

Results

Non-interacting bosons in driven-dissipative systems. The classical master equation (1) has recently been derived by Vorberg *et al.*[19] in the study of bosonic systems that are dissipative and driven by external sources. For a system of non-interacting bosons that is weakly coupled to a reservoir and driven by an external time-periodic force (a so-called Floquet system)[22–24], one can eliminate the reservoir degrees of freedom (Born and Markov approximation)[25,26] and the density matrix of the system becomes diagonal (see the Supplement of the work of Vorberg *et al.*[19]). The effective dynamics of the bosons become incoherent and are captured on a macroscopic level in terms of the coupled birth–death processes (1) with rates $\Gamma_{i\leftarrow j}=r_{ij}(N_i+1)N_j$ (that is all $s_{ij}=1$ in the rates (2)). These non-equilibrium set-ups may not only lead the bosons into a single, but also into multiple condensates[19].

For the incoherently driven-dissipative systems described above, the state E_i denotes a time-dependent Floquet state[22–24]. The total rate $\Gamma_{i\leftarrow j}$ for the transition of a boson from state E_j to E_i depends linearly on the number of bosons in the departure state (N_j) and the arrival state (N_i+1). The latter factor stems from the indistinguishability of bosons and reflects their tendency to congregate. Although we refer to equation (1) as a classical master equation and coherence does not build up, the quantum statistics of bosons is still encoded in the functional form of $\Gamma_{i\leftarrow j}$. The rate constant r_{ij} is determined by the microscopic properties of the system and the reservoir.

Condensation in the above set-up is to be distinguished from Bose–Einstein condensation. Typically, studies on Bose–Einstein condensation focus on the existence of long-range phase coherence in thermal equilibrium[10–15], its kinetic formation[13,14,27–32] and the fragmentation of a coherent condensate into multiple condensates (for example, when the equilibrium ground state is degenerate)[13,16]. In contrast, the classical birth–death processes (1) with rates (2) describe condensation in bosonic systems that are externally driven by a continuing supply of energy, dissipate into the environment and exhibit decoherence.

Equations of type (1) may also arise in atomic physics and quantum optics and are known as Pauli master equations[33–35]. They describe how the population of S non-degenerate energy levels changes over time when a system harbours N indistinguishable, non-interacting bosonic atoms. Such changes may occur by interactions with a radiation field that induces transitions between energy levels. A theoretical description of these transitions in terms of a Pauli master equation is appropriate if coherence is negligible. As in the previous example, the system then approaches a state in which some of the energy levels are macroscopically occupied (condensates), whereas the others are depleted. More generally, whenever a rate constant r_{ij} governs the transition of a single boson from a state E_j to E_i, the rates (2) with $s_{ij}=1$ for all i and j apply if N non-interacting bosons are brought into the system[19].

Strategy selection in evolutionary game theory. The classical master equation (1) also occurs in evolutionary game theory (EGT). Historically, EGT was developed to study the evolutionary processes that are driven by selection and mutation[36,37] and seeks to identify optimal strategies for competitive interactions. For example, EGT has been applied in the study of the prominent 'rock–paper–scissors' (RPS) game, which was proposed as a facilitator of species coexistence and has inspired both experimental and theoretical research[38–42]. Furthermore, the 'prisoner's dilemma' game serves as a paradigmatic model to explore the evolution and maintenance of cooperation[43,44]. The interplay between non-linear and stochastic effects underlies the dynamics of such evolutionary games[45–50].

In EGT, one typically considers a system of N interacting agents (classical particles) who repeatedly play one fixed strategy E_i out of the S possible choices $E_1, E_2, \ldots, E_S$. In each succeeding interaction, the defeated agent adopts the strategy of its opponent. Since N_j agents playing strategy E_j can potentially be defeated by one of the N_i agents playing strategy E_i, the rate of change is $\Gamma_{i\leftarrow j}=r_{ij}N_iN_j$. If an agent who plays E_j can also spontaneously mutate into an agent who plays E_i (with rate $\mu_{ij}=r_{ij}s_{ij}$), one recovers the classical master equation (1) with rates (2).

Thus, there exists a correspondence between condensation in incoherently driven-dissipative bosonic systems and strategy selection in EGT—the transition of bosons between states can be interpreted in terms of the interaction and mutation of agents employing evolutionary strategies. In effect, the states in an incoherently driven-dissipative set-up play an evolutionary game and the winning states form the condensates.

After having introduced the above examples, we now proceed with the mathematical and numerical analysis of the classical master equation (1). We show that the dynamics of condensation change on two distinct timescales. At the leading-order timescale, the dynamics are described by a set of non-linearly coupled, ordinary differential equations (see equation (3) below), which determine the states that become condensates. We identify these states by applying concepts from EGT. After an exposition of the physical principles that underlie the condensation dynamics, implications of our general results for incoherently driven-dissipative systems are discussed.

The antisymmetric Lotka–Volterra equation. The total number of particles needed for condensation phenomena to occur is large ($N\gg 1$). To detect macroscopic occupancies, it is also assumed that the particle density N/S is large. Therefore, one may approximate the classical master equation (1) by a Langevin equation for the state concentrations $x_i(t)=N_i(t)/N$ (details of the derivation are provided in Supplementary Note 1). Originally proposed for Brownian particles suspended in a liquid, the Langevin equation decomposes the dynamics of a sample trajectory of the random process into two contributions—into a deterministic drift and into noise stemming from the discreteness of particle numbers ('demographic fluctuations'). Both the demographic fluctuations and the contribution to the deterministic drift that corresponds to mutations in the EGT setting are suppressed by a small prefactor $1/N$. Therefore, these terms change the dynamics only slowly. The deterministic drift that corresponds to interactions between agents is, however, not suppressed. It thus governs the dynamics to leading order.

Hence, we find that the leading-order dynamics of the condensation process (1)–(2) are described by the differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = x_i(A\mathbf{x})_i\,. \tag{3}$$

The matrix A is antisymmetric and encodes the effective transition rates between states ($a_{ij}=r_{ij}-r_{ji}$). The constants s_{ij} that occur in the definition of the rates (2) do not change the leading-order dynamics, but they become relevant on subleading-order timescales.

We refer to equation (3) as the antisymmetric Lotka–Volterra equation (ALVE). It provides a description of pairwise interactions that preserve the total number of particles. Therefore, the

Table 1 | Condensation processes described by the ALVE in different fields of research.

Field of research	Entity	State	Process	Dynamics
Quantum physics[19]	Boson	Quantum state	Incoherent transition	Condensation/depletion
Evolutionary game theory[59,60]	Agent	Strategy	Game (+mutation)	Win/loss
Population dynamics[47,51–53]	Individual	Species	Competition (+mutation)	Survival/extinction
Chemical kinetics[56–58]	Molecule	Chemical species	Reaction (+conversion)	Production/consumption
Plasma physics[54,55]	Plasmon	Jet	Scattering	Increase/decrease

ALVE, antisymmetric Lotka–Volterra equation.
The ALVE (3) governs condensation processes in diverse fields of research. For example, for incoherently driven-dissipative bosonic systems, the ALVE describes condensation and depletion of states by incoherent transitions of non-interacting bosons. In EGT, the ALVE occurs in the context of winning and losing strategies played by agents.

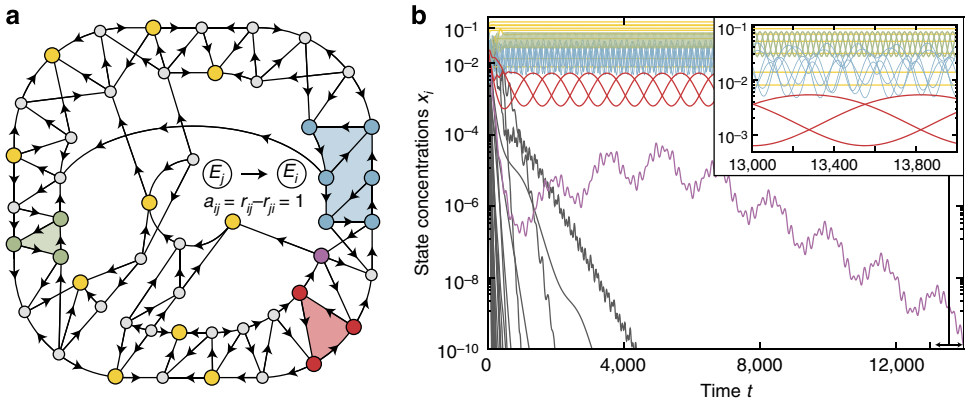


Figure 2 | Fragmentation of an exemplary system into multiple condensates with oscillating state concentrations. (**a**) Randomly sampled network of 50 states. Disks represent states. An arrow from state E_j to state E_i represents an effective rate constant $a_{ij} = r_{ij} - r_{ji} = 1$ (a missing arrow indicates a forbidden transition with $a_{ij} = 0$). Computation of a condensate vector c predicted relaxation into 10 isolated condensates (yellow), one interacting subsystem with six condensates (blue) and two RPS cycles (red and green). All other states become depleted. The complete network also comprises RPS cycles of which some states become depleted. Knowledge of the network topology alone is thus insufficient to determine condensates. (**b**) Temporal evolution of state concentrations x_i (logarithmic scale). Colours in accordance with **a**. Numerical integration of the ALVE (3) confirmed the selection of states based on the condensate vector c. Subsystems with six (blue) and three (red and green) condensates exhibit oscillations of concentrations with non-vanishing particle flow. Depletion of states occurs exponentially fast. Identifying condensates from a condensate vector c is more reliable than through numerical integration: The concentration of the state associated to the purple disk in **a** decays exponentially to a concentration of 1.5×10^{-7} before recovering transiently. Numerical integration cannot rule out permanent recovery at later times. Supplementary Fig. 2 demonstrates such a case.

ALVE finds a broad range of applications in diverse fields of research, in addition to the aforementioned condensation of bosons far from equilibrium. It was first studied by Volterra[51] in the context of predator–prey oscillations in population biology[47,52,53]. In plasma physics, the ALVE describes the spectra of plasma oscillations (Langmuir waves)[54,55], and in chemical kinetics it captures the dynamics of bimolecular autocatalytic reactions[18,56–58]. In EGT, the ALVE is known as the replicator equation of zero-sum games such as the RPS game[47,59–61]. Table 1 summarizes all of the above analogies.

Despite the simple structure of the ALVE, it exhibits a rich and complex behaviour. In the following, we show how the mathematical analysis of the ALVE explains condensation into multiple states (condensate selection). To this end, we extend an approach for the analysis of the ALVE that was introduced in the context of EGT[59,60].

Production of relative entropy and condensate selection. Our analysis starts from a theorem in linear programming theory[62]. Given an antisymmetric matrix A, it is always possible to find a vector $\mathbf{c}$ that fulfils the following conditions: its entries are positive for indices in $I \subseteq \{1, \ldots, S\}$ and zero for indices in $\bar{I} = \{1, \ldots, S\} - I$, whereas the entries of $A\mathbf{c}$ are zero for indices in I and negative for indices in $\bar{I}$ (Fig. 1b). Although several vectors $\mathbf{c}$ with these properties may exist, the index set I is unique and, thus, determined by the antisymmetric matrix A. Finding such a 'condensate vector' $\mathbf{c}$ is crucial for the understanding of condensate selection and of the condensation dynamics. The condensate vector has the following physical interpretations.

All condensate vectors yield fixed points of the ALVE (3). Because of the antisymmetry of matrix A, a linear stability analysis of these fixed points does not yield insight into the global dynamics (Supplementary Note 1). However, the global stability properties can be inferred by showing that the relative entropy of a condensate vector to the state concentrations,

$$D(\mathbf{c}||\mathbf{x}) = \sum_{i \in I} c_i \log(c_i/x_i), \qquad (4)$$

is a Lyapunov function (note that we do not consider the relative entropy of the state concentrations to the condensate vector, but define the relative entropy vice versa). The relative entropy (4) decreases with time and is bounded from below (see Methods and Supplementary Fig. 1). Therefore, the dynamics relax to a subsystem in which the relative entropy production is zero. The relaxation of relative entropy production is reminiscent of Prigogine's study of open systems in non-equilibrium thermodynamics. Indeed, we find that the system, to cite Prigogine's phrase, 'settles down to the state of least dissipation'[63].

This state of least dissipation is characterized even further by the condensate vector $\mathbf{c}$. Considering the definition of the relative entropy (4) and its boundedness, it follows that every concentration x_i with $i \in I$ remains larger than a positive constant. On the

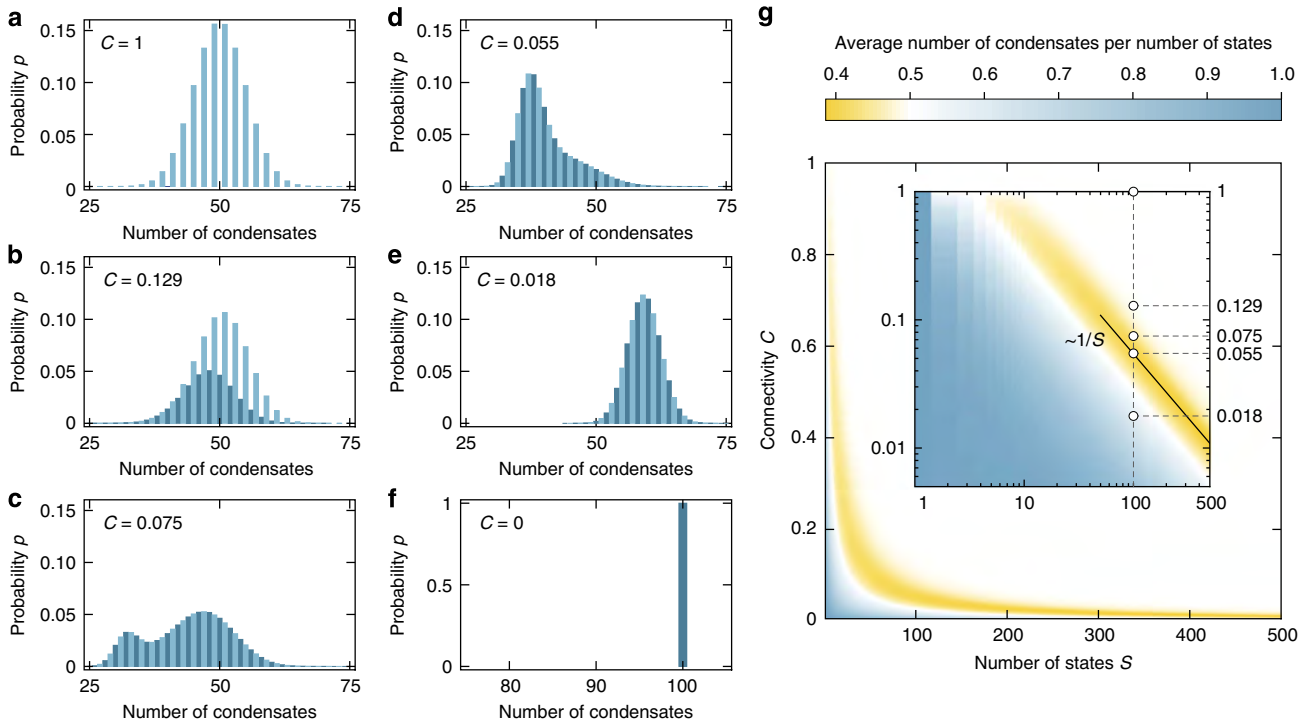


Figure 3 | Dependence of the number of condensates on the connectivity of states in random networks. (**a**–**f**) Measured probability p of finding a particular number of condensates for a system with $S=100$ states and with connectivity C (5×10^6 systems analysed per histogram). The connectivity specifies the percentage of states between which transitions of particles occur with a non-zero effective rate constant $a_{ij}=r_{ij}-r_{ji}$. Effective rate constants a_{ij} were sampled from a Gaussian distribution (zero mean, unit variance). (**a**) At full connectivity, the distribution is pseudobinomial with only odd numbers of condensates ($C=1$; light blue bars)[19,60,66]. (**b**) As the connectivity is reduced, even numbers of condensates become possible when systems decouple into even numbers of subsystems ($C=0.129$; dark blue bars). (**c**) The distribution exhibits bimodality ($C=0.075$) and (**d**) approaches a minimal average number of 40.2 condensates ($C=0.055$). (**e**) This average subsequently increases ($C=0.018$) because isolated states are trivially selected as condensates ($C=0$) as shown in **f**. (**g**) Average number of condensates per number of states (colour coded) plotted against the number of states S and the connectivity C (log–log graph in inset; $\geq10^4$ systems per data point, see Supplementary Fig. 4 for the reliability of the linear programming algorithm). White circles correspond to distributions shown in (**a**–**e**). The minimal relative number of condensates conforms to the power law $C\sim1/S^{\gamma}$ with $\gamma=0.998\pm0.008$ (s.e.m.) and can be related to the criticality of random networks[20].

other hand, states with indices in $\bar{I}$ become depleted for long times (see Methods). Therefore, we find that the condensate vector determines the selection of condensates. Positive entries of $\mathbf{c}$ correspond to states that become condensates, whereas zero entries of $\mathbf{c}$ correspond to states that become depleted. Both the set of condensates and the set of depleted states are unique (Fig. 1b) and independent of the initial conditions. Generically, the entries of the condensate vector are also unique upon normalization (its entries sum up to one) and yield the rate $|(A\mathbf{c})_i|$ at which a state E_i becomes depleted. The condensate selection occurs exponentially fast (see Fig. 2 and Methods).

After relaxation, the dynamics of the system are restricted to the condensates. In other words, the condensates form the attractor of the dynamics. However, the dynamics in this subsystem do not come to rest. The state of least dissipation is a dynamic state with a perpetually changing number of particles in the condensates—periodic, quasiperiodic and non-periodic oscillations are observed (Fig. 2b and Supplementary Fig. 1). In the generic case, the entries of the condensate vector represent the temporal average of condensate concentrations according to the ALVE (3). After condensate selection, the dynamics of these active condensates take place on a high-dimensional, deformed sphere[61].

An algebraic algorithm to find the condensates. Numerical integration of the ALVE (3) is neither a feasible nor a reliable method for identifying condensates (Fig. 2, Supplementary Figs 1 and 2). Instead, we determine these states by numerically searching for a condensate vector $\mathbf{c}$. To this end, we reformulate the above conditions on $\mathbf{c}$ in terms of two linear inequalities[62]:

$$A\mathbf{c}\leq 0 \quad \text{and} \quad \mathbf{c}-A\mathbf{c}>0\,. \tag{5}$$

We solve these inequalities with a linear programming algorithm that is both reliable and efficient. The time to find a condensate vector scales only polynomially with the number of states S (see Supplementary Fig. 3 and Methods for details).

Condensation in large random networks of states. We used our combined analytical and numerical approach to study how the connectivity of a random network of states affects the selection of condensates under the dynamics of the ALVE (3). The connectivity specifies the percentage of states between which particle transitions occur[20,64]. After having generated a network with a given connectivity, the strength and direction of an allowed transition between states E_j and E_i were determined by randomly sampling the corresponding effective rate constant $a_{ij}=r_{ij}-r_{ji}$.

Our results for condensation in large random networks of states are summarized in Fig. 3. When the connectivity of a network is zero, all of its states are isolated. Particles are not exchanged between states and none of the states becomes depleted. For an increased connectivity, isolated pairs of states are sampled in a random network. One state in an isolated pair is always depleted and the average number of condensates decreases

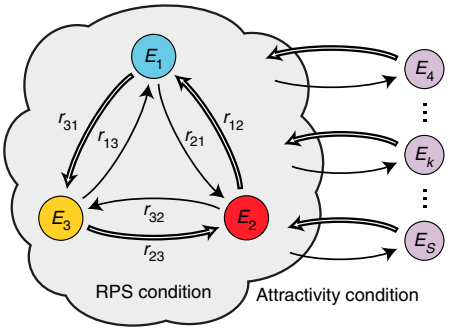


Figure 4 | Conditions for the emergence of a RPS cycle of condensates. Three particular states E_1, E_2 and E_3 (blue, red and yellow disks) of a network condense into a RPS cycle if, and only if, two conditions are fulfilled: First, the 'RPS condition' requires that the rate constants r_{ij} between the three states form a RPS network: $r_{i-1,i+1} > r_{i+1,i-1}$ (indices are counted modulo 3, for example, $r_{42} = r_{12}$ (framed arrows denote rate constants that are larger than rate constants for the respective reverse direction). Differences between these rate constants define the entries $c_i = r_{i-1,i+1} - r_{i+1,i-1}$ of an admissible condensate vector c. Second, the 'attractivity condition' requires that the weighted sum of rates from any exterior state E_k (purple disks) into the RPS cycle, $\sum_{j=1}^{3} c_j r_{jk}$ (framed arrows), is larger than the weighted sum of outbound rates, $\sum_{j=1}^{3} c_j r_{kj}$ (black arrows). In other words, the inflow of particles into the RPS cycle from any exterior state needs to be greater than the outflow to that state.

rapidly. On approaching a critical connectivity, cycles and trees of all orders become embedded in a random network. This critical connectivity scales inversely with the number of states[20]. We observe that, under the dynamics of the ALVE (3), the average number of condensates becomes minimal for a connectivity that also scales inversely with the number of states (see Fig. 3g). We attribute this minimum to the interplay between the criticality of random networks and condensate selection on connected components of the network[61,65]. Embedded directed cycles are a recurring motif[21] in the remaining network of condensates after condensate selection. Above the critical connectivity, a single giant cluster is formed. On average, half the number of states in this giant cluster become condensates once the network is fully connected ($C = 1$; Fig. 3a)[19,60,66]. Thus, our analysis emphasizes the importance of critical properties of random networks for condensate selection.

Design of active condensates. Our understanding of the condensate selection can be used to design systems that condense into a particular network of states, a game of condensates. We exemplify this procedure by formulating conditions under which a system relaxes into a RPS game of condensates[40]. Three particular states E_1, E_2 and E_3 in a system become a RPS cycle of condensates if, and only if, the following two conditions are fulfilled (Fig. 4). First, the 'RPS condition' requires that the rate constants between the three states form a RPS network (for example, $r_{12} > r_{21}$, $r_{23} > r_{32}$ and $r_{31} > r_{13}$). Second, the 'attractivity condition' requires that the inflow of particles into the RPS cycle from any other state E_k is greater than the outflow to that state E_k (for all $k = 4, \ldots, S$). The values of the rate constants between the states that become depleted are irrelevant. More complex games of condensates can be designed by formulating similar conditions on the rate constants. These conditions are formulated as inequalities that depend on Pfaffians of the antisymmetric matrix A and its submatrices (see Methods)[61]. The flow of particles between states in these systems causes condensate concentrations to oscillate (Fig. 2b).

Discussion

Our findings thus suggest intriguing dynamics of condensates in systems whose temporal evolutions are captured by the classical master equation (1) with rates (2), for example, in driven-dissipative systems of non-interacting bosons. Condensates observed on the leading-order timescale are metastable. For longer times, relaxation into a steady state occurs[19,67]. When detailed balance is broken in the system of condensates, the net probability current between at least two states does not vanish and a non-equilibrium steady state is approached[68,69]. The simplest way of designing such condensates is illustrated by the above RPS game. In this game, detailed balance is broken, for example, when the transition of particles is unidirectional (with totally asymmetric rate constants $r_{12} > r_{21} = 0$, $r_{23} > r_{32} = 0$, and $r_{31} > r_{13} = 0$). For non-interacting bosons in driven-dissipative systems, the continuing supply with energy through the external time-periodic driving force (Floquet system) and the dissipation of energy into the environment may, therefore, prevent the system from reaching equilibrium. How such systems may be realized in an experiment poses an interesting question for future research.

The transition of particles between condensates in the here-studied coupled birth–death processes parallels the interaction and mutation of winning agents in evolutionary game theory, reflecting an 'evolutionary game of condensates'. Our results suggest the possibility of creating novel bosonic systems with an oscillating occupation of condensates. Non-interacting bosons in incoherently driven-dissipative systems are promising candidates. Since the antisymmetric Lotka-Volterra equation also arises in population biology, chemical kinetics and plasma physics, all of our mathematical results apply to these fields as well.

Methods

Asymptotics of the ALVE. The asymptotic behaviour of the ALVE (3) can be characterized as follows: for every antisymmetric matrix A there exists a unique subset of states $I \subseteq \{1, \ldots, S\}$ whose concentrations stay away from zero for all times, that is,

$$x_i(t) \geq Const(A, \mathbf{x}_0) > 0 \text{ for all } t \geq 0 \text{ and for every } i \in I\,. \tag{6}$$

The set I is the set of condensates. All of the other states with indices in $\bar{I} = \{1, \ldots, S\} - I$ become depleted as $t \to \infty$, that is,

$$x_i(t) \to 0 \text{ as } t \to \infty \text{ for every } i \in \bar{I}\,. \tag{7}$$

The set of condensates can be determined algebraically from the antisymmetric matrix A and does not depend on the initial conditions $\mathbf{x}_0 \in \Delta_{S-1} = \{\mathbf{x} \in \mathbb{R}^S \mid x_i > 0 \text{ for all } i,\ \sum_{i=1}^{S} x_i = 1\}$.

To show this result, the time-dependent entropy $D(\mathbf{c}||\mathbf{x})(t)$ of a condensate vector $\mathbf{c} = (c_1, \ldots, c_n) \in \overline{\Delta}_{S-1}$ ($c_i \geq 0$ for all i and $\sum_i c_i = 1$) relative to the trajectory $\mathbf{x}(t)$ is considered (that is, the Kullback–Leibler divergence of $\mathbf{x}(t)$ from $\mathbf{c}$), see equation (4). A condensate vector is defined via the properties (see Fig. 1b):

$$c_i > 0 \text{ and } (A\mathbf{c})_i = 0 \quad \text{for } i \in I, \text{ and} \tag{8}$$

$$c_i = 0 \text{ and } (A\mathbf{c})_i < 0 \quad \text{for } i \in \bar{I}\,. \tag{9}$$

Such a vector can always be found for an antisymmetric matrix[62]. Notably, the index set I is unique although more than one condensate vector may exist.

Considering the time derivative of the relative entropy $D(\mathbf{c}||\mathbf{x})(t)$ and employing equations (3) and (8) yields:

$$\frac{\mathrm{d}}{\mathrm{d}t} D(\mathbf{c}||\mathbf{x})(t) = -\sum_{i=1}^{S} c_i \frac{\partial_t x_i}{x_i} = -\sum_{i=1}^{S} c_i (A\mathbf{x})_i = \sum_{i=1}^{S} (A\mathbf{c})_i x_i$$
$$= \sum_{i \in \bar{I}} (A\mathbf{c})_i x_i\,. \tag{10}$$

Since $(A\mathbf{c})_{\bar{I}} < 0$ and $\mathbf{x} > 0$, it follows that $\partial_t D(\mathbf{c}||\mathbf{x})(t) < 0$ (please note the overbars in subscripts, which may be lost when read in low resolution). Therefore, the relative entropy $D(\mathbf{c}||\mathbf{x})$ is a Lyapunov function if $\mathbf{c}$ is chosen in accordance with equations (8) and (9). Moreover, $D(\mathbf{c}||\mathbf{x})$ is bounded from above by $D(\mathbf{c}||\mathbf{x})(0)$ and from below by zero for all times. This can be seen from the definition of D, and

from the integration of equation (10) (using that $(A\mathbf{c})_{\bar{I}} < 0$ and $\mathbf{x} > 0$):

$$0 \le D(\mathbf{c}||\mathbf{x})(t) = D(\mathbf{c}||\mathbf{x})(0) + \int_0^t ds \sum_{i\in I}(A\mathbf{c})_i x_i(s) \le D(\mathbf{c}||\mathbf{x})(0) . \quad (11)$$

From the definition of the relative entropy in equation (4), it follows that every concentration x_i with $i \in I$ remains larger than a positive constant, that is, $x_i(t) \ge Const(A, \mathbf{x}_0) > 0$ for all times t (if $x_i(t) \to 0$ for $i \in I$, it follows that $D \to \infty$, which contradicts the boundedness of D).

Furthermore, equation (11) implies that,

$$-\int_0^t \mathrm{d}s\,(A\mathbf{c})_i x_i(s) \le -\int_0^t \mathrm{d}s \sum_{i\in\bar{I}}(A\mathbf{c})_i x_i(s) \le D(\mathbf{c}||\mathbf{x})(0) , \quad (12)$$

for every $i \in \bar{I}$ and for all t. Therefore, concentration x_i is integrable for every $i \in \bar{I}$ ($x_i \in L^1(0,\infty)$) with the bound:

$$0 < \int_0^\infty \mathrm{d}s\, x_i(s) \le \frac{D(\mathbf{c}||\mathbf{x})(0)}{-(A\mathbf{c})_i} = Const(A, \mathbf{x}_0) \quad \text{for every } i \in \bar{I} . \quad (13)$$

Since the derivative of the concentrations is bounded from above and below, $|\partial_t x_i| = |x_i(A\mathbf{x})_i| \le ||(A\mathbf{x})||_\infty \le ||A||_{\infty\to\infty} \le \mathrm{Const}(A)$, one concludes that x_i is uniformly continuous ($||A||_{\infty\to\infty}$ denotes the operator norm of A induced by the maximum norm on $\mathbb{R}^S$). Together with the integrability (13), it follows that the states with indices in $\bar{I}$ become depleted as $t \to \infty$, that is, $x_i(t) \to 0$ for $i \in \bar{I}$.

In conclusion, given an antisymmetric matrix $A = R - R^T$ via a rate constant matrix $R = \{r_{ij}\}_{i,j}$, one finds a condensate vector $\mathbf{c}$ that satisfies inequalities (8)–(9). The index set I, for which the entries of $\mathbf{c}$ are positive, represents condensates. The index set $\bar{I}$, for which entries of $\mathbf{c}$ are zero, represents states that become depleted. Moreover, equation (10) implies that the relative entropy becomes a conserved quantity in the subsystem of condensates (Supplementary Fig. 1).

Temporal average of condensate concentrations. The ALVE (3) is solved implicitly by,

$$x_i(t) = x_i(0)e^{t\cdot\left(A\langle\mathbf{x}\rangle_t\right)_i} , \quad (14)$$

with the time average of the trajectory $\langle\mathbf{x}\rangle_t$ defined as:

$$\langle\mathbf{x}\rangle_t = \frac{1}{t}\int_0^t \mathrm{d}s\, \mathbf{x}(s) . \quad (15)$$

It is shown above that $0 < Const(A, \mathbf{x}_0) \le x_i(t) \le 1$ holds for the states that become condensates ($i \in I$). By rearranging equation (14), one thus obtains:

$$\left|\left(A\langle\mathbf{x}\rangle_t\right)_i\right| \le \frac{1}{t}\left|\log\left(\frac{x_i(t)}{x_i(0)}\right)\right| \le \frac{Const(A, \mathbf{x}_0)}{t} \quad \text{for all } i \in I . \quad (16)$$

Note that *Const* is used to denote arbitrary positive, time-independent constants. Therefore, the right-hand side of equation (16) vanishes for $t \to \infty$. On the other hand, x_i is integrable for $i \in \bar{I}$ (equation (13)). Thus, the corresponding component of the time average converges to zero,

$$\langle x_i\rangle_t \le \frac{Const(A, \mathbf{x}_0)}{t} \to 0 \text{ as } t \to \infty \text{ for every } i \in \bar{I} . \quad (17)$$

Hence, the distance of the time average $\langle\mathbf{x}\rangle_t$ to the kernel of the antisymmetric submatrix A^I converges to zero (the submatrix A^I corresponds to the system of condensates with indices in I).

Structure of a generic antisymmetric matrix. For systems with an even number of states S, the antisymmetric matrix $A = R - R^T$ generically has a trivial kernel, whereas for systems with an odd number of states, the kernel of A is generically one dimensional. A higher dimensional kernel of A only occurs if the matrix entries are tuned[52,61,70]. As a consequence, when all of the entries above the diagonal of A are, for example, randomly drawn from a continuous probability distribution (for example from a Gaussian distribution), all 2^S submatrices of A have a kernel with dimension of less than or equal to one.

The projection of $\mathbf{x} \in \mathbb{R}^S$ to the subspace $\mathbb{R}^{(J)} \subseteq \mathbb{R}^S$ for an arbitrary index set $J \subseteq \{1, \ldots, S\}$ is defined as $\mathbf{x}_J := P_J\mathbf{x} := (x_j)_{j\in J}$. In other words, entries of $\mathbf{x}_J$ are zero for indices in the complement $\bar{J}$. In the following, the short notation $A^J := P_J A P_J$ is also used (see above). Furthermore, the set of antisymmetric matrices whose submatrices have a kernel with dimension ≤ 1 is defined:

$$\Omega := \left\{A \in \mathbb{R}^{S\times S} \mid A \text{ is antisymmetric and } \dim\ker A^J \le 1 \text{ for all } J \subseteq \{1, \ldots, S\}\right\} . \quad (18)$$

The complement $\bar{\Omega}$ has measure zero with respect to the flat measure $\mathrm{d}A$ on antisymmetric matrices (the translation invariant measure, which is sigma-finite and not trivial).

For antisymmetric matrices $A \in \Omega$, the kernel can be characterized as follows[61,70]. If the number of states S is even, the kernel of A is trivial: $\ker A = \{\mathbf{0}\}$. If the number of states is odd, the kernel is one-dimensional: $\ker A = \{\mathbf{v}\}$. This kernel element can be computed analytically in terms of Pfaffians of submatrices of A:

$$\mathbf{v} = \left(\mathbf{Pf}(A_{\hat{1}}), -\mathbf{Pf}(A_{\hat{2}}), \ldots, \mathbf{Pf}(A_{\hat{S}})\right) . \quad (19)$$

The submatrix $A_{\hat{k}} \in \mathbb{R}^{(S-1)\times(S-1)}$ denotes the matrix for which the k-th column and row are removed from A.

For antisymmetric matrices $A \in \Omega$, the normalized condensate vector $\mathbf{c}$ with $\sum_i c_i = 1$ and with properties (8), (9) is unique. The latter follows from $A^I\mathbf{c} = 0$ (equations (8) and (18)). Therefore, the condensate vector is the unique kernel vector of the subsystem of condensates whose interactions are characterized by the matrix A^I. Furthermore, I contains an odd number of elements. To determine the condensate vector for $A \in \Omega$, one can proceed as follows. For each odd-dimensional submatrix A^I with $I \subseteq \{1, \ldots, S\}$, one computes the kernel element $\mathbf{v}$ according to equation (19) and defines the vector $\mathbf{w} \in \mathbb{R}^S$ by setting $\mathbf{w}_I = \mathbf{v}$ and $\mathbf{w}_{\bar{I}} = 0$. There exists exactly one set I for which $(A\mathbf{w})_{\bar{I}} < 0$. The corresponding vector $\mathbf{w}$ is the unique condensate vector upon normalization.

Temporal average of condensate concentrations (generic case). It was shown above that the temporal average of condensate concentrations $\langle\mathbf{x}\rangle_t$ converges to a non-negative kernel element of the antisymmetric matrix A^I. In the generic case, the condensate vector $\mathbf{c}$ is the unique kernel element of A^I upon normalization. Therefore, positive entries of $\mathbf{c}$ represent the asymptotic temporal average of condensate concentrations,

$$\left\|\langle\mathbf{x}\rangle_t - \mathbf{c}\right\|_\infty \le \frac{Const(A, \mathbf{x}_0)}{t} \to 0 \text{ as } t \to \infty. \quad (20)$$

Exponentially fast depletion of states (generic case). On inserting equation (20) into the implicit solution (14) of the ALVE, the exponentially fast depletion of states with $i \in \bar{I}$ can be seen as follows (note that $(A\mathbf{c})_i < 0$ according to the choice of the condensate vector in equations (8) and (9)):

$$x_i(t) = x_i(0)e^{t\cdot\left(A\langle\mathbf{x}\rangle_t\right)_i} \quad (21)$$

$$\le x_i(0)e^{t\cdot\left((A\mathbf{c})_i + \|A(\langle\mathbf{x}\rangle_t - \mathbf{c})\|_\infty\right)} \quad (22)$$

$$\le x_i(0)e^{t\cdot(A\mathbf{c})_i + Const(A,\mathbf{x}_0)} \quad (23)$$

$$= Const(A, \mathbf{x}_0)e^{t\cdot(A\mathbf{c})_i} , \quad (24)$$

and analogously,

$$x_i(t) = x_i(0)e^{t\cdot\left(A\langle\mathbf{x}\rangle_t\right)_i} \ge Const(A, \mathbf{x}_0)e^{t\cdot(A\mathbf{c})_i} . \quad (25)$$

Therefore, condensate selection occurs exponentially fast at depletion rate $|(A\mathbf{c})_i|$. The dynamics of cases for non-generic antisymmetric matrices are discussed in Supplementary Note 2.

Linear programming algorithm. For the numeric computation of condensate vectors $\mathbf{c}$, a finite threshold $\delta > 0$ was introduced into the inequalities (5): $A\mathbf{c} \le 0$ and $\mathbf{c} - A\mathbf{c} \ge \delta > 0$. Its value was set to $\delta = 1$ by rescaling of $\mathbf{c}$. Numerical solution of the inequalities was performed by using the IBM ILOG CPLEX Optimization Studio 12.5 and its interface to the C++ language. The software Mathematica 9.0 from Wolfram Research was also found to be applicable. Further information on the calibration of the linear programming algorithm and a simplified Mathematica algorithm are provided in Supplementary Note 3.

References

1. Krapivsky, P. L., Redner, S. & Leyvraz, F. Connectivity of growing random networks. *Phys. Rev. Lett.* **85,** 4629–4632 (2000).
2. Bianconi, G. & Barabási, A.-L. Bose-Einstein condensation in complex networks. *Phys. Rev. Lett.* **86,** 5632–5635 (2001).
3. Evans, M. R. & Hanney, T. Nonequilibrium statistical mechanics of the zero-range process and related models. *J. Phys. A: Math. Gen.* **38,** R195–R240 (2005).
4. Evans, M. R. Bose-Einstein condensation in disordered exclusion models and relation to traffic flow. *Europhys. Lett.* **36,** 13–18 (1996).
5. Krug, J. & Ferrari, P. A. Phase transitions in driven diffusive systems with random rates. *J. Phys. A: Math. Gen.* **29,** L465–L471 (1996).
6. Chowdhury, D., Santen, L. & Schadschneider, A. Statistical physics of vehicular traffic and some related systems. *Phys. Rep.* **329,** 199–329 (2000).
7. Kaupužs, J., Mahnke, R. & Harris, R. J. Zero-range model of traffic flow. *Phys. Rev. E* **72,** 056125 (2005).
8. Spitzer, F. Interaction of Markov processes. *Adv. Math.* **5,** 246–290 (1970).
9. Evans, M. R. & Waclaw, B. Condensation in stochastic mass transport models: beyond the zero-range process. *J. Phys. A: Math. Theor.* **47,** 095001 (2014).
10. Bose, S. N. Plancks Gesetz und Lichtquantenhypothese. *Z. Phys.* **26,** 178–181 (1924).
11. Einstein, A. Quantentheorie des einatomigen idealen Gases. *Sitzb. d. Preuss. Akad. d. Wiss* 261–267 (1924).

12. Einstein, A. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzb. d. Preuss. Akad. d. Wiss* 3–14 (1925).
13. Griffin, A., Snoke, D. & Stringari, G. *Bose Einstein Condensation* (Cambridge Univ. Press, 1995).
14. Anglin, J. R. & Ketterle, W. Bose-Einstein condensation of atomic gases. *Nature* **416,** 211–218 (2002).
15. Penrose, O. & Onsager, L. Bose-Einstein condensation and liquid helium. *Phys. Rev.* **104,** 576–584 (1956).
16. Mueller, E. J., Ho, T.-L., Ueda, M. & Baym, G. Fragmentation of Bose-Einstein condensates. *Phys. Rev. A* **74,** 033612 (2006).
17. Gardiner, C. *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, 2009).
18. Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (Elsevier, 2007).
19. Vorberg, D., Wustmann, W., Ketzmerick, R. & Eckardt, A. Generalized Bose-Einstein condensation into multiple states in driven-dissipative systems. *Phys. Rev. Lett.* **111,** 240405 (2013).
20. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74,** 47–97 (2002).
21. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298,** 824–827 (2002).
22. Blümel, R. *et al.* Dynamical localization in the microwave interaction of Rydberg atoms: the influence of noise. *Phys. Rev. A* **44,** 4521–4540 (1991).
23. Kohler, S., Dittrich, T. & Hänggi, P. Floquet-Markovian description of the parametrically driven, dissipative harmonic quantum oscillator. *Phys. Rev. E* **55,** 300–313 (1997).
24. Breuer, H.-P., Huber, W. & Petruccione, F. Quasistationary distributions of dissipative nonlinear quantum oscillators in strong periodic driving fields. *Phys. Rev. E* **61,** 4883–4889 (2000).
25. Grifoni, M. & Hänggi, P. Driven quantum tunneling. *Phys. Rep.* **304,** 229–354 (1998).
26. Breuer, H.-P. & Petruccione, F. *The Theory of Open Quantum Systems* (Oxford Univ. Press, 2002).
27. Gardiner, C. W. & Zoller, P. Quantum kinetic theory: a quantum kinetic master equation for condensation of a weakly interacting Bose gas without a trapping potential. *Phys. Rev. A* **55,** 2902–2921 (1997).
28. Kagan, Y. & Svistunov, B. V. Evolution of correlation properties and appearance of broken symmetry in the process of Bose-Einstein condensation. *Phys. Rev. Lett.* **79,** 3331–3334 (1997).
29. Bijlsma, M. J., Zaremba, E. & Stoof, H. T. C. Condensate growth in trapped Bose gases. *Phys. Rev. A* **62,** 063609 (2000).
30. Gardiner, C. W., Lee, M. D., Ballagh, R. J., Davis, M. J. & Zoller, P. Quantum kinetic theory of condensate growth: comparison of experiment and theory. *Phys. Rev. Lett.* **81,** 5266–5269 (1998).
31. Walser, R., Williams, J., Cooper, J. & Holland, M. Quantum kinetic theory for a condensed bosonic gas. *Phys. Rev. A* **59,** 3878–3889 (1999).
32. Kocharovsky, V. V., Scully, M. O., Zhu, S.-Y. & Suhail Zubairy, M. Condensation of N bosons. II. Nonequilibrium analysis of an ideal Bose gas and the laser phase-transition analogy. *Phys. Rev. A* **61,** 023609 (2000).
33. Pauli, W. *Festschrift zum 60. Geburtstage A. Sommerfeld* (Hirzel, 1928).
34. Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, 1995).
35. Gardiner, C. W. & Zoller, P. *Quantum Noise* (Springer, 2004).
36. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, 1982).
37. Nowak, M. A. & Sigmund, K. Evolutionary dynamics of biological games. *Science* **303,** 793–799 (2004).
38. Sinervo, B. & Lively, C. M. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380,** 240–243 (1996).
39. Kerr, B., Riley, M., Feldman, M. & Bohannan, B. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418,** 171–174 (2002).
40. Reichenbach, T., Mobilia, M. & Frey, E. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448,** 1046–1049 (2007).
41. Weber, M. F., Poxleitner, G., Hebisch, E., Frey, E. & Opitz, M. Chemical warfare and survival strategies in bacterial range expansions. *J. R. Soc. Interface* **11,** 20140172 (2014).
42. Szolnoki, A. *et al.* Cyclic dominance in evolutionary games: a review. *J. R. Soc. Interface* **11,** 20140735 (2014).
43. Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428,** 646–650 (2004).
44. Szolnoki, A., Antonioni, A., Tomassini, M. & Perc, M. Binary birth-death dynamics and the expansion of cooperation by means of self-organized growth. *EPL* **105,** 48001 (2014).
45. McKane, A. J. & Newman, T. J. Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* **94,** 218102 (2005).
46. Traulsen, A., Claussen, J. C. & Hauert, C. Coevolutionary dynamics: from finite to infinite populations. *Phys. Rev. Lett.* **95,** 238701 (2005).
47. Reichenbach, T., Mobilia, M. & Frey, E. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E* **74,** 51907 (2006).
48. Melbinger, A., Cremer, J. & Frey, E. Evolutionary game theory in growing populations. *Phys. Rev. Lett.* **105,** 178101 (2010).
49. Biancalani, T., Dyson, L. & McKane, A. J. Noise-induced bistable states and their mean switching time in foraging colonies. *Phys. Rev. Lett.* **112,** 038101 (2014).
50. Rulands, S., Jahn, D. & Frey, E. Specialization and bet hedging in heterogeneous populations. *Phys. Rev. Lett.* **113,** 108102 (2014).
51. Volterra, V. *Leçons sur la Théorie Mathématique de la Lutte pour la Vie* (Gauthier-Villars, 1931).
52. Goel, N. S., Maitra, S. C. & Montroll, E. W. On the Volterra and other nonlinear models of interacting populations. *Rev. Mod. Phys.* **43,** 231–276 (1971).
53. May, R. M. *Stability and Complexity in Model Ecosystems* (Princeton Univ. Press, 1973).
54. Zakharov, V., Musher, S. & Rubenchik, A. Nonlinear stage of parametric wave excitation in a plasma. *JETP Lett.* **19,** 151–152 (1974).
55. Manakov, S. Complete integrability and stochastization of discrete dynamical systems. *Sov. Phys.-JETP* **40,** 269–274 (1975).
56. Itoh, Y. Boltzmann equation on some algebraic structure concerning struggle for existence. *Proc. Jpn Acad.* **47,** 854–858 (1971).
57. Di Cera, E., Phillipson, P. E. & Wyman, J. Chemical oscillations in closed macromolecular systems. *Proc. Natl Acad. Sci. USA* **85,** 5923–5926 (1988).
58. Di Cera, E., Phillipson, P. E. & Wyman, J. Limit-cycle oscillations and chaos in reaction networks subject to conservation of mass. *Proc. Natl Acad. Sci. USA* **86,** 142–146 (1989).
59. Akin, E. & Losert, V. Evolutionary dynamics of zero-sum games. *J. Math. Biol.* **20,** 231–258 (1984).
60. Chawanya, T. & Tokita, K. Large-dimensional replicator equations with antisymmetric random interactions. *J. Phys. Soc. Jpn* **71,** 429–431 (2002).
61. Knebel, J., Krüger, T., Weber, M. F. & Frey, E. Coexistence and survival in conservative Lotka-Volterra networks. *Phys. Rev. Lett.* **110,** 168106 (2013).
62. Kuhn, H. & Tucker, A. *Linear Inequalities and Related Systems* (Princeton Univ. Press, 1956).
63. Prigogine, I. Time, structure, and fluctuations. *Science* **201,** 777–785 (1978).
64. Gardner, M. R. & Ashby, W. R. Connectance of large dynamic (cybernetic) systems: critical values for stability. *Nature* **228,** 784 (1970).
65. Durney, C. H., Case, S. O., Pleimling, M. & Zia, R. K. P. Saddles, arrows, and spirals: deterministic trajectories in cyclic competition of four species. *Phys. Rev. E* **83,** 051108 (2011).
66. Allesina, S. & Levine, J. M. A competitive network theory of species diversity. *Proc. Natl Acad. Sci. USA* **108,** 5638–5642 (2011).
67. Eisert, J., Friesdorf, M. & Gogolin, C. Quantum many-body systems out of equilibrium. *Nat. Phys.* **11,** 124–130 (2015).
68. Zia, R. K. P. & Schmittmann, B. Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states. *J. Stat. Mech* **2007,** P07012 (2007).
69. Kriecherbauer, T. & Krug, J. A pedestrian's view on interacting particle systems, KPZ universality and random matrices. *J. Phys. A: Math. Theor.* **43,** 403001 (2010).
70. Cullis, C. E. *Matrices and Determinoids* vol. I and II (Cambridge Univ. Press, 1913).

Acknowledgements

We are grateful for fruitful discussions with Peter Zoller, Ulrich Schollwöck, Immanuel Bloch, Wilhelm Zwerger, André Eckardt, Daniel Vorberg, Alexander Schnell, Marianne Bauer, Brendan Osberg, Jacob Halatek, Matthias Bauer and Sebastian Koch. This work was supported by the Deutsche Forschungsgemeinschaft as project A7 of the SFB TR 12 'Symmetry and Universality in Mesoscopic Systems', and by the German Excellence Initiative via the programme 'Nanosystems Initiative Munich' (NIM). J.K. acknowledges funding from the Studienstiftung des Deutschen Volkes.

Author contributions

J.K., M.F.W., T.K. and E.F. designed, discussed and planned the study. T.K., J.K. and M.F.W. developed the analytical results. M.F.W., T.K. and J.K. developed the numerical algorithms and generated the data. J.K., M.F.W., T.K. and E.F. interpreted the results and wrote the manuscript.

Additional information

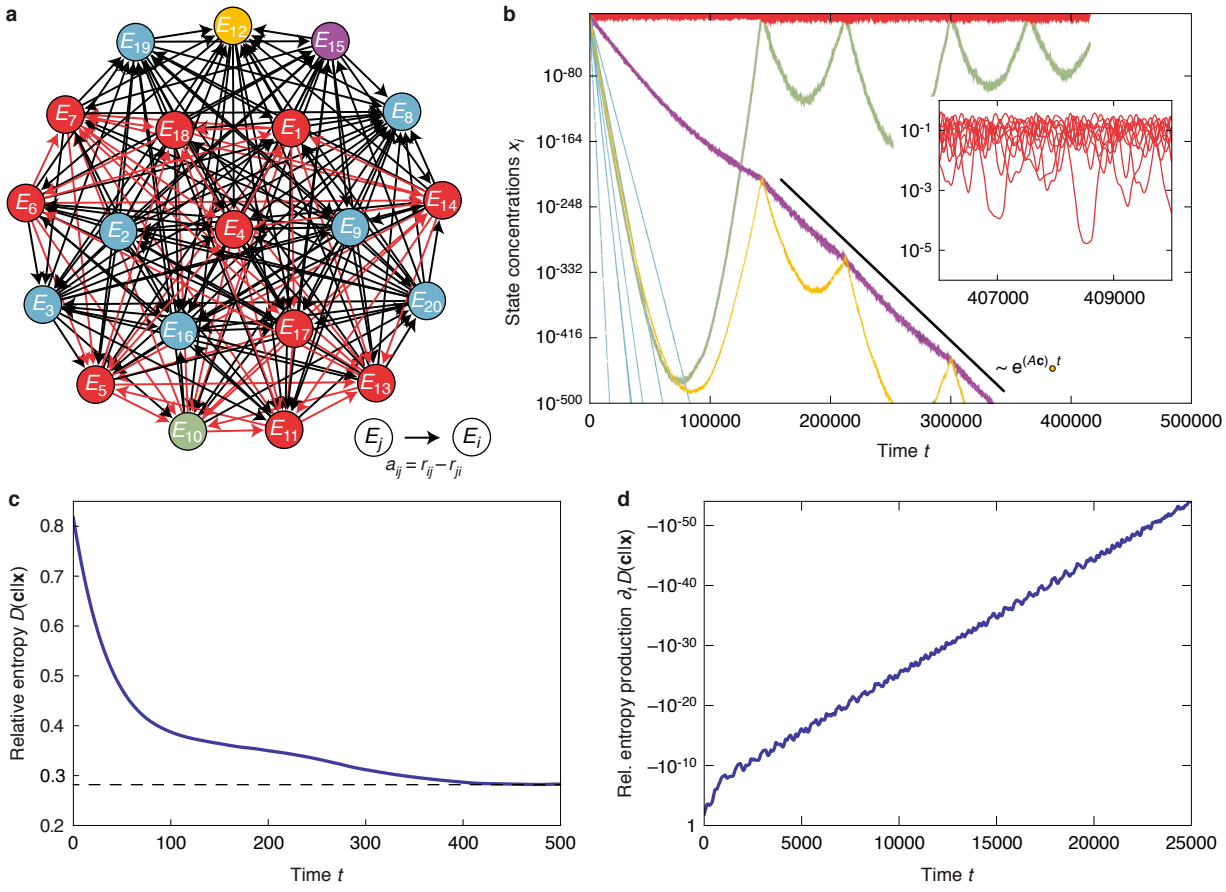
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

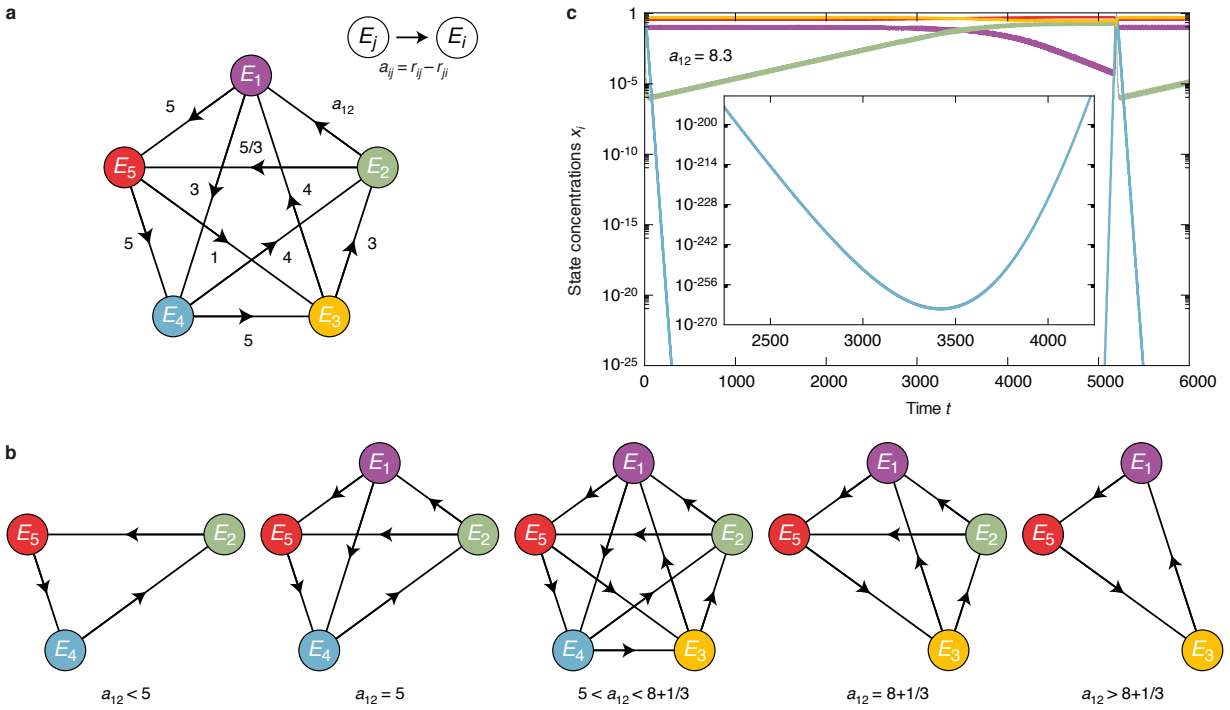
How to cite this article: Knebel, J. *et al.* Evolutionary games of condensates in coupled birth–death processes. *Nat. Commun.* 6:6977 doi: 10.1038/ncomms7977 (2015).

SUPPLEMENTARY FIGURE 1



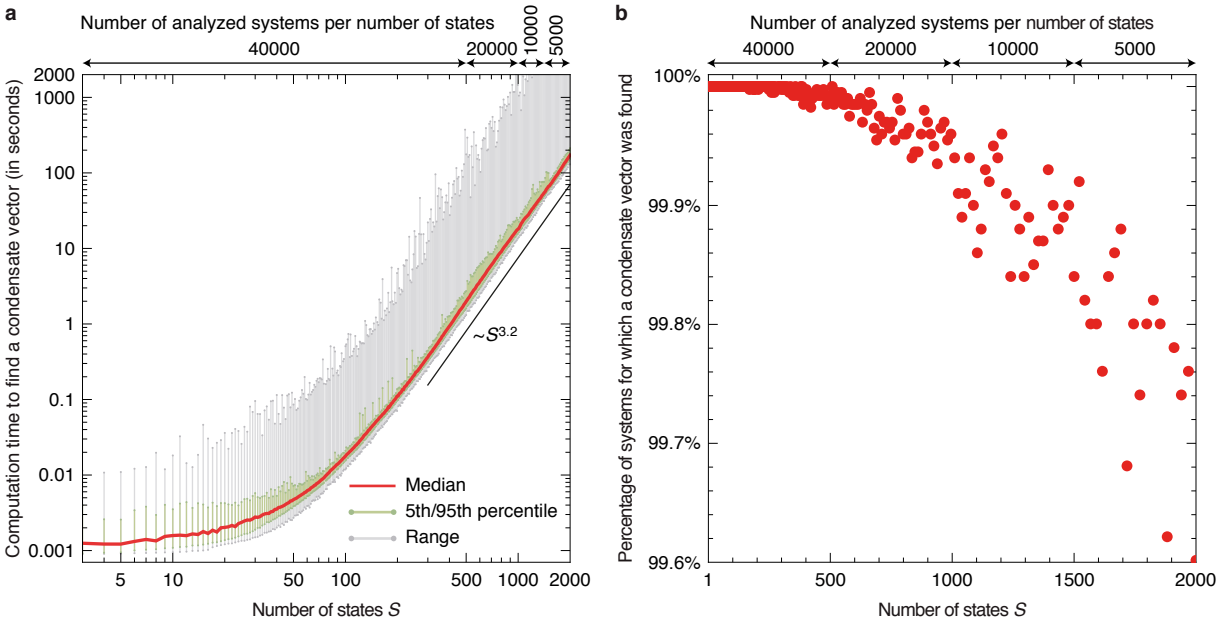
Supplementary Figure 1. Advantage of the algebraic algorithm over numerical integration of the antisymmetric Lotka-Volterra equation (ALVE). (**a**) A fully connected network with 20 states (colored disks). Effective transition rates a_{ij} (arrows) are listed in Supplementary Note 3 and were sampled from a Gaussian distribution (zero mean, unit variance). Computation of the unique normalized condensate vector $\mathbf{c}$ identifies 11 states as condensates (red and green disks) and depletion of 9 states (blue, yellow, and purple disks). (**b**) Trajectories show the temporal evolution of state concentrations x_i (colors in accordance with (**a**)). Numerical integration of the ALVE is highly unstable. Only the routine NDSolve with method "StiffnessSwitching" of Mathematica from Wolfram Research was able to track the concentrations for a sufficiently long time. Routines offered by the GNU Scientific Library failed (Dormand-Prince and Runge-Kutta-Fehlberg). The concentration corresponding to the green state transiently decreases to a value of $4{\cdot}10^{-473}$ before recovering. Its asymptotic temporal average is given by the corresponding entry of the condensate vector $\mathbf{c}$ as $1.6 \cdot 10^{-4}$. This average could, however, not be verified by integration due to numerical failure at $t = 414960$. Negative entries of $A\mathbf{c}$ determine the rates of exponentially fast depletion as illustrated by the black line for the yellow state (note the logarithmic scaling). (**c**) Temporal evolution of the relative entropy $D(\mathbf{c}||\mathbf{x})(t)$ (blue line). The relative entropy decreases towards a non-zero asymptotic value. (**d**) The production of relative entropy $\partial_t D(\mathbf{c}||\mathbf{x})(t)$ (blue line) is, therefore, negative and vanishes for large times.

SUPPLEMENTARY FIGURE 2



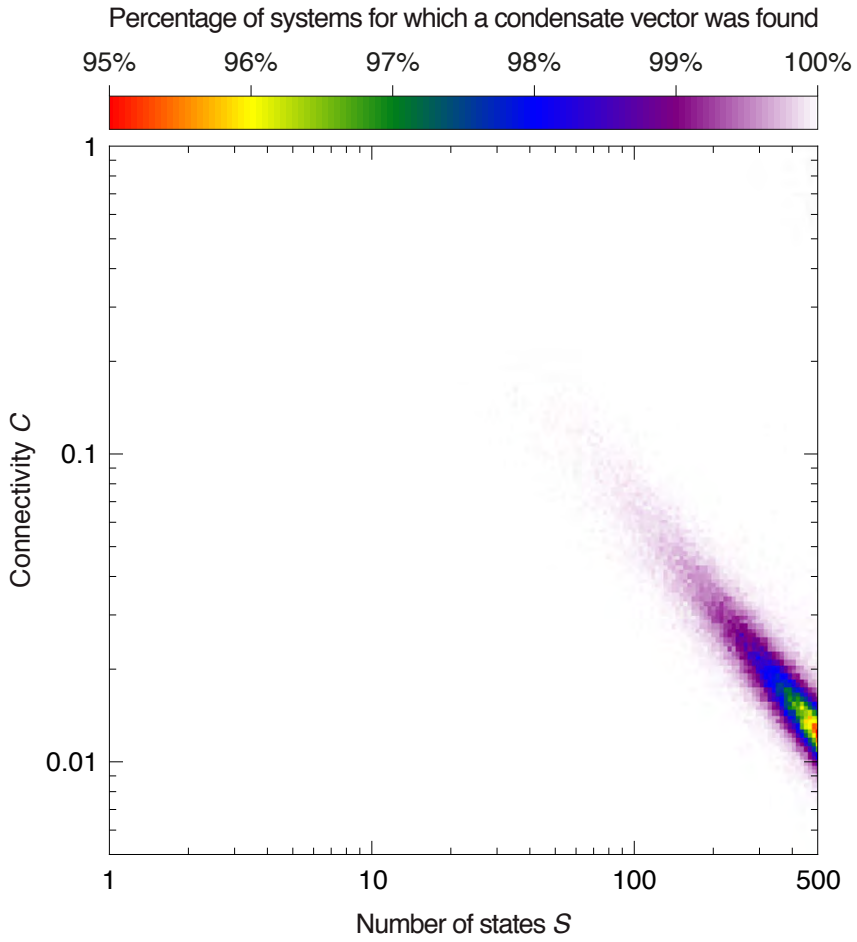
Supplementary Figure 2. Identification of condensates for a system with five states. (**a**) Colored disks represent states E_i. An arrow from E_j to E_i denotes an effective rate constant $a_{ij} = r_{ij} - r_{ji}$. (**b**) Computation of condensate vectors for different values of a_{12} yields depletion of states E_1 and E_3 for $a_{12} < 5$, depletion of state E_3 for $a_{12} = 5$, condensation of all states for $5 < a_{12} < 8 + 1/3$, depletion of state E_4 for $a_{12} = 8 + 1/3$, and depletion of states E_2 and E_4 for $a_{12} > 8 + 1/3$. These results can be verified by using the Mathematica code supplied in Supplementary Note 3. (**c**) Numerical integration of the ALVE confirms the selection of condensates, but becomes error-prone in the vicinity of values of a_{12} at which the set of condensates changes. The trajectories were obtained for $a_{12} = 8.3$, which is slightly smaller than the value at which E_4 becomes depleted. Identification of condensates from trajectories requires the introduction of a threshold for concentrations below which states are considered as depleted. If such a threshold is set to values larger than $\sim 2.5 \cdot 10^{-265}$, state E_4 is erroneously considered to be depleted, despite its periodic recovery.

SUPPLEMENTARY FIGURE 3



Supplementary Figure 3. Run-time analysis of the linear programming algorithm for fully connected random networks of states. The computation time to find a condensate vector was measured for systems with up to $S = 2000$ states. The systems were constructed by sampling effective rate constants $a_{ij} = r_{ij} - r_{ji}$ from a Gaussian distribution (zero mean, unit variance). The numbers of systems (ensemble size) that were analyzed for different numbers of states S are indicated at the top. (**a**) The median of computation times is shown in red, the regime between the 5th and 95th percentiles in green, and the range of computation times in grey (log-log graph). Computation time increases at most polynomially as indicated by the black line (linear fit of the median with exponent 3.221 ± 0.008 (s.e.m.)). (**b**) Percentage of networks for which a condensate vector was found (red dots). The lowest percentage was 99.6%. All computations were performed on machines with 10 Intel Xeon E5-2670v2 cores (2.50 GHz) and 128 GB RAM. Parallelization was not used in CPLEX.

SUPPLEMENTARY FIGURE 4



Supplementary Figure 4. Reliability of the linear programming algorithm for large random networks. Color-coded representation of the percentage of random networks with S states and connectivity C for which the linear programming algorithm found a suitable condensate vector ($\geq 10^4$ analyzed systems per data point; the figure accompanies Fig. 3b). The percentage was close to 100% for most parameters but decreased in the vicinity of the power law shown in Fig. 3b (lowest percentage: 94.9% for networks with $S = 500$ states and connectivity $C = 0.013$).

SUPPLEMENTARY NOTE 1: APPROXIMATION OF THE CLASSICAL MASTER EQUATION BY STOCHASTIC DIFFERENTIAL EQUATIONS AND OCCURRENCE OF THE ANTISYMMETRIC LOTKA-VOLTERRA EQUATION

In the following, we supplement the discussion of the classical master equation that governs condensation of non-interacting bosons in an incoherently driven-dissipative system. When the total number of particles in the system is large ($N \gg 1$), the classical master equation can be approximated by a Fokker-Planck equation. This Fokker-Planck equation is rewritten as an Itō stochastic differential equation, which is equivalent to a Langevin equation. From the analysis of the stochastic differential equation, we find that the leading order dynamics of the condensation process is governed by the antisymmetric Lotka-Volterra equation (equation (3) in the main text).

Classical master equation for the coupled birth-death processes with conservation of total particle number

We consider a system of S non-degenerate states $E_i, i = 1, \ldots, S$, each of which is occupied by N_i indistinguishable particles (Fig. 1a). The configuration of the system at time t is fully characterized by the vector of occupation numbers $\mathbf{N} = (N_1, N_2, \ldots, N_S) \in \mathbb{Z}^S$ with $N_i \geq 0$. It changes only due to coupled creation (birth) and annihilation (death) processes between two connected states E_i and E_j:

$$(N_1, \ldots, N_i - 1, \ldots, N_j + 1, \ldots, N_S) \leftrightarrows (N_1, \ldots, N_i, \ldots, N_j, \ldots, N_S) \leftrightarrows (N_1, \ldots, N_i + 1, \ldots, N_j - 1, \ldots, N_S)\,. \tag{1}$$

We introduce the following short-hand notation for these processes:

$$\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j \quad \leftrightarrows \quad \mathbf{N} \quad \leftrightarrows \quad \mathbf{N} + \mathbf{e}_i - \mathbf{e}_j\,. \tag{2}$$

Here, the vector $\mathbf{e}_i \in \mathbb{Z}^S$ denotes the unit vector in direction i (equal to one at index i, otherwise zero). The above creation and annihilation processes conserve the total number of particles $N = \sum_i N_i$. We are interested in the probability $P(\mathbf{N}, t)$ of finding the system in configuration $\mathbf{N}$ at time t, given that it was initially in configuration $\mathbf{N}_0$ at time t_0. The temporal evolution of the probability distribution $P(\mathbf{N}, t)$ is governed by the classical master equation (equation (1) in the main text):

$$\partial_t P(\mathbf{N}, t) = \sum_{\substack{i,j=1 \\ j \neq i}}^{S} \Big(\Gamma_{i \leftarrow j}(N_i - 1, N_j + 1) P(\mathbf{N} - \mathbf{e}_i + \mathbf{e}_j, t) - \Gamma_{i \leftarrow j}(N_i, N_j) P(\mathbf{N}, t) \Big)\,. \tag{3}$$

In our work, we consider the following transition rate from configuration $\mathbf{N}$ to configuration $\mathbf{N} + \mathbf{e}_i - \mathbf{e}_j$:

$$\Gamma_{i \leftarrow j}(N_i, N_j) = r_{ij}(N_i + s_{ij}) N_j, \text{ with } s_{ij} \geq 0 \text{ and } r_{ij} \geq 0\,. \tag{4}$$

This transition rate encompasses the two model classes of condensate selection in bosonic systems and in evolutionary game theory (EGT) as described in the main text. In the context of bosonic condensation, the parameters s_{ij} are equal to 1 for all i and j, whereas it may assume any non-negative value in the context of EGT. In EGT, s_{ij} contributes to the mutation or switching rate [1].

For the study of incoherently driven-dissipative systems of non-interacting bosons, the above description in terms of a classical master equation requires that the system under consideration is weakly coupled to a reservoir [2–5]. The reservoir has to be such that correlations in it decay rapidly. In particular, the Born-Markov and the rotating wave approximation are assumed in the derivation of the classical master equation. These assumptions imply that off-diagonal entries in the reduced density operator of the system decay fast enough such that coherence in the quantum system is negligible. In addition, the initial states of the system and the reservoir should not be correlated after their preparation.

These conditions are, for example, fulfilled in systems of non-interacting bosons that are both periodically driven in time (Floquet systems) and weakly coupled to a thermal bath (see Vorberg et al. [2] for a detailed discussion).

Derivation of the Fokker-Planck equation

In the following, we approximate the classical master equation (3) in the limit of a large number of particles ($N \gg 1$) [6]. For that purpose, we introduce state concentrations $\mathbf{x} = (x_1, \ldots, x_S)$ with $x_i = N_i/N$. The concentrations are intensive variables and elements of the $(S-1)$-dimensional open simplex $\Delta_{S-1} = \{\mathbf{x} \in \mathbb{R}^S \mid x_i > 0 \text{ for all } i,\ \sum_{i=1}^S x_i = 1\}$. In the limit $N \to \infty$, they become continuous variables. We denote their corresponding probability distribution by $p(\mathbf{x}, t)$. Furthermore, we rescale time by $t \to t/N$. It can be straightforwardly seen that the following formulation of the classical master equation for $p(\mathbf{x}, t)$ is equivalent to the form given in equation (3):

$$\partial_t p(\mathbf{x},t) = \sum_{\substack{i,j=1\\ j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} N^{1-k} \Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i - \Delta x_i, x_j - \Delta x_j)\, p(\mathbf{x} - \Delta x_i \mathbf{e}_i - \Delta x_j \mathbf{e}_j, t) - \gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j)\, p(\mathbf{x}, t) \Big) , \tag{5}$$

$$\text{with } \gamma^{(0)}(\Delta x_i, \Delta x_j; x_i, x_j) = r_{ij} x_i x_j\, \delta(\Delta x_i - 1/N)\delta(\Delta x_j + 1/N) , \tag{6}$$

$$\text{and } \gamma^{(1)}(\Delta x_i, \Delta x_j; x_i, x_j) = r_{ij}\, s_{ij}\, x_j\, \delta(\Delta x_i - 1/N)\delta(\Delta x_j + 1/N) . \tag{7}$$

This classical master equation can be approximated by performing a Kramers-Moyal expansion. Truncation of the expansion at second order, leads to the Fokker-Planck equation:

$$\partial_t p(\mathbf{x},t) = -\sum_{i=1}^{S} \partial_i \big(\alpha_i(\mathbf{x}) p(\mathbf{x},t)\big) + \frac{1}{2N} \sum_{i,j=1}^{S} \partial_i \partial_j \big(\beta_{ij}(\mathbf{x}) p(\mathbf{x},t)\big) . \tag{8}$$

Here, $\alpha(\mathbf{x})$ is the drift vector and $\beta(\mathbf{x})$ the diffusion matrix.

The drift vector is given by:

$$\alpha_i(\mathbf{x}) = \sum_{\substack{j=1\\ j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k} \Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i) \Big) N \Delta x_i , \tag{9}$$

$$= x_i \sum_{j=1}^{S} (r_{ij} - r_{ji}) x_j + \frac{1}{N} \sum_{j=1}^{S} (r_{ij} s_{ij} x_j - r_{ji} s_{ji} x_i) =: \alpha_{i,0}(\mathbf{x}) + \frac{1}{N} \alpha_{i,1}(\mathbf{x}) . \tag{10}$$

At leading order in $1/N$ (that is, at order $\mathcal{O}(1)$), the drift vector is determined by the antisymmetric part $A = R - R^T$ of the rate constant matrix $R = \{r_{ij}\}_{i,j}$. In other words, the matrix A is antisymmetric with entries $a_{ij} = r_{ij} - r_{ji}$.

For the diffusion matrix, we have to distinguish between its diagonal entries,

$$\beta_{ii}(\mathbf{x}) = \sum_{\substack{j=1\\j\neq i}}^{S} \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k}\Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i)\Big) N^2 \Delta x_i^2 \;, \tag{11}$$

$$= \sum_{\substack{j=1\\j\neq i}}^{S} (r_{ij} + r_{ji}) x_i x_j + \frac{1}{N} \sum_{\substack{j=1\\j\neq i}}^{S} (r_{ij} s_{ij} x_j + r_{ji} s_{ji} x_i) =: \beta_{ii,0}(\mathbf{x}) + \frac{1}{N}\beta_{ii,1}(\mathbf{x}) \;, \tag{12}$$

and its off-diagonal entries ($i \neq j$):

$$\beta_{ij}(\mathbf{x}) = \int_{\mathbb{R}} d\Delta x_i \int_{\mathbb{R}} d\Delta x_j \sum_{k\in\{0,1\}} \frac{1}{N^k}\Big(\gamma^{(k)}(\Delta x_i, \Delta x_j; x_i, x_j) + \gamma^{(k)}(\Delta x_j, \Delta x_i; x_j, x_i)\Big) N^2 \Delta x_i \Delta x_j \;, \tag{13}$$

$$= -(r_{ij} + r_{ji}) x_i x_j - \frac{1}{N}\big(r_{ij} s_{ij} x_j + r_{ji} s_{ji} x_i\big) =: \beta_{ij,0}(\mathbf{x}) + \frac{1}{N}\beta_{ij,1}(\mathbf{x}) \;. \tag{14}$$

Both diagonal and off-diagonal entries of the diffusion matrix are determined by the symmetric part of the rate constant matrix R at leading order in $1/N$.

Derivation of the stochastic differential equations

The Fokker-Planck equation (8) can be transformed into a system of Itō stochastic differential equations (SDEs) [6]:

$$\mathrm{d}x_i = \alpha_i(\mathbf{x})\,\mathrm{d}t + \frac{1}{\sqrt{N}} \sum_{j=1}^{S} \zeta_{ij}(\mathbf{x})\,\mathrm{d}W_j \;. \tag{15}$$

Here, $\mathrm{d}W_j$ represents a Wiener increment of zero mean and unit variance. The matrix $\zeta(\mathbf{x})$ is a square root of the diffusion matrix $\beta(\mathbf{x})$ in the sense that $\zeta\zeta^T = \beta$ (the diffusion matrix β is positive semi-definite). Although ζ is not unique, its choice does not change the stochastic nature of the process (an orthogonal transformation $\zeta \to \zeta\mathcal{T}$ with $\mathcal{T}\mathcal{T}^T = \mathbb{I}_S$ does not change the corresponding Fokker-Planck equation). The decomposition $\beta_{ij} = \beta_{ij,0} + \frac{1}{N}\beta_{ij,1}$ in equations (11) and (13) implies that ζ can be written as $\zeta_{ij} = \zeta_{ij,0} + \mathcal{O}\big(\frac{1}{N}\big)$.

The Itō SDEs (15) can also be written in Langevin form, which are often used in the physics literature [6, 7]:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = \alpha_i(\mathbf{x}) \; + \frac{1}{\sqrt{N}} \sum_{j=1}^{S} \zeta_{ij}(\mathbf{x})\,\eta_j \;. \tag{16}$$

Here, η_j represents uncorrelated Gaussian white noise.

The antisymmetric Lotka-Volterra equation

We identify a leading (fast) and a subleading (slow) timescale of the SDE (15). On the leading timescale ($t \sim \mathcal{O}(1)$), only the drift term $\alpha_{i,0}$ is relevant, whereas on the subleading timescale ($t \sim \mathcal{O}(N)$), the terms $\alpha_{i,1}$ and $\zeta_{ij,0}$ compete. The latter terms cause only slow changes on the leading $\mathcal{O}(1)$-timescale. In other words, the dynamics on the subleading timescale cause only slow changes of the $\mathcal{O}(1)$-trajectory.

More specifically, we find that at order $t \sim \mathcal{O}(1)$, only $\alpha_{i,0}$ determines the change in concentrations such that:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = x_i \sum_j (r_{ij} - r_{ji})x_j = x_i(A\mathbf{x})_i \, . \tag{17}$$

As stated in the main text, we refer to this equation as the antisymmetric Lotka-Volterra equation (ALVE). The initial concentrations are assumed to lie in the open simplex Δ_{S-1}, that is $\mathbf{x}(t=0) =: \mathbf{x}_0 \in \Delta_{S-1}$. We note that the dynamics defined by equation (17) cannot leave the simplex, that is $\mathbf{x}(t) \in \Delta_{S-1}$ for all times [8].

We note that the van Kampen system size expansion [9] of the master equation (3) yields the same deterministic equation (17) as our derivation via Fokker-Planck and Langevin equation at the leading order timescale.

Stability in the linear approximation around fixed points of the ALVE

We discuss the fixed points $\mathbf{x}^* \in \overline{\Delta}_{S-1}$ ($x_i^* \geq 0$ and $\sum_i x_1^* = 1$) of the ALVE (17), that is the points for which the dynamics is stationary ($\partial_t x_i^* = x_i^*(A\mathbf{x}^*)_i = 0$). In the following, we show that a linear stability analysis of these fixed points does not yield insight into the global dynamics of the ALVE.

First, every condensate vector $\mathbf{c}$ (normalized such that $\sum_i c_i = 1$) of the antisymmetric matrix A yields a fixed point of the ALVE. This can be seen from the properties of a condensate vector $\mathbf{c}$ (see Methods section of the main text) [10]:

$$c_i > 0 \;\text{ and } (A\mathbf{c})_i = 0 \text{ for } i \in I, \text{ and} \tag{18}$$

$$c_i = 0 \;\text{ and } (A\mathbf{c})_i < 0 \text{ for } i \in \bar{I} \, . \tag{19}$$

Notably, the index set I is unique although more than one condensate vector may exist. Furthermore, there exist fixed points $\mathbf{x}^* \in \overline{\Delta}_{S-1}$ and a different index set $J \neq I$ for which $x_j^* > 0$ and $(A\mathbf{x}^*)_j = 0$ for $j \in J$, and $x_j^* = 0$ for $j \in \bar{J}$ but $(A\mathbf{x}^*)_j < 0$ does not hold for all $j \in \bar{J}$ (in other words, condition (18) is fulfilled, but condition (19) is not).

We first study the stability in the linear approximation around the fixed points that are given by condensate vectors. Upon introducing the distance $\Delta\mathbf{x}$ of a normalized condensate vector $\mathbf{x}^* = \mathbf{c}$ from the concentrations $\mathbf{x}$ as a new variable, $\Delta\mathbf{x} \coloneqq \mathbf{x} - \mathbf{c}$, one obtains from the ALVE (17) the temporal behavior of that translated variable as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = \Delta x_i (A\mathbf{c})_i + c_i (A\Delta\mathbf{x})_i + \Delta x_i (A\Delta\mathbf{x})_i \;\text{ , that is,} \tag{20}$$

$$\text{for } i \in I: \;\; \frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = \sum_{j=1}^{S} c_i a_{ij} \Delta x_j + R(\Delta\mathbf{x}) \, , \tag{21}$$

$$\text{for } i \in \bar{I}: \;\; \frac{\mathrm{d}}{\mathrm{d}t}\Delta x_i = (A\mathbf{c})_i \Delta x_i + R(\Delta\mathbf{x}) \, , \tag{22}$$

with $R(\Delta\mathbf{x}) = \Delta x_i (A\Delta\mathbf{x})_i = \mathcal{O}(\|\Delta\mathbf{x}\|^2)$.

Next, we discuss the stability of the condensate vectors in the linear approximation (linear stability analysis of fixed points). The cardinality of the set I is referred to as $|I|$. After relabeling of the indices, one obtains up to linear order in $\|\Delta\mathbf{x}\|$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\Delta\mathbf{x}_I\\ \Delta\mathbf{x}_{\bar{I}}\end{pmatrix} = \begin{pmatrix}\tilde{A}_c & B\\ 0 & \tilde{A}_s\end{pmatrix}\begin{pmatrix}\Delta\mathbf{x}_I\\ \Delta\mathbf{x}_{\bar{I}}\end{pmatrix} =: \tilde{A}\Delta\mathbf{x}\,, \tag{23}$$

with $\tilde{A}_c$ denoting the $(|I| \times |I|)$-dimensional matrix with elements $(\tilde{A}_c)_{i,j} = c_i a_{ij}$ for $i, j \in I$. $\tilde{A}_s$ denotes the diagonal, $((S-|I|)\times(S-|I|))$-dimensional matrix with entries $(\tilde{A}_s)_{i,j} = (A\mathbf{c})_i\delta_{ij}$ for $i, j \in \bar{I}$. The matrix B is of dimension $|I| \times (S-|I|)$ with elements $B_{ij} = c_i a_{ij}$ for $i \in I$ and $j \in \bar{I}$.

The eigenvalues of the matrix $\tilde{A}$ determine the linear stability of the fixed points. Because of the block upper triangular structure of the matrix $\tilde{A}$, its eigenvalues are given by the eigenvalues of the matrices $\tilde{A}_s$ and $\tilde{A}_c$. All eigenvalues of the diagonal matrix $\tilde{A}_s$ are negative because $(A\mathbf{c})_i < 0$ for $i \in \bar{I}$. All eigenvalues of the matrix $\tilde{A}_c$ have vanishing real part. The latter can be seen from defining the nonsingular, $(|I| \times |I|)$-dimensional matrix V with elements $(V)_{i,j} = c_i\delta_{ij}$ for $i, j \in I$. Since the matrix A^I is an antisymmetric matrix, all eigenvalues of $\tilde{A}_c = VA^I$ are purely imaginary as well (the $(|I| \times |I|)$-dimensional submatrix A^I corresponds to the system of condensates with indices in I; see Methods section of the main text). However, the matrix $\tilde{A}_c$ is not antisymmetric in general. This argument can be seen as follows (see for example the Appendix in [11]): Consider the diagonal, nonsingular matrix $V^{1/2}$, whose square is the matrix V, and whose inverse is the matrix $V^{-1/2}$. The matrix $V^{-1/2}(VA^I)V^{1/2}$ has the same eigenvalues as the matrix VA^I. Since the matrix $V^{-1/2}(VA^I)V^{1/2} = V^{1/2}A^IV^{1/2}$ is antisymmetric and, thus, has purely imaginary eigenvalues, also the matrix VA^I has purely imaginary eigenvalues.

Consequently, the fixed points of the ALVE (17) that are given by the condensate vectors possess a $(S-|I|)$-dimensional local, invariant stable manifold $\mathcal{M}_s$ and a $|I|$-dimensional local, invariant center manifold $\mathcal{M}_c$ (see for example Theorem 3.2.1 in [12]). Concentrations with initial conditions chosen in $\mathcal{M}_s$ decay to zero exponentially fast. However, for initial concentrations that do not lie in $\mathcal{M}_s$, the temporal behavior cannot be inferred ("linearly stable solutions may be nonlinearly unstable" [12]).

We note that the above linear stability analysis applies to any fixed point of the ALVE. Any fixed point of the ALVE possesses an at least $|J|$-dimensional local, invariant center manifold $\mathcal{M}_c$ (see definition of J above). Therefore, a linear stability analysis of the fixed points of the ALVE (17) does not yield insight into the global dynamics of the ALVE, at least not in the straightforward fashion.

SUPPLEMENTARY NOTE 2: DYNAMICS OF CASES FOR NON-GENERIC ANTISYMMETRIC MATRICES

Some of the results presented in the main text for generic antisymmetric matrices ($A \in \Omega$) can be extended to matrices with higher dimensional kernels. Here, Ω is defined as in equation (18) of the main text as the set of antisymmetric matrices whose submatrices have a kernel with dimension less than one or equal to one:

$$\Omega = \left\{ A \in \mathbb{R}^{S\times S} \mid A \text{ is antisymmetric and } \dim\ker A^J \leq 1 \text{ for all } J \subseteq \{1,\ldots,S\} \right\} . \tag{24}$$

When submatrices of A have a kernel of dimension greater than or equal to two ($A \notin \Omega$) the statements generalize as follows: The temporal average of the projection to the surviving concentrations converges to the positive kernel of the attractive subsystem. This convergence takes place on a time scale which is not slower than $1/t$. More precisely, the distance between the time average of the concentrations of the selected states and the kernel of the surviving subsystem tends to zero for large times:

$$\text{dist}\left(\langle \mathbf{x}_I \rangle_t , \ker A^I\right) \leq \frac{Const(A, \mathbf{x}_0)}{t} . \tag{25}$$

Typically, one still finds exponentially fast depletion of states. The effective bounds on the depletion rates may depend on initial conditions.

SUPPLEMENTARY NOTE 3: LINEAR PROGRAMMING ALGORITHMS

In the following, we supplement the description and discussion of the linear programming algorithm from the Methods section in the main text. We detail on the CPEX algorithm and discuss its calibration, and provide a simplified Mathematica code.

CPLEX algorithm

The IBM ILOG CPLEX Optimization Studio 12.5 was used to numerically search for a condensate vector $\mathbf{c}$ for a given antisymmetric matrix A. Direct solution of the inequalities $A\mathbf{c} \leq 0$ and $\mathbf{c} - A\mathbf{c} \geq 1$ turned out to be numerically infeasible for systems with a large number of states S. Therefore, condensate vectors were primarily determined by solving the following linear programming problem: minimize $\epsilon_1 + \ldots + \epsilon_S$, subject to $-A\mathbf{c} + \epsilon \geq 0$ and $\mathbf{c} - A\mathbf{c} \geq 1$, with non-negative auxiliary variables $\epsilon \coloneqq (\epsilon_1, \ldots, \epsilon_S)$. The resulting vector $\mathbf{c}$ was used to define the set $I \coloneqq \{i \,|\, c_i > 5 \cdot 10^{-8}\}$ and its complement $\bar{I} = \{1, \ldots, S\} - I = \{i \,|\, c_i \leq 5 \cdot 10^{-8}\}$. The vector $\mathbf{c}$ was accepted as condensate vector, the set I as set of condensates, and the set $\bar{I}$ as set of depleted states if: $|(A\mathbf{c})_i| < 10^{-6}$ for all $i \in I$, and $-(A\mathbf{c})_i > 10^{-6}$ for all $i \in \bar{I}$. Minor extensions were added to the CPLEX algorithm to handle matrices for which an appropriate condensate vector could not be found. The Mathematica code in the following section exemplifies one of these extensions.

The above numerical thresholds were optimized by comparing inferred sets of condensates and of depleted states to sets that were derived using an alternative method. This alternative method is based on an analytical expression for kernel vectors (see equation (19) in the Methods section of the main text and [13, 14]). It is reliable but restricted to systems in which the network of states has connectivity $C = 1$ and in which the number of states is small (computational complexity grows exponentially with S). The correct identification of condensates and depleted states by the CPLEX algorithm was validated for 10^6 randomly sampled networks of states with connectivity $C = 1$ and $S = 18$ states. A detailed evaluation of the reliability of the CPLEX linear programming algorithm is provided in Supplementary Figs. 3 and 4.

Mathematica code

The following code for Mathematica 9.0 from Wolfram Research determines a condensate vector $\mathbf{c}$ by minimizing $\epsilon_1 + \ldots + \epsilon_S$, subject to $-A\mathbf{c} \geq 0$ and $\mathbf{c} - A\mathbf{c} + \epsilon \geq 1$. The resulting vector is used to infer the set of condensates I and the set of depleted states $\bar{I}$. The code can be used to verify the selection of states for the systems shown in Supplementary Figs. 1 and 2.

```
lpAlgorithm[noOfStates_, matrix_] :=
  Block[{condensateVector, condensates, depletedStates},
    Block[
    {
    solution
    (* solution: first half: condensate vector b, second half: aux. vector ε *),
    vector = Join[ConstantArray[0., noOfStates], ConstantArray[1., noOfStates]]
    (* vector: first half: w.r.t condensate vector b, second half: w.r.t. aux. vector ε *),
    lhs = ConstantArray[0., {2*noOfStates, 2*noOfStates}]
    (* lhs: upper left: -Ab, lower half: b-Ab+ε *),
    rhs = ConstantArray[0., 2*noOfStates]
    (* rhs: upper left: 0, lower half: 1 *)
    },

    Do[
      (* -Ab >= 0 *)
      Do[lhs[[i, j]] = -matrix[[i, j]], {j, 1, noOfStates}];
      rhs[[i]] = 0.;

      (* b-Ab+ε >= 1 *)
      lhs[[noOfStates + i, i]] = 1.;
      Do[lhs[[noOfStates + i, j]] -= matrix[[i, j]], {j, 1, noOfStates}];
      lhs[[noOfStates + i, noOfStates + i]] = 1.;
      rhs[[noOfStates + i]] = 1.;
    , {i, 1, noOfStates}];

    solution = Check[LinearProgramming[vector, lhs, rhs], {}];
    condensateVector = solution[[1;;noOfStates]];
    ];

  condensates = Flatten[Position[condensateVector, _?(# > 0.1&)]];
  depletedStates = Complement[Range[noOfStates], condensates];
  Return[{condensates, depletedStates, condensateVector/Total[condensateVector]}];
  ];
```

The following code can be used to verify the selection of states for the system with five states shown in Supplementary Fig. 2:

```
noOfStates = 5;
a12 = 8.3;
matrix = {{0,a12,4,-3,-5},{-a12,0,3,4,-5/3},{-4,-3,0,5,1},{3,-4,-5,0,5},{5,5/3,-1,-5,0}};

{condensates, depletedStates, condensateVector} = lpAlgorithm[noOfStates, matrix];

Print["Condensates: ", Length[condensates]];
Print["   Indices: ", condensates];
Print["   b: ", condensateVector[[condensates]]];
Print["   A.b: ", (matrix.condensateVector)[[condensates]]];
Print[];

Print["Depleted states: ", Length[depletedStates]];
Print["   Indices: ", depletedStates];
Print["   b: ", condensateVector[[depletedStates]]];
Print["   A.b: ", (matrix.condensateVector)[[depletedStates]]];
Print[];
```

To verify the selection of states for the system with 20 states that is shown in Supplementary Fig. 1, the first two lines of the above code have to be changed to:

```
noOfStates = 20;
matrix=
{{0.0000000000,0.1012965582,0.0960864501,0.1257833702,-0.0595764929,0.0924501179,-0.1016301739,
-0.0795618281,0.0487881199,-0.1122071087,-0.0286971727,-0.1000496191,-0.0103185636,0.0130714180,
-0.0572150115,-0.0129797149,-0.0706576396,-0.0389344063,0.0663966936,0.1467041892},{-0.1012965582,
0.0000000000,-0.0712011514,0.0415422138,-0.0019111200,-0.0816168047,-0.0477211059,-0.1006350936,
0.1522277516,0.0720296603,-0.2058646002,-0.1253355555,0.0717377587,-0.1063033227,-0.0942343547,
-0.0391532296,0.0049849208,-0.0929232005,-0.0840063075,0.0559612167},{-0.0960864501,0.0712011514,
0.0000000000,-0.0304803123,-0.1741821153,0.0076812872,0.0297497936,-0.0075143775,-0.0120359985,
-0.0110515825,-0.0808005342,0.2738320344,0.1170451029,-0.0357777508,-0.0896471841,-0.1577036387,
-0.3757310353,0.0296054422,-0.2259424452,0.0989073494},{-0.1257833702,-0.0415422138,0.0304803123,
0.0000000000,0.0510755453,0.0115416481,0.0985481106,0.0272202852,0.0213164598,-0.0535439544,
-0.0456534129,0.1362532620,0.1074372756,0.0375029473,-0.0533083982,0.1648089686,-0.0526653130,
0.0389051602,0.0438611333,0.0539352012},{0.0595764929,0.0019111200,0.1741821153,-0.0510755453,
0.0000000000,0.0109896214,0.1353304974,0.0912349277,0.0296142827,-0.1538255140,0.0431260101,
-0.0502994390,-0.0163964089,-0.1759377604,0.0481992186,0.0664062093,0.1736770927,-0.0448630119,
0.0535605474,-0.0257754999},{-0.0924501179,0.0816168047,-0.0076812872,-0.0115416481,-0.0109896214,
0.0000000000,-0.1090515383,-0.1188694334,0.0514000882,-0.0331443533,0.1623742129,0.0492738763,
0.0176787814,-0.1341072593,-0.0009543378,-0.0789222831,-0.0579314512,0.0892386351,-0.0686414708,
0.0492364011},{0.1016301739,0.0477211059,-0.0297497936,-0.0985481106,-0.1353304974,0.1090515383,
0.0000000000,-0.0303370160,0.0506580949,0.0225254369,0.1119589132,-0.2732763845,0.0903284019,
0.0780506743,0.1487517615,-0.0050831919,-0.0357202770,-0.1006725919,-0.0014275275,0.0744309213},
{0.0795618281,0.1006350936,0.0075143775,-0.0272202852,-0.0912349277,0.1188694334,0.0303370160,
0.0000000000,0.0881041025,0.0210453129,-0.0131581374,-0.0515644614,0.0300418276,0.0765770257,
0.1482013668,0.0876706565,-0.1600022303,-0.1501954314,0.0381309952,0.1069018065},{-0.0487881199,
-0.1522277516,0.0120359985,-0.0213164598,-0.0296142827,-0.0514000882,-0.0506580949,-0.0881041025,
0.0000000000,-0.0937867616,-0.1253362044,0.1051027292,-0.0160557361,0.0120747605,0.0410327424,
-0.1178120937,-0.0104974723,0.1001865178,0.0915443356,-0.0590317396},{0.1122071087,-0.0720296603,
0.0110515825,0.0535439544,0.1538255140,0.0331443533,-0.0225254369,-0.0210453129,0.0937867616,
0.0000000000,-0.1718419085,0.0842948432,0.1084407671,-0.1297238294,0.0768833880,-0.0866723482,
0.0062786219,-0.0986826408,-0.1352434973,-0.1425316892},{0.0286971727,0.2058646002,0.0808005342,
0.0456534129,-0.0431260101,-0.1623742129,-0.1119589132,0.0131581374,0.1253362044,0.1718419085,
0.0000000000,0.1936649126,-0.0428450371,0.0782428143,0.0592244942,-0.0554995964,0.0105651402,
-0.0184545141,0.1317087624,-0.1175718037},{0.1000496191,0.1253355555,-0.2738320344,-0.1362532620,
0.0502994390,-0.0492738763,0.2732763845,0.0515644614,-0.1051027292,-0.0842948432,-0.1936649126,
0.0000000000,0.0668839360,0.0174093226,0.0573713882,-0.0610992396,-0.0280347431,0.1058623608,
0.1781275929,0.0152443634},{0.0103185636,-0.0717377587,-0.1170451029,-0.1074372756,0.0163964089,
-0.0176787814,-0.0903284019,-0.0300418276,0.0160557361,-0.1084407671,0.0428450371,-0.0668839360,
0.0000000000,0.0571450290,0.1871354534,0.0147123474,0.1010276308,0.0366350187,0.0630761367,
-0.0728212225},{-0.0130714180,0.1063033227,0.0357777508,-0.0375029473,0.1759377604,0.1341072593,
-0.0780506743,-0.0765770257,-0.0120747605,0.1297238294,-0.0782428143,-0.0174093226,-0.0571450290,
0.0000000000,0.0158510770,-0.0301637492,0.0379895572,0.0353221008,-0.0410300505,0.0399902646},
{0.0572150115,0.0942343547,0.0896471841,0.0533083982,-0.0481992186,0.0009543378,-0.1487517615,
-0.1482013668,-0.0410327424,-0.0768833880,-0.0592244942,-0.0573713882,-0.1871354534,-0.0158510770,
0.0000000000,-0.1410043405,0.0473724443,0.1164556594,0.0120263929,0.0383652365},{0.0129797149,
0.0391532296,0.1577036387,-0.1648089686,-0.0664062093,0.0789222831,0.0050831919,-0.0876706565,
0.1178120937,0.0866723482,0.0554995964,0.0610992396,-0.0147123474,0.0301637492,0.1410043405,
0.0000000000,-0.0051486466,0.0770482371,-0.0619160030,0.1041763643},{0.0706576396,-0.0049849208,
0.3757310353,0.0526653130,-0.1736770927,0.0579314512,0.0357202770,0.1600022303,0.0104974723,
-0.0062786219,-0.0105651402,0.0280347431,-0.1010276308,-0.0379895572,-0.0473724443,0.0051486466,
0.0000000000,0.0771683333,0.0494219197,0.3405228485},{0.0389344063,0.0929232005,-0.0296054422,
-0.0389051602,0.0448630119,-0.0892386351,0.1006725919,0.1501954314,-0.1001865178,0.0986826408,
0.0184545141,-0.1058623608,-0.0366350187,-0.0353221008,-0.1164556594,-0.0770482371,-0.0771683333,
0.0000000000,-0.1856676553,0.0357968997},{-0.0663966936,0.0840063075,0.2259424452,-0.0438611333,
-0.0535605474,0.0686414708,0.0014275275,-0.0381309952,-0.0915443356,0.1352434973,-0.1317087624,
-0.1781275929,-0.0630761367,0.0410300505,-0.0120263929,0.0619160030,-0.0494219197,0.1856676553,
0.0000000000,0.1384008887},{-0.1467041892,-0.0559612167,-0.0989073494,-0.0539352012,0.0257754999,
-0.0492364011,-0.0744309213,-0.1069018065,0.0590317396,0.1425316892,0.1175718037,-0.0152443634,
0.0728212225,-0.0399902646,-0.0383652365,-0.1041763643,-0.3405228485,-0.0357968997,-0.1384008887,
0.0000000000}};
```

SUPPLEMENTARY REFERENCES

[1] M. A. Nowak and K. Sigmund, Science **303**, 793 (2004).

[2] D. Vorberg, W. Wustmann, R. Ketzmerick, and A. Eckardt, Phys. Rev. Lett. **111**, 240405 (2013).

[3] M. Grifoni and P. Hänggi, Phys. Rep. **304**, 229 (1998).

[4] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer, Berlin Heidelberg, 2004).

[5] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[6] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2009).

[7] U. C. Täuber, *Critical Dynamics* (Cambridge University Press, Cambridge, MA, 2014).

[8] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, 1998).

[9] N. G. Van Kampen, *Stochastic Process in Physics and Chemistry* (Elsevier, Amsterdam, 2007).

[10] H. Kuhn and A. Tucker, *Linear Inequalities and Related Systems* (Princeton University Press, Princeton, NJ, 1956).

[11] R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, NJ, 1973).

[12] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Springer-Verlag, New York, 2003).

[13] C. E. Cullis, *Matrices and Determinoids*, Vol. I and II (Cambridge University Press, Cambridge, 1913).

[14] J. Knebel, T. Krüger, M. F. Weber, and E. Frey, Phys. Rev. Lett. **110**, 168106 (2013).

III The stochastic dynamics of bacterial range expansions

1 Introduction

When environmental conditions change or resources are depleted, the survival of a species often depends on its ability to colonize new territory. The understanding of these abilities is of pivotal importance in many fields of biological research. On the macroscopic scale, global changes are currently being observed in the distribution of animals and plants due to anthropogenic climate change [452]. On the microscopic scale, the ability of pathogenic bacteria to colonize human tissue and catheter surfaces poses threats to immunocompromised patients [453, 454].

The colonization of new territory by a species is called a range expansion. In the following, we study such range expansions for a bacterial model system of three *Escherichia coli* strains. Together with the microbiology group of M. Opitz, we explored the determinants of the strains' coexistence during range expansions, both from an experimental and a theoretical perspective. The following text serves as an introduction to our publication "Chemical warfare and survival strategies in bacterial range expansions" [5], which is reprinted in section 2 of this chapter. Author contributions are listed on page 8 of the publication (page 203 of this thesis).

Throughout the recent years, microbial systems have received increasing attention in the study of range expansions [4, 5, 455–465]. Although observations made on the microscopic scale may not translate directly to the macroscopic scale, certain characteristics of range expansions may well be universal. That is, for example, the case for their decrease of genetic diversity [457, 463, 464, 466] and for the "surfing" of genes (or alleles) along expanding population fronts [457, 467–471]. When a population expands spatially, offspring in newly colonized territory are more likely to have descended from an individual along the population's front than from an individual in the population's bulk (provided that mixing within the population is slow). This "founder effect" [472] reduces the genetic diversity in newly colonized territory and also results in the spatial demixing of genetic traits [457]. Both effects have been observed in expanding microbial colonies [457, 463, 464] and in data on the genetic diversity of modern humans [466, 473, 474]. Furthermore, the founder effect acts as a continual population bottleneck and amplifies genetic drift along an expanding front. Selectively neutral and even deleterious mutations thereby gain the ability to surf along the front and can reach high frequencies [457, 467–471].

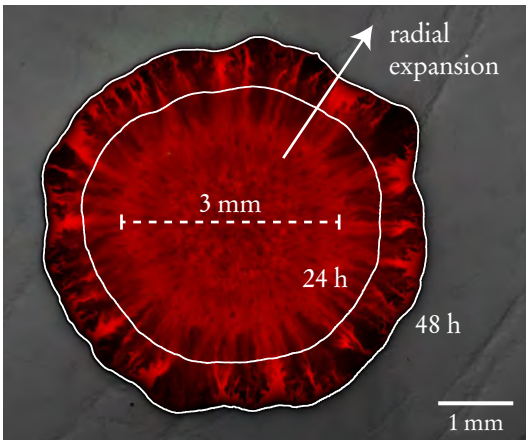
Experiments on expanding microbial populations are also increasingly performed to test principles of theoretical ecology and game theory [4, 5, 461–464]. Thus far, most of these studies have been limited to communities of two bacterial strains [461–464]. It was, for example, shown that the genetic demixing caused by range expansions promotes the evolution of cooperation in a system of two strains of the budding yeast *Saccharomyces cerevisiae*. In this system, range expansions cause a spatial separation of invertase producers (the "cooperators") and non-producers (the "defectors"), protecting the cooperators from exploitation. Furthermore, genetic demixing was found to impede weak or asymmetric mutualism between two cross-feeding strains of *S. cerevisiae* [464]. In our work [5], we explored the role of an ecological pattern that has been linked to the maintenance of biodiversity in communities of at least three strains (or species): intransitive, cyclic dominance [4, 102, 172, 475–488].

1.1 Bacterial model system of three *Escherichia coli* strains

A perfect cyclic dominance holds for the strategies of the rock-paper-scissors game: rock crushes scissors, scissors cut paper, paper covers rock. In biological systems, cyclic dominance is hardly realized in such a perfect way and only holds with respect to appropriately chosen observables. For example, Sinervo and Lively observed a cyclic dominance between the mating strategies of three morphs of male lizards with respect to estimates of their fitnesses [102]. Each morph population was found to be susceptible to invasion by lizards of other morphotypes in a cyclic way.

The possibility of cyclic dominance between three strains of *E. coli* was proposed on theoretical grounds in [478, 480] and was tested experimentally both *in vitro* [4] and *in vivo* [482]. In these experiments, one of the strains was a producer of a toxin (strain C), one strain was resistant to the toxin (strain R), and one strain was sensitive to it (strain S). The toxin was either colicin E2, which cleaves the DNA of sensitive cells [4, 482], or colicin E1, which depolarizes their cytoplasmic membranes [482]. Pairwise competition experiments established that the sensitive strain outgrew the resistant strain and that the resistant strain outgrew the toxin-producing strain — thus completing the cycle. In the *in vitro* experiments, Kerr and co-workers conducted competitions between small single-strain colonies and found that the three strains can coexist in a static spatial environment [4].

In our work [5], we conducted *in vitro* experiments and theoretical modelling to understand the role of cyclic dominance and of two other "ecological scenarios" in maintaining coexistence during range expansions (for the colicin E2 system). A range expansion experiment can be started by inoculating a droplet of bacterial suspension onto an agar plate containing appropriate growth medium [457]. After a lag phase [489, 490], the bacterial cells start to divide and to deplete the



C (no fluorescent protein) | R (mCherry)

Figure 1 Range expansion of a bacterial colony that initially comprised three strains of *Escherichia coli* (equal initial strain ratios; cyclic dominance scenario). The strain C produced the DNase colicin E2, the strain R was resistant to this colicin and was labelled with the red fluorescent protein mCherry, and the strain S was sensitive to the endonuclease activity of colicin E2 and was labelled with the green fluorescent protein (GFP). The colony was inoculated from a 1 µl mixture of overnight liquid cultures as explained in our article [5] (reprinted in section 2). The diameter of the colony's homeland measured approximately 3 mm as indicated by the dashed white line. Monoclonal sectors were founded by toxin-producing and resistant cells. The sensitive strain S was already suppressed in the colony's homeland and failed to establish sectors. The solid white lines denote the colony's boundary after 24 and 48 hours (overlay with bright-field image). The original picture was taken by G. Poxleitner and is included in our article [5].

resources in the colony's "homeland". Nutrients become limited to the colony's boundary and fuel its radial expansion. Figure 1 illustrates the expansion of a colony that initially comprised all of our three *E. coli* strains. Besides radial, effectively two-dimensional expansions, researchers have also studied expansions from linear inoculations [457] and are now turning towards three-dimensional expansions [491–493]. Three-dimensional range expansions are of clinical relevance because they occur, for example, when a tumour invades healthy tissue [494].

By assigning the fluorescent proteins mCherry and GFP to our strains, we were able to design three different ecological scenarios: a cyclic dominance scenario, a hierarchical scenario, and an intermediate scenario (cf. figure 2 and figure 1 of [5]). Besides allowing us to distinguish the strains by their fluorescence, the fluorescent proteins gave us control over the strains' growth rates. These growth rates were identified with the maximal radial expansion velocities of single-strain colonies (see the Electronic Supplementary Material of our publication [5] on pages 204–217 of this thesis). We observed that strains labelled with mCherry suffered a significant

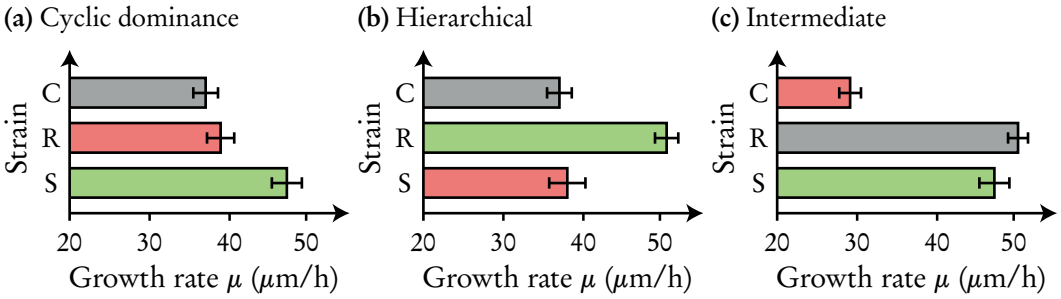


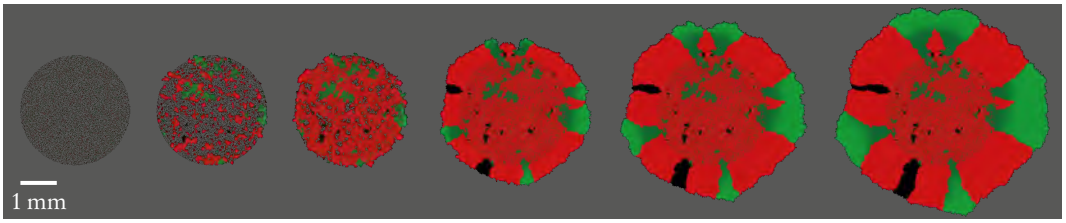
Figure 2 Growth rates of our *E. coli* strains in three ecological scenarios (colicin-producing strain C, resistant strain R, and sensitive strain S; slow growth conditions). The growth rates represent the maximal radial expansion velocities of single-strain colonies (see the Electronic Supplementary Material of [5]). The colour of a bar indicates the fluorescent protein assigned to the strain (grey: no fluorescent protein, red: red fluorescent protein mCherry, green: green fluorescent protein (GFP)). **(a)** Cyclic dominance scenario with growth rate hierarchy $\mu_S > \mu_R > \mu_C$. The cycle is closed by the action of the colicin on the sensitive strain S. **(b)** Hierarchical scenario with growth rate hierarchy $\mu_R > \mu_S \geq \mu_C$. Colonies of the resistant strain R outgrew colonies of both other strains. **(c)** Intermediate scenario with the growth rates μ_R and μ_S being significantly larger than μ_C.

decrease in the speed at which they colonized new territory. Making use of this finding, we implemented a scenario with cyclic dominance between the strains by labelling the resistant strain R with mCherry. Assigning mCherry to the sensitive strain S instead resulted in a hierarchical scenario. In this scenario, single-strain colonies of strain R outgrew colonies of both C and S. An intermediate scenario was obtained by assigning mCherry to strain C. Here, the growth rates of both R and S were significantly larger than the growth rate of strain C. The growth rate hierarchies of these scenarios are shown in figure 2.

1.2 Computational model of bacterial range expansions

Upon inoculating droplets containing equal amounts of the three *E. coli* strains onto agar plates (M63 medium), we found that only the resistant strain survived range expansions in the hierarchical and intermediate scenarios. For the cyclic dominance scenario, the toxin-producing strain survived as well, but three-strain coexistence was not observed (cf. figure 2 of our publication [5]). To understand which strains survive along the expanding front of a colony, we devised a computational model of range expansions.

Whether a bacterial strain manages to establish and to maintain an outward sector depends on chance effects during the reproduction of cells (random genetic drift). To incorporate such effects into a model, we described the proliferation of cells and their long-range interaction via colicin in terms of a continuous-time



C (no fluorescent protein) | S (GFP) | R (mCherry)

Figure 3 Dynamics of a simulated range expansion. The simulation was performed using the stochastic simulation algorithm (SSA) as explained in the Electronic Supplementary Material of [5] (see pages 204–217 of this thesis). The simulation parameters were adjusted to the cyclic dominance scenario whose strain and marker combinations are shown below the time-lapse. The grey background of the figure represents uncolonized agar with M63 growth medium. Black, green, and red regions indicate a colonization by the respective strain (dark green: lysed sensitive strain). At the beginning of the simulation, strains were randomly distributed over a circular region of the simulation lattice whose radius represented 1.5 mm. The initial filling fraction of this homeland was assumed to be 25 % and the initial strain ratios were chosen as S:R:C = 1:1:0.1. These ratios allowed for a coexistence of all three strains until the simulation was stopped at a colony radius of 3 mm.

Markov process with discontinuous sample paths (see chapter I). Thus, our model was based on a master equation. Due to the model's complexity, it is, however, impossible to write this equation down. Individual *in silico* range expansions were generated using the stochastic simulation algorithm (SSA) of Gillespie [80, 81] (cf. section 1.4 of chapter I). Figure 3 demonstrates the expansion of a simulated colony for the cyclic dominance scenario with a low initial amount of the toxin-producing strain. Unlike generation-based stepping-stone models [495, 496], our continuous-time description of the expansion process enabled us to consider growth rate differences between the individual strains. Their growth dynamics was modelled in terms of hopping processes on a discrete spatial lattice. Every combination of strain and fluorescent marker was characterized by a growth rate parameter and a lag time parameter for this purpose (an uncolonized lattice site could only be colonized from a neighbouring site after the lag time of the respective strain had passed). The toxin interaction between cells was modelled in terms of a source and degradation process. Since the diffusion of colicin happens on a much faster time scale than the division of individual *E. coli* cells (the diffusion constant of colicin is on the order of 10^{-7} cm^2/s [497]), we approximated the solution of the source and degradation process in terms of quasi-static, exponentially decreasing colicin profiles around toxin-producing sites. The shape of the profiles depended on a strength and on a range parameter. Further details on the model can be found in the Electronic Supplementary Material of our publication (reprinted on pages 204–217 of this thesis). Let us emphasize that the values of the above parameters were

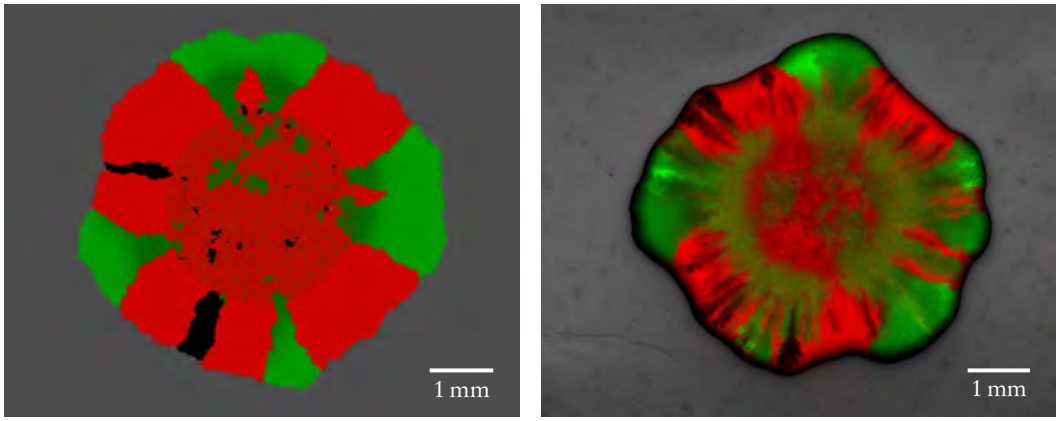


Figure 4 Comparison of the simulated colony from figure 3 on the left to an experimentally observed colony on the right (cyclic dominance scenario; initial strain ratios S:R:C = 1:1:0.1). The experiment was conducted by G. Poxleitner as part of our publication [5]. Both images were part of figure 4 in the publication. The experimental image was overlaid with a bright-field image to help the distinction between sectors of the toxin-producing strain and the background. Both the simulated colony and the experimentally observed colony exhibit a coexistence of all three strains at their fronts with faint sectors of the toxin-producing strain C.

calibrated using data from independent experiments on the growth of single-strain colonies. The values of the mesoscopic growth rates and lag times were calibrated so as to reproduce experimentally determined growth curves. The strength and range of the toxin interaction were determined by comparing inhibition zones between colliding sensitive and colicin-producing colonies.

The fully-calibrated model both reproduced the surviving strains of range expansion experiments and allowed us to formulate phenomenological "biodiversity laws" (cf. figure 5 of our publication [5]). These laws described experimental parameter regimes around which three-strain coexistence was expected to be maximal. We found that cyclic dominance was not a prerequisite for three-strain coexistence. Instead, the strains' coexistence depended on the right balance between strain growth rates, initial strain ratios, and the range of the toxin. For the cyclic dominance scenario, the model predicted three-strain coexistence around initial strain ratios of S:R:C = 1:1:0.1 in the inoculum (cf. figure 3 of [5]). Figure 4 compares a simulated and an experimentally observed colony for these strain ratios, which both exhibit the coexistence of all strains. Moreover, our model predicted the possibility of three-strain coexistence at equal initial strain ratios of S:R:C = 1:1:1 but for a characteristic toxin range that was reduced from 125 µm to approximately 50 µm. The experimental test of this prediction remains for future work.

2 Publication

Chemical warfare and survival strategies in bacterial range expansions

by

M. F. Weber,*[,1,2] G. Poxleitner,*[,2] E. Hebisch[2], E. Frey[1,2], and M. Opitz[2]

*equal contribution, [1]Arnold Sommerfeld Center for Theoretical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, Munich 80333, Germany, [2]Center for NanoScience, Faculty of Physics, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, Munich 80539, Germany

reprinted on pages 196–203

from

***J. R. Soc. Interface* 11(96), 20140172 (2014),**

DOI: 10.1098/rsif.2014.0172.

Electronic Supplementary Material reproduced on pages 204–235.

JOURNAL OF THE ROYAL SOCIETY Interface

rsif.royalsocietypublishing.org

Research

Cite this article: Weber MF, Poxleitner G, Hebisch E, Frey E, Opitz M. 2014 Chemical warfare and survival strategies in bacterial range expansions. *J. R. Soc. Interface* **11**: 20140172.
http://dx.doi.org/10.1098/rsif.2014.0172

Received: 18 February 2014
Accepted: 16 April 2014

Subject Areas:
biophysics, computational biology, biocomplexity

Keywords:
microbial ecology, range expansions, bacterial model systems, maintenance of biodiversity, species coexistence, *Escherichia coli*

Author for correspondence:
Madeleine Opitz
e-mail: opitz@physik.uni-muenchen.de

†These authors contributed equally to this study.
‡Present address: Department of NanoBiophotonics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsif.2014.0172 or via http://rsif.royalsocietypublishing.org.

Royal Society Publishing

Chemical warfare and survival strategies in bacterial range expansions

Markus F. Weber[1,2,†], Gabriele Poxleitner[2,†], Elke Hebisch[2,‡], Erwin Frey[1,2] and Madeleine Opitz[2]

[1]Arnold Sommerfeld Center for Theoretical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, Munich 80333, Germany
[2]Center for NanoScience, Faculty of Physics, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, Munich 80539, Germany

Dispersal of species is a fundamental ecological process in the evolution and maintenance of biodiversity. Limited control over ecological parameters has hindered progress in understanding of what enables species to colonize new areas, as well as the importance of interspecies interactions. Such control is necessary to construct reliable mathematical models of ecosystems. In our work, we studied dispersal in the context of bacterial range expansions and identified the major determinants of species coexistence for a bacterial model system of three *Escherichia coli* strains (toxin-producing, sensitive and resistant). Genetic engineering allowed us to tune strain growth rates and to design different ecological scenarios (cyclic and hierarchical). We found that coexistence of all strains depended on three strongly interdependent factors: composition of inoculum, relative strain growth rates and effective toxin range. Robust agreement between our experiments and a thoroughly calibrated computational model enabled us to extrapolate these intricate interdependencies in terms of phenomenological biodiversity laws. Our mathematical analysis also suggested that cyclic dominance between strains is not a prerequisite for coexistence in competitive range expansions. Instead, robust three-strain coexistence required a balance between growth rates and either a reduced initial ratio of the toxin-producing strain, or a sufficiently short toxin range.

1. Introduction

The fate of a species depends on the abilities of its members to colonize new areas and to outperform competitors [1,2]. A central theme of ecological research is to understand these abilities and to explain how many competing species still manage to live in lasting coexistence, especially during arms races over common resources [3–10]. Structured environments were theoretically proposed to be facilitators of biodiversity [3,11–19]. However, experimental verification of the proposed mechanisms promoting biodiversity is hard to come by. Ecological studies traditionally focused on systems of mammals and plants, but the long reproduction times and large spatial scales involved impede experimental progress [20]. To circumvent these problems, recent studies have turned to microbial model systems in which both spatial and temporal scales are experimentally better accessible [3,21–23]. New methods of genetic engineering even admit the possibility to modify the behaviour of test species. These methods stimulated further research on microbial systems and increased our knowledge about their transient and long-term dynamics [24]. For microbial life in well-mixed culture, for example, experimental and theoretical models have recently shown how transient processes can be amplified by recurring life cycles to change a system's long-term fate [25,26]. In spatial environments, long-term limits are more difficult to attain. We followed a previous study on competitions of three bacterial strains of *Escherichia coli* (toxin-producing, sensitive and resistant) in fixed spatial environments [3] and

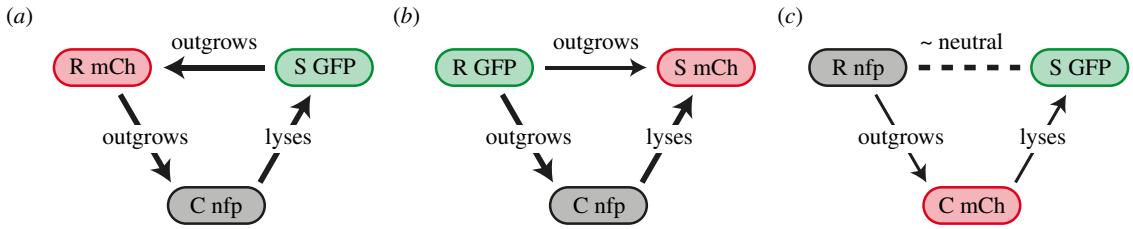


Figure 1. Three ecological scenarios that were designed by altering strain growth rates through the expression of the fluorescent protein mCherry. (*a*) In the cyclic scenario I, the hierarchy of single strain growth rates was $\mu_S > \mu_R > \mu_C$ (for numerical values, see the electronic supplementary material, table S1 and figure S1). Cyclic dominance held because colicin emitted by the toxin-producing C strain inhibited, and eventually lysed, cells of the sensitive S strain. (*b*) In the hierarchical scenario II, the resistant strain outperformed the two other strains and the growth rate hierarchy was $\mu_R > \mu_S \geq \mu_C$. (*c*) In the intermediate scenario III, colonies formed by either the S or the R strain expanded at roughly the same rate, and outgrew colonies formed by the toxin-producing C strain.

identified traits that ensure the transient coexistence of strains during the course of range expansions.

What determines whether a bacterial species thrives or falters as it explores new areas can be studied systematically after droplet inoculation on an agar plate [5,27–30]. Recent experimental studies have highlighted the importance of random genetic drift in driving population differentiation along the expanding fronts of bacterial colonies—an effect that gives rise to monoclonal sectoring patterns [5,31]. Natural microbial colonies and biofilms are characterized by a complex community structure [21,22,32], which is shaped by competition between strains for resources such as nutrients and space [2,5,27–30], interference competition through the production of toxins [3,8,22,29,33,34], and different forms of mutualism, cooperation and cheating [4,6,22,35]. Only a few recent studies, most of them theoretical, have explored the role of such interactions for expanding populations [36–38]. Experimental studies are appearing just recently [10,39,40] and are much needed to identify and characterize the key principles that drive population dynamics in expanding systems. In our work, we investigated range expansions for a bacterial model system comprising three *Escherichia coli* strains: a toxic strain, a sensitive strain (facing death upon the encounter of toxins) and a resistant strain. By genetically altering strain growth rates, we created three different ecological scenarios, including a hierarchical scenario and a scenario that mimicked a cyclic rock–paper–scissors game (figure 1) [3,18]. Control over strain growth rates also enabled us to acquire sufficient experimental data to construct and validate a computational model of the expansion process. The model was used to predict parameter regimes for which coexistence of all three strains was observed in experiments. Furthermore, we identified the factors that determined a strain's chance of survival (composition of the inoculum, relative growth rates and effective toxin range), and quantified the relationship between these factors in terms of phenomenological 'biodiversity laws'. Our work highlights the central importance of bacterial interactions in the evolution and maintenance of biological diversity, and pursues the theoretical aim to understand how interactions affect coexistence [41].

2. Material and methods

2.1. Bacterial strains and fluorescent proteins

The strains used in our study represent the *Escherichia coli* Colicin E2 system (BZB1011 (sensitive 'S' strain), $E2^C$-BZB1011 (toxic 'C' strain) and $E2^R$-BZB1011 (resistant 'R' strain)) [3]. For visualization of distinct strains, plasmids expressing either the green fluorescent protein (GFP), the red fluorescent protein mCherry (mCh) or no fluorescent protein (nfp) were introduced into S, R and C, respectively. The resulting strains were named: S_{GFP}, S_{mCh}, S_{nfp}, R_{GFP}, etc. All fluorescent proteins were expressed from the arabinose-inducible promoter pBAD as present in the plasmid pBAD24. Introduction of the fluorescent proteins resulted in the plasmids pBAD24-GFP [42] and pBAD24-mCherry [43]. To prevent plasmid loss, all plasmids, including the plasmid not expressing a fluorescent protein, carried an ampicillin antibiotic resistance.

2.2. Preparation of the system and growth conditions

Bacteria were grown in overnight cultures of liquid M63 medium at 37°C, supplemented with glycerol (0.2%), casein hydrolysate (0.2%) and arabinose (0.2%) for fluorescence induction, and with ampicillin (100 µg ml^{-1}). Analysis of colony development was performed on M63 agar plates (1.5% agar) that were prepared as above for the liquid culture.

Strain mixtures were diluted from the overnight culture to OD 0.1 at different initial ratios as indicated in the next sections. Ratios S : R : C (shorthand for $r_S : r_R : r_C$) of 1 : 1 : 1 represent an equal amount of all three strains. Ratios 5 : 1 : 1 indicate that the S strain was initially added five times more than the R and C strains, whereas ratios of 1 : 1 : 0.1 indicate that the C strain was added at one-tenth of the other two strains. Droplets of the resulting mixture (1 µl) were applied to M63 agar plates in triplicate. The time between mixing of strains and inoculation had to be kept short, because droplets of inoculum temporarily form well-mixed environments. Tuning the pH level of our agar plates resulted in slow colony growth at pH 6 (slow growth condition 'S') compared with fast colony growth at optimal pH 7 (fast growth condition 'F'). Each experiment (for slow and fast growth conditions) was performed two times and revealed qualitatively the same result.

2.3. Analysis of colony development

Colony development was recorded using an upright microscope (90i, Nikon, Düsseldorf, Germany). Fluorescence was analysed using filter sets with 472/30 nm excitation for GFP (DM: 495, BA: 520/35 BP), whereas excitation for mCherry was 562/25 nm (DM: 593, BA: 641/45 BP). Images were taken with a DS-Qi1MC digital camera (Nikon). Background correction and image analysis were performed using the free software ImageJ.

In order to quantify the growth dynamics of the three strains, we recorded the expansion of single-strain colonies for each combination of strain, fluorescent marker and growth condition in parallel by taking bright field images. For slow growth conditions, these pictures were recorded every 2 h from 4 to 34 h after inoculation. A final picture was recorded 48 h after inoculation. For fast growth conditions, the pictures were recorded

every hour from 4 to 16 h and every 2 h from 16 to 30 h after inoculation. A final picture was recorded 50 h after inoculation.

Bright field images were also recorded for the expansion of multi-strain colonies 48 h after inoculation, together with images for the two fluorescent proteins GFP and mCherry. Only colour overlays of the two fluorescence channels are shown. We chose a natural representation of colours for the visualization of fluorescent strains. Strains expressing the GFP are, therefore, shown in green, strains expressing the red fluorescent protein mCherry are shown in red. The choice helps to identify the strain that suffered a significant decrease in growth rate owing to the expression of mCherry. For slow growth conditions, expression of mCherry by the C strain caused a decrease of its growth rate of 21.5% when compared with its non-fluorescent state, for the R strain, a decrease of 22.6%, and for the S strain, a decrease of 24.3%. Growth rate decrease was less for fast growth conditions: 13.0% for C, 16.4% for R and 15.7% for S. All growth rates are listed and displayed in the electronic supplementary material, table S1 and figure S1. A strain not expressing a fluorescent protein (nfp) was present in black areas of a colour overlay. For cases in which the non-fluorescent strain could not be distinguished from surrounding agar, we used a bright field image to delineate the boundary of the colony (cf. figure 2).

2.4. Computational model of competitive range expansions

Our theoretical model rested on a coarse-grained, mesoscopic description of the bacterial expansion process. The model was agent-based, and movement of agents was restricted to a two-dimensional lattice, following previous stochastic simulations of range expansions [31]. Owing to the exceedingly large number of bacterial cells in experimentally observed colonies, it was not possible to treat individual bacterial cells as agents. Instead, each agent (a colonized lattice site) represented the bacterial strain that locally dominated a certain area (patch) of a colony (further details on the physical size represented by lattice sites can be found in the electronic supplementary material, appendix S1 and table S1). Our model thereby coarse-grained the microscopic dynamics and reduced both high cell numbers in the lateral direction as well as the increasing number of cell layers in the vertical direction to a lattice of patches.

The growth dynamics of the expansion process was modelled as 'hopping processes' from colonized to free patches. The speed of these processes depended on two strain-dependent parameters: mesoscopic growth rate μ_m and mesoscopic lag-time τ_m. These parameters were adjusted such that our simulation reproduced experimental data on the radial growth of single-strain colonies. The mesoscopic lag-time helped us to effectively add a time dependence to the mesoscopic growth rate (a colonized patch could proliferate only after its lag-time τ_m had passed). This time dependence also allowed us to reproduce the lag phase and the gradual expansion of colonies during acceleration phase that were observed in growth curves (see electronic supplementary material, figure S8). In order to match simulated growth curves to experimental data, we sampled τ_m on initially colonized lattice sites from broad, strain- and marker-dependent Gaussian distributions. The calibration of these distributions is explained in the electronic supplementary material, appendix S1.

The temporal dynamics of range expansions was simulated employing a Gillespie algorithm [44]. The algorithm also governed the toxin interaction between sensitive and toxic lattice sites (we assumed that 3% of initially and newly colonized lattice sites dominated by the C strain were toxic [45]). Our inclusion of this interaction explicitly accounted for the long-range diffusion of colicins (a nearest-neighbour interaction would have been insufficient to recover experimental results). Because diffusion of colicins happens on a much faster time scale than consecutive cell divisions of *E. coli* (diffusion constant of colicins of the order of $10^{-7}\,\mathrm{cm^2\,s^{-1}}$ [46]), we approximated the colicin dynamics by a stationary source and degradation process. The process suggested exponentially decaying colicin profiles around toxin-producing lattice sites. The colicin profiles introduced two additional parameters into the simulation: their heights around sources (local colicin strength κ) and their widths (characteristic length scale λ). We adjusted the parameters using estimates from the literature [33,47] and by calibrating them to experiments on colliding sensitive and colicin producing colonies (see electronic supplementary material, figure S11).

3. Results and discussion

3.1. Design of distinct ecological scenarios

As detailed in the Material and methods section, we studied competitive microbial range expansion for a prominent model system that comprises three genetically distinct strains of *Escherichia coli* [3]: a toxin-producing strain (C), a sensitive strain (S) and a resistant strain (R). During growth, around 3% [45] of the C cells undergo lysis and release colicin E2 (diffusion constant of the order of $10^{-7}\,\mathrm{cm^2\,s^{-1}}$ [46]). The colicins subsequently diffuse through the extracellular fluid around bacterial cells until possibly being absorbed by sensitive *E. coli* cells. The sensitive cells are prone to the endonuclease activity of colicin and suffer a degeneration of their DNA, which inhibits further cell divisions [48]. Eventually, the cells lyse. Inhibition zones around toxic C cells may be as large as 100–400 μm in radius [33,47]. Because colicin production is costly, the growth rate of these cells is significantly lower than those of the other two strains. We genetically engineered two of our three strains to express either GFP or the red fluorescent protein mCherry (the strain not expressing a fluorescent protein is marked as nfp). We observed that while a strain expressing GFP could expand at roughly the same speed as its non-fluorescent counterpart, this did not hold for strains expressing mCherry. We discovered that their growth rate was significantly reduced by the expression of mCherry (see the electronic supplementary material, figure S1*a*). This effect was also observed for growth in liquid culture [43]. The fluorescent proteins thereby allowed us to design different ecological scenarios by changing the order in which the proteins were assigned to our strains (figure 1). Every scenario differed from one another by changes in relative strain growth rates as described in the following. Furthermore, the fluorescent proteins allowed us to visualize each strain independently during its expansion in range (see Material and methods).

The control over the growth rates of our three strains enabled us to design three different ecological scenarios and to study how the composition of expanding colonies depended on the interplay between resource and interference competition. In a first scenario (I), we arranged the bacterial growth rates such that $\mu_S > \mu_R > \mu_C$ (mCherry expressed by the R strain). As detailed in the electronic supplementary material, appendix S1, we determined these growth rates by measuring the maximal radial expansion velocity ($\mu\mathrm{m\,h^{-1}}$) of single-strain S, R, and C colonies. These rates were thus independent of the toxin action of C on S, which was quantified independently as described further below. It followed from the above hierarchy $\mu_S > \mu_R > \mu_C$ and from the toxin action of C on S that our first ecological system exhibited a *cyclic* (*non-transitive*) dominance: C dominated S by killing

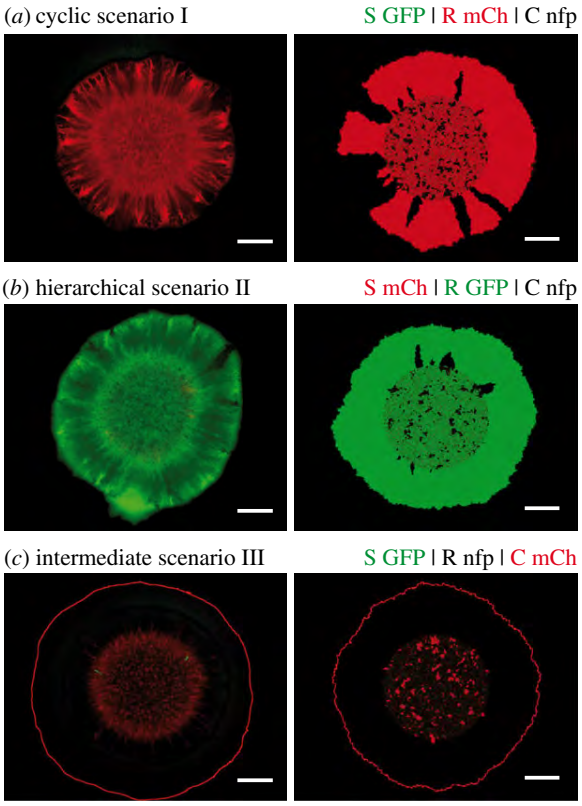


Figure 2. Segregation patterns arising in range expansions initiated at initial strain ratios S : R : C = 1 : 1 : 1. Experimental outcomes are displayed in the left column (images obtained 48 h after inoculation), simulation results in the right column (simulations stopped after colony had reached a radius of 3 mm). Different strain/fluorescent marker combinations were used for visualization and to implement distinct ecological scenarios. The combinations are indicated above individual rows (GFP, green fluorescent protein; mCh, red fluorescent protein mCherry; nfp, no fluorescent protein). For further information on the robustness of experimental as well as theoretical results, see the electronic supplementary material, appendix S1 and figures S3 and S4. White scale bars represent 1 mm. (*a*) Transient coexistence of the R and the C strain was maintained in the cyclic scenario I. (*b*,*c*) The R strain outcompeted both the S and the C strain in the strictly hierarchical scenario II and in the intermediate scenario III.

it, S outgrew R, which, in turn, outgrew C (figure 1*a*). This hierarchy resembled the order of strategies in the children's game rock–paper–scissors [3,18]. In the second scenario (II), the ordering of growth rates was chosen as: $\mu_R > \mu_S \geq \mu_C$ (mCherry expressed by the S strain). Hence, the interaction network was *strictly hierarchical* (*transitive*), with R displacing C because of its growth advantage, and C displacing S through its allelopathic effect on S (figure 1*b*). In a third intermediate scenario (III), the toxic strain had by far the lowest growth rate, whereas those of R and S were nearly equal (mCherry expressed by the C strain). Under these conditions, the competition network was neither cyclic nor strictly hierarchical: R dominated C by resource competition, and C dominated S by interference competition, but the interaction between R and S was selectively nearly neutral (figure 1*c*). After droplet inoculation of 1 µl mixtures on agar plates (supplemented with minimal M63 medium; see Material and methods), we tracked the composition of bacterial colonies over 48 h and identified the strains that coexisted along expanding fronts of colonies. The strains that were present along these fronts after 48 h were considered as survivors of the range expansion. It was sufficient that a strain had established at least one stable sector that touched the edge of an expanding colony to be considered a survivor. Our notion of survival and coexistence did not evaluate the number of stable sectors or the relative frequency of strains along the fronts of colonies.

We developed a theoretical model to explain the outcome of bacterial competitions and to predict growth parameters for which a maximal degree of coexistence could be observed in experiments along the expanding fronts of colonies. The predictions were verified experimentally as described in the following. Let us note that we focused on the transient coexistence of bacterial strains on time spans that were accessible to experiments. Korolev *et al*. [49] developed methods to determine when such transient coexistence is eventually lost. However, the approach does not consider toxin interaction between strains and can only be applied to cases in which either the S or the C strain has already disappeared from a colony's front. In such situations, the strain that eventually dominates may be inferred from the radial expansion velocities of single-strain colonies that are listed in the electronic supplementary material, table S1.

3.2. Cyclic dominance is not sufficient to ensure coexistence of all strains

We first sought to determine the surviving strains when a droplet of inoculum that contained an equal number of all three strains (initial ratios S : R : C = 1 : 1 : 1) expanded in range. Surprisingly, in the cyclic rock–paper–scissors scenario I, we found no evidence for coexistence of all three strains. In a previous report, such three-strain coexistence was observed for spatially extended systems with a regular arrangement of neighbouring single-strain colonies [3]. Competitive exclusion with dominance of the fastest-growing strain (S) was not observed either. Instead, we found that S was driven to extinction, whereas strains R and C dominated the colony front, where they formed monoclonal sectors (figure 2*a*). Notably, in the non-cyclic scenarios II and III, coexistence was completely lost. Here, the R strain outcompeted both S and C, and was the only survivor with access to uncolonized area (figure 2*b*,*c*). Hence, 'survival of the fastest' [10] could only be observed in hierarchical scenario II, whereas who survived in the other two scenarios was more subtle and was heavily affected by the long-range toxin action. The outcomes of our bacterial competitions were shown to be robust against small changes in relative growth rates of the strains (induced by reassigning the fluorescent protein GFP while keeping the assignment of mCherry; see the electronic supplementary material, appendix S1), and robust against changes in overall growth conditions (slow growth at pH 6, fast growth at pH 7; see Material and methods and the electronic supplementary material, figure S1). The results of supporting experiments are listed in the electronic supplementary material, figures S3 and S4.

3.3. Identification of biodiversity zones

To elucidate the above findings and to identify the factors that promote or jeopardize survival of the competing strains, we developed a stochastic agent-based model to capture the system dynamics on a coarse-grained scale (see Material and methods and the electronic supplementary material, appendix S1). Our theoretical approach rested on a lattice-based

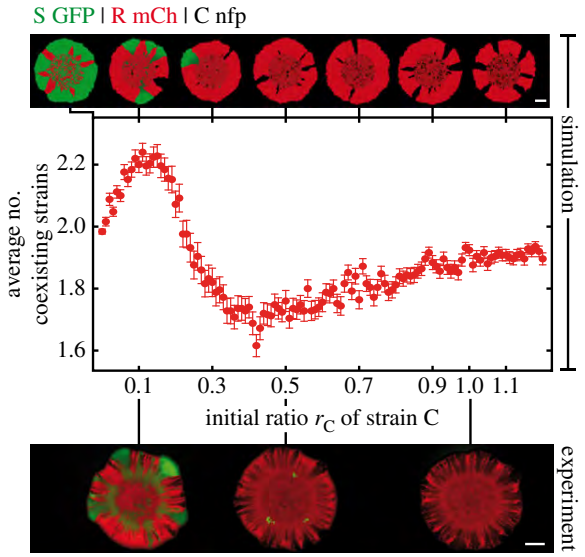


Figure 3. Dependence of the degree of coexistence on the relative initial amount of the C strain. With initial ratios of S : R : C = 1 : 1 : 1 (r_S: r_R: r_C) in the cyclic scenario I, the R and C strain came to dominate the expanding front. As the initial ratio of the C strain r_C was reduced, R strain sectors became broader, in agreement with experimental observations (depicted below the theoretical results; see also the electronic supplementary material, figure S5). Further reduction of the initial ratio of C weakened its allelopathic effect on S such that expanding S sectors emerged. A maximal number of coexisting strains along colony fronts was observed around S : R : C = 1 : 1 : 0.1 (24% of simulations exhibited three-strain coexistence; error bars: s.e.m., n = 250). The three-strain coexistence at this initial ratio was also observed in experiments. White scale bars represent 1 mm.

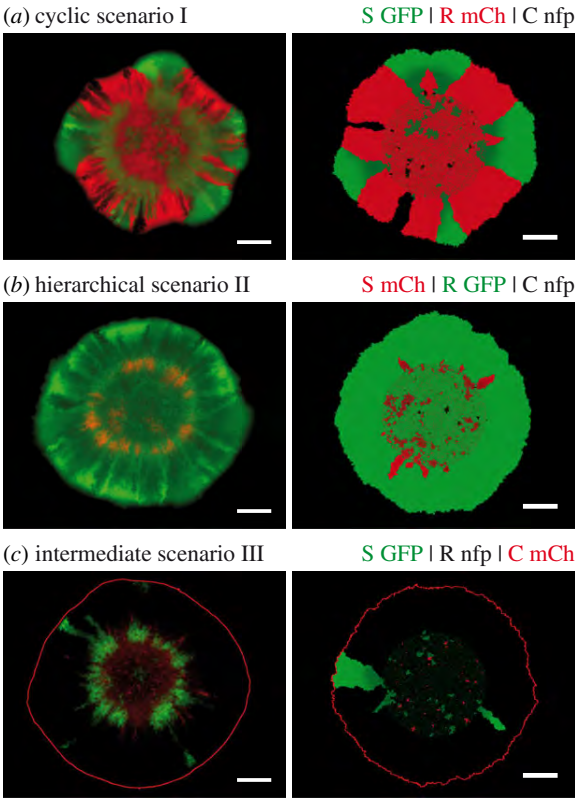


Figure 4. Range expansions at initial ratios S : R : C = 1 : 1 : 0.1. See the legend of figure 2 for additional information. (*a*) A transient coexistence of all three strains was observed along the colony's expanding rim in the cyclic scenario I. The distinct protrusions formed by S_{GFP} hint at its selective advantage over the other two strains (see the electronic supplementary material, table S1). Simulations and deterministic analysis anticipated an eventual dominance of S_{GFP} on longer time scales [49]. (*b*) In the strictly hierarchical scenario II, the growth rate of the S strain was slowed by the expression of mCherry such that it could not compete against R, despite the low initial ratio of C. (*c*) Both R and S strains survived the range expansion in the intermediate scenario III, with the former strain being dominant over the latter.

description of range expansions and extended previous models [31] by considering the long-range nature of the toxin interaction. We performed additional experiments on the expansion of single-strain colonies to adjust the model's parameters. Comparisons between experimental and simulated growth curves enabled us to determine all model parameters except for the toxin interaction. We modelled this interaction based on a source and degradation process, and estimated its range and strength by measuring the distance between the front of a growing C colony and the front of a neighbouring S colony (see the electronic supplementary material, appendix S1). The estimate complies with literature values [33,47]. Even though our theoretical model simplified the bacterial dynamics (e.g. by considering only a single bacterial layer, whereas the real colony piled up in hundreds of them in its interior), the model captured the essential parameters. We successfully applied the model to reproduce experimentally observed segregation patterns and to predict the strains that survived a range expansion (figure 2). Let us emphasize that the model's parametrization rested on independent experiments as described above.

After having established and validated a reliable theoretical model that reproduced our experimental observations, we investigated whether it was possible to rescue the S strain. As the survival of the S strain was directly coupled to the C strain's presence, we analysed how reductions in the initial ratio of the C strain affected the other strains' survival (in particular of the S strain). Simulation data for the cyclic ecological scenario I predicted that reduction of its initial ratio should lead to the formation of broader R sectors at the expense of C (figure 3). The same effect was seen in experiments with initial ratios of S : R : C = 1 : 1 : 0.5 (see the electronic supplementary material, figure S5). Further reduction of the initial ratio of C in our simulations revealed a regime of three-strain coexistence centred around S : R : C = 1 : 1 : 0.1 (figure 3). This permissive zone of biodiversity in parameter space coincided remarkably well with our experimental observations of transient three-strain coexistence at ratios 1 : 1 : 0.1 (figure 4*a*). For ecological scenarios with a more hierarchical interaction relationship between strains (scenarios II and III), the R strain was clearly dominant (figure 4*b*,*c*). Hence, toxin resistance is apparently a more effective survival strategy than either rapid growth or toxin production if the hierarchical order in the competition network is enhanced.

Whether a bacterial strain manages to survive a range expansion and to populate a colony's expanding front depends on two aspects: first, on its ability to form initial clusters in the inoculum from which outward sectors may emerge; second, on the stability of the arising sectors to the annihilation of neighbouring sector boundaries [5]. Both of these aspects are subjected to random genetic drift and may prevent the establishment of stable sectors in a simulation (figure 3). However, whether a specific outcome of the

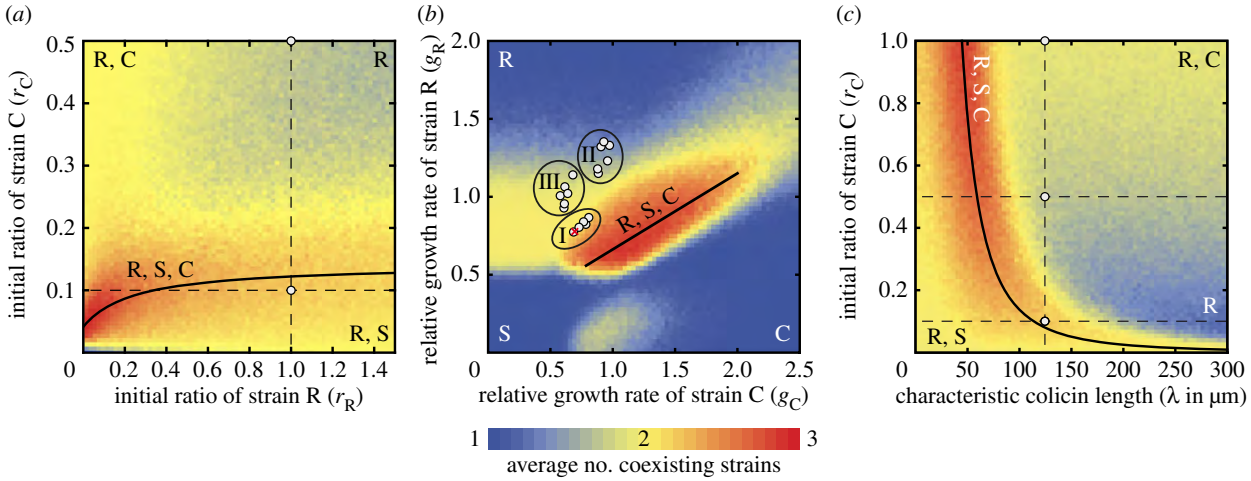


Figure 5. Coexistence diagrams and formulation of 'biodiversity laws'. We extended the simulations to study the individual impact of initial strain ratios, relative growth rates and toxin range. Parameters not under consideration were set to their values in the cyclic scenario I (see the electronic supplementary material, table S1). Letters within the figure indicate strains that survived a range expansion in the corresponding parameter regime. The colour scale on the bottom visualizes the level of coexistence averaged over 250 simulations. Solid lines delineate regimes of maximal three-strain coexistence and are referred to as 'biodiversity laws'. (*a*) Studying the initial ratios of strain C (r_C) and of strain R (r_R) revealed a regime of maximal coexistence that followed the *saturation law* $r_C = (0.01 + 0.14r_R)/(0.24 + r_R)$. Three-strain coexistence at the lower white circle was supported by experimental realizations. At the upper white circle, survival of R and C was seen in experiments. (*b*) Varying the relative growth rates of strains R and C with respect to that of S revealed that cyclic dominance is not a prerequisite for the maintenance of biodiversity (initial ratios S : R : C = 1 : 1 : 0.1). Maximal coexistence follows the *linear law* $g_R = 0.17 + 0.49g_C$ (straight line). White circles represent experimental results that were in accordance with our theoretical predictions for the indicated ecological scenario (including experimental results for fast growth conditions and for small growth rate variations; see the electronic supplementary material, appendix S1). The red cross indicates an experiment in which stable coexistence was not observed ($F_{0.1}$ in the electronic supplementary material, figure S7). Surviving strains were also required to have expanded by at least half the distance of the leading strain out of an inoculum. The criterion was only relevant for highly diverging growth rates (e.g. in the lower left). (*c*) Coexistence diagram for the influence of the C strain's toxicity. The colicin interaction with characteristic length scale of $\lambda = 125$ μm led to survival of both R and C at initial ratios S : R : C = 1 : 1 : 1 and 1 : 1 : 0.5, and to three-strain coexistence at 1 : 1 : 0.1 in experiments (white circles). Simulations predicted an increased level of coexistence along the *power law* $r_C \sim 1/\lambda^{2.46}$.

bacterial competition is possible in principle depended solely on the interplay between three factors in our experiments: (i) on the initial strain ratios in the inoculum (demographic noise owing to low absolute cell numbers was of minor importance), (ii) on the relative growth rates of the three strains, and (iii) on the effective range of colicin toxicity. On the other hand, differences in lag-times between strains played only a minor role in deciding whether a particular strain survived along the expanding front of a colony. To gain insights into the mechanisms responsible for the dependence of biodiversity on the three factors (i)–(iii) and into how they are correlated with each other, we extended our simulations to explore broad parameter ranges.

If the initial ratios of R and C were varied with respect to the initial ratio of S in cyclic ecological scenario I, then our simulations showed that biodiversity was most pronounced when the initial ratio r_C of the C strain was reduced to 5–20% of that of the S strain (figure 5*a*). Higher initial ratios of C suppressed growth of the S strain completely, but the R strain ended up dominating the expanding front. In this case, toxin resistance may be seen as a 'cheating' strategy: the R strain could profit from colicin production by the 'cooperating' C strain without having to pay the associated metabolic costs. By cheating, the R strain managed to beat S, even though it would have been the loser in a direct pairwise competition. Conversely, at lower initial ratios of the C strain, the S strain could still bear the incurred costs. Both R and S outgrew the C strain and eventually shared the expanding front. Our results indicated that a narrow range of initial ratios delineated a regime of maximal biodiversity. Biodiversity required that increases in the initial ratio of C were compensated for by even larger increases in the initial ratio of R. The correlation was quantified by the saturation law—a 'biodiversity law'—shown in figure 5*a*. We attributed the saturation to the finite range of colicin toxicity: dense swathes of R cells were needed to shield sensitive cells from the toxin emitted by the C strain. Behind these barriers, surviving S cells could give rise to sectors, leading to the eventual coexistence of all three strains.

Subsequently, we set the initial ratios of the three strains to the rescue window S : R : C = 1 : 1 : 0.1 and investigated how changes in the relative growth rates of the strains (i.e. changes in the interaction hierarchy) affected the degree of coexistence. Our simulations showed that three-strain coexistence was most pronounced when the growth rates were of comparable size and when the growth rates of strains C and R were varied in a correlated fashion: $\mu_R \sim \mu_C$ (figure 5*b*). As our model predicted two- and three-strain coexistence (as well as its absence) in full accordance with experimental results (R and S in the intermediate scenario III, all three strains in the cyclic scenario I, but only the R strain in the hierarchical scenario II), we expect our theoretical predictions to be highly relevant for future experimental studies. Moreover, our simulations revealed that cyclic dominance is not a necessary prerequisite for biodiversity. For range expansion ecologies, biological diversity can even be maintained if the toxin-producing C strain grows fastest. This result seems paradoxical at first sight, but it demonstrates that both the initial ratios and the growth rates of competing strains are equally important ecological

parameters. During the initial phase of expansion after inoculation, the combined effect of the two parameters determines which strain is more likely to establish sector-like domains. In order to avoid being overgrown by the other two strains, the C strain must compensate for its lower initial ratio by growing at a faster rate. A phase diagram that resembled the one in figure 5*b* was computed for range expansions of selectively neutral, non-interacting strains at equal initial strain ratios. The biodiversity window of this null model disappeared in the presence of toxin interaction, but was recovered upon reducing the initial C strain ratio. The changes to the null model were crucial for predicting the surviving strains in our experiments. It would be highly interesting to study how other kinds of bacterial interactions affect the coexistence diagram of the null model.

Finally, to understand the role of colicin in maintaining biodiversity during range expansions, we analysed the importance of the toxin's effective range (see the electronic supplementary material, appendix S1). Our *in silico* studies revealed that maintenance of biodiversity required a strong inverse correlation between the initial ratio of the toxic strain and the length scale of colicin toxicity: $r_C \sim 1/\lambda^{2.46}$ (figure 5*c*). A long-range toxin interaction (length scale of $\lambda \approx 125\ \mu m$) was, therefore, optimal for species coexistence around the initial strain ratios $S:R:C = 1:1:0.1$. However, our simulations suggested that a more circumscribed radius of toxin action ($\lambda \approx 50\ \mu m$) would be necessary to sustain coexistence at equal strain ratios $1:1:1$. The reduction in colicin range weakened the allelopathic effect of C on the fast-growing S strain to a level at which all strains could coexist along the expanding front, despite equal initial ratios in the inoculum. In conclusion, the coexistence diagram in figure 5*c* revealed that changes in the range of colicin toxicity have a strong impact on biodiversity. The maintenance of coexistence relied on the fine-tuning of the interference competition via colicin between the strains. In more general terms, the biodiversity law encodes how coexistence depends on the balance between the amount of the producers of an interaction agent and the range of the agent. We expect that the inverse correlation between the two can also be observed in other systems in which an inhibiting interaction is mediated by an agent. Future studies should explore how the law changes for other kinds of interactions.

4. Conclusion

Range expansion experiments provide a new perspective on the significance of competition between species in spatially extended ecological systems. Neither strength of numbers, nor growth rate differences, nor choice of competition strategy alone determines success of their dispersal. The right balance between these factors must be struck. We identified this balance for range expansions of a bacterial model system of three *Escherichia coli* strains and experimentally validated theoretical predictions on strain coexistence. We used the model to extrapolate in parameter space and described the regimes of maximal coexistence in terms of phenomenological 'biodiversity laws'. The laws showed how changes in the interaction between bacterial strains can have subtle but lasting effects on the eventual composition of a microbial ecosystem. Our approach may help to understand more complex ecosystems whose dynamics cannot be formulated in terms of simplistic rules.

Acknowledgements. We thank M. Riley for the kind gift of the original S, R and C strains, and K. Jung for the kind gift of plasmid pBAD24-GFP and pmCherry. For fruitful discussions and technical support, we thank A. Boschini, J.-T. Kuhr, J. Landsberg, G. Schwake, B. Nickel, J. O. Rädler and T. Reichenbach. The authors also thank D. Braun, J. Knebel, B. Osberg and P. B. Rainey for discussions and critical reading of the manuscript. M.O. and E.F. designed the research. G.P. and E.H. performed the experiments. M.F.W. conceived and performed the simulations. G.P., M.F.W., E.F. and M.O. interpreted the data. M.F.W., E.F. and M.O. wrote the paper.

Funding statement. This work was supported by the Deutsche Forschungsgemeinschaft through grant nos. FR 850/9-1 and RA 655/5-1, the NanoSystems Initiative Munich (NIM) and the Center for Nanoscience (CeNS).

References

1. Sakai AK *et al*. 2001 The population biology of invasive species. *Annu. Rev. Ecol. Syst.* **32**, 305–332. (doi:10.1146/annurev.ecolsys.32.081501.114037)
2. Hastings A *et al*. 2005 The spatial spread of invasions: new developments in theory and evidence. *Ecol. Lett.* **8**, 91–101. (doi:10.1111/j.1461-0248.2004.00687.x)
3. Kerr B, Riley MA, Feldman MW, Bohannan BJM. 2002 Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418**, 171–174. (doi:10.1038/nature00823)
4. Greig D, Travisano M. 2004 The prisoner's dilemma and polymorphism in yeast SUC genes. *Proc. R. Soc. Lond. B* **271**, S25–S26. (doi:10.1098/rsbl.2003.0083)
5. Hallatschek O, Hersen P, Ramanathan S, Nelson DR. 2007 Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl Acad. Sci. USA* **104**, 19 926–19 930. (doi:10.1073/pnas.0710150104)
6. Gore J, Youk H, Van Oudenaarden A. 2009 Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**, 253–256. (doi:10.1038/nature07921)
7. Mougi A, Kondoh M. 2012 Diversity of interaction types and ecological community stability. *Science* **337**, 349–351. (doi:10.1126/science.1220529)
8. Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F, Le Roux F, Mincer T, Polz MF. 2012 Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* **337**, 1228–1231. (doi:10.1126/science.1219385)
9. Datta MS, Korolev KS, Cvijovic I, Dudley C, Gore J. 2013 Range expansion promotes cooperation in an experimental microbial metapopulation. *Proc. Natl Acad. Sci. USA* **110**, 7354–7359. (doi:10.1073/pnas.1217517110)
10. Van Dyken JD, Müller MJI, Mack KML, Desai MM. 2013 Spatial population expansion promotes the evolution of cooperation in an experimental prisoner's dilemma. *Curr. Biol.* **23**, 919–923. (doi:10.1016/j.cub.2013.04.026)
11. Nowak MA, May RM. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)
12. Frank SA. 1994 Spatial polymorphism of bacteriocins and other allelopathic traits. *Evol. Ecol.* **8**, 369–386. (doi:10.1007/BF01238189)
13. Hassell MP, Comins HN, May RM. 1994 Species coexistence and self-organizing spatial dynamics. *Nature* **370**, 290–292. (doi:10.1038/370290a0)
14. Durrett R, Levin S. 1997 Allelopathy in spatially distributed populations. *J. Theor. Biol.* **185**, 165–171. (doi:10.1006/jtbi.1996.0292)
15. Czárán TL, Hoekstra RF, Pagie L. 2002 Chemical warfare between microbes promotes biodiversity. *Proc. Natl Acad. Sci. USA* **99**, 786–790. (doi:10.1073/pnas.012399899)

16. Amarasekare P. 2003 Competitive coexistence in spatially structured environments: a synthesis. *Ecol. Lett.* **6**, 1109–1122. (doi:10.1046/j.1461-0248.2003.00530.x)
17. Szabó G, Fáth G. 2007 Evolutionary games on graphs. *Phys. Rep.* **446**, 97–216. (doi:10.1016/j.physrep.2007.04.004)
18. Reichenbach T, Mobilia M, Frey E. 2007 Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448**, 1046–1049. (doi:10.1038/nature06095)
19. Wakano JY, Nowak MA, Hauert C. 2009 Spatial dynamics of ecological public goods. *Proc. Natl Acad. Sci. USA* **106**, 7910–7914. (doi:10.1073/pnas.0812644106)
20. Sinervo B, Lively CM. 1996 The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380**, 240–243. (doi:10.1038/380240a0)
21. Jessup CM, Kassen R, Forde SE, Kerr B, Buckling A, Rainey PB, Bohannan BJM. 2004 Big questions, small worlds: microbial model systems in ecology. *Trends Ecol. Evol.* **19**, 189–197. (doi:10.1016/j.tree.2004.01.008)
22. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010 Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **8**, 15–25. (doi:10.1038/nrmicro2259)
23. Rainey PB, Travisano M. 1998 Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72. (doi:10.1038/27900)
24. Hastings A. 2004 Transients: the key to long-term ecological understanding? *Trends Ecol. Evol.* **19**, 39–45. (doi:10.1016/j.tree.2003.09.007)
25. Chuang JS, Rivoire O, Leibler S. 2009 Simpson's paradox in a synthetic microbial system. *Science* **323**, 272–275. (doi:10.1126/science.1166739)
26. Cremer J, Melbinger A, Frey E. 2012 Growth dynamics and the evolution of cooperation in microbial populations. *Sci. Rep.* **2**, 281. (doi:10.1038/srep00281)
27. Ben-Jacob E, Schochet O, Tenenbaum A, Cohen I, Czirok A, Vicsek T. 1994 Generic modeling of cooperative growth-patterns in bacterial colonies. *Nature* **368**, 46–49. (doi:10.1038/368046a0)
28. Golding I, Cohen I, Ben-Jacob E. 1999 Studies of sector formation in expanding bacterial colonies. *Europhys. Lett.* **48**, 587–593. (doi:10.1209/epl/I1999-00524-7)
29. Be'er A *et al*. 2010 Lethal protein produced in response to competition between sibling bacterial colonies. *Proc. Natl Acad. Sci. USA* **107**, 6258–6263. (doi:10.1073/pnas.1001062107)
30. Korolev KS, Avlund M, Hallatschek O, Nelson DR. 2010 Genetic demixing and evolution in linear stepping stone models. *Rev. Mod. Phys.* **82**, 1691–1718. (doi:10.1103/RevmodPhys.82.1691)
31. Hallatschek O, Nelson DR. 2010 Life at the front of an expanding population. *Evolution* **64**, 193–206. (doi:10.1111/j.1558-5646.2009.00809.X)
32. Hansen SK, Rainey PB, Haagensen JAJ, Molin S. 2007 Evolution of species interactions in a biofilm community. *Nature* **445**, 533–536. (doi:10.1038/nature05514)
33. Chao L, Levin BR. 1981 Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proc. Natl Acad. Sci. USA* **78**, 6324–6328. (doi:10.1073/pnas.78.10.6324)
34. Majeed H, Gillor O, Kerr B, Riley MA. 2011 Competitive interactions in *Escherichia coli* populations: the role of bacteriocins. *ISME J.* **5**, 71–81. (doi:10.1038/Ismej.2010.90)
35. Rainey PB, Rainey K. 2003 Evolution of cooperation and conflict in experimental bacterial populations. *Nature* **425**, 72–74. (doi:10.1038/nature01906)
36. Nadell CD, Foster KR, Xavier JB. 2010 Emergence of spatial structure in cell groups and the evolution of cooperation. *PLoS Comput. Biol.* **6**, e1000716. (doi:10.1371/journal.pcbi.1000716)
37. Korolev KS, Nelson DR. 2011 Competition and cooperation in one-dimensional stepping-stone models. *Phys. Rev. Lett.* **107**, 088103. (doi:10.1103/physrevlett.107.088103)
38. Kuhr JT, Leisner M, Frey E. 2011 Range expansion with mutation and selection: dynamical phase transition in a two-species Eden model. *New J. Phys.* **13**, 113013. (doi:10.1088/1367-2630/13/11/113013)
39. Momeni B, Brileya KA, Fields MW, Shou W. 2013 Strong inter-population cooperation leads to partner intermixing in microbial communities. *eLife* **2**, e00230. (doi:10.7554/eLife.00230)
40. Müller MJI, Neugeboren BI, Nelson DR, Murray AW. 2014 Genetic drift opposes mutualism during spatial population expansion. *Proc. Natl Acad. Sci. USA* **111**, 1037–1042. (doi:10.1073/pnas.1313285111)
41. Knebel J, Krüger T, Weber MF, Frey E. 2013 Coexistence and survival in conservative Lotka–Volterra networks. *Phys. Rev. Lett.* **110**, 168106. (doi:10.1103/physrevlett.110.168106)
42. Megerle JA, Fritz G, Gerland U, Jung K, Rädler JO. 2008 Timing and dynamics of single cell gene expression in the arabinose utilization system. *Biophys. J.* **95**, 2103–2115. (doi:10.1529/biophysj.107.127191)
43. Hebisch E, Knebel J, Landsberg J, Frey E, Leisner M. 2013 High variation of fluorescence protein maturation times in closely related *Escherichia coli* strains. *PLoS ONE* **8**, e75991. (doi:10.1371/journal.pone.0075991)
44. Gillespie DT. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434. (doi:10.1016/0021-9991(76)90041-3)
45. Cascales E, Buchanan SK, Duche D, Kleanthous C, Lloubes R, Postle K, Riley M, Slatin S, Cavard D. 2007 Colicin biology. *Microbiol. Mol. Biol. Rev.* **71**, 158–229. (doi:10.1128/MMBR.00036-06)
46. Liang SM, Xu J, Weng LH, Dai HJ, Zhang XL, Zhang LN. 2006 Protein diffusion in agarose hydrogel in situ measured by improved refractive index method. *J. Control. Release* **115**, 189–196. (doi:10.1016/j.jconrel.2006.08.006)
47. Ozeki H, Stocker BAD, De Margerie H. 1959 Production of colicine by single bacteria. *Nature* **184**, 337–339. (doi:10.1038/184337a0)
48. Schaller K, Nomura M. 1976 Colicin-E2 is a DNA endonuclease. *Proc. Natl Acad. Sci. USA* **73**, 3989–3993. (doi:10.1073/pnas.73.11.3989)
49. Korolev KS, Müller MJI, Karahan N, Murray AW, Hallatschek O, Nelson DR. 2012 Selective sweeps in growing microbial colonies. *Phys. Biol.* **9**, 026008. (doi:10.1088/1478-3975/9/2/026008)

Electronic Supplementary Material for
Chemical Warfare and Survival Strategies in Bacterial Range Expansions

Markus F. Weber[1,2,#], Gabriele Poxleitner[2,#], Elke Hebisch[2,‡], Erwin Frey[1,2], Madeleine Opitz[2,*]

[1]Arnold-Sommerfeld-Center for Theoretical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, D-80333 Munich, Germany.
[2]Center for NanoScience, Faculty of Physics, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, D-80539 Munich, Germany.

[#]These authors contributed equally to this study.

[*]Opitz@physik.uni-muenchen.de.

Appendix S1. Supporting Information on experimental and theoretical methods.

The following text comprises an experimental and a theoretical part. The experimental section includes additional experimental results, further information about experimental procedures and data analysis, and a description of the experimentally derived parameters that were used in the accompanying theoretical model. The subsequent section gives a detailed report on our theoretical model of competitive range expansions.

Table of Contents

Abbreviations

Bacterial strains of *Escherichia coli*:
- S = sensitive
- R = resistant
- C = colicin producing

Ecological scenarios (see main text):
- I and I* = cyclic (non-transitive)
- II and II* = strictly hierarchical (transitive)
- III and III* = intermediate (neither cyclic nor strictly hierarchical)

Growth conditions:
- S = slow
- F = fast

Initial strain ratios S:R:C (shorthand for r_S:r_R:r_C):
- 1 = 1:1:1
- 0.5 = 1:1:0.5
- 0.1 = 1:1:0.1

Fluorescent proteins:
- nfp = no fluorescent protein
- GFP = green fluorescent protein
- mCh = red fluorescent protein mCherry

Statistics:
- s.d. = standard deviation
- s.e.m. = standard error of the mean
- r.s.e.m. = relative standard error of the mean

1. Experimental Part

1.1. Results

1.1.1. Tuning of Growth Conditions

As described in *Materials and Methods* (main text), plasmids expressing either no fluorescent protein (nfp), the green fluorescent protein (GFP), or the red fluorescent protein mCherry (mCh) were introduced into the sensitive (S), resistant (R), and colicin producing (C) strains that were used in this study. Expression of a fluorescent protein affected the growth of single-strain colonies in two ways: it altered its growth rate μ (figure S1) and/or its lag time τ (figure S2). We refer the reader to section 1.2 for definitions of these two parameters. That section also gives a detailed description of how the respective numerical values in table S1 (depicted in figures S1 and S2) were determined. We obtained data for two different growth conditions: slow (S) and fast (F). These growth conditions are described in *Materials and Methods*.

1.1.2. Robustness of Results for Equal Initial Strain Ratios

We analysed three distinct ecological scenarios (cyclic scenario I, hierarchical scenario II, and intermediate scenario III) with respect to resource and interference competition (see main text). Our experiments were performed as triplicates and clearly showed that under the same growth conditions, only little variation in experimental outcomes could be observed (figure S3).

Furthermore, we analysed the effect of small growth rate differences on coexistence patterns as described in the following. Only expression of the red fluorescent protein mCherry led to a significant reduction of colony growth rate (figure S1). We obtained the differences between ecological scenarios I, II, and III by using a different assignment of mCherry to our three strains. We complemented these three scenarios by exchanging the assignment of GFP between the two strains not expressing mCherry. Expression of GFP resulted in only small changes in growth rate and lag time (figures S1 and S2). In this way, we obtained the new auxiliary scenarios I*, II*, and III*, where I/I*, II/II*, and III/III* are pairs of scenarios in which the same strain expressed mCherry. Similar coexistence patterns as observed for scenarios I-III were observed for the new scenarios I*-III*. A stable coexistence of strains R and C was observed in cyclic scenario I*, and survival of only strain R in both the hierarchical scenario II* and in the intermediate scenario III* (figure S3). This finding demonstrated the robustness of our experimental results for scenarios I-III. Furthermore, the analysis confirmed that no plasmid loss or spontaneous mutations had occurred during our experiments, effects that might have affected coexistence patterns due to a loss of fluorescence. Hence, small variations in growth rates did not affect our essential experimental outcomes (figure S3).

We next assessed the impact of overall growth conditions on coexistence patterns. We compared the data obtained for slow growth conditions to data obtained for fast growth conditions (figure S4; see *Materials and Methods* for details). Under slow growth conditions (S) (growth rate of 42 μm/h averaged over all strain/marker combinations; cf. table S1), the resistant strain R outcompeted the S and the C strain, with the exception of cyclic scenarios I and I* for which a stable coexistence of R and C was observed (over the experimental time frame of 48 h). This finding agreed with our expectations based on experimentally observed growth rates and lag times. Fast growth conditions (F) (average growth rate of 46 μm/h) led to a significant

increase of colony sizes, but similar segregation patterns as for slow growth conditions were observed. This observation demonstrated the robustness of our findings with respect to environmental conditions. Only cyclic scenario I* differed by a significantly reduced amount of the C strain at the edge of colonies. The R and the C strain had similar growth rates in this scenario (figure S1), but differed in their lag times (figure S2). Under slow growth conditions, the C strain was able to establish a stable coexistence with R, despite its longer lag time. However, this coexistence was reduced under fast growth conditions, and the C strain was only present in a few sectors at the colony front. Stable coexistence between R and C could be established in cyclic scenario I under fast growth conditions.

1.1.3. Influence of the Initial Ratio of the C Strain

To study the influence of the toxin colicin on strain coexistence, we repeated the experiments under slow growth conditions but with different initial ratios of the C strain. At reduced initial ratios of the C strain, less colicin was present in the system. We compared our data obtained for equal initial ratios (S:R:C = 1:1:1) with data for the ratios 1:1:0.5 and 1:1:0.1 (figure S5). Experiments with initial ratios of 1:1:0.5 showed little qualitative differences as compared to results from experiments with ratios 1:1:1, in accordance with our theoretical findings (figure S6). For the cyclic scenarios I and I*, we observed a shift of sector sizes towards larger sectors of the R strain. This finding was captured by simulations (figure 3 in the main text). A further reduction of the initial C strain ratio to 0.1 resulted in a qualitatively different outcome: a transient coexistence of all three strains was observed in cyclic scenarios I and I*, although less pronounced in the latter (figure S5). This finding indicated the presence of a threshold: below a specific colicin concentration, coexistence of all strains was possible in cyclic scenarios, while above this threshold, coexistence of only two strains, namely R and C, could be observed. The outcomes of the strictly hierarchical scenarios II and II* remained the same as compared to the 1:1:1 experiments, whereas traces of S could be detected at the edge of colonies in the intermediate scenarios III and III*. Therefore, it seemed that the C strain's initial ratio had no effect on strictly hierarchical ecological systems, and that it had only a small influence on systems whose competition network is intermediate (as defined in figure 1).

Subsequently, we analysed whether under fast growth conditions, reduction of the C strain's initial ratio also had an effect in comparison to results for equal initial ratios (figures S4 and S7). Again, similar results as for slow growth conditions were detected, with an increased presence of the S strain in cyclic scenarios I and I*, and in intermediate scenarios III and III*. However, the pronounced coexistence of all three strains observed in cyclic scenario I under slow growth conditions was not present under fast growth conditions. In particular, after a short phase of coexistence of the R and the S strain, S was outgrown by the R strain. Furthermore, C strain sectors arose at a certain distance from the inoculum, resulting in the eventual coexistence of R and C at the colony front. This observation indicated that a stable equilibrium between resource and interference competition could only be established under slow growth conditions. As this finding could not be observed in our theoretical simulations (see section 2), we attributed it to additional factors on the microscopic level, such as the presence of multiple bacterial layers that are not captured by the theoretical modelling. However, both experiment and theory indicated that under fast growth conditions, three-strain coexistence was principally possible in the cyclic scenario I* (figure S7).

1.2. Analysis of Experimentally Derived Parameters

Bacterial growth in liquid medium is well studied and can be divided into six phases [1]. In the first phase, called the lag phase, bacteria adapt to their new environment and to new nutrient conditions. Second, growth rate increases during the acceleration phase until the third phase, the exponential phase, is reached. Here, growth conditions are optimal and growth rate maximal. Afterwards, growth decelerates during the fourth phase, called retardation phase. In the fifth phase, growth ceases. This stage is called the stationary phase. Unfavourable conditions such as nutrient limitation or toxin accumulation may be the reason. Finally, death of bacteria may cause a negative growth rate and, thereby, the sixth phase, where growth declines.

Bacteria that form colonies on solid surfaces undergo the same phases of life as described above. However, one must distinguish between the growth of single bacteria within the colony (microscopic level) and the growth of the colony itself (macroscopic level). In our experiments, 1 µl droplets of bacterial mixture were inoculated on agar surface. These droplets of optical density 0.1 partly evaporated and left a single layer of cells on the surface. The layer covered only part of the inoculation area. When the first bacteria in such an inoculation area have overcome their lag phase, these bacteria start to multiply and fill out free space in the area (acceleration phase on the microscopic level). Growth in the vertical direction occurs as well, until the macroscopic colony has reached a height of approximately 200 µm [2]. Both the filling out of free space and the growth in the vertical direction sum up to the macroscopic lag phase of a colony. After this macroscopic lag phase has passed, a significant two-dimensional expansion can first be observed (acceleration phase on the macroscopic level). However, the bacterial colony does not reach the phase of exponential growth, but a phase of linear growth (if growth conditions are good). In the centre of the colony, several hundred bacterial layers can be present and most cells enter the stationary growth phase. The number of bacterial layers decreases towards the expanding front of the colony [3] where cells are still in the exponential phase. If growth conditions worsen, both bacterial cells at the front of the colony, and the colony itself, may enter the retardation and stationary phases.

Bacterial colony growth was monitored over time using an upright microscope and appropriate filter sets as described in *Materials and Methods* (main text). The sizes of colonies were measured in terms of their two-dimensional areas (in units mm^2). The data were obtained via thresholding, using analysis tools offered by the freeware ImageJ. The determined areas were converted to effective colony radii $r = \sqrt{A/\pi}$ (figure S8*a*). The resulting radial growth curves were used to obtain radial growth rates and lag times of single-strain colonies. For this purpose, we fitted the growt curves to the sigmoidal five-parameter logistic function (figure S8*a*):

$$r(t) = a + \frac{r_0 - a}{(1 + (t/c)^b)^d} \quad . \tag{1}$$

This function provided a good approximation of all our data sets and made it easy to formulate analytic expressions for characteristic observables (e.g. for the growth rate). We used the software Mathematica 8.0.1 by Wolfram Research to find best-fit parameters (based on a Levenberg-Marquardt algorithm). We defined the growth rate µ of colonies as the maximal slope of the logistic fit (measured in µm/h). For the lag time τ (measured in h), we adapted a definition of Lodge and Hinshelwood [4] as stated in the classic review of Monod on growth in liquid culture [1] to the expansion

of colonies. We defined τ as the difference between the time when a colony reached the radius r_μ where its radial growth rate is maximal (i.e. equal to μ), and the time at which an ideal colony (no lag phase, radius that increases linearly in time) expanding at growth rate μ would reach radius r_μ (figure S8*a*). When the radial growth curve of the ideal colony is shifted by the lag time τ, it touches the real colony's growth curve tangentially at the radius r_μ. The values of the growth rate μ and of the lag time τ depend on the specific strain, on the expressed fluorescent protein, and on the growth condition that is considered. They were, therefore, determined for each combination of strain, fluorescent marker and growth condition. Mean values that were each averaged over three experimental data sets are listed in table S1.

2. Theoretical Part

In the following, we introduce a mathematical model that was used to reproduce our experimental observations of bacterial segregation patterns that arose in range expansions of competing sensitive (S), resistant (R), and colicin producing (C) *Escherichia coli* strains. Maintenance and extinction of bacterial competitors was found to depend on the strains' growth properties, as well as on the colicin condition. The latter comprises the fraction of C cells producing and releasing colicin, the effective diffusion range of colicin, and the rate at which sensitive cells become growth inhibited (and are subsequently lysed) due to colicin uptake. Our theoretical model reproduced essential features of spatial segregation patterns that we experimentally observed in competitive range expansions. The model correctly predicted the surviving strains for the different ecological scenarios discussed in the main text. Our mesoscopic description of bacterial dynamics may serve as reference for future research on the coexistence of bacterial strains in expanding populations.

2.1. Experimentally Motivated Simulation Parameters

Our theoretical model coarse-grained the microscopic bacterial dynamics and gives a mesoscopic description of competitive range expansions. It was based on a set of parameters that we found to be essential for reproducing experimentally observed segregation patterns. The parameters governing the radial expansion of simulated colonies were called "mesoscopic growth rate" μ_m and "mesoscopic lag time" τ_m. They were determined by comparing experimentally determined growth curves and simulated growth curves as described in the next section. The parameter τ_m was considered a Gaussian random variable with standard deviation $\Delta\tau_m$. As discussed in the next section, it enabled us to implement an effective time dependence of the mesoscopic growth rate μ_m, and, as a consequence, to model the time dependence of colony growth rates (during the lag, acceleration and linear growth phases). In order to calibrate the parameters, we introduced an auxiliary observable for technical purposes. We used the logistic fit (1) to measure the times over which experimentally determined growth curves exhibited positive curvature. We refer to the corresponding parameter as "time frame" Δt. In order to obtain numerical estimates for Δt, we evaluated the curvature of experimentally determined growth curves. We defined Δt as the distance between which the curvature was equal to its maximal value, times a factor of exp(-0.5). We made the observation that regimes of positive curvature could be well approximated by Gaussian functions of the form (figure S8*b*):

$$M * \exp\left(-\frac{(t-t_M)^2}{2*(\Delta t/2)^2}\right) \qquad (2)$$

(with maximal curvature M and time of maximal curvature t_M). Numerical estimates of Δt were averaged over values obtained from three independent experiments (for each combination of strain, marker, and growth condition). In the next section, we describe how Δt was used to initialise the standard deviation $\Delta\tau_m$ of the random variable τ_m.

Our theoretical model required two additional parameters to incorporate the toxin interaction between the colicin producing and the sensitive strain. We captured the production and diffusion of colicin, as well as the rate at which sensitive cells became growth inhibited due to colicin uptake, in terms of an effective interaction length scale λ and an effective interaction strength κ. These parameters were adjusted to reproduce qualitative features that we saw in experiments on competing single-strain S and C colonies, and also comply with estimates from literature [5, 6]. Further details are given in the next section.

2.2. Modelling of Single-Strain Colony Growth

The following two sections propose a theoretical model that reproduced bacterial segregation patterns seen in competitive range expansions of sensitive (S), resistant (R), and colicin producing (C) *E.coli* strains. We combined previous coarse-grained approaches to the stochastic modelling of segregation patterns of two non-interacting bacterial strains [7, 8] with models of the interference competition between sensitive and toxic strains [9, 10]. We explicitly accounted for the long-range nature of the toxin interaction between S and C cells, with allelopathic effects being perceptible over distances larger than 100 µm [5, 6]. Our theoretical approach modelled the toxin interaction via colicin on the basis of stationary solutions of source and degradation processes as explained in the next section. Furthermore, the mesoscopic lag times already introduced above enabled us to reproduce experimentally determined growth curves of single-strain colonies. This aspect of the model will be discussed in the present section. We emphasise that the values of all simulation parameters were motivated by experimental data obtained from either the growth of single-strain colonies, or from competitions of S and C strain colonies. They also comply with literature estimates [5, 6] on the range of toxins.

We assumed that bacterial dispersal to uncolonised area was dominated by steric exclusion following cell division, and that dynamics inside colonised domains were negligible. Spatial segregation patterns that arose during the expansion of bacterial colonies were frozen in time over the course of experimental observation. Considering the high number of bacterial cells that are already present in droplets of inoculum from liquid culture (on the order of 10^6-10^7 cells), we resorted to a mesoscopic description of the growth dynamics. Fast simulations were needed for both the adjustment of simulation parameters, and for the computation of extrapolating diagrams as shown in figure 5 (main text). We chose a stochastic lattice-based description with 750x750 lattice sites (see table S1 for parameters used in the simulation). The number of lattice sites was a compromise between a low impact of small number fluctuations on the results of simulations, and the possibility both to calibrate our model, and to perform statistical analyses (figure 5 in the main text). It became increasingly difficult to achieve sufficient sample sizes when the number of lattice sites was increased (mainly due to the explicit modelling of the long-range toxin interaction described below). However, our prediction of transient three-strain coexistence in cyclic scenario I around a reduced C strain ratio of r_C=0.1 (figure 3 in the main text) remained valid for lattices with 1500x1500 sites. The

images of simulated bacterial colonies presented in our study stem from simulations using triangular lattices (lattice sites with 6 nearest neighbours). Only negligibly small changes were observed when using square lattices with Moore neighbourhoods (lattice sites with 8 nearest neighbours; growth rates along diagonal directions were reduced by a factor of $1/\sqrt{2}$. The length scale of our model was fixed by interpreting the lattice's side length as 7.14 mm of physical length. Each lattice site carried an integer valued simulation parameter to denote the locally dominant strain of either sensitive (S), resistant (R), or colicin producing (C) *E.coli* cells, or to denote uncolonised area (agar site A). In the following, we occasionally refer to S, R, or C lattice sites by which we mean the associated parameter that described the locally dominant strain. We did not consider dynamics on sub-lattice scales but only between different lattice sites. As only around 3% of bacterial C cells actually produce and emit colicins [11], our theoretical model needed to distinguish between lattice sites dominated by non-toxic wild-type cells (C^{wild}) or by toxic cells (C^{toxic}). Colicin-producing *E.coli* cells lyse upon the emission of colicins, but we did not find it necessary to consider lattice sites dominated by lysed C cells in order to reproduce bacterial segregation patterns. However, sensitive *E.coli* cells suffer DNA degradation and growth inhibition upon the uptake of diffusing colicin, leading to subsequent lysis. Our model accounted for the growth inhibition by distinguishing between lattice sites dominated by active (S^{active}) or by growth inhibited cells ($S^{inhib.}$). The mesoscopic bacterial colonisation of agar sites by adjacent colonised sites is described by the following interaction scheme:

$$
\begin{array}{lcllcll}
C^{wild} & + & A & \rightarrow & C^{wild} & + & C^{wild} \quad (0.97\ \mu_{m,C}) \\
C^{wild} & + & A & \rightarrow & C^{wild} & + & C^{toxic} \quad (0.03\ \mu_{m,C}) \\
R & + & A & \rightarrow & R & + & R \quad (\mu_{m,R}) \\
S^{active} & + & A & \rightarrow & S^{active} & + & S^{active} \quad (\mu_{m,S})
\end{array}
$$

These processes occured at the mesoscopic growth rates listed in brackets. As apparent from these interactions, we assumed that 3% of agar lattice sites colonised by adjacent C^{wild} sites turned out to be dominated by toxic cells. Toxic C sites could not proliferate to adjacent agar sites because colicin-emitting cells are lysed during the emission [11]. Lattice sites that were dominated by growth inhibited S cells were also excluded from proliferation. Lattice sites could not be recolonised. Toxic C sites were depicted in slightly lighter colours than wild C sites in pictures of simulated bacterial colonies, inhibited S sites in slightly darker colours than active S sites.

"Chemical reactions" of the above form can be directly translated into stochastic Master equations and can be evaluated with the help of the Gillespie algorithm [12]. Coexistence diagrams as shown in figure 5 (main text) required more than 10^6 individual realizations of the expansion process (statistics over 250 runs for each of the 75x75 parameter points). In order to generate these diagrams, an efficient implementation of the Gillespie algorithm was needed (we used a binary reaction rate tree [13]). We decided for the use of the Gillespie algorithm because it preserves the correct temporal dynamics of stochastic processes. Correct temporal dynamics were important for the inclusion of mesoscopic lag times and of the toxin interaction. The latter acted on a multitude of time scales, depending on the distance between sensitive and toxic lattice sites.

As explained in section 1.2, only a part of inoculation areas was initially filled with bacteria. The remaining free space was colonised during the acceleration phase of

growth after bacteria awoke from their lag phase. We assumed that *in silico* inoculation areas were disc-shaped with radii of 1.5 mm and approximated that 25% of the areas were initially filled with bacteria. The reduced initial density enabled us to capture patterns that emerged inside the inoculation area of colonies. Given an initial strain ratio of $r_S:r_R:r_C$, we initialised lattice sites within a distance of 1.5 mm of the simulation lattice's centre to be dominated by either S, R, C^{wild}, or C^{toxic} with probabilities $0.25 * (r_S, r_R, 0.97*r_C, 0.03*r_C) / (r_S+r_R+r_C)$, respectively. The remaining vacant lattice sites were interpreted as agar. Simulations were stopped after a simulated colony had reached a radius of 3 mm. Our model did not consider that experimentally recorded growth curves levelled off at later times (retardation phase). This effect was due to dehydration of agar and possibly also nutrient depletion (cf. figure S9). However, the effect did not influence bacterial segregation patterns because these were mainly shaped during early growth phases.

In our experiments, droplets of inoculum were taken from overnight culture in which the stationary growth phase had been reached. After inoculation, bacterial colonies required a certain time before lateral growth could be detected (figure S8*a*). The length of this time period differed from strain to strain and also depended on the expression of fluorescent markers and on the overall growth condition (figure S2). We reproduced the time dependence of radial colony growth during lag and acceleration phase with the help of the mesoscopic lag time τ_m in the following way. We required that the colonisation of agar lattice sites by adjacent colonised sites could only occur after a certain time τ_m had passed. The time differed from site to site and was drawn from a strain and marker dependent Gaussian distribution $p(\tau_m)$ (we use the same letter τ_m to denote the mean and the samples of the distribution, and also the random variable itself). The calibration of the distribution is described below. Different samples of τ_m on individual lattice sites introduced an effective time dependence of the local mesoscopic growth rate μ_m (it was zero on a lattice site before its lag time τ_m had passed). Such a time dependence of reaction rates is commonly beyond the scope of the classical Gillespie algorithm and extensions of the algorithm [14, 15] are numerically inadequate for large spatial systems. However, our (local) mesoscopic growth rates were simple piecewise constant functions of time and could be modelled by implementing a "stochastic clock". This clock was a stochastic process that was executed at a high ticking rate, but did not change the system's state. In this manner, we enforced frequent updates of the Gillespie algorithm's event driven simulation time. After the simulation time had exceeded the mesoscopic lag time on a lattice site, the growth rate on that lattice site was set to its finite, strain and marker dependent value μ_m. Without the stochastic clock, it could have occurred that a local growth rate remained zero for too long. A high ticking rate of the stochastic clock ensured that time increments were small and reduced numerical errors due to deprecated zero growth rates. This way of incorporating time dependent growth rates incurred computational costs due to a larger number of exponentially distributed time increments that were needed to reach a specific simulation time. However, we found that these costs were negligible as compared to costs that arose from the Gillespie algorithm's search of reaction channels. Furthermore, the stochastic clock was deactivated after all mesoscopic lag times had passed.

We were now left with three parameters that needed to be adjusted to recover experimentally determined single-strain growth curves: the mesoscopic growth rate μ_m, the mean τ_m of the mesoscopic lag time's Gaussian distribution $p(\tau_m)$, and

the distribution's standard deviation $\Delta\tau_m$. In order to reproduce the lag and acceleration phases of experimentally determined growth curves, we found it necessary to choose $\Delta\tau_m$ to be sufficiently large. We set the value of $\Delta\tau_m$ to the value of the time frame Δt (depending on the specific strain/marker combination). This time frame measured the range over which experimentally determined growth curves of single-strain colonies exhibited positive curvature (see section 2.1).

As the next step, we adjusted μ_m and τ_m. For this purpose, we performed simulations for a wide range of values of these parameters and computed macroscopic growth rates μ and lag times τ from the resulting simulated growth curves (cf. section 1.2). By means of polynomial interpolation, we determined functional relations that connected μ and τ to μ_m and τ_m. By numerically inverting these relations, we determined values of μ_m and τ_m such that the resulting μ and τ matched values that had been determined from our experimental data (cf. section 1.2). We thereby obtained simulation parameters that made it possible to closely approximate experimentally determined growth curves. We compare such growth curves to averaged simulated growth curves in figure S9*a* (expansion of R_{mCh} colonies; fast growth conditions).

We also tested whether a model with fixed (i.e. not randomly distributed), strain/marker dependent mesoscopic lag times could approximate experimentally determined growth curves and segregation patterns. For this purpose, we assumed equal lag times τ_m for every lattice site initially colonised by a specific strain and expressed marker. Averaged simulated growth curves based on this model approximated experimental data reasonably well, with only small deviations during the acceleration phase (figure S9*b*). These differences originated from the simulated colonies' abrupt radial expansion after the mesoscopic lag time had passed (simultaneously on every lattice site). However, we did not manage to reproduce experimentally observed segregation patterns with this model (after including the toxin interaction described in the next section). At equal initial strain ratios, the R strain was found to be the only survivor in all ecological scenarios (figure S10*a*). Due to the short lag time of R as compared to that of C (cf. figure S2), and also due to growth inhibition of S by colicin, lattice sites dominated by the R strain rapidly multiplied and barred other strains from uncolonised sites. At a low initial ratio of the C strain, at most two-strain coexistence of species was observed using this model.

Furthermore, we tested a model that neglected the lag phase. For that purpose, we used the mesoscopic growth rates determined for the model with fixed, strain/marker dependent mesoscopic lag times, but subsequently set all lag times to zero. We thereby lose the possibility to reproduce experimentally observed growth curves. However, it turned out that the reduced model was able to reproduce bacterial segregation patterns seen in experiments. For cyclic scenario I and equal initial strain ratios, experimental and theoretical patterns matched even closer. A larger number of C_{nfp} sectors was observed (figure S10). The reduced model led to similar predictions on biodiversity zones, with three-strain coexistence around the reduced initial C strain ratio of r_C=0.1 in cyclic scenario I. However, our original model based on Gaussian distributed mesoscopic lag times could reproduce both experimentally determined growth curves and predict the correct survivors of range expansions. The broadness of Gaussian distributions in this model significantly overlapped for different strain/marker combinations. This aspect appeared to be important for achieving the latter agreement.

2.3. Modelling of the Toxin Interaction

The following section complements our model of single-strain colony growth with a theoretical approach to the toxin interaction between sensitive and toxic bacterial cells. In a schematic way, the interaction can be written as:

$$C^{toxic} + S \rightarrow C^{toxic} + S^{inhib.} .$$

Since the size of colicin lies on the order of a few nanometres (60 kDa [11]) and is, therefore, a factor 1000 smaller than individual *E.coli* cells, we believed an effective mesoscopic theory to be more suitable than a microscopic approach. In the experimental setting, colicin accumulates inside of toxic C cells and is emitted once lysis proteins trigger the carrier's cell death [11]. After colicins are released, they diffuse through the extracellular fluid surrounding bacterial cells until possibly being absorbed by sensitive *E.coli*. The sensitive cells are prone to the endonuclease activity of colicin and suffer a degeneration of their DNA. Thereby, they loose the ability to undergo cell divisions [16]. Eventually, these cells lyse. We resorted to an effective description of the process and neglected the time dependence of the emission and diffusion of colicins. This left us with a source and degradation process in which toxic C cells acted as sources of colicin. The diffusion constant of colicin is on the order of 10^{-7} cm^2/s [17]. Hence, the diffusion of colicin happens on a much faster time scale than the time scale on which *E.coli* undergo cell division (approximately 1 h). Therefore, we assumed that sources and sinks in our model were essentially static in space. It followed from the source and degradation process that toxin producing C cells were surrounded by quasi-static, exponentially decreasing colicin concentration profiles [18, 19]. Therefore, we incorporated the toxin interaction on the mesoscopic scale of our theoretical model by assuming exponentially decreasing colicin profiles $\kappa * \exp(-x/\lambda)$ around toxic C sites, with approximated characteristic scale λ = 125 µm, and approximated strength κ = 3 h^{-1}. The approach was robust against changes in the colicin concentration profiles, as long as growth inhibition of sensitive sites acted strongly around sources, and concentration profiles fell off over distances between λ = 100–175 µm (in consistency with literature values [5, 6]; note that these references give maximal colicin ranges, not characteristic length scales). The concentration profiles defined the rates at which sensitive lattice sites at a distance x from a toxic C site absorbed colicin and became growth inhibited. Our results depended only weakly on κ for values between 1.5 h^{-1} and 3.5 h^{-1}. For larger lattices, the value of κ had to be decreased in order to compensate for a larger number of toxic lattice sites (note that sites of our simulation lattice do not correspond to individual bacteria). The parameters chosen for our model made it possible to recover qualitative features that had been observed during the competition of neighbouring single-strain C_{nfp} and S_{mCh} colonies (figure S11).

2.4. Comparison of Segregation Patterns

The previous two sections outline how the parameters of our theoretical model of range expansions were adjusted to reproduce experimentally determined growth curves of single-strain colonies. It also describes the way in which we modelled the allelopathic interaction between toxic and sensitive *E.coli*. Combining these two aspects of our model, and initializing inoculation areas with multiple bacterial strains made it possible to recover bacterial segregation patterns seen in experiments. As explained in the main text, our first theoretical intent was to test whether our fully calibrated model was able to reproduce patterns observed at equal initial strain ratios (S:R:C = 1:1:1). Figures S3 and S4 give a comparison between experimental and

simulated colonies for these ratios. These two figures exemplify experimentally observed and simulated bacterial colonies for the ecological scenarios I-III and for the auxiliary scenarios I*-III*. The auxiliary scenarios were obtained by exchanging the strains expressing nfp and GFP (see section 1.1.2). We tested the stability of segregation patterns that had been observed for slow growth conditions (figure S3) by performing experiments and simulations for fast growth conditions (figure S4). Our theoretical model recovered the essential features of expanding colonies that we saw in competitive range expansions of S, R, and C at equal initial ratios. It correctly predicted the strains that populated the fronts of colonies. Even subtle details such as the strains dominating the areas of inoculation showed agreement between experiment and theory for most of the scenarios. However, our simulation could not capture the fine and evenly distributed sectors formed by the R and C strains in the cyclic scenarios I and I*. We attributed this observation to the fact that our simulation was limited to a two-dimensional lattice, whereas natural colonies grew in several bacterial layers. Our theoretical model could not account for layer-dependent alterations of growth parameters.

Our experimental and theoretical studies agreed on the observation that for equal initial strain ratios, at most two strains survived a competitive range expansion. One of these two was always the R strain, irrespective of the ecological scenario. The R strain was, however, weakened by the expression of mCherry in cyclic scenarios I and I*, such that it had to share the expanding fronts of colonies with the C strain. As outlined in the main text, we asked ourselves whether it was possible to tune initial strain ratios to allow for coexistence of all three strains. Based on the two-strain coexistence observed in cyclic scenario I, we used our theoretical model to study how bacterial segregation patterns changed upon reducing the C strain's initial ratio. Following a transient increase of the width of R strain sectors, we observed qualitative changes around an initial C strain ratio of 0.1 (figure 3 in the main text). Approximately 24% of our simulations showed coexistence of all three strains for this initial ratio, with C strain sectors being the smallest (figure S6). In order to test whether this prediction also held true in experiments, we performed range expansion experiments for cyclic scenario I at initial strain ratios S:R:C = 1:1:0.1. A comparison between figures S5 and S6 showed adequate agreement. It was, furthermore, tested whether the results obtained from our simulations agreed with experimentally observed segregation patterns for the other ecological scenarios (figures S5 and S6), and whether the agreement was robust against changes of overall growth conditions (figure S7). We found agreement between all patterns except for scenario I under fast growth conditions and, to a lesser degree, for scenario I* under slow growth conditions. In the former case, experimentally observed colonies displayed an increased amount of sensitive bacteria around their areas of inoculation, but the sensitive bacteria's radial growth was inhibited and small C strain sectors eventually emerged. A similar phenomenon occurred in the second case (cyclic scenario I*, slow growth conditions) where expanding C strain sectors arose at a certain distance from the area of inoculation and impeded the growth of S (small amounts of S could still be found along the expanding front).

3. Supporting References

1. Monod, J. 1949 The Growth of Bacterial Cultures. *Annu. Rev. Microbiol.* **3**, 371-394. (doi:10.1146/Annurev.Mi.03.100149.002103)

2. Wimpenny, J.W.T. 1979 The growth and form of bacterial colonies. *J. Gen. Microbiol.* **114**, 483-486. (doi:10.1099/00221287-114-2-483)

3. Gómez-Aguado, F., Gomez-Lus, M.L., Corcuera, M.T., Alou, L., Alonso, M.J., Sevillano, D., Val, D., Palmeiro, A., Iglesias, N. & Prieto, J. 2009 Colonial architecture and growth dynamics of Staphylococcus aureus resistant to methicillin. *Rev. Esp. Quim.* **22**, 224-227.

4. Lodge, R.M. & Hinshelwood, C.N. 1943 51. Physicochemical aspects of bacterial growth. Part IX. The lag phase of Bact. lactis aerogenes. *J. Chem. Soc.*, 213-219. (doi:10.1039/JR9430000213)

5. Ozeki, H., Stocker, B.A.D. & De Margerie, H. 1959 Production of colicine by single bacteria. *Nature* **184**, 337-339. (doi:10.1038/184337a0)

6. Chao, L. & Levin, B.R. 1981 Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proc. Natl. Acad. Sci. USA* **78**, 6324-6328. (doi:10.1073/Pnas.78.10.6324)

7. Hallatschek, O. & Nelson, D.R. 2010 Life at the front of an expanding population. *Evolution* **64**, 193-206. (doi:10.1111/J.1558-5646.2009.00809.X)

8. Korolev, K.S., Xavier, J.B., Nelson, D.R. & Foster, K.R. 2011 A quantitative test of population genetics using spatiogenetic patterns in bacterial colonies. *Am. Nat.* **178**, 538-552. (doi:10.1086/661897)

9. Iwasa, Y., Nakamaru, M. & Levin, S.A. 1998 Allelopathy of bacteria in a lattice population: Competition between colicin-sensitive and colicin-producing strains. *Evol. Ecol.* **12**, 785-802. (doi:10.1023/A:1006590431483)

10. Nakamaru, M. & Iwasa, Y. 2000 Competition by allelopathy proceeds in traveling waves: Colicin-immune strain aids colicin-sensitive strain. *Theor. Popul. Biol.* **57**, 131-144. (doi:10.1006/Tpbi.1999.1448)

11. Cascales, E., Buchanan, S.K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S. & Cavard, D. 2007 Colicin biology. *Microbiol. Mol. Biol. R.* **71**, 158-229. (doi:10.1128/Mmbr.00036-06)

12. Gillespie, D.T. 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403-434. (doi:10.1016/0021-9991(76)90041-3)

13. Slepoy, A., Thompson, A.P. & Plimpton, S.J. 2008 A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks. *J. Chem. Phys.* **128**. (doi:10.1063/1.2919546)

14. Prados, A., Brey, J.J. & SanchezRey, B. 1997 A dynamical Monte Carlo algorithm for master equations with time-dependent transition rates. *J. Stat. Phys.* **89**, 709-734. (doi:10.1007/Bf02765541)

15. Lu, T., Volfson, D., Tsimring, L. & Hasty, J. 2004 Cellular growth and division in the Gillespie algorithm. *Syst. Biol.* **1**, 121-128. (doi:10.1049/Sb:20045016)

16. Schaller, K. & Nomura, M. 1976 Colicin-E2 Is a DNA Endonuclease. *Proc. Natl. Acad. Sci. USA* **73**, 3989-3993. (doi:10.1073/Pnas.73.11.3989)

17. Liang, S.M., Xu, J., Weng, L.H., Dai, H.J., Zhang, X.L. & Zhang, L.N. 2006 Protein diffusion in agarose hydrogel in situ measured by improved refractive index method. *J. Control. Release* **115**, 189-196. (doi:10.1016/J.Jconrel.2006.08.006)

18. Kicheva, A., Pantazis, P., Bollenbach, T., Kalaidzidis, Y., Bittig, T., Julicher, F. & Gonzalez-Gaitan, M. 2007 Kinetics of morphogen gradient formation. *Science* **315**, 521-525. (doi:10.1126/Science.1135774)

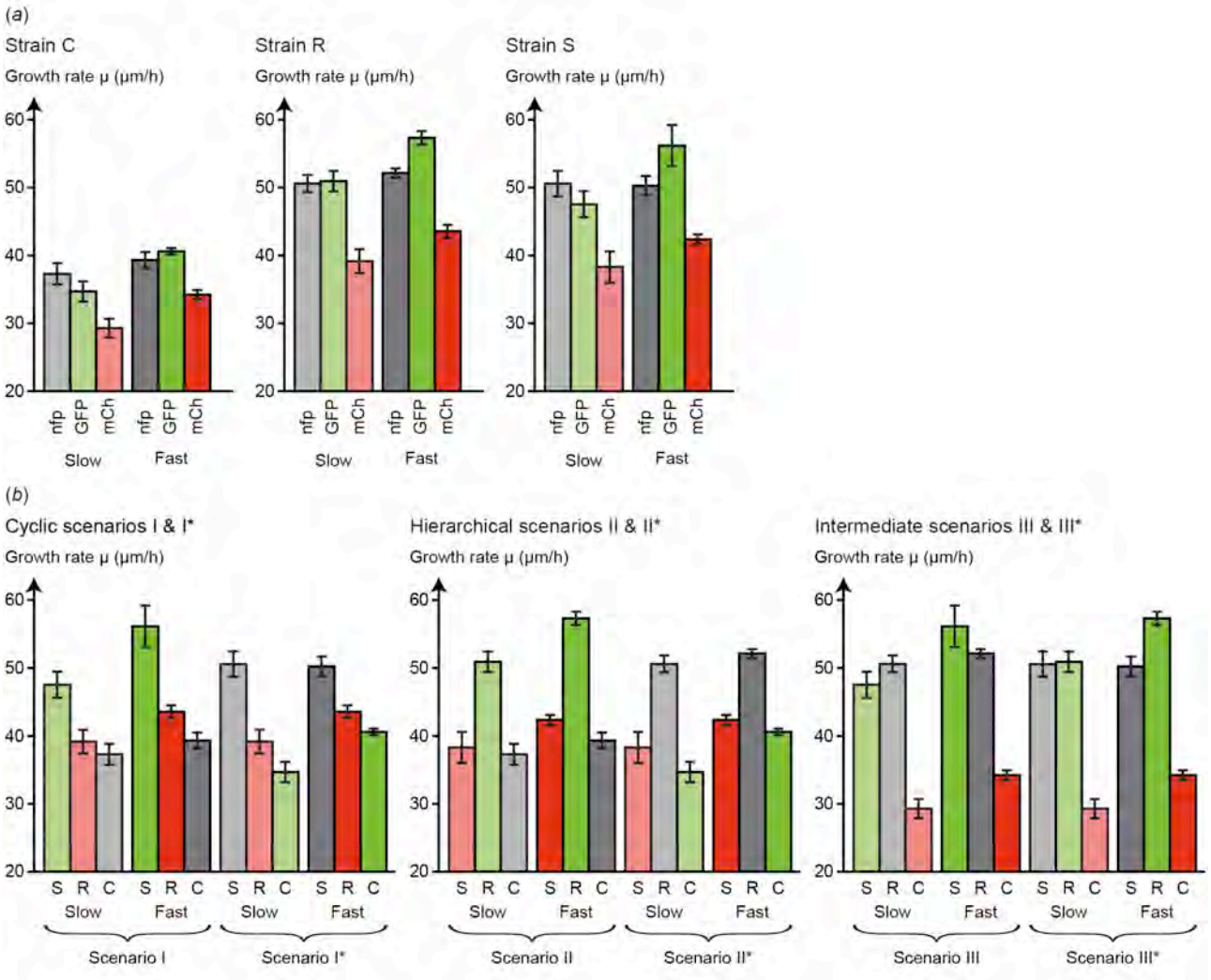
19. Wartlick, O., Kicheva, A. & Gonzalez-Gaitan, M. 2009 Morphogen gradient formation. *Cold Spring Harb. Perspect. Biol.* **1**. (doi:10.1101/cshperspect.a001255)

Table S1. Simulation parameters and data on single-strain colony growth.

Dimensions of the lattice							750 x 750		
Physical side length of the lattice							7.14 mm		
Tessellation of the lattice							triangular		
Radius of the initial colony							1.5 mm		
Colony radius at which the simulation is stopped							3.0 mm		
Filling fraction of the inoculation area							0.25		
Fraction of C^{toxic} among initial C sites and among offspring of C^{wild} sites							0.03		
Characteristic length scale λ of the toxin interaction							125 µm		
Local strength κ of the colicin interaction							3 h^{-1}		
Strain and marker dependent observables of expanding single-strain colonies									
	C_{nfp}	C_{GFP}	C_{mCh}	R_{nfp}	R_{GFP}	R_{mCh}	S_{nfp}	S_{GFP}	S_{mCh}
Slow growth conditions (S)									
µ (µm/h)	37.3 (4.2%)	34.7 (4.3%)	29.3 (4.7%)	50.6 (2.5%)	50.9 (2.9%)	39.2 (4.4%)	50.6 (3.7%)	47.6 (4.0%)	38.3 (6.0%)
τ (h)	11.12 (3.7%)	10.71 (2.0%)	10.97 (2.6%)	7.78 (4.3%)	8.19 (2.8%)	8.73 (2.6%)	7.96 (3.7%)	8.20 (3.4%)	8.69 (2.8%)
Δt (h)	3.11 (9.7%)	3.42 (2.9%)	3.43 (6.1%)	2.89 (10.1%)	3.50 (7.5%)	3.01 (6.1%)	3.87 (5.8%)	4.43 (2.4%)	4.24 (7.0%)
Fast growth conditions (F)									
µ (µm/h)	39.3 (3.0%)	40.6 (1.2%)	34.2 (2.2%)	52.1 (1.3%)	57.3 (1.7%)	43.6 (2.2%)	50.3 (2.9%)	56.2 (5.4%)	42.4 (1.7%)
τ (h)	6.94 (2.1%)	8.17 (1.8%)	8.95 (2.4%)	5.33 (2.9%)	5.89 (1.8%)	6.15 (1.8%)	5.39 (2.2%)	5.49 (4.2%)	6.14 (2.3%)
Δt (h)	4.55 (0.9%)	5.02 (1.4%)	5.37 (2.8%)	2.52 (6.0%)	2.28 (3.5%)	3.00 (6.9%)	2.30 (4.2%)	2.75 (11.4%)	3.45 (6.8%)

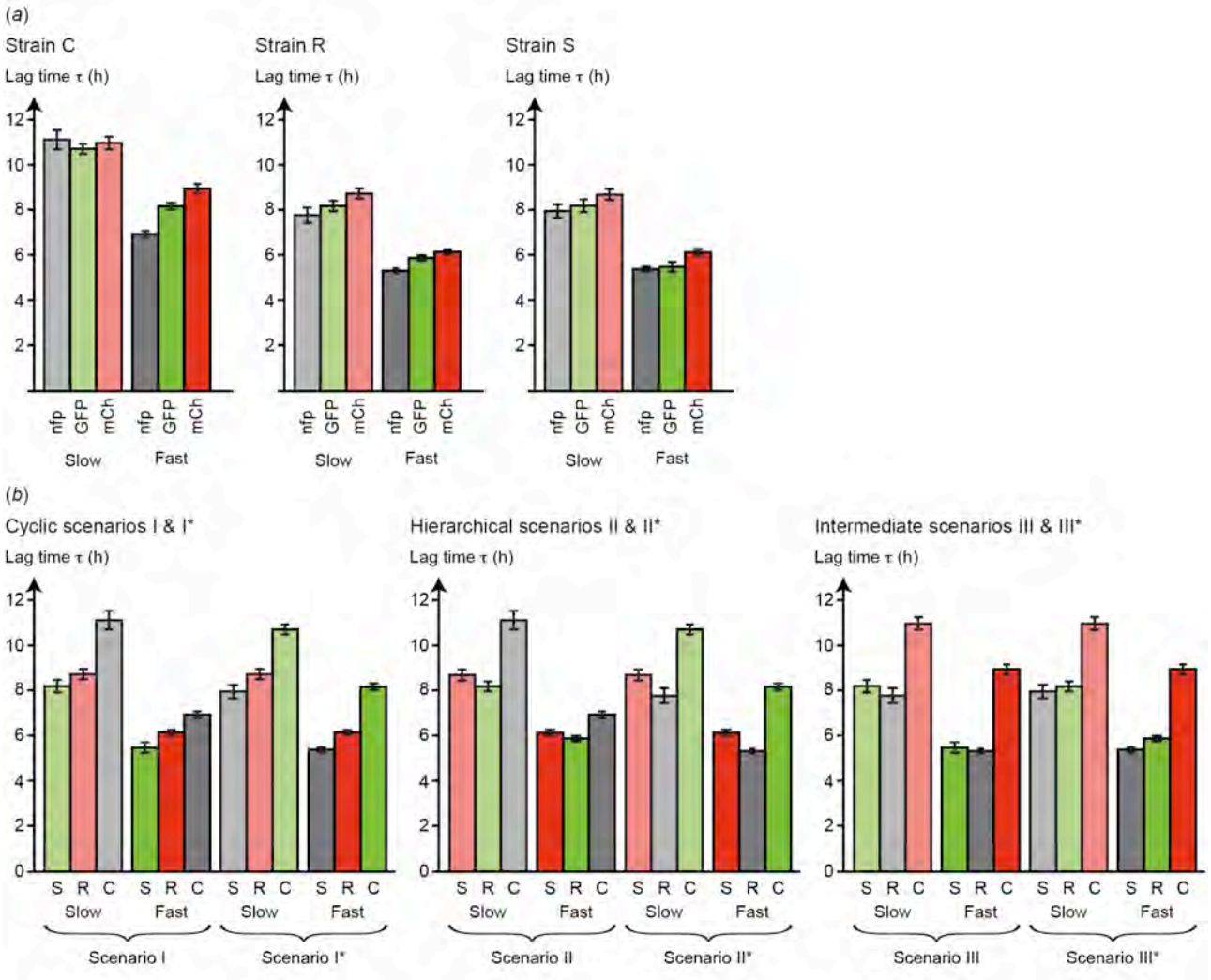
The upper half of the table lists important parameter values of our theoretical model. We refer the reader to section 2 of appendix S1 in which each of these parameters is explained in detail. The lower half of the table lists maximal radial growth rates (µ), lag times (τ), and time frames (Δt) that were obtained from logistic fits of experimental data on the growth of single-strain colonies (each averaged over three independent realizations). Values in brackets represent relative standard errors of the mean (r.s.e.m.). For a definition of parameters, we refer the reader to sections 1.2 and 2.1 in appendix S1. The data were used to calibrate our simulations as explained in section 2.1.

Figure S1. Impact of fluorescent protein expression on single-strain colony growth rates.



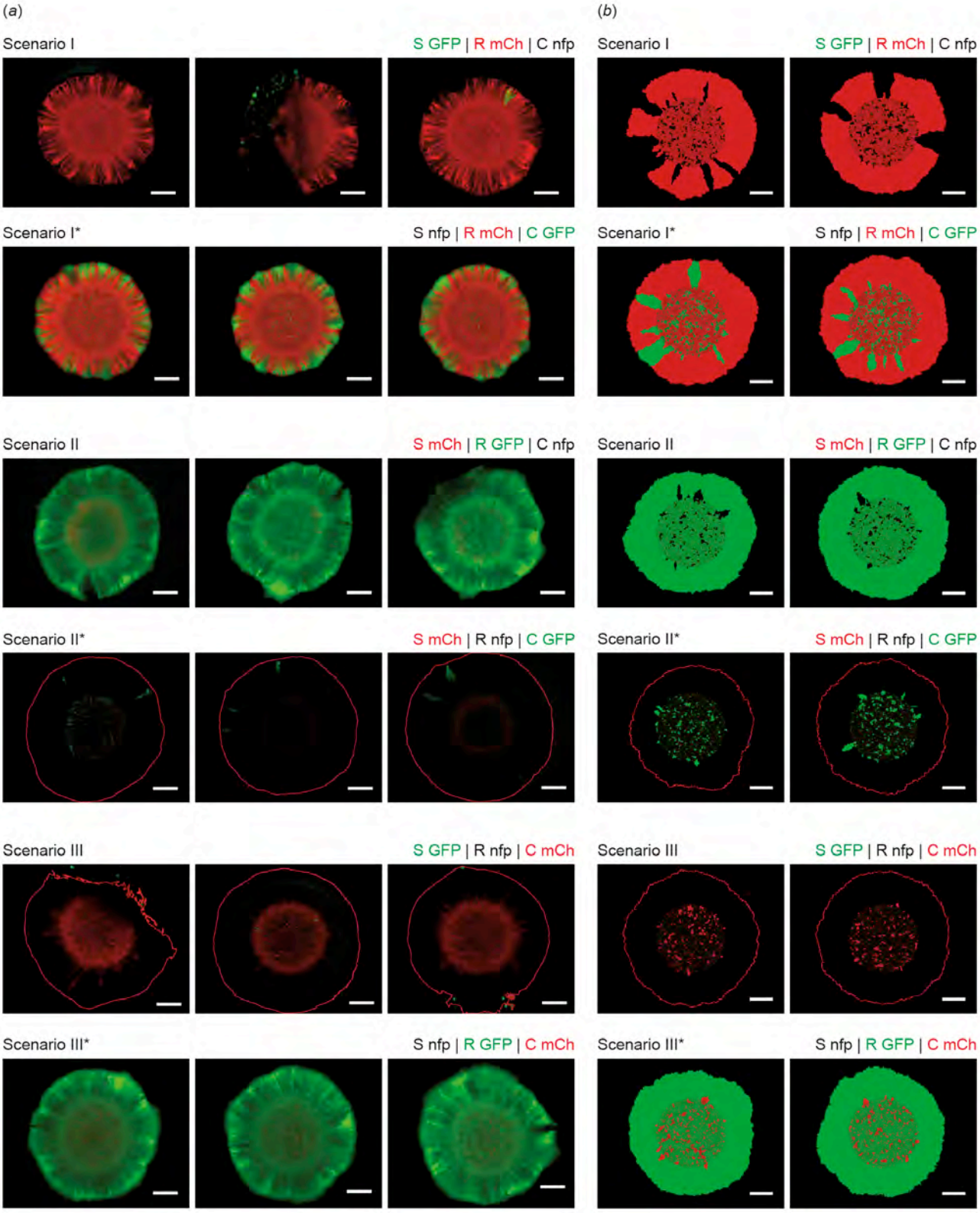
This figure illustrates the growth rates μ (μm/h) of single-strain S, R, and C colonies when cells expressed the green fluorescent protein GFP (green), the red fluorescent protein mCherry (red), or no fluorescent protein (grey) for the different growth conditions slow (light colours) and fast (dark colours). Data represent averages over three different experiments. Error bars give standard errors of the mean (s.e.m.). (*a*) The data are grouped by strain and growth condition. While expression of the green fluorescent protein had only a small effect on bacterial growth rates, they were significantly reduced when the red fluorescent protein mCherry was expressed. Independent of the expressed fluorescent protein, the C strain showed a reduced growth rate as compared to the S and the R strain whose growth rates were of comparable magnitude. The growth rates are also listed in table S1. (*b*) The data are grouped by growth condition and ecological scenario (including the auxiliary scenarios I*-III* introduced in appendix S1, section 1.1.2)

Figure S2. Impact of fluorescent protein expression on single-strain colony lag times.



This figure illustrates the lag times τ (h) of single-strain S, R, and C colonies when cells expressed the green fluorescent protein GFP (green), the red fluorescent protein mCherry (red), or no fluorescent protein (grey) for the different growth conditions slow (light colours) and fast (dark colours). Data represent averages over three different experiments. Error bars give standard errors of the mean (s.e.m.). (*a*) The data are grouped by strain and growth condition. The lag time of the C strain was significantly longer as compared to those of the R and the S strain. Expression of the green or red fluorescent protein resulted in no, or only a small, change in lag time (significant change only for the strain C with fast growth conditions). The lag times are also listed in table S1. (*b*) The data are grouped by growth condition and ecological scenario (including the auxiliary scenarios I*-III* introduced in appendix S1, section 1.1.2)

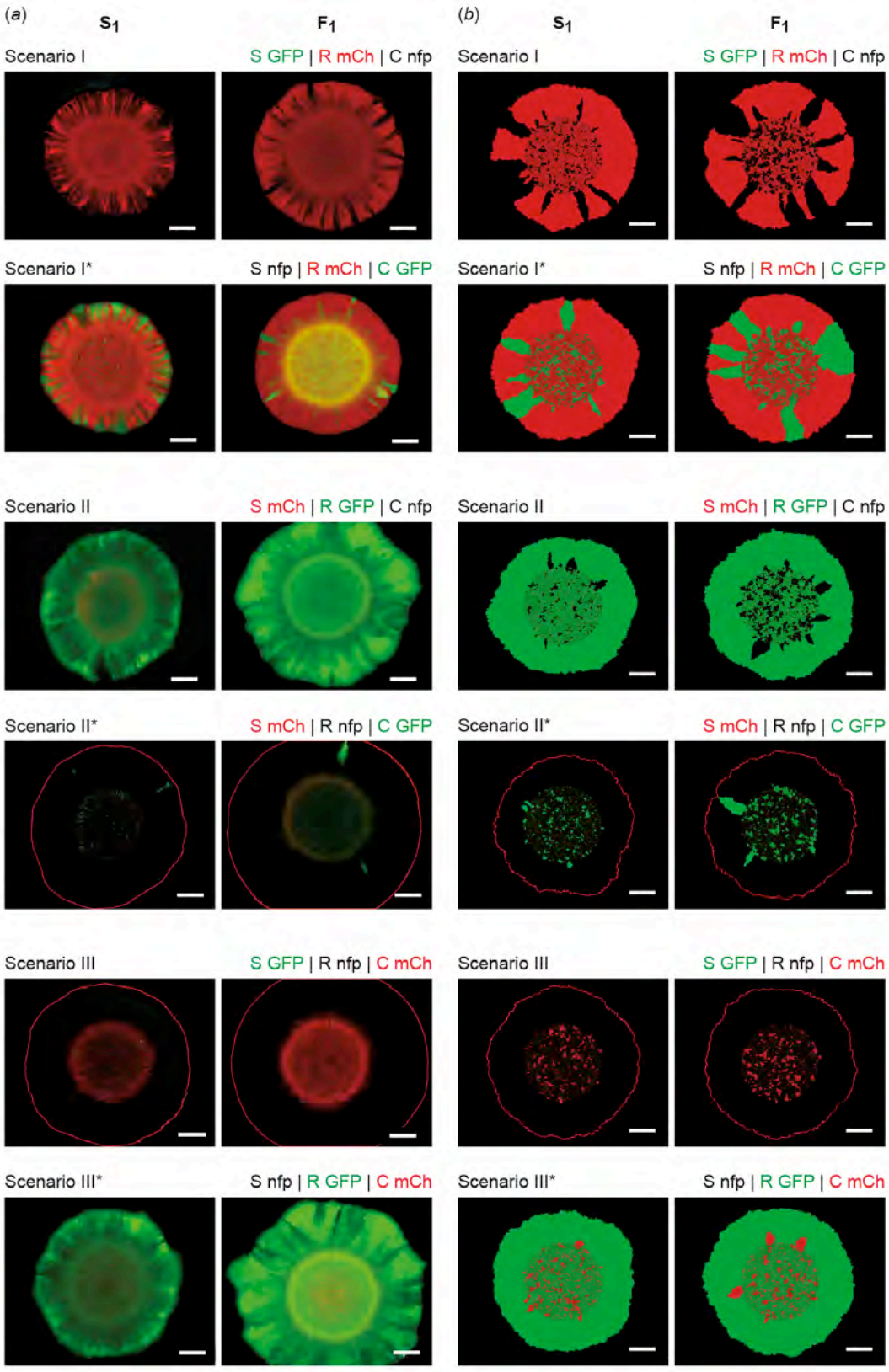
Figure S3. Colony development for slow growth conditions (initial strain ratios S:R:C = 1:1:1, observation after 48 h).



(*a*) Experimental data, (*b*) Simulation results. Triplicates of experimental colonies were analysed for each ecological scenario. The S, R, and C strains expressed fluorescent proteins as depicted above the individual rows: black = nfp, green = GFP, red = mCherry. Red lines depict the edge of the colony in scenarios II* and III. White scale bars represent 1 mm. Scenarios marked by an asterisk (*) are auxiliary scenarios that were

obtained by exchanging the assignment of GFP between the two strains not expressing mCherry (see appendix S1, section 1.1.2).

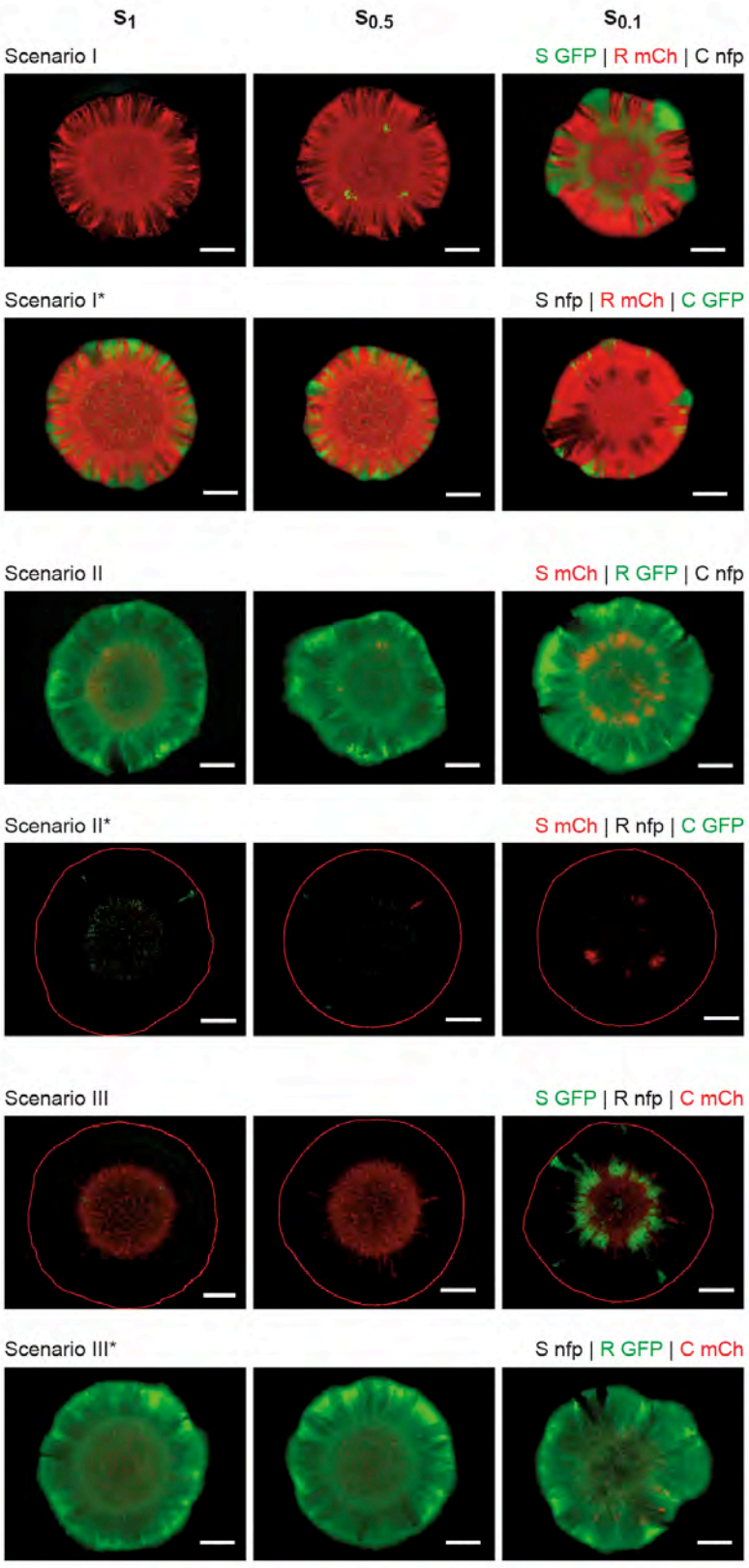
Figure S4. Colony development for slow (S_1) and fast (F_1) growth conditions (initial strain ratios S:R:C = 1:1:1, observation after 48 h).



(*a*) Experimental data, (*b*) Simulation results. We analysed experimental colonies obtained for each ecological scenario. The S, R, and C strains expressed fluorescent proteins as depicted above the individual rows: black = nfp, green = GFP, red = mCh. Red lines depict the edge of the colony in scenarios II* and III. White scale bars represent 1 mm.

Scenarios marked by an asterisk (*) are auxiliary scenarios that were obtained by exchanging the assignment of GFP between the two strains not expressing mCherry (see appendix S1, section 1.1.2).

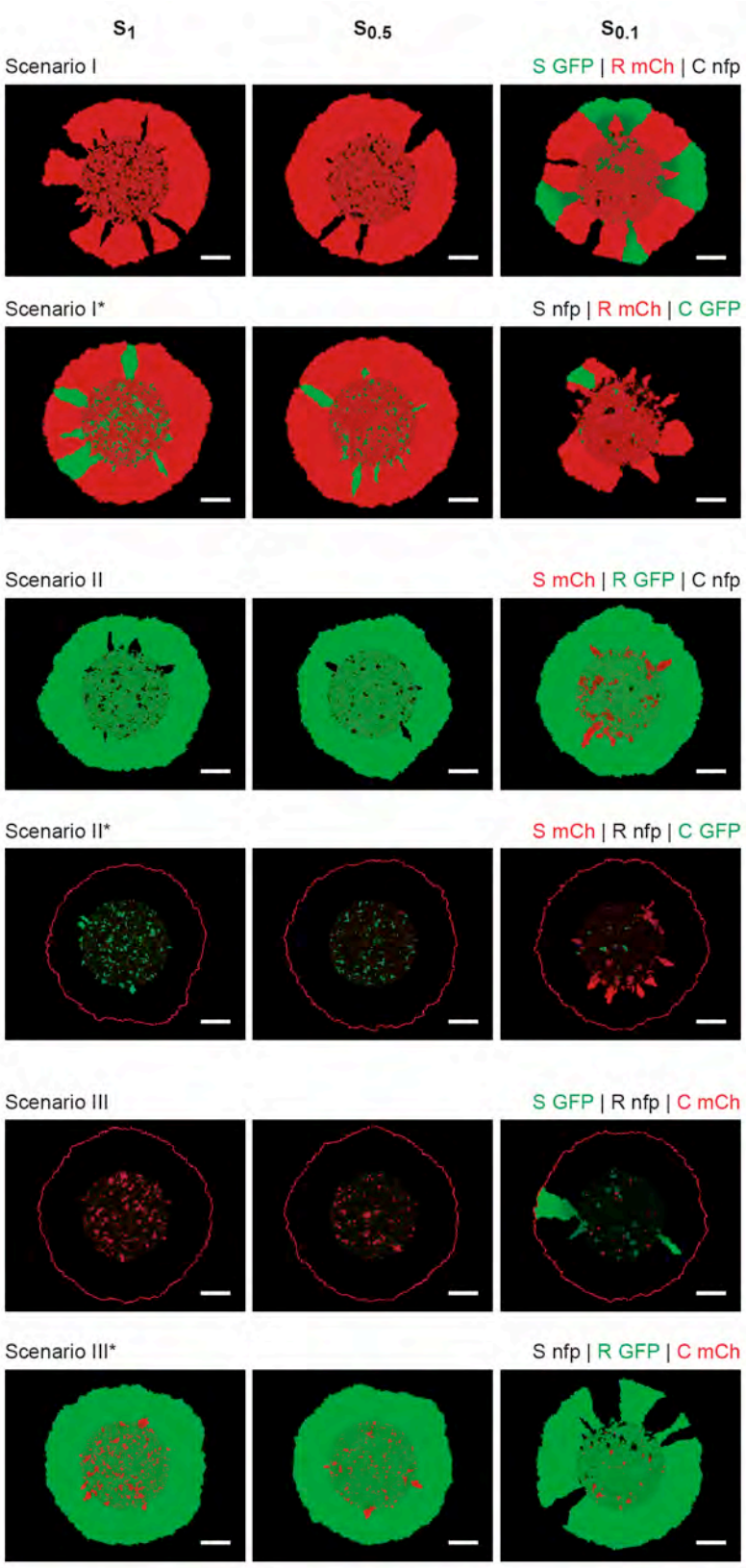
Figure S5. Colony development for different initial ratios of the C strain (experimental data, observation after 48 h).



For slow growth conditions, different initial ratios of the colicin producing strain C were added at the time of inoculation t_0, resulting in different colicin concentrations. We analysed experimental colonies obtained for each ecological scenario. The S, R, and C strains expressed fluorescent proteins as depicted above the individual rows: black = nfp,

green = GFP, red = mCh. Initial ratios S:R:C from left to right: S_1 = 1:1:1, $S_{0.5}$ = 1:1:0.5, $S_{0.1}$ = 1:1:0.1. Corresponding colony images obtained from simulations are shown in figure S6. Red lines depict the edge of the colony in scenarios II* and III. White scale bars represent 1 mm. Scenarios marked by an asterisk (*) are auxiliary scenarios that were obtained by exchanging the assignment of GFP between the two strains not expressing mCherry (see appendix S1, section 1.1.2).

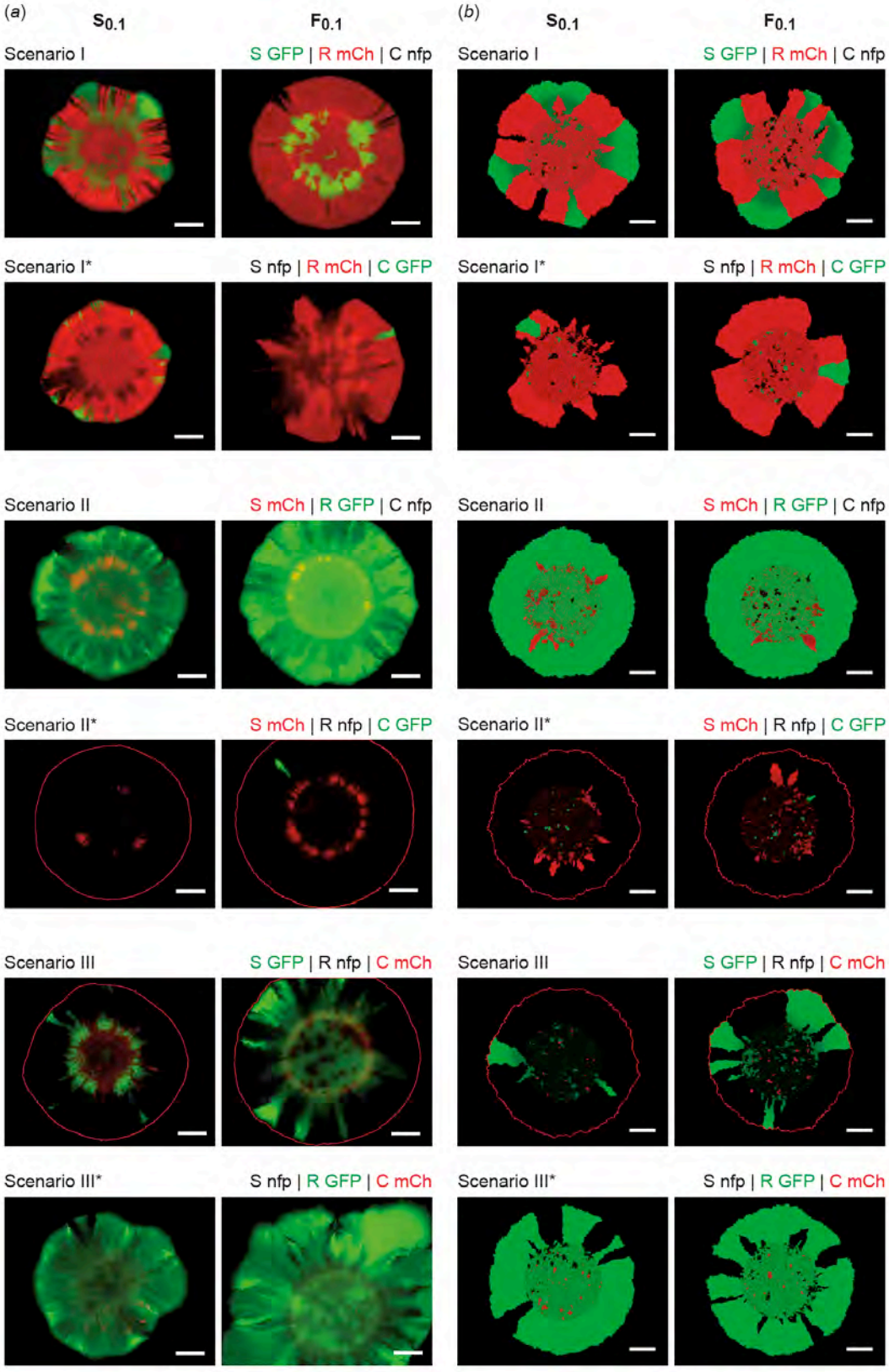
Figure S6. Colony development for different initial ratios of the C strain (simulation results).



Simulations of colony growth under slow growth conditions were performed for different initial ratios of the colicin producing strain C. The figure displays final colonies acquired from simulations for each ecological scenario and accompanies the experimentally observed colonies shown in figure S5 (see that figure for further information on the

presentation; initial strain ratio S_X = 1:1:X for S, R, and C, respectively). Simulation parameters were chosen as listed in table S1. White scale bars represent 1 mm. Scenarios marked by an asterisk (*) are auxiliary scenarios that were obtained by exchanging the assignment of GFP between the two strains not expressing mCherry (see appendix S1, section 1.1.2).

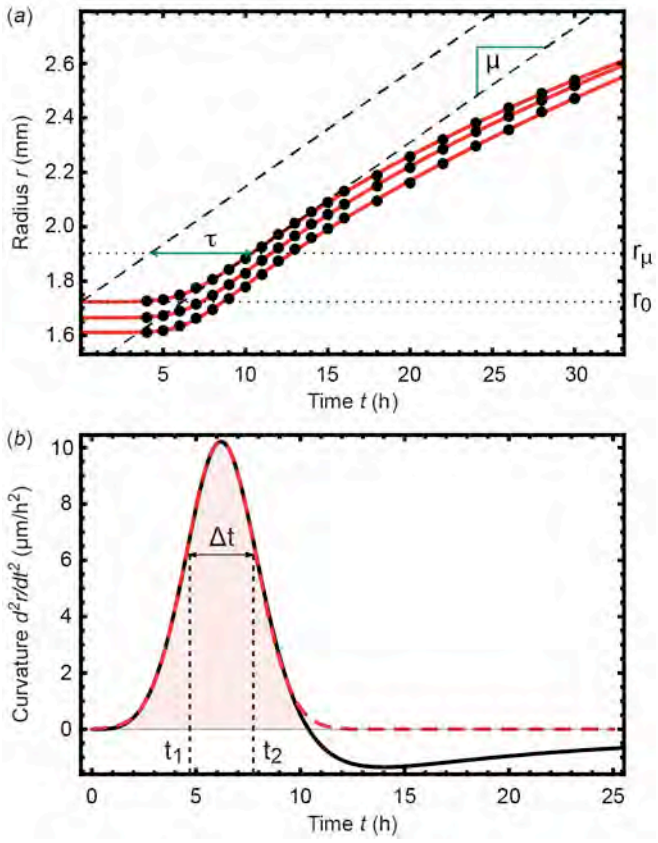
Figure S7. Colony development at a low initial ratio of the C strain for slow ($S_{0.1}$) and fast ($F_{0.1}$) growth conditions (initial strain ratios S:R:C = 1:1:0.1, observation after 48 h).



(*a*) Experimental data, (*b*) Simulation results. We analysed experimental colonies obtained for each ecological scenario. The S, R, and the C strain expressed fluorescent proteins as depicted above the individual rows: black = nfp, green = GFP, red = mCh. Red lines depict the edge of the colony in scenarios II* and III. White scale bars represent 1 mm. Scenarios

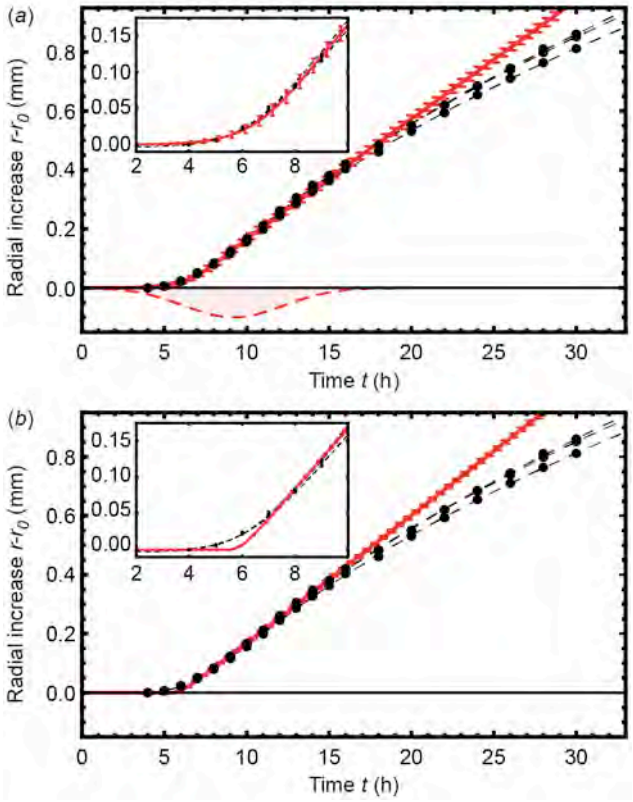
marked by an asterisk (*) are auxiliary scenarios that were obtained by exchanging the assignment of GFP between the two strains not expressing mCherry (see appendix S1, section 1.1.2).

Figure S8. Definition of the growth rate μ, of the lag time τ and of the time frame Δt.



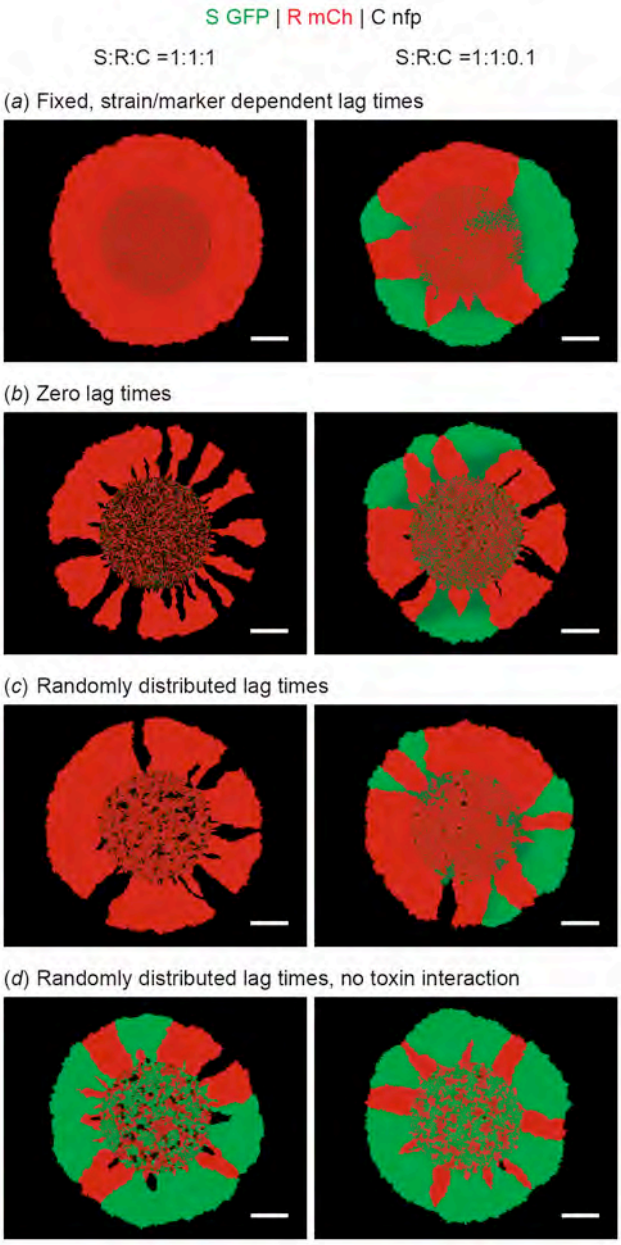
(*a*) Three experimental data sets (black dots) exemplify the radial expansion of R mCherry (R_{mCh}) colonies. The R strain expressed mCherry and proliferated under fast growth conditions (F). Red solid lines: logistic fits of the experimental data (cf. equation (1)). Green solid lines: definition of the maximal growth rate μ and of the lag time τ. Black lines with respect to the uppermost growth curve: (right dashed line) linear fit through the point of maximal growth rate μ, (left dashed line) linear fit shifted by the lag time τ, (upper dotted line) radius r_μ attained at the point of maximal growth rate, (lower dotted line) initial colony radius r_0. (*b*) The black solid line exemplifies the curvature of the logistic fit of an experimentally determined growth curve (using the uppermost red curve shown in (*a*)). The time frame $\Delta t = t_2 - t_1$ measures the time over which the growth curve exhibits positive curvature. The times t_1 and t_2 mark the locations at which the curvature assumes its maximal value multiplied by exp(-0.5). The Gaussian function defined in equation (2) of appendix S1 provided a good approximation of the growth curve's regime of positive curvature (red dashed line).

Figure S9. Comparison between experimentally determined and simulated growth curves for the growth of R mCherry (R_{mCh}) colonies under fast growth condition (F).



The experimentally determined growth curves (black dots) in figure S8*a* were shifted by their initial colony radii and were compared with averaged simulated growth curves. Dashed black lines represent logistic fits of experimental data (equation (1)) as discussed in section 1.2 of appendix S1. The mean growth curve obtained from simulations is shown as red line and was acquired by averaging over 1000 simulated growth processes. Error bars depict one standard deviation of the sample of simulated growth curves. The insets focus on the initial regimes of maximal positive curvature (acceleration phase).
(*a*) The averaged simulated growth curve was based on a model with Gaussian distributed mesoscopic lag times τ_m (see section 2.2 of appendix S1). The rescaled Gaussian distribution that was used for sampling of τ_m on individual lattice sites is illustrated on the negative y-axis. The lag and acceleration phase of experimental growth curves were approximated reasonably well. We attributed the deviations that arose at later times ($t > 17.5$ h) to the dehydration of agar and to nutrient depletion in the experimental system.
(*b*) When fixed (i.e. not randomly distributed), strain/marker dependent mesoscopic lag times were used on initially colonised lattice sites, we observe a larger deviation between simulated growth curves and experimental data during the acceleration phase. The corresponding model also failed to reproduce experimentally observed segregation patterns.

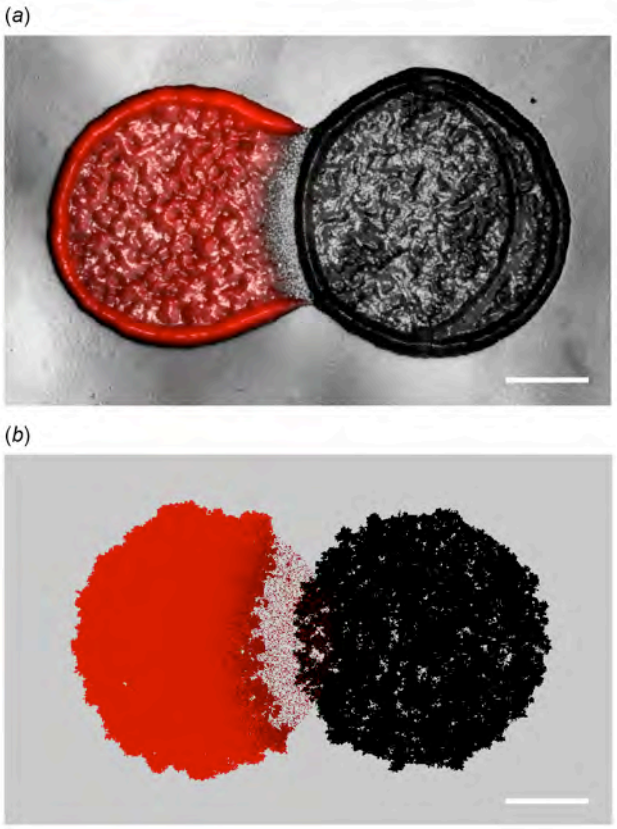
Figure S10. Simulated colonies based on different implementations of mesoscopic lag times, and visualisation of the toxin interaction's impact on segregation patterns.



All figures correspond to cyclic scenario I and slow growth conditions. The images may be compared to their corresponding experimental counterparts in figures 2*a* and 4*a* (main text). Initial strain ratios and the strain/fluorescent marker combination are depicted above the colonies. White scale bars represent 1 mm. (*a*) As discussed in appendix S1, a model based on fixed (i.e. not randomly distributed), strain/maker dependent mesoscopic lag times failed to reproduce experimentally observed segregation patterns. At equal initial strain ratios, only the R strain persisted, whereas R and C survived in experiments. The model also failed to capture the experimentally observed coexistence of all three strains when the C strain ratio was reduced to r_C=0.1. (*b*) A model that neglected the lag phase could capture experimentally observed segregation patterns both at r_C=0.1 and r_C=1. The model could not be used to reproduce experimentally determined radial growth curves. (*c*) A model based on randomly distributed mesoscopic lag times could reproduce both experimentally determined growth curves and predict the correct surviving strains of competitive range expansions. A fewer number of C_{nfp} sectors was observed as compared

to the results shown in (*b*). Predictions on biodiversity zones (figure 3) were similar.
(*d*) Without toxin interaction, simulations predicted that the sensitive S strain survived even at equal initial strain ratios. The patterns shown for C strain ratios of r_C=0.1 and r_C=1.0 were at odds with the experimental results shown in figures 2*a* and 4*a*. Colicin was always present in these experiments.

Figure S11. Effective range of colicin.



Colicin produced by toxic C cells diffuses through extracellular fluid until possibly being absorbed by sensitive (S) cells. Upon absorption, the endonuclease activity of colicin causes DNA degeneration of the S cells and inhibits further cell divisions. Eventually, the S cells are lysed. These figures exemplify the interaction between single-strain C_{nfp} (black) and S_{mCh} (red) colonies under slow growth conditions. White scale bars represent 1 mm. (*a*) Experimental results. A droplet of inoculum that contained 1 µl of S_{mCh} was placed next to a droplet that contained 1 µl of C_{nfp} (with a small overlap). The picture was taken after 12 hours of incubation at 37°C. Up to this time, growth was mainly observed in the inside of the areas of inoculation, accompanied by an increasing thickness of the colonies. The colonies' effective radii increased only slightly. (*b*) Simulation of the experiment. The initial colony radii were set to 1.5 mm and the simulation was stopped after 12 hours of simulated time. Darkened red regions represent lattice sites with a majority of lysed S cells. The sparsely populated area between the two colonies could not be colonised by S cells due to inhibition by the toxin. Simulation parameters were chosen in accordance with table S1.

Bibliography

[1] D. Vorberg, W. Wustmann, R. Ketzmerick and A. Eckardt. Generalized Bose-Einstein condensation into multiple states in driven-dissipative systems. *Phys. Rev. Lett.* **111**(24), 240405, 2013. DOI: 10.1103/PhysRevLett.111.240405.

[2] J. Knebel, T. Krüger, M. F. Weber and E. Frey. Coexistence and survival in conservative Lotka-Volterra networks. *Phys. Rev. Lett.* **110**(16), 168106, 2013. DOI: 10.1103/PhysRevLett.110.168106.

[3] J. Knebel, M. F. Weber, T. Krüger and E. Frey. Evolutionary games of condensates in coupled birth-death processes. *Nat. Commun.* **6**, 6977, 2015. DOI: 10.1038/ncomms7977.

[4] B. Kerr, M. A. Riley, M. W. Feldman and B. J. M. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* **418**(6894), 171–174, 2002. DOI: 10.1038/nature00823.

[5] M. F. Weber, G. Poxleitner, E. Hebisch, E. Frey and M. Opitz. Chemical warfare and survival strategies in bacterial range expansions. *J. R. Soc. Interface* **11**(96), 20140172, 2014. DOI: 10.1098/rsif.2014.0172.

[6] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys. (Berlin)* **322**(8), 549–560, 1905. DOI: 10.1002/andp.19053220806.

[7] A. D. Fokker. Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld. *Ann. Phys. (Berlin)* **348**(5), 810–820, 1914. DOI: 10.1002/andp.19143480507.

[8] M. Planck. Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie. *Sitz.-ber. Preuß. Akad. Wiss.* 324–341, 1917.

[9] A. Kolmogoroff. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104**(1), 415–458, 1931. DOI: 10.1007/BF01457949.

[10] A. Nordsieck, W. E. Lamb Jr. and G. E. Uhlenbeck. On the theory of cosmic-ray showers I: The Furry model and the fluctuation problem. *Physica* **7**(4), 344–360, 1940. DOI: 10.1016/S0031-8914(40)90102-1.

[11] M. Doi. Second quantization representation for classical many-particle system. *J. Phys. A: Math. Gen.* **9**(9), 1465–1477, 1976. DOI: 10.1088/0305-4470/9/9/008.

[12] M. Doi. Stochastic theory of diffusion-controlled reaction. *J. Phys. A: Math. Gen.* **9**(9), 1479–1495, 1976. DOI: 10.1088/0305-4470/9/9/009.

[13] Y. B. Zel'dovich and A. A. Ovchinnikov. The mass action law and the kinetics of chemical reactions with allowance for thermodynamic fluctuations of the density. *Sov. Phys. JETP* **47**(5), 829–834, 1978.

[14] H. A. Rose. Renormalized kinetic theory of nonequilibrium many-particle classical systems. *J. Stat. Phys.* **20**(4), 415–447, 1979. DOI: 10.1007/bf01011780.

[15] P. Grassberger and M. Scheunert. Fock-Space methods for identical classical objects. *Fortschr. Phys.* **28**(10), 547–578, 1980. DOI: 10.1002/prop.19800281004.

[16] A. S. Mikhailov. Path integrals in chemical kinetics I. *Phys. Lett. A* **85**(4), 214–216, 1981. DOI: 10.1016/0375-9601(81)90017-7.

[17] A. S. Mikhailov. Path integrals in chemical kinetics II. *Phys. Lett. A* **85**(8–9), 427–429, 1981. DOI: 10.1016/0375-9601(81)90429-1.

[18] N. Goldenfeld. Kinetics of a model for nucleation-controlled polymer crystal growth. *J. Phys. A: Math. Gen.* **17**(14), 2807–2821, 1984. DOI: 10.1088/0305-4470/17/14/024.

[19] A. S. Mikhailov and V. V. Yashin. Quantum-field methods in the theory of diffusion-controlled reactions. *J. Stat. Phys.* **38**(1-2), 347–359, 1985. DOI: 10.1007/BF01017866.

[20] L. Peliti. Path integral approach to birth-death processes on a lattice. *J. Phys. (Paris)* **46**(9), 1469–1483, 1985. DOI: 10.1051/jphys:019850046090146900.

[21] P. C. Martin, E. D. Siggia and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev. A* **8**(1), 423–437, 1973. DOI: 10.1103/PhysRevA.8.423.

[22] C. de Dominicis. Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques. *J. Phys. (Paris)* **37**(C1), 247–253, 1976. DOI: 10.1051/jphyscol:1976138.

[23] H.-K. Janssen. On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties. *Z. Phys. B* **23**(4), 377–380, 1976. DOI: 10.1007/BF01316547.

[24] R. Bausch, H.-K. Janssen and H. Wagner. Renormalized field theory of critical dynamics. *Z. Phys. B* **24**(1), 113–127, 1976. DOI: 10.1007/BF01312880.

[25] V. Fock. Konfigurationsraum und zweite Quantelung. *Z. Phys.* **75**(9), 622–647, 1932. DOI: 10.1007/BF01344458.

[26] P. A. M. Dirac. A new notation for quantum mechanics. *Math. Proc. Camb. Phil. Soc.* **35**(3), 416–418, 1939. DOI: 10.1017/S0305004100021162.

[27] E. Schrödinger. Der stetige Übergang von der Mikro- zur Makromechanik. *Naturwissenschaften* **14**(28), 664–666, 1926. DOI: 10.1007/BF01507634.

[28] E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys. Rev. Lett.* **10**(7), 277–279, 1963. DOI: 10.1103/PhysRevLett.10.277.

[29] R. J. Glauber. Coherent and incoherent states of the radiation field. *Phys. Rev.* **131**(6), 2766–2788, 1963. DOI: 10.1103/PhysRev.131.2766.

[30] G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.* **80**(2), 268–272, 1950. DOI: 10.1103/PhysRev.80.268.

[31] S. Sandow and S. Trimper. Aggregation processes in a master-equation approach. *Europhys. Lett.* **21**(8), 799–804, 1993. DOI: 10.1209/0295-5075/21/8/001.

[32] H. Patzlaff, S. Sandow and S. Trimper. Diffusion and correlation in a coherent representation. *Z. Phys. B* **95**(3), 357–362, 1994. DOI: 10.1007/BF01343964.

[33] P.-A. Bares and M. Mobilia. Diffusion-limited reactions of hard-core particles in one dimension. *Phys. Rev. E* **59**(2), 1996–2009, 1999. DOI: 10.1103/PhysRevE.59.1996.

[34] V. Brunel, K. Oerding and F. van Wijland. Fermionic field theory for directed percolation in (1 + 1)-dimensions. *J. Phys. A: Math. Gen.* **33**(6), 1085–1097, 2000. DOI: 10.1088/0305-4470/33/6/301.

[35] M. Schulz and P. Reineker. Exact substitute processes for diffusion-reaction systems with local complete exclusion rules. *New J. Phys.* **7**, 31, 2005. DOI: 10.1088/1367-2630/7/1/031.

[36] É. M. Silva, P. T. Muzy and A. E. Santana. Fock space for fermion-like lattices and the linear Glauber model. *Physica A* **387**(21), 5101–5109, 2008. DOI: 10.1016/j.physa.2008.04.015.

[37] C. W. Gardiner and S. Chaturvedi. The Poisson representation. I. A new technique for chemical master equations. *J. Stat. Phys.* **17**(6), 429–468, 1977. DOI: 10.1007/BF01014349.

[38] S. Chaturvedi and C. W. Gardiner. The Poisson representation. II. Two-time correlation functions. *J. Stat. Phys.* **18**(5), 501–522, 1978. DOI: 10.1007/BF01014520.

[39] A. M. Walczak, A. Mugler and C. H. Wiggins. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl. Acad. Sci. USA* **106**(16), 6529–6534, 2009. DOI: 10.1073/pnas.0811999106.

[40] V. Elgart and A. Kamenev. Rare event statistics in reaction-diffusion systems. *Phys. Rev. E* **70**(4), 041106, 2004. DOI: 10.1103/PhysRevE.70.041106.

[41] M. Assaf and B. Meerson. Spectral formulation and WKB approximation for rare-event statistics in reaction systems. *Phys. Rev. E* **74**(4), 041115, 2006. DOI: 10.1103/PhysRevE.74.041115.

[42] M. Assaf and B. Meerson. Spectral theory of metastability and extinction in birth-death systems. *Phys. Rev. Lett.* **97**(20), 200602, 2006. DOI: 10.1103/PhysRevLett.97.200602.

[43] M. Assaf, A. Kamenev and B. Meerson. Population extinction in a time-modulated environment. *Phys. Rev. E* **78**(4), 041123, 2008. DOI: 10.1103/PhysRevE.78.041123.

[44] M. Sasai and P. G. Wolynes. Stochastic gene expression as a many-body problem. *Proc. Natl. Acad. Sci. USA* **100**(5), 2374–2379, 2003. DOI: 10.1073/pnas.2627987100.

[45] S. F. Edwards. The statistical dynamics of homogeneous turbulence. *J. Fluid Mech.* **18**(2), 293–273, 1964. DOI: 10.1017/S0022112064000180.

[46] K. R. Sreenivasan and G. L. Eyink. Sam Edwards and the turbulence theory. *Stealing the Gold*. Ed. by P. Goldbart, N. Goldenfeld and D. Sherrington. Oxford University Press, 2005, DOI: 10.1093/acprof:oso/9780198528531.001.0001.

[47] H. Haken. Generalized Onsager-Machlup function and classes of path integral solutions of the Fokker-Planck equation and the master equation. *Z. Phys. B* **24**(3), 321–326, 1976. DOI: 10.1007/BF01360904.

[48] K. Wódkiewicz. Functional representation of a non-Markovian probability distribution in statistical mechanics. *Phys. Lett.* **84**(2), 56–58, 1981. DOI: 10.1016/0375-9601(81)90589-2.

[49] L. Pesquera, M. A. Rodriguez and E. Santos. Path integrals for non-Markovian processes. *Phys. Lett. A* **94**(6-7), 287–289, 1983. DOI: 10.1016/0375-9601(83)90719-3.

[50] R. F. Fox. Functional-calculus approach to stochastic differential equations. *Phys. Rev. A* **33**(1), 467–476, 1986. DOI: 10.1103/PhysRevA.33.467.

[51] R. F. Fox. Uniform convergence to an effective Fokker-Planck equation for weakly colored noise. *Phys. Rev. A* **34**(5), 4525–4527, 1986. DOI: 10.1103/PhysRevA.34.4525.

[52] J. F. Luciani and A. D. Verga. Functional integral approach to bistability in the presence of correlated noise. *Europhys. Lett.* **4**(3), 255–261, 1987. DOI: 10.1209/0295-5075/4/3/001.

[53] A. Förster and A. S. Mikhailov. Optimal fluctuations leading to transitions in bistable systems. *Phys. Lett. A* **126**(8-9), 459–462, 1988. DOI: 10.1016/0375-9601(88)90039-4.

[54] J. F. Luciani and A. D. Verga. Bistability driven by correlated noise: Functional integral treatment. *J. Stat. Phys.* **50**(3), 567–597, 1988. DOI: 10.1007/BF01026491.

[55] A. J. Bray and A. J. McKane. Instanton calculation of the escape rate for activation over a potential barrier driven by colored noise. *Phys. Rev. Lett.* **62**(5), 493–496, 1989. DOI: 10.1103/PhysRevLett.62.493.

[56] P. Hänggi. Path integral solutions for non-Markovian processes. *Z. Phys. B* **75**(2), 275–281, 1989. DOI: 10.1007/BF01308011.

[57] A. J. McKane. Noise-induced escape rate over a potential barrier: Results for general noise. *Phys. Rev. A* **40**(7), 4050–4053, 1989. DOI: 10.1103/PhysRevA.40.4050.

[58] A. J. McKane, H. C. Luckock and A. J. Bray. Path integrals and non-Markov processes. I. General formalism. *Phys. Rev. A* **41**(2), 644–656, 1990. DOI: 10.1103/PhysRevA.41.644.

[59] A. J. Bray, A. J. McKane and T. J. Newman. Path integrals and non-Markov processes. II. Escape rates and stationary distributions in the weak-noise limit. *Phys. Rev. A* **41**(2), 657–667, 1990. DOI: 10.1103/PhysRevA.41.657.

[60] H. C. Luckock and A. J. McKane. Path integrals and non-Markov processes. III. Calculation of the escape-rate prefactor in the weak-noise limit. *Phys. Rev. A* **42**(4), 1982–1996, 1990. DOI: 10.1103/PhysRevA.42.1982.

[61] T. G. Venkatesh and L. M. Patnaik. Effective Fokker-Planck equation: Path-integral formalism. *Phys. Rev. E* **48**(4), 2402–2412, 1993. DOI: 10.1103/PhysRevE.48.2402.

[62] S. J. B. Einchcomb and A. J. McKane. Using path-integral methods to calculate noise-induced escape rates in bistable systems: The case of quasi-monochromatic noise. *Fluctuations and Order - The New Synthesis*. Ed. by M. Millonas. Institute for Nonlinear Science. Springer, New York, NY, 1996, pp. 139–154. DOI: 10.1007/978-1-4612-3992-5_10.

[63] C. Mahanta and T. G. Venkatesh. Damped stochastic system driven by colored noise: Analytical solution by a path integral approach. *Phys. Rev. E* **62**(2), 1509–1520, 2000. DOI: 10.1103/PhysRevE.62.1509.

[64] F. W. Wiegel. *Introduction to Path-Integral Methods in Physics and Polymer Science*. World Scientific Publishing Company, Singapore, 1986.

[65] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. International Series of Monographs on Physics 73. Clarendon Press, Oxford, 1988.

[66] T. A. Vilgis. Polymer theory: path integrals and scaling. *Phys. Rep.* **336**(3), 167–254, 2000. DOI: 10.1016/S0370-1573(99)00122-2.

[67] H. Kleinert. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*. 5th edition. World Scientific Publishing Company, New Jersey, 2009. Chap. 15.

[68] P. C. Bressloff and J. M. Newby. Path integrals and large deviations in stochastic hybrid systems. *Phys. Rev. E* **89**(4), 042701, 2014. DOI: 10.1103/PhysRevE.89.042701.

[69] P. C. Bressloff. *Stochastic Processes in Cell Biology*. Interdisciplinary Applied Mathematics. Vol. 41. Springer, Cham, 2014. Chap. 10. DOI: 10.1007/978-3-319-08488-6.

[70] P. C. Bressloff. Path-Integral methods for analyzing the effects of fluctuations in stochastic hybrid neural networks. *J. Math. Neurosc.* **5**(4), 2015. DOI: 10.1186/s13408-014-0016-z.

[71] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Adv. Phys.* **49**(7), 815–958, 2000. DOI: 10.1080/00018730050198152.

[72] G. Ódor. Universality classes in nonequilibrium lattice systems. *Rev. Mod. Phys.* **76**(3), 663–724, 2004. DOI: 10.1103/RevModPhys.76.663.

[73] M. Henkel, H. Hinrichsen and S. Lübeck. *Non-Equilibrium Phase Transitions - Volume I: Absorbing Phase Transitions*. Theoretical and Mathematical Physics. Springer, Dordrecht, 2008. DOI: 10.1007/978-1-4020-8765-3.

[74] U. C. Täuber, M. Howard and B. P. Vollmayr-Lee. Applications of field-theoretic renormalization group methods to reaction-diffusion problems. *J. Phys. A: Math. Gen.* **38**(17), R79–R131, 2005. DOI: 10.1088/0305-4470/38/17/R01.

[75] U. C. Täuber. *Critical Dynamics - A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior*. Cambridge University Press, Cambridge, 2014.

[76] R. Zwanzig. Memory effects in irreversible thermodynamics. *Phys. Rev.* **124**(4), 983–992, 1961. DOI: 10.1103/PhysRev.124.983.

[77] R. Zwanzig. On the identity of three generalized master equations. *Physica* **30**(6), 1109–1123, 1964. DOI: 10.1016/0031-8914(64)90102-8.

[78] H. Mori. Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.* **33**(3), 423–455, 1965. DOI: 10.1143/PTP.33.423.

[79] I. Oppenheim and K. E. Shuler. Master equations and Markov processes. *Phys. Rev.* **138**(4B), B1007–B1011, 1965. DOI: 10.1103/PhysRev.138.B1007.

[80] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**(4), 403–434, 1976. DOI: 10.1016/0021-9991(76)90041-3.

[81] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**(25), 2340–2361, 1977. DOI: 10.1021/j100540a008.

[82] D. T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A* **188**(1), 404–425, 1992. DOI: 10.1016/0378-4371(92)90283-V.

[83] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**, 35–55, 2007. DOI: 10.1146/annurev.physchem.58.032806.104637.

[84] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.* **91**(6), 1505–1512, 1953. DOI: 10.1103/PhysRev.91.1505.

[85] N. Wiener. The average of an analytic functional. *Proc. Natl. Acad. Sci. USA* **7**(9), 253–260, 1921. DOI: 10.1073/pnas.7.9.253.

[86] N. Wiener. The average of an analytic functional and the Brownian movement. *Proc. Natl. Acad. Sci. USA* **7**(10), 294–298, 1921. DOI: 10.1073/pnas.7.10.294.

[87] R. Brown. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philos. Mag. Ser. 2* **4**(21), 161–173, 1828. DOI: 10.1080/14786442808674769.

[88] G. L. Eyink. Action principle in nonequilibrium statistical dynamics. *Phys. Rev. E* **54**(4), 3419–3435, 1996. DOI: 10.1103/PhysRevE.54.3419.

[89] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers I*. Springer, New York, NY, 1999. Chap. 10. DOI: 10.1007/978-1-4757-3069-2.

[90] M. Kac. On distributions of certain Wiener functionals. *Trans. Amer. Math. Soc.* **65**(1), 1–13, 1949. DOI: 10.2307/1990512.

[91] K. Itô. Stochastic integral. *Proc. Imp. Acad.* **20**(8), 519–524, 1944. DOI: 10.3792/pia/1195572786.

[92] K. Itô. On a stochastic integral equation. *Proc. Japan Acad.* **22**(2), 32–35, 1946. DOI: 10.3792/pja/1195572371.

[93] K. Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.* **1**, 35–47, 1950.

[94] D. S. Lemons. Paul Langevin's 1908 paper "On the theory of Brownian motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530–533 (1908)]. *Am. J. Phys.* **65**(11), 1079–1081, 1997. DOI: 10.1119/1.18725.

[95] F. van Wijland. Field theory for reaction-diffusion processes with hard-core particles. *Phys. Rev. E* **63**(2), 022101, 2001. DOI: 10.1103/PhysRevE.63.022101.

[96] J. Cardy. Reaction-diffusion processes. *Non-equilibrium Statistical Mechanics and Turbulence*. Ed. by S. Nazarenko and O. V. Zaboronski. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 2008, pp. 108–131. DOI: 10.1017/CBO9780511812149.

[97] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* **28**(6), 1049–1070, 1926. DOI: 10.1103/PhysRev.28.1049.

[98] W. Heisenberg. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Phys.* **33**(1), 879–893, 1925. DOI: 10.1007/BF01328377.

[99] M. Born and P. Jordan. Zur Quantenmechanik. *Z. Phys.* **34**(1), 858–888, 1925. DOI: 10.1007/BF01328531.

[100] M. Born, W. Heisenberg and P. Jordan. Zur Quantenmechanik. II. *Z. Phys.* **35**(8-9), 557–615, 1926. DOI: 10.1007/BF01379806.

[101] J. L. Cardy and U. C. Täuber. Field theory of branching and annihilating random walks. *J. Stat. Phys.* **90**(1/2), 1–56, 1998. DOI: 10.1023/A:1023233431588.

[102] B. Sinervo and C. M. Lively. The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380**(6571), 240–243, 1996. DOI: 10.1038/380240a0.

[103] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. 3rd edition. North-Holland Personal Library. Elsevier, Amsterdam, 2007.

[104] R. Huang, I. Chavez, K. M. Taute, B. Lukić, S. Jeney, M. G. Raizen and E.-L. Florin. Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid. *Nat. Phys.* **7**(7), 576–580, 2011. DOI: 10.1038/nphys1953.

[105] T. Franosch, M. Grimm, M. Belushkin, F. M. Mor, G. Foffi, L. Forró and S. Jeney. Resonances arising from hydrodynamic memory in Brownian motion. *Nature* **478**(7367), 85–88, 2011. DOI: 10.1038/nature10498.

[106] M. von Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. Phys. (Berlin)* **326**(14), 756–780, 1906. DOI: 10.1002/andp.19063261405.

[107] O. Klein. Zur statistischen Theorie der Suspensionen und Lösungen. *Ark. Mat. Astr. Fys.* **16**(5), PhD thesis, 1921.

[108] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**(4), 284–304, 1940. DOI: 10.1016/S0031-8914(40)90098-2.

[109] H. Risken. *The Fokker-Planck Equation - Methods of Solution and Applications*. 2nd edition. Springer Series in Synergetics. Vol. 18. Springer, Berlin, Heidelberg, 1996. DOI: 10.1007/978-3-642-61544-3.

[110] M. Kimura. Some problems of stochastic-processes in genetics. *Ann. Math. Statist.* **28**(4), 882–901, 1957. DOI: 10.1214/aoms/1177706791.

[111] H. P. de Vladar and N. H. Barton. The contribution of statistical physics to evolutionary biology. *Trends Ecol. Evol.* **26**(8), 424–432, 2011. DOI: 10.1016/j.tree.2011.04.002.

[112] R. P. Feynman. The principle of least action in quantum mechanics. *Feynman's Thesis: A New Approach to Quantum Theory*. Ed. by L. M. Brown. World Scientific Publishing Company, Singapore, 2005, pp. 1–69.

[113] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics - Stochastic Modelling and Applied Probability. Vol. 23. Springer, Berlin, Heidelberg, 1992. DOI: 10.1007/978-3-662-12616-5.

[114] S. N. Majumdar. Brownian functionals in physics and computer science. *Curr. Sci.* **89**(12), 2076–2092, 2005.

[115] C. Gardiner. *Stochastic Methods - A Handbook for the Natural and Social Sciences*. 4th edition. Springer Series in Synergetics. Vol. 13. Springer, Berlin, Heidelberg, 2009.

[116] D. A. Harville. *Matrix Algebra From a Statistician's Perspective*. Springer, New York, NY, 1997. Chap. 21.9. DOI: 10.1007/b98818.

[117] G. Volpe and J. Wehr. Effective drifts in dynamical systems with multiplicative noise: a review of recent progress. *Rep. Prog. Phys.* **79**(5), 2016. DOI: 10.1088/0034-4885/79/5/053901.

[118] R. Lipowsky, S. Klumpp and T. M. Nieuwenhuizen. Random walks of cytoskeletal motors in open and closed compartments. *Phys. Rev. Lett.* **87**(10), 108101, 2001. DOI: 10.1103/PhysRevLett.87.108101.

[119] A. Parmeggiani, T. Franosch and E. Frey. Phase coexistence in driven one-dimensional transport. *Phys. Rev. Lett.* **90**(8), 086601, 2003. DOI: 10.1103/PhysRevLett.90.086601.

[120] A. Parmeggiani, T. Franosch and E. Frey. Totally asymmetric simple exclusion process with Langmuir kinetics. *Phys. Rev. E* **70**(4), 046101, 2004. DOI: 10.1103/PhysRevE.70.046101.

[121] J. Howard. Molecular motors: structural adaptations to cellular functions. *Nature* **389**(6651), 561–567, 1997. DOI: 10.1038/39247.

[122] W. J. Anderson. *Continuous-Time Markov Chains - An Applications-Oriented Approach*. Springer Series in Statistics - Probability and its Applications. Springer, New York, NY, 1991. DOI: 10.1007/978-1-4612-3038-0.

[123] A. Kolmogoroff. Zur Theorie der Markoffschen Ketten. *Math. Ann.* **112**(1), 155–160, 1936. DOI: 10.1007/BF01565412.

[124] W. H. Furry. On fluctuation phenomena in the passage of high energy electrons through lead. *Phys. Rev.* **52**(6), 569–581, 1937. DOI: 10.1103/PhysRev.52.569.

[125] W. Feller. Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheor.* **5**(1), 11–40, 1939. DOI: 10.1007/BF01602932.

[126] M. Delbrück. Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **8**(1), 120–124, 1940. DOI: 10.1063/1.1750549.

[127] K. Singer. Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. *J. R. Stat. Soc. B* **15**(1), 92–106, 1953.

[128] A. F. Bartholomay. Stochastic models for chemical reactions: I. Theory of the unimolecular reaction process. *B. Math. Biophys.* **20**(3), 175–190, 1958. DOI: 10.1007/bf02478297.

[129] I. M. Krieger and P. J. Gans. First-order stochastic processes. *J. Chem. Phys.* **32**(1), 247–250, 1960. DOI: 10.1063/1.1700909.

[130] K. Ishida. Stochastic model for bimolecular reaction. *J. Chem. Phys.* **41**(8), 2472–2478, 1964. DOI: 10.1063/1.1726290.

[131] D. A. McQuarrie. Kinetics of small systems. I. *J. Chem. Phys.* **38**(2), 433–436, 1963. DOI: 10.1063/1.1733676.

[132] D. A. McQuarrie, C. J. Jachimowski and M. E. Russell. Kinetics of small systems. II. *J. Chem. Phys.* **40**(10), 2914–2921, 1964. DOI: 10.1063/1.1724926.

[133] D. A. McQuarrie. Stochastic approach to chemical kinetics. *J. Appl. Probab.* **4**(3), 413–478, 1967. DOI: 10.2307/3212214.

[134] G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems - From Dissipative Structures to Order Through Fluctuations*. John Wiley & Sons, New York, NY, 1977. Chap. 10.

[135] B. P. Lee. Renormalization group calculation for the reaction $kA \to \emptyset$. *J. Phys. A: Math. Gen.* **27**(8), 2633–2652, 1994. DOI: 10.1088/0305-4470/27/8/004.

[136] B. P. Lee and J. Cardy. Renormalization group study of the $A + B \to \emptyset$ diffusion-limited reaction. *J. Stat. Phys.* **80**(5-6), 971–1007, 1995. DOI: 10.1007/BF02179861.

[137] J. Cardy and U. C. Täuber. Theory of branching and annihilating random walks. *Phys. Rev. Lett.* **77**(23), 4780–4783, 1996. DOI: 10.1103/PhysRevLett.77.4780.

[138] H.-K. Janssen and U. C. Täuber. The field theory approach to percolation processes. *Ann. Phys.* **315**(1), 147–192, 2005. DOI: 10.1016/j.aop.2004.09.011.

[139] R. J. Glauber. Time-dependent statistics of the Ising model. *J. Math. Phys.* **4**(2), 294–307, 1963. DOI: 10.1063/1.1703954.

[140] K. Kawasaki. Diffusion constants near the critical point for time-dependent Ising models. I. *Phys. Rev.* **145**(1), 224–230, 1966. DOI: 10.1103/PhysRev.145.224.

[141] K. Kawasaki. Diffusion constants near the critical point for time-dependent Ising models. II. *Phys. Rev.* **148**(1), 375–381, 1966. DOI: 10.1103/PhysRev.148.375.

[142] K. Kawasaki. Diffusion constants near the critical point for time-dependent Ising models. III. Self-diffusion constant. *Phys. Rev.* **150**(1), 285–290, 1966. DOI: 10.1103/physrev.150.285.

[143] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**(6767), 335–338, 2000. DOI: 10.1038/35002125.

[144] M. Thattai and A. van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA* **98**(15), 8614–8619, 2001. DOI: 10.1073/pnas.151588598.

[145] C. V. Rao, D. M. Wolf and A. P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature* **420**(6912), 231–237, 2002. DOI: 10.1038/nature01258.

[146] G. Karlebach and R. Shamir. Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell. Biol.* **9**(10), 770–780, 2008. DOI: 10.1038/nrm2503.

[147] V. Shahrezaei and P. S. Swain. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. USA* **105**(45), 17256–17261, 2008. DOI: 10.1073/pnas.0803850105.

[148] L. S. Tsimring. Noise in biology. *Rep. Prog. Phys.* **77**(2), 026601, 2014. DOI: 10.1088/0034-4885/77/2/026601.

[149] N. T. J. Bailey. A simple stochastic epidemic. *Biometrika* **37**(3/4), 193–202, 1950. DOI: 10.2307/2332371.

[150] M. J. Keeling and J. V. Ross. On methods for studying stochastic disease dynamics. *J. R. Soc. Interface* **5**(19), 171–181, 2008. DOI: 10.1098/rsif.2007.1106.

[151] A. J. Black and A. J. McKane. WKB calculation of an epidemic outbreak distribution. *J. Stat. Mech. Theor. Exp.* **2011**(12), P12006, 2011. DOI: 10.1088/1742-5468/2011/12/P12006.

[152] K. Rock, S. Brand, J. Moir and M. J. Keeling. Dynamics of infectious diseases. *Rep. Prog. Phys.* **77**(2), 026602, 2014. DOI: 10.1088/0034-4885/77/2/026602.

[153] E. Clayton, D. P. Doupé, A. M. Klein, D. J. Winton, B. D. Simons and P. H. Jones. A single type of progenitor cell maintains normal epidermis. *Nature* **446**(7132), 185–189, 2007. DOI: 10.1038/nature05574.

[154] T. Chou. Peeling and sliding in nucleosome repositioning. *Phys. Rev. Lett.* **99**(5), 058105, 2007. DOI: 10.1103/PhysRevLett.99.058105.

[155] I. Volkov, J. R. Banavar, S. P. Hubbell and A. Maritan. Neutral theory and relative species abundance in ecology. *Nature* **424**(6952), 1035–1037, 2003. DOI: 10.1038/nature01883.

[156] A. J. McKane and T. J. Newman. Stochastic models in population biology and their deterministic analogs. *Phys. Rev. E* **70**(4), 041902, 2004. DOI: 10.1103/PhysRevE.70.041902.

[157] A. J. McKane and T. J. Newman. Predator-prey cycles from resonant amplification of demographic stochasticity. *Phys. Rev. Lett.* **94**(21), 218102, 2005. DOI: 10.1103/PhysRevLett.94.218102.

[158] M. Mobilia, I. T. Georgiev and U. C. Täuber. Phase transitions and spatio-temporal fluctuations in stochastic lattice Lotka-Volterra models. *J. Stat. Phys.* **128**(1), 447–483, 2006. DOI: 10.1007/s10955-006-9146-3.

[159] T. Reichenbach, M. Mobilia and E. Frey. Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model. *Phys. Rev. E* **74**(5), 051907, 2006. DOI: 10.1103/PhysRevE.74.051907.

[160] I. Volkov, J. R. Banavar, S. P. Hubbell and A. Maritan. Patterns of relative species abundance in rainforests and coral reefs. *Nature* **450**(7166), 45–49, 2007. DOI: 10.1038/nature06197.

[161] T. Butler and D. Reynolds. Predator-prey quasicycles from a path-integral formalism. *Phys. Rev. E* **79**(3), 032901, 2009. DOI: 10.1103/PhysRevE.79.032901.

[162] T. Butler and N. Goldenfeld. Robust ecological pattern formation induced by demographic noise. *Phys. Rev. E* **80**(3), 030902(R), 2009. DOI: 10.1103/PhysRevE.80.030902.

[163] J.-M. Park, E. Muñoz and M. W. Deem. Quasispecies theory for finite populations. *Phys. Rev. E* **81**(1), 011902, 2010. DOI: 10.1103/PhysRevE.81.011902.

[164] A. E. Noble, A. Hastings and W. F. Fagan. Multivariate Moran process with Lotka-Volterra phenomenology. *Phys. Rev. Lett.* **107**(22), 228101, 2011. DOI: 10.1103/PhysRevLett.107.228101.

[165] A. Dobrinevski and E. Frey. Extinction in neutrally stable stochastic Lotka-Volterra models. *Phys. Rev. E* **85**(5), 051903, 2012. DOI: 10.1103/PhysRevE.85.051903.

[166] A. J. Black and A. J. McKane. Stochastic formulation of ecological models and their applications. *Trends Ecol. Evol.* **27**(6), 337–345, 2012. DOI: 10.1016/j.tree.2012.01.014.

[167] U. C. Täuber. Population oscillations in spatial stochastic Lotka–Volterra models: a field-theoretic perturbational analysis. *J. Phys. A: Math. Theor.* **45**(40), 405002, 2012. DOI: 10.1088/1751-8113/45/40/405002.

[168] H.-Y. Shih and N. Goldenfeld. Path-integral calculation for the emergence of rapid evolution from demographic stochasticity. *Phys. Rev. E* **90**(5), 050702(R), 2014. DOI: 10.1103/PhysRevE.90.050702.

[169] J. Cremer, A. Melbinger and E. Frey. Growth dynamics and the evolution of cooperation in microbial populations. *Sci. Rep.* **2**, 281, 2012. DOI: 10.1038/srep00281.

[170] M. Reiter, S. Rulands and E. Frey. Range expansion of heterogeneous populations. *Phys. Rev. Lett.* **112**(14), 148103, 2014. DOI: 10.1103/PhysRevLett.112.148103.

[171] K. Wienand, M. Lechner, F. Becker, H. Jung and E. Frey. Non-selective evolution of growing populations. *PLoS ONE* **10**(8), e0134300, 2015. DOI: 10.1371/journal.pone.0134300.

[172] T. Reichenbach, M. Mobilia and E. Frey. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. *Nature* **448**(7157), 1046–1049, 2007. DOI: 10.1038/nature06095.

[173] T. Reichenbach, M. Mobilia and E. Frey. Noise and correlations in a spatial population model with cyclic competition. *Phys. Rev. Lett.* **99**(23), 238105, 2007. DOI: 10.1103/PhysRevLett.99.238105.

[174] T. Reichenbach, M. Mobilia and E. Frey. Self-organization of mobile populations in cyclic competition. *J. Theor. Biol.* **254**(2), 368–383, 2008. DOI: 10.1016/j.jtbi.2008.05.014.

[175] M. Mobilia and M. Assaf. Fixation in evolutionary games under non-vanishing selection. *Europhys. Lett.* **91**(1), 10002, 2010. DOI: 10.1209/0295-5075/91/10002.

[176] M. Assaf and M. Mobilia. Large fluctuations and fixation in evolutionary games. *J. Stat. Mech. Theor. Exp.* **2010**(09), P09009, 2010. DOI: 10.1088/1742-5468/2010/09/P09009.

[177] E. Frey. Evolutionary game theory: Theoretical concepts and applications to microbial communities. *Physica A* **389**(20), 4265–4298, 2010. DOI: 10.1016/j.physa.2010.02.047.

[178] A. Melbinger, J. Cremer and E. Frey. Evolutionary game theory in growing populations. *Phys. Rev. Lett.* **105**(17), 178101, 2010. DOI: 10.1103/PhysRevLett.105.178101.

[179] J. Cremer, A. Melbinger and E. Frey. Evolutionary and population dynamics: A coupled approach. *Phys. Rev. E* **84**(5), 051921, 2011. DOI: 10.1103/PhysRevE.84.051921.

[180] A. Traulsen, J. C. Claussen and C. Hauert. Stochastic differential equations for evolutionary dynamics with demographic noise and mutations. *Phys. Rev. E* **85**(4), 041901, 2012. DOI: 10.1103/PhysRevE.85.041901.

[181] A. J. Black, A. Traulsen and T. Galla. Mixing times in evolutionary game dynamics. *Phys. Rev. Lett.* **109**(2), 028101, 2012. DOI: 10.1103/PhysRevLett.109.028101.

[182] A. Melbinger, J. Cremer and E. Frey. The emergence of cooperation from a single mutant during microbial life cycles. *J. R. Soc. Interface* **12**(108), 20150171, 2015. DOI: 10.1098/rsif.2015.0171.

[183] J. Krug and H. Spohn. Kinetic roughening of growing surfaces. *Solids Far from Equilibrium*. Ed. by C. Godrèche. Collection Alea-Saclay: Monographs and Texts in Statistical Physics. Cambridge University Press, Cambridge, 1991,

[184] W. Weidlich. Physics and social science - the approach of synergetics. *Phys. Rep.* **204**(1), 1–163, 1991. DOI: 10.1016/0370-1573(91)90024-G.

[185] W. Weidlich and M. Braun. The master equation approach to nonlinear economics. *J. Evol. Econ.* **2**(3), 233–265, 1992. DOI: 10.1007/bf01202420.

[186] T. Lux. Herd behaviour, bubbles and crashes. *Econ. J.* **105**(431), 881–896, 1995. DOI: 10.2307/2235156.

[187] C. Castellano, S. Fortunato and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**(2), 591–646, 2009. DOI: 10.1103/RevModPhys.81.591.

[188] D. Gross, J. F. Shortle, J. M. Thompson and C. M. Harris. *Fundamentals of Queueing Theory*. 4th edition. John Wiley & Sons, Hoboken, NJ, 2008.

[189] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *J. Phys. I France* **2**(12), 2221–2229, 1992. DOI: 10.1051/jp1:1992277.

[190] D. Helbing. Traffic and related self-driven many-particle systems. *Rev. Mod. Phys.* **73**(4), 1067–1141, 2001. DOI: 10.1103/RevModPhys.73.1067.

[191] T. Reichenbach, T. Franosch and E. Frey. Exclusion processes with internal states. *Phys. Rev. Lett.* **97**(5), 050603, 2006. DOI: 10.1103/PhysRevLett.97.050603.

[192] T. Kriecherbauer and J. Krug. A pedestrian's view on interacting particle systems, KPZ universality and random matrices. *J. Phys. A: Math. Theor.* **43**(40), 403001, 2010. DOI: 10.1088/1751-8113/43/40/403001.

[193] T. Chou, K. Mallick and R. K. P. Zia. Non-equilibrium statistical mechanics: from a paradigmatic model to biological transport. *Rep. Prog. Phys.* **74**(11), 116601, 2011. DOI: 10.1088/0034-4885/74/11/116601.

[194] D. Vorberg, W. Wustmann, H. Schomerus, R. Ketzmerick and A. Eckardt. Nonequilibrium steady states of ideal bosonic and fermionic quantum gases. *Phys. Rev. E* **92**(6), 062119, 2015. DOI: 10.1103/PhysRevE.92.062119.

[195] M. R. Evans and B. Waclaw. Condensation in stochastic mass transport models: beyond the zero-range process. *J. Phys. A: Math. Theor.* **47**(9), 095001, 2014. DOI: 10.1088/1751-8113/47/9/095001.

[196] S. Grosskinsky, F. Redig and K. Vafayi. Condensation in the inclusion process and related models. *J. Stat. Phys.* **142**(5), 952–974, 2011. DOI: 10.1007/s10955-011-0151-9.

[197] C. Giardinà, F. Redig and K. Vafayi. Correlation inequalities for interacting particle systems with duality. *J. Stat. Phys.* **141**(2), 242–263, 2010. DOI: 10.1007/s10955-010-0055-0.

[198] C. Giardinà, J. Kurchan and F. Redig. Duality and exact correlations for a model of heat conduction. *J. Math. Phys.* **48**(3), 033301, 2007. DOI: 10.1063/1.2711373.

[199] F. Spitzer. Interaction of Markov processes. *Adv. Math.* **5**(2), 246–290, 1970. DOI: 10.1016/0001-8708(70)90034-4.

[200] H. Spohn. *Large Scale Dynamics of Interacting Particles*. Texts and Monographs in Physics. Springer, Berlin, Heidelberg, 1991. DOI: 10.1007/978-3-642-84371-6.

[201] T. M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Grundlehren der mathematischen Wissenschaften. Vol. 324. Springer, Berlin, Heidelberg, 1999. DOI: 10.1007/978-3-662-03990-8.

[202] G. M. Schütz. Exactly solvable models for many-body systems far from equilibrium. *Phase Transitions and Critical Phenomena*. Ed. by C. Domb and J. L. Lebowitz. Vol. 19. Academic Press, San Diego, CA, 2000,

[203] S. A. Janowsky and J. L. Lebowitz. Finite-size effects and shock fluctuations in the asymmetric simple-exclusion process. *Phys. Rev. A* **45**(2), 618–625, 1992. DOI: 10.1103/PhysRevA.45.618.

[204] B. Derrida, E. Domany and D. Mukamel. An exact solution of a one-dimensional asymmetric exclusion model with open boundaries. *J. Stat. Phys.* **69**(3-4), 667–687, 1992. DOI: 10.1007/BF01050430.

[205] B. Derrida, M. R. Evans, V. Hakim and V. Pasquier. Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A: Math. Gen.* **26**(7), 1493–1517, 1993. DOI: 10.1088/0305-4470/26/7/011.

[206] G. Schütz and E. Domany. Phase transitions in an exactly soluble one-dimensional exclusion process. *J. Stat. Phys.* **72**(1-2), 277–296, 1993. DOI: 10.1007/BF01048050.

[207] G. M. Schütz. Exact solution of the master equation for the asymmetric exclusion process. *J. Stat. Phys.* **88**(1-2), 427–445, 1997. DOI: 10.1007/BF02508478.

[208] K. Johansson. Shape fluctuations and random matrices. *Commun. Math. Phys* **209**(2), 437–476, 2000. DOI: 10.1007/s002200050027.

[209] V. B. Priezzhev. Exact nonstationary probabilities in the asymmetric exclusion process on a ring. *Phys. Rev. Lett.* **91**(5), 050601, 2003. DOI: 10.1103/PhysRevLett.91.050601.

[210] V. B. Priezzhev. Non-stationary probabilities for the asymmetric exclusion process on a ring. *Pramana J. Phys.* **64**(6), 915–925, 2005. DOI: 10.1007/BF02704153.

[211] C. A. Tracy and H. Widom. Integral formulas for the asymmetric simple exclusion process. *Commun. Math. Phys* **279**(3), 815–844, 2008. DOI: 10.1007/s00220-008-0443-3.

[212] C. A. Tracy and H. Widom. A Fredholm determinant representation in ASEP. *J. Stat. Phys.* **132**(2), 291–300, 2008. DOI: 10.1007/s10955-008-9562-7.

[213] C. A. Tracy and H. Widom. Asymptotics in ASEP with step initial condition. *Commun. Math. Phys* **290**(1), 129–154, 2009a. DOI: 10.1007/s00220-009-0761-0.

[214] C. A. Tracy and H. Widom. Total current fluctuations in the asymmetric simple exclusion process. *J. Math. Phys.* **50**(9), 095204, 2009. DOI: 10.1063/1.3136630.

[215] A. Melbinger, L. Reese and E. Frey. Microtubule length regulation by molecular motors. *Phys. Rev. Lett.* **108**(25), 258104, 2012. DOI: 10.1103/PhysRevLett.108.258104.

[216] Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F. U. Hartl, M. J. Kerner and D. Frishman. Protein abundance profiling of the Escherichia coli cytosol. *BMC Genomics* **9**, 102, 2008. DOI: 10.1186/1471-2164-9-102.

[217] C. R. Doering, K. V. Sargsyan and L. M. Sander. Extinction times for birth-death processes: exact results, continuum asymptotics, and the failure of the Fokker-Planck approximation. *Multiscale Model. Simul.* **3**(2), 283–299, 2005. DOI: 10.1137/030602800.

[218] J. R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1998. Chap. 2–3. DOI: 10.2277/0521633966.

[219] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.* **45**(1), 3–49, 2003. DOI: 10.1137/S00361445024180.

[220] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Rev.* **20**(4), 801–836, 1978. DOI: 10.1137/1020098.

[221] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.* **124**(4), 044104–1–044104–13, 2006. DOI: 10.1063/1.2145882.

[222] K. Burrage, M. Hegland, S. MacNamara and R. B. Sidje. A Krylov-based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems. *Proc. of The A. A. Markov 150th Anniversary Meeting*. Ed. by A. N. Langville and W. J. Stewart. Charleston, SC, 12-14 June 2006, pp. 21–37.

[223] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.* **7**(4), 649–673, 1954. DOI: 10.1002/cpa.3160070404.

[224] S. Blanes, F. Casas, J. A. Oteo and J. Ros. The Magnus expansion and some of its applications. *Phys. Rep.* **470**(5-6), 151–238, 2009. DOI: 10.1016/j.physrep.2008.11.001.

[225] A. Klenke. *Probability Theory - A Comprehensive Course*. 2nd edition. Universitext. Springer, London, 2014. Chap. 17.3. DOI: 10.1007/978-1-4471-5361-0.

[226] A. Jensen. Markoff chains as an aid in the study of Markoff processes. *Scand. Actuar. J.* **1953**(Supp. 1), 87–91, 1953. DOI: 10.1080/03461238.1953.10419459.

[227] A. Hellander. Efficient computation of transient solutions of the chemical master equation based on uniformization and quasi-Monte Carlo. *J. Chem. Phys.* **128**(15), 154109, 2008. DOI: 10.1063/1.2897976.

[228] F. Didier, T. A. Henzinger, M. Mateescu and V. Wolf. Fast adaptive uniformization of the chemical master equation. *High Performance Computational Systems Biology, HIBI '09*. IEEE, Trento, 14-16 Oct. 2009, pp. 118–127. DOI: 10.1109/HiBi.2009.23.

[229] D. Gross and D. R. Miller. The randomization technique as a modeling tool and solution procedure for transient Markov processes. *Oper. Res.* **32**(2), 343–361, 1984. DOI: 10.1287/opre.32.2.343.

[230] A. Reibman and K. Trivedi. Numerical transient analysis of Markov-models. *Comput. Opns Res.* **15**(1), 19–36, 1988. DOI: 10.1016/0305-0548(88)90026-3.

[231] A. P. A. van Moorsel and W. H. Sanders. Adaptive uniformization. *Comm. Statist. Stochastic Models* **10**(3), 619–647, 1994. DOI: 10.1080/15326349408807313.

[232] W. Feller. On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.* **48**, 488–515, 1940. DOI: 10.2307/1990095.

[233] J. L. Doob. Markoff chains–denumerable case. *Trans. Amer. Math. Soc.* **58**(3), 455–473, 1945. DOI: 10.2307/1990339.

[234] J. L. Doob. Topics in the theory of Markoff chains. *Trans. Amer. Math. Soc.* **52**, 37–64, 1942. DOI: 10.1090/S0002-9947-1942-0006633-7.

[235] D. G. Kendall. Stochastic processes and population growth. *J. R. Stat. Soc. B* **11**(2), 230–282, 1949.

[236] D. G. Kendall. An artificial realization of a simple "birth-and-death" process. *J. R. Stat. Soc. B* **12**(1), 116–119, 1950.

[237] M. S. Bartlett. Stochastic processes or the statistics of change. *J. R. Stat. Soc. C* **2**(1), 44–64, 1953. DOI: 10.2307/2985327.

[238] A. B. Bortz, M. H. Kalos and J. L. Lebowitz. A new algorithm for Monte Carlo simulation of Ising spin systems. *J. Comput. Phys.* **17**(1), 10–18, 1975. DOI: 10.1016/0021-9991(75)90060-1.

[239] U. Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**(12), 126001, 2012. DOI: 10.1088/0034-4885/75/12/126001.

[240] N. Empacher. *Die Wegintegrallösung der Mastergleichung*. PhD thesis. University of Stuttgart, 1992.

[241] D. Helbing. A contracted path integral solution of the discrete master equation. *Phys. Lett. A* **195**(2), 128–134, 1994. DOI: 10.1016/0375-9601(94)90085-X.

[242] D. Helbing and R. Molini. Occurence probabilities of stochastic paths. *Phys. Lett. A* **212**(3), 130–137, 1996. DOI: 10.1016/0375-9601(96)00010-2.

[243] S. X. Sun. Path summation formulation of the master equation. *Phys. Rev. Lett.* **96**(21), 210602, 2006. DOI: 10.1103/PhysRevLett.96.210602.

[244] B. Harland and S. X. Sun. Path ensembles and path sampling in nonequilibrium stochastic systems. *J. Chem. Phys.* **127**(10), 104103, 2007. DOI: 10.1063/1.2775439.

[245] A. Jackson and S. Pigolotti. Statistics of trajectories in two-state master equations. *Phys. Rev. E* **79**(2), 021121, 2009. DOI: 10.1103/PhysRevE.79.021121.

[246] H. Jasiulewicz and W. Kordecki. Convolutions of Erlang and of Pascal distributions with applications to reliability. *Demonstratio Math.* **36**(1), 231–238, 2003.

[247] M. Akkouchi. On the convolution of exponential distributions. *J. Chungcheong Math. Soc.* **21**(4), 2008.

[248] H. E. Daniels. Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25**(4), 631–650, 1954. DOI: 10.1214/aoms/1177728652.

[249] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A* **104**(9), 1876–1889, 2000. DOI: 10.1021/jp993732q.

[250] A. Slepoy, A. P. Thompson and S. J. Plimpton. A constant-time kinetic Monte Carlo algorithm for simulation of large biochemical reaction networks. *J. Chem. Phys.* **128**(20), 205101, 2008. DOI: 10.1063/1.2919546.

[251] D. T. Gillespie, A. Hellander and L. R. Petzold. Perspective: Stochastic algorithms for chemical kinetics. *J. Chem. Phys.* **138**(17), 170901, 2013. DOI: 10.1063/1.4801941.

[252] M. Tomita *et al.* E-CELL: software environment for whole-cell simulation. *Bioinformatics* **15**(1), 72–84, 1999. DOI: 10.1093/bioinformatics/15.1.72.

[253] D. Adalsteinsson, D. McMillen and T. C. Elston. Biochemical network stochastic simulator (BioNetS): software for stochastic modeling of biochemical networks. *BMC Bioinform.* **5**(1), 24, 2004. DOI: 10.1186/1471-2105-5-24.

[254] M. Ander, P. Beltrao, B. Di Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomás-Oliveira and L. Serrano. SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *IEE P. Syst. Biol.* **1**(1), 129–138, 2004. DOI: 10.1049/sb:20045017.

[255] J. Hattne, D. Fange and J. Elf. Stochastic reaction-diffusion simulation with MesoRD. *Bioinformatics* **21**(12), 2923–2924, 2005. DOI: 10.1093/bioinformatics/bti431.

[256] K. R. Sanft, S. Wu, M. Roh, J. Fu, R. K. Lim and L. R. Petzold. StochKit2: software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics* **27**(17), 2457–2458, 2011. DOI: 10.1093/bioinformatics/btr401.

[257] I. Hepburn, W. Chen, S. Wils and E. De Schutter. STEPS: efficient simulation of stochastic reaction–diffusion models in realistic morphologies. *BMC Syst. Biol.* **6**, 36, 2012. DOI: 10.1186/1752-0509-6-36.

[258] B. Drawert, S. Engblom and A. Hellander. URDME: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC Syst. Biol.* **6**, 76, 2012. DOI: 10.1186/1752-0509-6-76.

[259] L. Petzold, C. Krintz and P. Lötstedt. StochSS: Stochastic Simulation Service. URL: http://www.stochss.org/.

[260] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* **115**(4), 1716–1733, 2001. DOI: 10.1063/1.1378322.

[261] D. T. Gillespie and L. R. Petzold. Improved leap-size selection for accelerated stochastic simulation. *J. Chem. Phys.* **119**(16), 8229, 2003. DOI: 10.1063/1.1613254.

[262] M. Rathinam, L. R. Petzold, Y. Cao and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chem. Phys.* **119**(24), 12784, 2003. DOI: 10.1063/1.1627296.

[263] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. *J. Chem. Phys.* **121**(21), 10356, 2004. DOI: 10.1063/1.1810475.

[264] A. Chatterjee, D. G. Vlachos and M. A. Katsoulakis. Binomial distribution based τ-leap accelerated stochastic simulation. *J. Chem. Phys.* **122**(2), 024112, 2005. DOI: 10.1063/1.1833357.

[265] Y. Cao, D. T. Gillespie and L. R. Petzold. Avoiding negative populations in explicit Poisson tau-leaping. *J. Chem. Phys.* **123**(5), 054104, 2005. DOI: 10.1063/1.1992473.

[266] Y. Cao, D. T. Gillespie and L. R. Petzold. Efficient step size selection for the tau-leaping simulation method. *J. Chem. Phys.* **124**(4), 044109, 2006. DOI: 10.1063/1.2159468.

[267] A. Auger, P. Chatelain and P. Koumoutsakos. R-leaping: Accelerating the stochastic simulation algorithm by reaction leaps. *J. Chem. Phys.* **125**(8), 084103, 2006. DOI: 10.1063/1.2218339.

[268] Y. Cao, D. T. Gillespie and L. R. Petzold. Adaptive explicit-implicit tau-leaping method with automatic tau selection. *J. Chem. Phys.* **126**(22), 224101, 2007. DOI: 10.1063/1.2745299.

[269] D. F. Anderson. Incorporating postleap checks in tau-leaping. *J. Chem. Phys.* **128**(5), 054103, 2008. DOI: 10.1063/1.2218339.

[270] T. Lu, D. Volfson, L. Tsimring and J. Hasty. Cellular growth and division in the Gillespie algorithm. *IEE P. Syst. Biol.* **1**(1), 121–128, 2004. DOI: 10.1049/sb:20045016.

[271] D. F. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J. Chem. Phys.* **127**(21), 214107-1–214107-10, 2007. DOI: 10.1063/1.2799998.

[272] E. Roberts, S. Be'er, C. Bohrer, R. Sharma and M. Assaf. Dynamics of simple gene-network motifs subject to extrinsic fluctuations. *Phys. Rev. E* **92**(6), 062717, 2015. DOI: 10.1103/PhysRevE.92.062717.

[273] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery. *Numerical Recipes - The Art of Scientific Computing*. 3rd edition. Cambridge University Press, Cambridge, 2007. Chap. 17.5, pp. 931–942.

[274] J. E. Moyal. Stochastic processes and statistical physics. *J. R. Stat. Soc. B* **11**(2), 150–210, 1949.

[275] N. G. van Kampen. A power series expansion of the master equation. *Can. J. Phys.* **39**(4), 551–567, 1961. DOI: 10.1139/p61-056.

[276] D. Schnoerr, G. Sanguinetti and R. Grima. Approximation and inference methods for stochastic biochemical kinetics - a tutorial review. arXiv:1608.06582, pp. 1–69, 2016.

[277] L. A. Goodman. Population growth of the sexes. *Biometrics* **9**(2), 212–225, 1953. DOI: 10.2307/3001852.

[278] P. Whittle. On the use of the normal approximation in the treatment of stochastic processes. *J. R. Stat. Soc. B* **19**(2), 268–281, 1957.

[279] M. J. Keeling. Multiplicative moments and measures of persistence in ecology. *J. Theor. Biol.* **205**(2), 269–281, 2000. DOI: 10.1006/jtbi.2000.2066.

[280] I. Nåsell. An extension of the moment closure method. *Theor. Popul. Biol.* **64**(2), 233–239, 2003. DOI: 10.1016/S0040-5809(03)00074-1.

[281] C. H. Lee, K.-H. Kim and P. Kim. A moment closure method for stochastic reaction networks. *J. Chem. Phys.* **130**(13), 134107, 2009. DOI: 10.1063/1.3103264.

[282] P. Smadbeck and Y. N. Kaznessis. A closure scheme for chemical master equations. *Proc. Natl. Acad. Sci. USA* **110**(35), 14261–14265, 2013. DOI: 10.1073/pnas.1306481110.

[283] D. Schnoerr, G. Sanguinetti and R. Grima. Validity conditions for moment closure approximations in stochastic chemical kinetics. *J. Chem. Phys.* **141**(8), 084103, 2014. DOI: 10.1063/1.4892838.

[284] E. Lakatos, A. Ale, P. D. W. Kirk and M. P. H. Stumpf. Multivariate moment closure techniques for stochastic kinetic models. *J. Chem. Phys.* **143**(9), 094107, 2015. DOI: 10.1063/1.4929837.

[285] D. Schnoerr, G. Sanguinetti and R. Grima. Comparison of different moment-closure approximations for stochastic chemical kinetics. *J. Chem. Phys.* **143**(18), 185101, 2015. DOI: 10.1063/1.4934990.

[286] F. Di Patti, S. Azaele, J. R. Banavar and A. Maritan. System size expansion for systems with an absorbing state. *Phys. Rev. E* **83**(1), 010102(R), 2011. DOI: 10.1103/PhysRevE.83.010102.

[287] P. A. M. Dirac. The Lagrangian in quantum mechanics. *Phys. Z. Sowjetunion* **3**, 64–72, 1933.

[288] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20**(2), 367–387, 1948. DOI: 10.1103/RevModPhys.20.367.

[289] R. P. Feynman and A. R. Hibbs. *Quantum Mechanics and Path Integrals: Emended Edition*. Ed. by D. F. Styer. Dover Books on Physics. Dover Publications, New York, NY, 2010.

[290] P. A. M. Dirac. *The Principles of Quantum Mechanics*. 4th edition. Oxford University Press, Oxford, 1958. Chap. VI.

[291] D. C. Mattis and M. L. Glasser. The uses of quantum field theory in diffusion-limited reactions. *Rev. Mod. Phys.* **70**(3), 979–1001, 1998. DOI: 10.1103/RevModPhys.70.979.

[292] A. Altland and B. D. Simons. *Condensed Matter Field Theory*. 2nd edition. Cambridge University Press, Cambridge, 2010. DOI: 10.1017/CBO9780511789984.

[293] J. C. Baez and J. Biamonte. Quantum Techniques for Stochastic Mechanics. arXiv:1209.3632, pp. 1–268, 2015.

[294] M. E. Fisher. Renormalization group theory: Its basis and formulation in statistical physics. *Rev. Mod. Phys.* **70**(2), 653–681, 1998. DOI: 10.1103/RevModPhys.70.653.

[295] J. Cardy. Renormalisation group approach to reaction-diffusion problems. *Mathematical Beauty of Physics*. Ed. by J.-B. Zuber. Vol. 24. Adv. Ser. Math. Phys. arXiv:cond-mat/9607163. 1997, pp. 113–128.

[296] L. Peliti. Renormalisation of fluctuation effects in the $A + A \to A$ reaction. *J. Phys. A: Math. Gen.* **19**(6), L365–L367, 1986. DOI: 10.1088/0305-4470/19/6/012.

[297] M. Howard and J. Cardy. Fluctuation effects and multiscaling of the reaction-diffusion front for $A + B \to \emptyset$. *J. Phys. A: Math. Gen.* **28**(13), 3599–3621, 1995. DOI: 10.1088/0305-4470/28/13/007.

[298] M. Howard. Fluctuation kinetics in a multispecies reaction-diffusion system. *J. Phys. A: Math. Gen.* **29**(13), 3437–3460, 1996. DOI: 10.1088/0305-4470/29/13/016.

[299] K. Oerding. The $A + B \to \emptyset$ annihilation reaction in a quenched random velocity field. *J. Phys. A: Math. Gen.* **29**(2), 7051–7065, 1996. DOI: 10.1088/0305-4470/29/22/009.

[300] P.-A. Rey and M. Droz. A renormalization group study of a class of reaction-diffusion models, with particles input. *J. Phys. A: Math. Gen.* **30**(4), 1101–1114, 1997. DOI: 10.1088/0305-4470/30/4/013.

[301] T. Sasamoto, S. Mori and M. Wadati. Universal properties of the $mA + nB \to \emptyset$ diffusion-limited reaction. *Physica A* **247**(1-4), 357–378, 1997. DOI: 10.1016/s0378-4371(97)00387-7.

[302] F. v. Wijland, K. Oerding and H. J. Hilhorst. Wilson renormalization of a reaction-diffusion process. *Physica A* **251**(1-2), 179–201, 1998. DOI: 10.1016/S0378-4371(97)00603-1.

[303] U. C. Täuber, M. J. Howard and H. Hinrichsen. Multicritical behavior in coupled directed percolation processes. *Phys. Rev. Lett.* **80**(10), 2165–2168, 1998. DOI: 10.1103/PhysRevLett.80.2165.

[304] H. Hinrichsen and M. Howard. A model for anomalous directed percolation. *Eur. Phys. J. B* **7**(4), 635–643, 1999. DOI: 10.1007/s100510050656.

[305] Y. Y. Goldschmidt, H. Hinrichsen, M. Howard and U. C. Täuber. Nonequilibrium critical behavior in unidirectionally coupled stochastic processes. *Phys. Rev. E* **59**(6), 6381–6408, 1999. DOI: 10.1103/PhysRevE.59.6381.

[306] M. Hnatich and J. Honkonen. Velocity-fluctuation-induced anomalous kinetics of the $A + A \to \emptyset$ reaction. *Phys. Rev. E* **61**(4), 3904–3911, 2000. DOI: 10.1103/PhysRevE.61.3904.

[307] D. C. Vernon. Long range hops and the pair annihilation reaction $A + A \to \emptyset$: renormalization group and simulation. *Phys. Rev. E* **68**(4), 041103, 2003. DOI: 10.1103/PhysRevE.68.041103.

[308] A. A. Winkler and E. Frey. Validity of the law of mass action in three-dimensional coagulation processes. *Phys. Rev. Lett.* **108**(10), 108301, 2012. DOI: 10.1103/PhysRevLett.108.108301.

[309] I. Homrighausen, A. A. Winkler and E. Frey. Fluctuation effects in the pair-annihilation process with Lévy dynamics. *Phys. Rev. E* **88**(1), 012111, 2013. DOI: 10.1103/PhysRevE.88.012111.

[310] M. Droz and L. Sasvari. Renormalization-group approach to simple reaction-diffusion phenomena. *Phys. Rev. E* **48**(4), R2343–R2346, 1993. DOI: 10.1103/PhysRevE.48.R2343.

[311] B. P. Lee and J. Cardy. Scaling of reaction zones in the $A + B \to 0$ diffusion-limited reaction. *Phys. Rev. E* **50**(5), R3287–R3290, 1994. DOI: 10.1103/PhysRevE.50.R3287.

[312] M. J. Howard and G. T. Barkema. Shear flows and segregation in the reaction $A + B \to \emptyset$. *Phys. Rev. E* **53**(6), 5949–5956, 1996. DOI: 10.1103/PhysRevE.53.5949.

[313] Z. Konkoli, H. Johannesson and B. P. Lee. Fluctuation effects in steric reaction-diffusion systems. *Phys. Rev. E* **59**(4), R3787–R3790, 1999. DOI: 10.1103/PhysRevE.59.R3787.

[314] R. Dickman and R. Vidigal. Path-integral representation for a stochastic sandpile. *J. Phys. A: Math. Gen.* **35**(34), 7269–7285, 2002. DOI: 10.1088/0305-4470/35/34/303.

[315] H. J. Hilhorst, O. Deloubrière, M. J. Washenberger and U. C. Täuber. Segregation in diffusion-limited multispecies pair annihilation. *J. Phys. A: Math. Gen.* **37**(28), 7063–7093, 2004. DOI: 10.1088/0305-4470/37/28/001.

[316] H. K. Janssen, F. v. Wijland, O. Deloubrière and U. C. Täuber. Pair contact process with diffusion: Failure of master equation field theory. *Phys. Rev. E* **70**(5), 056114, 2004. DOI: 10.1103/PhysRevE.70.056114.

[317] S. Whitelam, L. Berthier and J. P. Garrahan. Renormalization group study of a kinetically constrained model for strong glasses. *Phys. Rev. E* **71**(2), 026128, 2005. DOI: 10.1103/PhysRevE.71.026128.

[318] M. Hnatič, J. Honkonen and T. Lučivjanský. Two-loop calculation of anomalous kinetics of the reaction $A + A \to \emptyset$ in randomly stirred fluid. *Eur. Phys. J. B* **86**(5), 214, 2013. DOI: 10.1140/epjb/e2013-30982-9.

[319] L. Canet, B. Delamotte, O. Deloubrière and N. Wschebor. Nonperturbative renormalization-group study of reaction-diffusion processes. *Phys. Rev. Lett.* **92**(19), 195703, 2004. DOI: 10.1103/PhysRevLett.92.195703.

[320] L. Canet, H. Chaté and B. Delamotte. Quantitative phase diagrams of branching and annihilating random walks. *Phys. Rev. Lett.* **92**(25), 255703, 2004. DOI: 10.1103/PhysRevLett.92.255703.

[321] L. Canet, H. Chaté, B. Delamotte, I. Dornic and M. A. Muñoz. Nonperturbative fixed point in a nonequilibrium phase transition. *Phys. Rev. Lett.* **95**(10), 100601, 2005. DOI: 10.1103/PhysRevLett.95.100601.

[322] V. Elgart and A. Kamenev. Classification of phase transitions in reaction-diffusion models. *Phys. Rev. E* **74**(4), 041101, 2006. DOI: 10.1103/PhysRevE.74.041101.

[323] M. A. Buice and J. D. Cowan. Field-theoretic approach to fluctuation effects in neural networks. *Phys. Rev. E* **75**(5), 051919, 2007. DOI: 10.1103/PhysRevE.75.051919.

[324] M. A. Buice and J. D. Cowan. Statistical mechanics of the neocortex. *Prog. Biophys. Mol. Biol.* **99**(2–3), 53–86, 2009. DOI: 10.1016/j.pbiomolbio.2009.07.003.

[325] M. A. Buice and C. C. Chow. Beyond mean field theory: statistical field theory for neural networks. *J. Stat. Mech. Theor. Exp.* **2013**(03), P03003, 2013. DOI: 10.1088/1742-5468/2013/03/P03003.

[326] E. Bettelheim, O. Agam and N. M. Shnerb. "Quantum phase transitions" in classical nonequilibrium processes. *Physica E* **9**(3), 600–608, 2001. DOI: 10.1016/S1386-9477(00)00268-X.

[327] U. C. Täuber. Stochastic population oscillations in spatial predator-prey models. *J. Phys.: Conf. Ser.* **319**, 012019, 2011. DOI: 10.1088/1742-6596/319/1/012019.

[328] B. Zhang and P. G. Wolynes. Stem cell differentiation as a many-body problem. *Proc. Natl. Acad. Sci. USA* **111**(28), 10185–10190, 2014. DOI: 10.1073/pnas.1408561111.

[329] W. Feller. On the theory of stochastic processes, with particular reference to applications. *Proc. [First] Berkeley Symp. on Math. Statist. and Prob.* University of California Press, Berkeley, CA, 1949, pp. 403–432.

[330] A. M. Walczak, A. Mugler and C. H. Wiggins. Analytic methods for modeling stochastic regulatory networks. *Computational Modeling of Signaling Networks*. Ed. by X. Liu and M. D. Betterton. Vol. 880. Methods in Molecular Biology. Humana Press, New York, NY, 2012, pp. 273–322. DOI: 10.1007/978-1-61779-833-7_13.

[331] J. G. Skellam. The frequency distribution of the difference between two Poisson variates belonging to different populations. *J. R. Stat. Soc.* **109**(3), 296, 1946. DOI: 10.2307/2981372.

[332] H. Takayasu and A. Y. Tretyakov. Extinction, survival, and dynamical phase transition of branching annihilating random walk. *Phys. Rev. Lett.* **68**(20), 3060–3063, 1992. DOI: 10.1103/PhysRevLett.68.3060.

[333] I. Jensen. Critical behavior of branching annihilating random walks with an odd number of offsprings. *Phys. Rev. E* **47**(1), R1–R4, 1993. DOI: 10.1103/PhysRevE.47.R1.

[334] L. E. Ballentine. *Quantum Mechanics - A Modern Development*. World Scientific Publishing Company, Singapore, 1998. Chap. 6.2.

[335] M. E. Peskin and D. V. Schroeder. *An Introduction To Quantum Field Theory*. Westview Press, Boulder, CO, 1995.

[336] J. Ohkubo. One-parameter extension of the Doi–Peliti formalism and its relation with orthogonal polynomials. *Phys. Rev. E* **86**(4), 042102, 2012. DOI: 10.1103/PhysRevE.86.042102.

[337] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark, eds. *NIST Handbook of Mathematical Functions*. Print companion to http://dlmf.nist.gov/. Cambridge University Press, New York, NY, 2010.

[338] K. J. Wiese. Coherent-state path integral versus coarse-grained effective stochastic equation of motion: from reaction diffusion to stochastic sandpiles. *Phys. Rev. E* **93**(4), 042117, 2016. DOI: 10.1103/PhysRevE.93.042117.

[339] R. B. Stinchcombe, M. D. Grynberg and M. Barma. Diffusive dynamics of deposition-evaporation systems, jamming, and broken symmetries in related quantum-spin models. *Phys. Rev. E* **47**(6), 4018–4036, 1993. DOI: 10.1103/PhysRevE.47.4018.

[340] M. D. Grynberg, T. J. Newman and R. B. Stinchcombe. Exact-solutions for stochastic adsorption-desorption models and catalytic surface processes. *Phys. Rev. E* **50**(2), 957–971, 1994. DOI: 10.1103/PhysRevE.50.957.

[341] S. Goldstein. On diffusion by discontinuous movements, and on the telegraph equation. *Q. J. Mech. Appl. Math.* **4**(2), 129–156, 1951. DOI: 10.1093/qjmam/4.2.129.

[342] M. Kac. A stochastic model related to the telegrapher's equation. *Rocky Mt. J. Math.* **4**(3) Reprinting of an article published in 1956, 497–510, 1974. DOI: 10.1216/RMJ-1974-4-3-497.

[343] F. A. Berezin. *The Method of Second Quantization*. Academic Press, Orlando, FL, 1966. Chap. 1.3.

[344] M. Combescure and D. Robert. *Coherent States and Applications in Mathematical Physics*. Theoretical and Mathematical Physics. Springer, Dordrecht, 2012. Chap. 11.2. DOI: 10.1007/978-94-007-0196-0.

[345] M. A. Aliev. On the description of the Glauber dynamics of one dimensional disordered Ising chain. *Physica A* **277**(3-4), 261–273, 2000. DOI: 10.1016/S0378-4371(99)00442-2.

[346] M. Mobilia, R. K. P. Zia and B. Schmittmann. Complete solution of the kinetics in a far-from-equilibrium Ising chain. *J. Phys. A: Math. Gen.* **37**(32), L407–L413, 2004. DOI: 10.1088/0305-4470/37/32/L03.

[347] F. C. Alcaraz, M. Droz, M. Henkel and V. Rittenberg. Reaction-diffusion processes, critical dynamics, and quantum chains. *Ann. Phys.* **230**(2), 250–302, 1994. DOI: 10.1006/aphy.1994.1026.

[348] G. Schütz and S. Sandow. Non-abelian symmetries of stochastic processes: derivation of correlation functions for random-vertex models and disordered-interacting-particle systems. *Phys. Rev. E* **49**(4), 2726–2741, 1994. DOI: 10.1103/PhysRevE.49.2726.

[349] G. M. Schütz. Reaction-diffusion processes of hard-core particles. *J. Stat. Phys.* **79**(1-2), 243–264, 1995. DOI: 10.1007/BF02179389.

[350] M. Henkel, E. Orlandini and J. Santos. Reaction-diffusion processes from equivalent integrable quantum chains. *Ann. Phys.* **259**(2), 163–231, 1997. DOI: 10.1006/aphy.1997.5712.

[351] E. Carlon, M. Henkel and U. Schollwöck. Density matrix renormalization group and reaction-diffusion processes. *Eur. Phys. J. B* **12**(1), 99–114, 1999. DOI: 10.1007/s100510050983.

[352] E. Carlon, M. Henkel and U. Schollwöck. Critical properties of the reaction-diffusion model $2A \to 3A$, $2A \to 0$. *Phys. Rev. E* **63**(3), 036101, 2001. DOI: 10.1103/PhysRevE.63.036101.

[353] S.-C. Park, D. Kim and J.-M. Park. Path-Integral formulation of stochastic processes for the exclusive particle systems. *Phys. Rev. E* **62**(6), 7642–7645, 2000. DOI: 10.1103/PhysRevE.62.7642.

[354] S.-C. Park and J.-M. Park. Generating function, path integral representation, and equivalence for stochastic exclusive particle systems. *Phys. Rev. E* **71**(2), 026113, 2005. DOI: 10.1103/PhysRevE.71.026113.

[355] J. Tailleur, J. Kurchan and V. Lecomte. Mapping out-of-equilibrium into equilibrium in one-dimensional transport models. *J. Phys. A: Math. Theor.* **41**(50), 505001, 2008. DOI: 10.1088/1751-8113/41/50/505001.

[356] R. Kubo, M. Matsuo and K. Kitahara. Fluctuation and relaxation of macrovariables. *J. Stat. Phys.* **9**(1), 51–96, 1973. DOI: 10.1007/BF01016797.

[357] Hu Gang. Stationary solution of master equations in the large-system-size limit. *Phys. Rev. A* **36**(12), 5782–5790, 1987. DOI: 10.1103/PhysRevA.36.5782.

[358] M. I. Dykman, E. Mori, J. Ross and P. M. Hunt. Large fluctuations and optimal paths in chemical kinetics. *J. Chem. Phys.* **100**(8), 5735–5750, 1994. DOI: 10.1063/1.467139.

[359] D. A. Kessler and N. M. Shnerb. Extinction rates for fluctuation-induced metastabilities: a real-space WKB approach. *J. Stat. Phys.* **127**(5), 861–886, 2007. DOI: 10.1007/s10955-007-9312-2.

[360] B. Meerson and P. V. Sasorov. Noise-driven unlimited population growth. *Phys. Rev. E* **78**(6), 060103(R), 2008. DOI: 10.1103/PhysRevE.78.060103.

[361] C. Escudero and A. Kamenev. Switching rates of multistep reactions. *Phys. Rev. E* **79**(4), 041149, 2009. DOI: 10.1103/PhysRevE.79.041149.

[362] M. Assaf and B. Meerson. Extinction of metastable stochastic populations. *Phys. Rev. E* **81**(2), 021116, 2010. DOI: 10.1103/PhysRevE.81.021116.

[363] O. Ovaskainen and B. Meerson. Stochastic models of population extinction. *Trends Ecol. Evol.* **25**(11), 643–652, 2010. DOI: 10.1016/j.tree.2010.07.009.

[364] B. Meerson and P. V. Sasorov. Extinction rates of established spatial populations. *Phys. Rev. E* **83**(1), 011129, 2011. DOI: 10.1103/PhysRevE.83.011129.

[365] N. R. Smith and B. Meerson. Extinction of oscillating populations. *Phys. Rev. E* **93**(3), 032109, 2016. DOI: 10.1103/PhysRevE.93.032109.

[366] M. Assaf and B. Meerson. WKB theory of large deviations in stochastic populations. arXiv:1612.01470, 2016.

[367] J. W. Turner and M. Malek-Mansour. On the absorbing zero boundary problem in birth and death processes. *Physica A* **93**(3-4), 517–525, 1978. DOI: 10.1016/0378-4371(78)90172-3.

[368] M. Assaf and B. Meerson. Spectral theory of metastability and extinction in a branching-annihilation reaction. *Phys. Rev. E* **75**(3), 031122, 2007. DOI: 10.1103/PhysRevE.75.031122.

[369] M. Assaf, B. Meerson and P. V. Sasorov. Large fluctuations in stochastic population dynamics: momentum-space calculations. *J. Stat. Mech. Theor. Exp.* **2010**(07), P07018, 2010. DOI: 10.1088/1742-5468/2010/07/P07018.

[370] L. C. Evans. *Partial Differential Equations: Second Edition*. 2nd ed. edition. Graduate Studies in Mathematics. Vol. 19. American Mathematical Society, 2010.

[371] I. B. Schwartz, E. Forgoston, S. Bianco and L. B. Shaw. Converging towards the optimal path to extinction. *J. R. Soc. Interface* **8**(65), 1699–1707, 2011. DOI: 10.1098/rsif.2011.0159.

[372] M. Assaf, A. Kamenev and B. Meerson. Population extinction risk in the aftermath of a catastrophic event. *Phys. Rev. E* **79**(1), 011127, 2009. DOI: 10.1103/PhysRevE.79.011127.

[373] B. Meerson and O. Ovaskainen. Immigration-extinction dynamics of stochastic populations. *Phys. Rev. E* **88**(1), 012124, 2013. DOI: 10.1103/PhysRevE.88.012124.

[374] Y. Lan, P. G. Wolynes and G. A. Papoian. A variational approach to the stochastic aspects of cellular signal transduction. *J. Chem. Phys.* **125**(12), 124106, 2006. DOI: 10.1063/1.2353835.

[375] J. Ohkubo. Approximation scheme for master equations: Variational approach to multivariate case. *J. Chem. Phys.* **129**(4), 044108, 2008. DOI: 10.1063/1.2957462.

[376] P. D. Drummond, T. G. Vaughan and A. J. Drummond. Extinction times in autocatalytic systems. *J. Phys. Chem. A* **114**(39), 10481–10491, 2010. DOI: 10.1021/jp104471e.

[377] P. D. Drummond, C. W. Gardiner and D. F. Walls. Quasiprobability methods for nonlinear chemical and optical systems. *Phys. Rev. A* **24**(2), 914–926, 1981. DOI: 10.1103/PhysRevA.24.914.

[378] D. Elderfield. Exact macroscopic dynamics in non-equilibrium chemical-systems. *J. Phys. A: Math. Gen.* **18**(11), 2049–2059, 1985. DOI: 10.1088/0305-4470/18/11/026.

[379] M. Droz and A. McKane. Equivalence between Poisson representation and Fock space formalism for birth-death processes. *J. Phys. A: Math. Gen.* **27**(13), L467–L474, 1994. DOI: 10.1088/0305-4470/27/13/002.

[380] P. D. Drummond. Gauge Poisson representations for birth/death master equations. *Eur. Phys. J. B* **38**(4), 617–634, 2004. DOI: 10.1140/epjb/e2004-00157-2.

[381] O. Deloubrière, L. Frachebourg, H. J. Hilhorst and K. Kitahara. Imaginary noise and parity conservation in the reaction $A + A \leftrightharpoons 0$. *Physica A* **308**(1-4), 135–147, 2002. DOI: 10.1016/S0378-4371(02)00548-4.

[382] K. G. Petrosyan and C.-K. Hu. Nonequilibrium Lyapunov function and a fluctuation relation for stochastic systems: Poisson-representation approach. *Phys. Rev. E* **89**(4), 042132, 2014. DOI: 10.1103/PhysRevE.89.042132.

[383] J. Burnett and I. J. Ford. Coagulation kinetics beyond mean field theory using an optimised Poisson representation. *J. Chem. Phys.* **142**(19), 194112, 2015. DOI: 10.1063/1.4921350.

[384] M. S. Swanson. *Path Integrals and Quantum Processes*. Academic Press, San Diego, CA, 1992. Chap. 5.1.

[385] B. P. Lee. *Critical Behavior in Non-Equilibrium Systems*. PhD thesis. University of California, Santa Barbara, 1994.

[386] R. F. Pawula. Generalizations and extensions of the Fokker-Planck-Kolmogorov equations. *IEEE Trans. Inform. Theory* **13**(1), 33–41, 1967. DOI: 10.1109/TIT.1967.1053955.

[387] R. F. Pawula. Approximation of the linear Boltzmann equation by the Fokker-Planck equation. *Phys. Rev.* **162**(1), 186–188, 1967. DOI: 10.1103/PhysRev.162.186.

[388] P. Thomas, C. Fleck, R. Grima and N. Popović. System size expansion using Feynman rules and diagrams. *J. Phys. A: Math. Theor.* **47**(45), 455007, 2014. DOI: 10.1088/1751-8113/47/45/455007.

[389] A. Andreanov, G. Biroli, J.-P. Bouchaud and A. Lefèvre. Field theories and exact stochastic equations for interacting particle systems. *Phys. Rev. E* **74**(3), 030101(R), 2006. DOI: 10.1103/PhysRevE.74.030101.

[390] A. Lefèvre and G. Biroli. Dynamics of interacting particle systems: stochastic process and field theory. *J. Stat. Mech. Theor. Exp.* **2007**(07), P07024, 2007. DOI: 10.1088/1742-5468/2007/07/P07024.

[391] K. Itakura, J. Ohkubo and S.-i. Sasa. Two Langevin equations in the Doi-Peliti formalism. *J. Phys. A: Math. Theor.* **43**(12), 125001, 2010. DOI: 10.1088/1751-8113/43/12/125001.

[392] H. K. Janssen, U. C. Täuber and E. Frey. Exact results for the Kardar-Parisi-Zhang equation with spatially correlated noise. *Eur. Phys. J. B* **9**(3), 491–511, 1999. DOI: 10.1007/s100510050790.

[393] E. Frey, U. C. Täuber and H. K. Janssen. Scaling regimes and critical dimensions in the Kardar-Parisi-Zhang problem. *Europhys. Lett.* **47**(1), 14–20, 1999. DOI: 10.1209/epl/i1999-00343-4.

[394] R. L. Stratonovich. On a method of calculating quantum distribution functions. *Sov. Phys. Dokl.* **2**, 416, 1957.

[395] J. Hubbard. Calculation of partition functions. *Phys. Rev. Lett.* **3**(2), 77–78, 1959. DOI: 10.1103/PhysRevLett.3.77.

[396] W. Horsthemke and A. Bach. Onsager-Machlup Function for one dimensional nonlinear diffusion processes. *Z. Phys. B* **22**(2), 189–192, 1975. DOI: 10.1007/BF01322364.

[397] D. Dürr and A. Bach. The Onsager-Machlup function as Lagrangian for the most probable path of a diffusion process. *Commun. Math. Phys* **60**(2), 153–170, 1978. DOI: 10.1007/BF01609446.

[398] H. Ito. Probabilistic construction of Lagrangean of diffusion process and Its application. *Prog. Theor. Phys.* **59**(3), 725–741, 1978. DOI: 10.1143/PTP.59.725.

[399] Y. Takahashi and S. Watanabe. The probability functionals (Onsager-Machlup functions) of diffusion processes. *Stochastic Integrals*. Ed. by D. Williams. Vol. 851. Lecture Notes in Mathematics. Springer, Berlin, Heidelberg, 1981, pp. 433–463. DOI: 10.1007/BFb0088735.

[400] J. Zinn-Justin. *Quantum Field Theory and Critical Phenomena*. 4th edition. International Series of Monographs on Physics 113. Oxford University Press, Oxford, 2002. Chap. 4.6.

[401] Y. Tang, R. Yuan and P. Ao. Summing over trajectories of stochastic dynamics with multiplicative noise. *J. Chem. Phys.* **141**(4), 044125, 2014. DOI: 10.1063/1.4890968.

[402] M. Chaichian and A. Demichev. *Path Integrals in Physics - Volume I: Stochastic Processes and Quantum Mechanics*. Series in Mathematical and Computational Physics. Vol. 1. IOP Publishing, Bristol, 2001. Chap. 1.

[403] M. J. Howard and U. C. Täuber. 'Real' versus 'imaginary' noise in diffusion-limited reactions. *J. Phys. A: Math. Gen.* **30**(22), 7721–7731, 1997. DOI: 10.1088/0305-4470/30/22/011.

[404] D. Hochberg, M. P. Zorzano and F. Morán. Complex noise in diffusion-limited reactions of replicating and competing species. *Phys. Rev. E* **73**(6), 066109, 2006. DOI: 10.1103/PhysRevE.73.066109.

[405] D. C. Torney and H. M. McConnell. Diffusion-limited reactions in one dimension. *J. Phys. Chem.* **87**(11), 1941–1951, 1983. DOI: 10.1021/j100234a023.

[406] S. Redner. *A Guide to First-Passage Processes*. Cambridge University Press, Cambridge, 2001.

[407] I. Jensen. Critical exponents for branching annihilating random walks with an even number of offspring. *Phys. Rev. E* **50**(5), 3623–3633, 1994. DOI: 10.1103/PhysRevE.50.3623.

[408] M. Moshe. Recent developments in Reggeon field theory. *Phys. Rep.* **37**(3), 255–345, 1978. DOI: 10.1016/0370-1573(78)90098-4.

[409] E. Frey, U. C. Täuber and F. Schwabl. Crossover from self-similar to self-affine structures in percolation. *Europhys. Lett.* **26**(6), 413–418, 1994. DOI: 10.1209/0295-5075/26/6/003.

[410] E. Frey, U. C. Täuber and F. Schwabl. Crossover from isotropic to directed percolation. *Phys. Rev. E* **49**(6), 5058–5072, 1994. DOI: 10.1103/PhysRevE.49.5058.

[411] H. K. Janssen. On the nonequilibrium phase transition in reaction-diffusion systems with an absorbing stationary state. *Z. Phys. B* **42**(2), 151–154, 1981. DOI: 10.1007/BF01319549.

[412] P. Grassberger. On phase transitions in Schlögl's second model. *Z. Phys. B* **47**(4), 365–374, 1982. DOI: 10.1007/BF01313803.

[413] L. Canet, H. Chaté and B. Delamotte. General framework of the non-perturbative renormalization group for non-equilibrium steady states. *J. Phys. A: Math. Theor.* **44**(49), 495001, 2011. DOI: 10.1088/1751-8113/44/49/495001.

[414] J. Zinn-Justin. *Phase Transitions and Renormalization Group*. Oxford Graduate Texts. Oxford University Press, Oxford, 2007. DOI: 10.1093/acprof:oso/9780199227198.001.0001.

[415] R. Dickman and R. Vidigal. Path integrals and perturbation theory for stochastic processes. *Braz. J. Phys.* **33**(1), 73–93, 2003. DOI: 10.1590/S0103-97332003000100005.

[416] G. U. Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. R. Soc. B* **213**(402-410), 21–87, 1925. DOI: 10.1098/rstb.1925.0002.

[417] M. H. DeGroot and M. J. Schervish. *Probability and Statistics*. 4th edition. Pearson, Boston, 2012. Chap. 5.5.

[418] C. Wissel. Manifolds of equivalent path integral solutions of the Fokker-Planck equation. *Z. Phys. B* **35**(2), 185–191, 1979. DOI: 10.1007/BF01321245.

[419] T. Mansour and M. Schork. *Commutation Relations, Normal Ordering, and Stirling Numbers*. Discrete Mathematics and Its Applications. Chapman and Hall/CRC, Boca Raton, FL, 2015. P. 410.

[420] H. F. Trotter. On the product of semi-groups of operators. *Proc. Amer. Math. Soc.* **10**(4), 545–551, 1959. DOI: 10.1090/S0002-9939-1959-0108732-6.

[421] J. Touchard. Sur les cycles des substitutions. *Acta Math.* **70**(1), 243–297, 1939. DOI: 10.1007/BF02547349.

[422] G. Peccati and M. S. Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams - A survey with computer implementation*. Bocconi & Springer Series. Vol. 1. Springer, Milan, 2011. Chap. 3.3. DOI: 10.1007/978-88-470-1679-8.

[423] K. Kang and S. Redner. Fluctuation effects in Smoluchowski reaction-kinetics. *Phys. Rev. A* **30**(5), 2833–2836, 1984. DOI: 10.1103/PhysRevA.30.2833.

[424] F. Benitez and N. Wschebor. Branching-rate expansion around annihilating random walks. *Phys. Rev. E* **86**(1), 010104, 2012. DOI: 10.1103/PhysRevE.86.010104.

[425] F. Benitez and N. Wschebor. Branching and annihilating random walks: Exact results at low branching rate. *Phys. Rev. E* **87**(5), 052132, 2013. DOI: 10.1103/PhysRevE.87.052132.

[426] P.-A. Rey and J. Cardy. Asymptotic form of the approach to equilibrium in reversible recombination reactions. *J. Phys. A: Math. Gen.* **32**(9), 1585–1603, 1999. DOI: 10.1088/0305-4470/32/9/008.

[427] Z. Konkoli and H. Johannesson. Two-species reaction-diffusion system with equal diffusion constants: Anomalous density decay at large times. *Phys. Rev. E* **62**(3), 3276–3280, 2000. DOI: 10.1103/PhysRevE.62.3276.

[428] N. Bleistein and R. A. Handelsman. *Asymptotic Expansions of Integrals*. Dover Publications, New York, NY, 1986. Chap. 6.1.

[429] A. Mugler, A. M. Walczak and C. H. Wiggins. Spectral solutions to stochastic models of gene expression with bursts and regulation. *Phys. Rev. E* **80**(4), 041921, 2009. DOI: 10.1103/PhysRevE.80.041921.

[430] J. Ohkubo. Duality in interacting particle systems and boson representation. *J. Stat. Phys.* **139**(3), 454–465, 2010. DOI: 10.1007/s10955-009-9910-2.

[431] J. Ohkubo. Solving partial differential equation via stochastic process. *Unconventional Computation*. Ed. by C. S. Calude, M. Hagiya, K. Morita, G. Rozenberg and J. Timmis. Vol. 6079. Springer, Berlin, Heidelberg, 2010, pp. 105–114. DOI: 10.1007/978-3-642-13523-1_13.

[432] J. Ohkubo. Extended duality relations between birth-death processes and partial differential equations. *J. Phys. A: Math. Theor.* **46**(37), 375004, 2013. DOI: 10.1088/1751-8113/46/37/375004.

[433] J. Ohkubo. Duality-based calculations for transition probabilities in birth-death processes. arXiv:1509.07193, 2015.

[434] F. Benitez, C. Duclut, H. Chaté, B. Delamotte, I. Dornic and M. A. Muñoz. Langevin equations for reaction-diffusion processes. *Phys. Rev. Lett.* **117**(10), 100601, 2016. DOI: 10.1103/PhysRevLett.117.100601.

[435] J. Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK, 1982.

[436] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. 1st edition. Cambridge University Press, Cambridge, 1998. DOI: 10.1017/CBO9781139173179.

[437] M. A. Nowak and K. Sigmund. Evolutionary dynamics of biological games. *Science* **303**(5659), 793–799, 2004. DOI: 10.1126/science.1093411.

[438] A. Rapoport and A. M. Chammah. *Prisoner's Dilemma - A Study in Conflict and Cooperation*. With the collaboration of C. J. Orwant. The University of Michigan Press, 1965.

[439] A. Rapoport and A. M. Chammah. The game of chicken. *Am. Behav. Sci.* **10**(3), 10–28, 1966. DOI: 10.1177/000276426601000303.

[440] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. 3rd edition. Princeton University Press, Princeton, 1953.

[441] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Math. Ann.* **100**(1), 295–320, 1928. DOI: 10.1007/BF01448847.

[442] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science* **211**(4489), 1390–1396, 1981. DOI: 10.1126/science.7466396.

[443] W. H. Press and F. J. Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl. Acad. Sci. USA* **109**(26), 10409–10413, 2012. DOI: 10.1073/pnas.1206569109.

[444] J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature* **246**(5427), 15–18, 1973. DOI: 10.1038/246015a0.

[445] P. D. Taylor and L. B. Jonker. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* **40**(1-2), 145–156, 1978. DOI: 10.1016/0025-5564(78)90077-9.

[446] P. Schuster and K. Sigmund. Replicator dynamics. *J. Theor. Biol.* **100**(3), 533–538, 1983. DOI: 10.1016/0022-5193(83)90445-9.

[447] A. J. Lotka. Undamped oscillations derived from the law of mass action. *J. Am. Chem. Soc.* **42**(8), 1595–1599, 1920. DOI: 10.1021/ja01453a010.

[448] V. Volterra. *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*. 1st edition. Gauthier-Villars, Paris, 1931.

[449] N. S. Goel, S. C. Maitra and E. W. Montroll. On the Volterra and other nonlinear models of interacting populations. *Rev. Mod. Phys.* **43**(2), 231–276, 1971. DOI: 10.1103/RevModPhys.43.231.

[450] The Big Bang Theory. *The Lizard-Spock Expansion*. Season 2, Episode 8. 2008.

[451] H. W. Kuhn and A. W. Tucker. *Linear Inequalities and Related Systems*. Princeton University Press, Princeton, NJ, 1956.

[452] C. Parmesan. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669, 2006. DOI: 10.1146/annurev.ecolsys.37.091305.110100.

[453] J. W. Costerton, P. S. Stewart and E. P. Greenberg. Bacterial biofilms: a common cause of persistent infections. *Science* **284**(5418), 1318–1322, 1999. DOI: 10.1126/science.284.5418.1318.

[454] R. M. Donlan. Biofilms: microbial life on surfaces. *Emerg. Infect. Dis.* **8**(9), 881–890, 2002. DOI: 10.3201/eid0809.020063.

[455] E. Ben-Jacob, O. Schochet, A. Tenenbaum, I. Cohen, A. Czirók and T. Vicsek. Generic modeling of cooperative growth-patterns in bacterial colonies. *Nature* **368**(6466), 46–49, 1994. DOI: 10.1038/368046a0.

[456] I. Golding, I. Cohen and E. Ben-Jacob. Studies of sector formation in expanding bacterial colonies. *Europhys. Lett.* **48**(5), 587–593, 1999. DOI: 10.1209/epl/i1999-00524-7.

[457] O. Hallatschek, P. Hersen, S. Ramanathan and D. R. Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. USA* **104**(50), 19926–19930, 2007. DOI: 10.1073/pnas.0710150104.

[458] O. Hallatschek and D. R. Nelson. Life at the front of an expanding population. *Evolution* **64**(1), 193–206, 2010. DOI: 10.1111/j.1558-5646.2009.00809.x.

[459] A. Be'er, G. Ariel, O. Kalisman, Y. Helman, A. Sirota-Madi, H. P. Zhang, E. L. Florin, S. M. Payne, E. Ben-Jacob and H. L. Swinney. Lethal protein produced in response to competition between sibling bacterial colonies. *Proc. Natl. Acad. Sci. USA* **107**(14), 6258–6263, 2010. DOI: 10.1073/pnas.1001062107.

[460] K. S. Korolev, M. J. I. Müller, N. Karahan, A. W. Murray, O. Hallatschek and D. R. Nelson. Selective sweeps in growing microbial colonies. *Phys. Biol.* **9**(2), 026008, 2012. DOI: 10.1088/1478-3975/9/2/026008.

[461] B. Momeni, K. A. Brileya, M. W. Fields and W. Shou. Strong inter-population cooperation leads to partner intermixing in microbial communities. *eLife* **2**, e00230, 2013. DOI: 10.7554/eLife.00230.

[462] M. Sen Datta, K. S. Korolev, I. Cvijovic, C. Dudley and J. Gore. Range expansion promotes cooperation in an experimental microbial metapopulation. *Proc. Natl. Acad. Sci. USA* **110**(18), 7354–7359, 2013. DOI: 10.1073/pnas.1217517110.

[463] J. D. Van Dyken, M. J. I. Müller, K. M. L. Mack and M. M. Desai. Spatial population expansion promotes the evolution of cooperation in an experimental prisoner's dilemma. *Curr. Biol.* **23**(10), 919–923, 2013. DOI: 10.1016/j.cub.2013.04.026.

[464] M. J. I. Müller, B. I. Neugeboren, D. R. Nelson and A. W. Murray. Genetic drift opposes mutualism during spatial population expansion. *Proc. Natl. Acad. Sci. USA* **111**(3), 1037–1042, 2014. DOI: 10.1073/pnas.1313285111.

[465] W. Möbius, A. W. Murray and D. R. Nelson. How obstacles perturb population fronts and alter their genetic structure. *PLoS Comput. Biol.* **11**(12), e1004615, 2015. DOI: 10.1371/journal.pcbi.1004615.

[466] T. Hofer, N. Ray, D. Wegmann and L. Excoffier. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Ann. Hum. Genet.* **73**(1), 95–108, 2009. DOI: 10.1111/j.1469-1809.2008.00489.x.

[467] C. A. Edmonds, A. S. Lillie and L. L. Cavalli-Sforza. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* **101**(4), 975–979, 2004. DOI: 10.1073/pnas.0308064100.

[468] S. Klopfstein, M. Currat and L. Excoffier. The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* **23**(3), 482–490, 2006. DOI: 10.1093/molbev/msj057.

[469] O. Hallatschek and D. R. Nelson. Gene surfing in expanding populations. *Theor. Popul. Biol.* **73**(1), 158–170, 2008. DOI: 10.1016/j.tpb.2007.08.008.

[470] L. Excoffier and N. Ray. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **23**(7), 347–351, 2008. DOI: 10.1016/j.tree.2008.04.004.

[471] L. Excoffier, M. Foll and R. J. Petit. Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* **40**, 481–501, 2009. DOI: 10.1146/annurev.ecolsys.39.110707.173414.

[472] E. Mayr. *Systematics and the Origin of Species*. Cambridge University Press, New York, NY, 1942.

[473] N. A. Rosenberg, J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky and M. W. Feldman. Genetic structure of human populations. *Science* **298**(5602), 2381–2385, 2002. DOI: 10.1126/science.1078311.

[474] S. Ramachandran, O. Deshpande, C. C. Roseman, N. A. Rosenberg, M. W. Feldman and L. L. Cavalli-Sforza. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc. Natl. Acad. Sci. USA* **102**(44), 15942–15947, 2005. DOI: 10.1073/pnas.0507611102.

[475] R. M. May and W. J. Leonard. Nonlinear aspects of competition between three species. *SIAM J. Appl. Math.* **29**(2), 243–253, 1975. DOI: 10.1137/0129022.

[476] L. W. Buss and J. B. C. Jackson. Competitive networks: nontransitive competitive relationships in cryptic coral reef environments. *Am. Nat.* **113**(2), 223–234, 1979. DOI: 10.1086/283381.

[477] D. R. Taylor and L. W. Aarssen. Complex competitive relationships among genotypes of three perennial grasses: implications for species coexistence. *Am. Nat.* **136**(3), 305–327, 1990. DOI: 10.1086/285100.

[478] R. Durrett and S. Levin. Allelopathy in spatially distributed populations. *J. Theor. Biol.* **185**(2), 165–171, 1997. DOI: 10.1006/jtbi.1996.0292.

[479] R. Durrett and S. Levin. Spatial aspects of interspecific competition. *Theor. Popul. Biol.* **53**(1), 30–43, 1998. DOI: 10.1006/tpbi.1997.1338.

[480] G. Szabó and T. Czárán. Phase transition in a spatial Lotka-Volterra model. *Phys. Rev. E* **63**(6), 061904, 2001. DOI: 10.1103/PhysRevE.63.061904.

[481] T. L. Czaran, R. F. Hoekstra and L. Pagie. Chemical warfare between microbes promotes biodiversity. *Proc. Natl. Acad. Sci. USA* **99**(2), 786–790, 2002. DOI: 10.1073/pnas.012399899.

[482] B. C. Kirkup and M. A. Riley. Antibiotic-mediated antagonism leads to a bacterial game of rock-paper-scissors in vivo. *Nature* **428**(6981), 412–414, 2004. DOI: 10.1038/nature02429.

[483] R. A. Laird and B. S. Schamp. Competitive intransitivity promotes species coexistence. *Am. Nat.* **168**(2), 182–193, 2006. DOI: 10.1086/506259.

[484] R. A. Lankau and S. Y. Strauss. Mutual feedbacks maintain both genetic and species diversity in a plant community. *Science* **317**(5844), 1561–1563, 2007. DOI: 10.1126/science.1147455.

[485] R. A. Laird and B. S. Schamp. Species coexistence, intransitivity, and topological variation in competitive tournaments. *J. Theor. Biol.* **256**(1), 90–95, 2009. DOI: 10.1016/j.jtbi.2008.09.017.

[486] D. D. Cameron, A. White and J. Antonovics. Parasite-grass-forb interactions and rock-paper-scissor dynamics: predicting the effects of the parasitic plant *Rhinanthus minoron* host plant communities. *J. Ecol.* **97**(6), 1311–1319, 2009. DOI: 10.1111/j.1365-2745.2009.01568.x.

[487] J. R. Nahum, B. N. Harding and B. Kerr. Evolution of restraint in a structured rock-paper-scissors community. *Proc. Natl. Acad. Sci. USA* **108**(Suppl. 2), 10831–10838, 2011. DOI: 10.1073/pnas.1100296108.

[488] A. Szolnoki, M. Mobilia, L. L. Jiang, B. Szczesny, A. M. Rucklidge and M. Perc. Cyclic dominance in evolutionary games: a review. *J. R. Soc. Interface* **11**(100), 20140735, 2014. DOI: 10.1098/rsif.2014.0735.

[489] R. M. Lodge and C. N. Hinshelwood. 51. Physicochemical aspects of bacterial growth. Part IX. The lag phase of Bact. lactis aerogenes. *J. Chem. Phys.* 213–219, 1943.

[490] J. Monod. The growth of bacterial cultures. *Annu. Rev. Microbiol.* **3**, 371–394, 1949. DOI: 10.1146/annurev.mi.03.100149.002103.

[491] M. O. Lavrentovich and D. R. Nelson. Asymmetric mutualism in two- and three-dimensional range expansions. *Phys. Rev. Lett.* **112**(13), 138102, 2014. DOI: 10.1103/PhysRevLett.112.138102.

[492] C. Vulin, J.-M. Di Meglio, A. B. Lindner, A. Daerr, A. Murray and P. Hersen. Growing yeast into cylindrical colonies. *Biophys. J.* **106**(10), 2214–2221, 2014. DOI: 10.1016/j.bpj.2014.02.040.

[493] M. O. Lavrentovich. Critical fitness collapse in three-dimensional spatial population genetics. *J. Stat. Mech. Theor. Exp.* **2015**(5), P05027, 2015. DOI: 10.1088/1742-5468/2015/05/P05027.

[494] R. P. Araujo and D. L. S. McElwain. A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bull. Math. Biol.* **66**, 1039–1091, 2004. DOI: 10.1016/j.bulm.2003.11.002.

[495] M. Kimura and G. H. Weiss. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**(4), 561–576, 1964.

[496] K. S. Korolev, M. Avlund, O. Hallatschek and D. R. Nelson. Genetic demixing and evolution in linear stepping stone models. *Rev. Mod. Phys.* **82**(2), 1691–1718, 2010. DOI: 10.1103/RevModPhys.82.1691.

[497] S. Liang, J. Xu, L. Weng, H. Dai, X. Zhang and L. Zhang. Protein diffusion in agarose hydrogel in situ measured by improved refractive index method. *J. Control. Release* **115**(2), 189–196, 2006. DOI: 10.1016/j.jconrel.2006.08.006.

Acknowledgements

First, I would like to express my gratitude to my thesis adviser Erwin Frey for his support and for giving me the freedom to explore diverse research projects. You were always a great partner in developing new ideas and in transforming these ideas into publishable results. I would also like to thank you for creating such a great atmosphere in your group, making it a fantastic place for scientific research.

Along the same lines, I would like to thank all the present and former members of the "Statistical and Biological Physics" group at LMU Munich whom I had the chance to meet (particularly Matthias Rank, Anna Melbinger, Cornelius Weig, Matthias Lechner, Patrick Hillenbrand, Severin Schink, Marianne Bauer, Isabella Krämer, Steffen Rulands, Jacob Halatek, Jan-Timm Kuhr, Cristina Pop, ...). Erwin always ensured that the supporters of Borussia Dortmund are in the majority in his group (at least by volume), with two German championships being celebrated during my stay, as well as a 5:2 cup final defeat of some local club.

Very special thanks go to my current office mate Johannes Knebel, both for the great collaboration on Lotka-Volterra dynamics and condensation in driven-dissipative bosonic systems, and for the time you invested in proofreading my manuscripts. I very much enjoyed our scientific and non-scientific discussions and hope that you have a great time at Harvard! Further thanks go to Philipp Geiger, Andrea Fertl, Tobias Göppel, and Hannes Herrmann who I accompanied during their master and bachelor theses, together with Johannes Knebel. I also wish to thank Madeleine Opitz for our joint work on microbial range expansions, which, after some ups and downs, resulted in a great publication.

During my doctoral studies, I had the opportunity to present our work at various national and international conferences and universities, and to attend the summer research course "Microbial Strategies for Survival and Evolution" at the Kavli Institute for Theoretical Physics (KITP) in 2014. In this regard, I would like to thank Ned S. Wingreen (Princeton University), David R. Nelson (Harvard University), Ernst-Ludwig Florin (UT Austin), and Otto X. Coredero (formerly ETH Zurich) for their hospitality, and the Center for NanoScience (CeNS), the Nanosystems Initiative Munich (NIM), the German Academic Exchange Service (DAAD), and the KITP for funding. Special thanks go to Philippe Remigi and Paul B. Rainey (both Massey University) with whom I collaborated during the summer research course, and to Boris Shraiman (KITP) for organizing the course.

Finally, I would like to thank my parents Christel and Ulrich, as well as my sister Miriam and my brother Mathias for their love and support.