Master equations and the theory of stochastic path integrals

Markus F Weber and Erwin Frey

Arnold Sommerfeld Center for Theoretical Physics and Center for NanoScience, Department of Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany

E-mail: frey@lmu.de

Received 8 September 2016
Accepted for publication 20 January 2017
Published 17 March 2017

Abstract
This review provides a pedagogic and self-contained introduction to master equations and to their representation by path integrals. Since the 1930s, master equations have served as a fundamental tool to understand the role of fluctuations in complex biological, chemical, and physical systems. Despite their simple appearance, analyses of master equations most often rely on low-noise approximations such as the Kramers–Moyal or the system size expansion, or require ad-hoc closure schemes for the derivation of low-order moment equations. We focus on numerical and analytical methods going beyond the low-noise limit and provide a unified framework for the study of master equations. After deriving the forward and backward master equations from the Chapman–Kolmogorov equation, we show how the two master equations can be cast into either of four linear 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. Spectral methods, WKB approximations, and a variational approach have been proposed for the analysis of the PDE. 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. Subsequently, the solutions of the PDEs are 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. To make this review accessible to a broad community, we have used the language of probability theory rather than quantum (field) theory and do not assume any knowledge of the latter. The probabilistic structures underpinning various

1 Author to whom any correspondence should be addressed.
technical concepts, such as coherent states, the Doi-shift, and normal-ordered observables, are thereby made explicit.

Keywords: stochastic processes, Markov processes, master equations, path integrals, path summation, spectral analysis, rare event probabilities

(Some figures may appear in colour only in the online journal)

Contents

1. Introduction ......................................................... 3
   1.1. Scope of this review ........................................... 3
   1.2. Organization of this review ................................. 3
   1.3. Continuous-time Markov processes and the forward and backward master equations ...... 5
   1.4. Analytical and numerical methods for the solution of master equations ........................ 9
   1.5. History of stochastic path integrals .......................... 12
   1.6. Résumé .......................................................... 13
2. The probability generating function ........................................ 13
   2.1. Flow of the generating function ............................. 15
   2.2. Bases for particular stochastic processes ................. 17
      2.2.1. Random walks ............................................. 17
      2.2.2. Chemical reactions ....................................... 17
      2.2.3. Intermezzo: the unit vector basis ...................... 20
      2.2.4. Processes with locally excluding particles ........... 20
   2.3. Methods for the analysis of the generating function’s flow equation .......................... 22
      2.3.1. A spectral method for the computation of stationary distributions ...................... 22
      2.3.2. WKB approximations and related approaches .................................................. 23
      2.3.3. A variational method ..................................... 24
   2.4. Résumé .......................................................... 24
3. The marginalized distribution and the probability generating functional ...................................... 25
   3.1. Flow of the marginalized distribution ...................... 25
   3.2. Bases for particular stochastic processes ................. 26
      3.2.1. Random walks ............................................. 26
      3.2.2. Chemical reactions ....................................... 27
   3.3. Mean extinction times ......................................... 28
   3.4. Flow of the generating functional ........................... 29
   3.5. The Poisson representation .................................... 29
   3.6. Résumé .......................................................... 30
4. The backward path integral representation ......................... 31
   4.1. Derivation ...................................................... 32
   4.2. Simple growth and linear decay ............................. 33
   4.3. Feynman–Kac formula for jump processes ............... 33
   4.4. Intermezzo: the backward Kramers–Moyal expansion ................................................. 34
      4.4.1. Processes with continuous state ............................................. 34
      4.4.2. Path integral representation of the backward Kramers–Moyal expansion .... 35
   4.4.3. Path integral representation of the backward Fokker–Planck equation ............... 35
   4.4.4. The Onsager–Machlup function ......................... 36
   4.4.5. Alternative discretization schemes ....................... 36
   4.4.6. Wiener’s path integral ..................................... 37
5. The forward path integral representation ............................. 41
   5.1. Derivation ...................................................... 41
   5.2. Linear processes ............................................... 42
   5.3. Intermezzo: the forward Kramers–Moyal expansion ................................................. 42
   5.4. Résumé .......................................................... 43
6. Path integral representation of averaged observables ................................. 43
   6.1. Derivation ...................................................... 44
   6.2. Intermezzo: alternative derivation based on coherent states .................................... 45
   6.3. Perturbation expansions ..................................... 46
   6.4. Coagulation .................................................... 46
   6.5. Résumé .......................................................... 48
7. Stationary paths .................................................... 48
   7.1. Forward path integral approach ............................. 48
   7.2. Binary annihilation ........................................... 50
   7.3. Backward path integral approach ........................... 51
   7.4. Binary annihilation ........................................... 52
      7.4.1. Leading order ............................................. 52
      7.4.2. Beyond leading order ..................................... 53
   7.5. Résumé .......................................................... 54
8. Summary and outlook ................................................. 54
   Acknowledgments ..................................................... 56
   Appendix A. Proof of the Feynman–Kac formula in section 1.3 ........................................ 56
   Appendix B. Proof of the path summation representation in section 1.4 ............................. 57
   Appendix C. Solution of the random walk in sections 2.2.1 and 3.2.1 .............................. 57
   Appendix D. Details on the evaluation of the backward path integral representation in section 4.3 57
   References .............................................................. 58
1. Introduction

1.1. Scope of this review

The theory of continuous-time Markov processes is largely built on two equations: the Fokker–Planck [1–4] and the master equation [4, 5]. 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 [6–15], shortly after such representations had been derived for the Fokker–Planck equation [16–19]. Both approaches were heavily influenced by quantum theory, introducing such concepts as the Fock space [20] with its ‘bras’ and ‘kets’ [21], coherent states [22–24], and ‘normal-ordering’ [25] 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 [26–31]. 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 review. 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 [32, 33]. The framework presented in this review 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 [34], WKB approximations and other ‘semi-classical’ methods for the computation of rare event probabilities [35–38], and a variational approach that was proposed in the context of stochastic gene expression [39]. 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 review, 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 [40, 41], a path integral representation of Haken [42], path integral representations of non-Markov processes [43–58] and of polymers [59–62], and a representation of ‘hybrid’ processes [63–65]. 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 [66–68] are provided by the review of Täuber, Howard, and Vollmayr-Lee [69] as well as the book by Täuber [70]. 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 [4]. It may be interesting to extend the discussed methods to ‘generalized’ or ‘physical’ master equations with memory kernels [71–74].

1.2. Organization of this review

The organization of this review 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 at time , given that the process has been in state at time , is represented by the conditional probability distribution . Whereas the forward master equation evolves this probability distribution forward in time, starting out at time , the backward master equation evolves the distribution backward in time, starting out at time . Both master equations can be derived from the Chapman–Kolmogorov equation (see 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 review 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 review. 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 explicit choice of the basis functions depends on the process being studied. We discuss different choices of the basis function 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 [83] and of Sasai and Wolynes [39], the WKB approximations [84] and spectral methods of Elgart and Kamenev [35] and of Assaf and Meerson [36–38], and the spectral method of Walczak, Mugler, and Wiggins [34]. 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 \( p(\tau|n,t_0,n_0) \) over a set of basis functions \( \{ |n_0 \rangle \} \) 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 \( \{ \langle n \rangle \} \), 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 [32, 33].

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 (see 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 4.5.4 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 [16–19] 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 [79] and, for purely diffusive Brownian motion [82], the representation simplifies to the path integral of Wiener [80, 81]. Moreover, we recover a Feynman–Kac like formula [85], which solves the (backward) Fokker–Planck equation in terms of an average over the paths of an Itô stochastic differential equation [86–88] (or of a Langevin equation [89]).

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 (see 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 5.3. 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 integrals are performed over the real line. Grassmann path integrals [26–31] for systems whose particles locally exclude one another are not considered. It is, however, explained in section 2.2.4 how such systems can be treated without the need for Grassmann variables, based on a method recently proposed by van Wijland [90]. Second, transformations of the path integral variables such as the ‘Doi-shift’ [91] 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 (see right side of figure 1). This representation can be derived both from the backward and forward path integral representations (see section 6.1), and by representing the forward master equation in terms of the eigenvectors of creation and annihilation matrices (‘coherent states’; see section 6.2). The duality between these two approaches resembles the duality between the wave [92] and matrix [93–95] formulations of quantum mechanics. Let us note that our resulting path integral does not involve a ‘second-quantized’ or ‘normal-ordered’ observable [96]. 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 $2A \rightarrow A$ 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 [35]. 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 $2A \rightarrow \emptyset$, improving an earlier approximation of the process by Elgart and Kamenev [35] 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 review 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 \( \tau \) 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 \( n \in \mathbb{R}_{\geq 0} \) the price of a stock between the opening and closing times of an exchange. The assumption of a continuous time parameter \( \tau \) 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 [97]. By allowing \( \tau \) 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 \( \tau \) 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 \( \tau \) 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 \( p(t, n|\tau_1, m_1; \ldots; \tau_i, m_i; t_0, n_0) \) must agree with \( p(t, n|\tau_i, m_i) \) whenever \( t > \tau_i > \cdots > \tau_1 \). 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 [98].

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 [82]. Over a very short time scale, the motion of such a particle is ballistic and its velocity highly auto-correlated [99]. 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 [100]). Thus, on a sufficiently coarse time scale, the motion of the particle looks diffusive 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 \( \tau \) 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 [101]. 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 [102, 103] (the book of Risken [104] 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 [1–4, 104]. For a single random variable \( x \in \mathbb{R} \), e.g. the position of the Brownian particle, this equation has the generic form

\[
\frac{\partial}{\partial \tau} p(\tau, x|t_0, x_0) = -\frac{\partial}{\partial x} [\alpha_x(x)p] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\beta_x(x)p].
\]

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 \( \alpha_x \) is often called a drift coefficient and \( \beta_x \) a diffusion coefficient (note, however, that in the context of population genetics, \( \beta_x \) describes the strength of random genetic ‘drift’ [105, 106]). 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, \( \beta_x \) represents a positive-semidefinite matrix). In the mathematical community, the Fokker–Planck equation is better known as the Kolmogorov forward equation [105], honouring Kolmogorov’s fundamental contributions to the theory of continuous-time Markov processes [4]. Whereas the above Fokker–Planck equation evokes 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

\[
\frac{\partial}{\partial \tau} p(\tau, x|t, x_0) = \alpha_x(x_0) \frac{\partial}{\partial x} p + \frac{1}{2} \beta_x(x_0) \frac{\partial^2}{\partial x^2} p.
\]

The forward and backward Fokker–Planck equations provide information about the conditional probability distribution but not about the individual paths of a Brownian particle. The 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 [85, 107]2. 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 [108–110]:

\[
p(t, x|\tau, x_0) = \langle \delta(x - x(t)) \rangle_W.
\]

The brackets \( \langle \rangle_W \) represent an average over realizations of a Wiener process \( W \), which evolves through uncorrelated

---

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. [4].

2 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.
Gaussian increments \( dW \). 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) \([36-88]\):

\[
dx(s) = \alpha(x(s))ds + \sqrt{\beta(x(s))}dW(s).
\]

The diffusion coefficient \( \beta \) 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, \( \sqrt{\beta} \) may be chosen as the unique symmetric and positive-semidefinite square root of \( \beta \) [111]). Algorithms for the numerical solution of SDEs are provided in [108]. In the physical sciences, SDEs are often written as Langevin equations [89]. For a discussion of stochastic differential equations the reader may refer to a recent report on progress [112].

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 [113–115]. 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 [116]. 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 \( \tau \) is therefore characterized by the occupation of its binding sites. With only a single cytoskeletal filament whose binding sites are labelled as \( \{0, 1, 2, \ldots\} = \mathbb{L} \), the variable \( n = (n_0, n_1, n_2, \ldots) \in \{0, 1\}^\mathbb{L} \) 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 sites 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 [110], 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 being chosen as the set of integers \( \mathbb{Z} \), \( n \) could, for example, represent the position of a molecular motor along a cytoskeletal filament. On the other hand, a non-negative integer \( n \in \mathbb{N}_0 \) 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. \( n \equiv (n^A, n^n, n^C) \in \mathbb{N}_0^3 \)), 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 \( \tau (t \geq \tau \geq t_0) \), resulting in

\[
p(t, n|t_0, n_0) = \sum_m p(t, n|\tau, m|t_0, n_0).
\]

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

\[
p(t, n|t_0, n_0) = \sum_m p(t, n|\tau, m|t_0, n_0).
\]

Letting \( p(\tau|t_0) \) denote the matrix with elements \( p(t, n|t_0, n_0) \), the Chapman–Kolmogorov equation can also be written as \( p(\tau|t_0) = p(t|\tau)p(\tau|t_0) \) (semigroup property). Note that the matrix notation requires a mapping between the state space of \( n \) and \( n_0 \) and an appropriate index set \( I \subseteq \mathbb{N} \). 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 (6), we define

\[
Q_{e, \Delta t}(n, m) := p(\tau + \Delta t, n|\tau, m) - \delta_{n,m} / \Delta t
\]

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

\[
Q_{e}(n, m) := \lim_{\Delta t \to 0} Q_{e, \Delta t}(n, m).
\]

These are the elements of the transition (rate) matrix \( Q_{e} \), which is also called the infinitesimal generator of the Markov process or is simply referred to as the \( Q_{e} \)-matrix. Its off-diagonal elements \( w_{i}(m, n) := Q_{e}(n, m) \) denote the rates at which probability flows from a state \( m \) to a state \( n \equiv n \). The ‘exit rates’ \( w_{i}(m) := -Q_{e}(m, m) \), on the other hand, describe the rates at which probability leaves state \( m \). Both \( w_{i}(m, n) \) and \( w_{i}(m) \) are non-negative for all \( n \) and \( m \). All of the processes considered here shall conserve the total probability, requiring that \( \sum_{n} Q_{e}(n, m) = 0 \) or, equivalently, \( w_{i}(m) = \sum_{n} w_{i}(n, m) \) (with \( w_{i}(m, m) := 0 \)). The finiteness of the exit rate \( w_{i}(m) \) and the conservation of total probability imply that we consider a *stable* and *conservative* Markov process [117]. In the natural sciences, the master equation is commonly written in terms of \( w_{i} \), but most mathematicians prefer \( Q_{e} \). These matrices can be converted into one another by employing

\[1\text{The Langevin equation corresponding to the SDE (4) reads} \]

\[\delta x(s) = \alpha(x(s)) + \sqrt{\beta(x(s))} \eta(s), \text{ with the Gaussian white noise} \ \eta(s) \]

\[\text{having zero mean and the auto-correlation function} \ (\eta(s)\eta(s')) = \delta(s-s'). \]
\[ Q_t(n,m) = \delta_{m,m} W_t(m). \] (9)

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

Assuming that the limit in (8) interchanges with a sum over the state \( m \), the (forward) master equation now follows from the Chapman–Kolmogorov equation (6) as

\[
\begin{align*}
\frac{\partial}{\partial \tau} p(\tau,n|t_0,n_0) &= \lim_{\Delta t \to 0} \frac{p(\tau + \Delta t, n|t_0, n_0) - p(\tau, n|t_0, n_0)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \sum_m \frac{p(\tau + \Delta t, n|m) - b_{n,m} p(\tau, m|t_0, n_0)}{\Delta t} \\
&= \sum_m Q_t(n,m) p(\tau, m|t_0, n_0). \quad (10)
\end{align*}
\]

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 \( \frac{\partial}{\partial \tau} p(\tau|t_0) = Q_t p(\tau|t_0) \). In terms of \( w_\tau \), it assumes the intuitive gain-loss form

\[
\begin{align*}
\frac{\partial}{\partial \tau} w_\tau(n,m) &= \sum_m \left[ w_\tau(n,m) p(\tau, m|\cdot) - w_\tau(m,n) p(\tau, n|\cdot) \right]. \quad (11)
\end{align*}
\]

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 [70, 74]. The master equation (10) or (11) 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 [118] for detailed balance [70].

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 [70]. Whether or not detailed balance is actually fulfilled is, however, not relevant for the methods discussed in this review. Information on the existence and uniqueness of an asymptotic stationary distribution of the master equation can be found in [117].

The name ‘master equation’ was originally coined by Nordsieck, Lamb, and Uhlenbeck [5] in their study of the Furry model of cosmic rain showers [119]. Shortly before, Feller applied an equation of the same structure to the growth of populations [120] and Delbrück to well-mixed, auto-catalytic chemical reactions [121]. Delbrück’s line of research was followed by several others [122–125], most notably by McQuarrie [126–128] (see also the books [98, 110, 129]). In these articles, several elementary chemical reactions are solved by methods that also appear later in this review. 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, [96, 130–133] 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 [134–137], gene regulatory networks [34, 138–143], the spreading of diseases [144–147], epidermal homeostasis [148], nucleosome repositioning [149], ecological [150–163] and bacterial dynamics [158, 164–167], evolutionary game theory [168–179], surface growth [180], and social and economic processes [181–184]. Queuing processes are also often modelled in terms of master equations, but in this context, the equations are typically referred to as Kolmogorov equations [185]. Moreover, master equations and the SSA have helped to understand the formation of traffic jams on highways [186, 187], the walks of molecular motors along cytoskeletal filaments [113–115, 188–190], and the condensation of bosons in driven-dissipative quantum systems [178, 191, 192]. 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 [193–196]. Transport processes are commonly modelled in terms of the (totally) asymmetric simple exclusion process (ASEP or TASEP) [197–200]. 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 [201–212]. Some of these results were established by applying the Bethe ansatz to the master equation of the ASEP [205, 209]. The review of Chou, Mallick, and Zia provides a comprehensive account of the ASEP and of its variants [190]. The master equation of the TASEP with Langmuir kinetics was recently used to understand the length regulation of microtubules [213].

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 [214]. 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 (10) 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 (1) is complemented by the backward Fokker–Planck equation (2), the (forward) master equation (10) is complemented by a backward master equation. This equation can be derived from the Chapman–Kolmogorov equation (6) as

\[ \partial_t p(t|\tau, n_0) = \lim_{\Delta t \to 0} \frac{p(t, n|\tau - \Delta t, n_0) - p(t|\tau, n_0)}{\Delta t} = \lim_{\Delta t \to 0} \sum_m p(t|\tau, m) \frac{p(t, m|\tau - \Delta t, n_0) - \delta_{m,n_0}}{\Delta t} = \sum_m p(t|\tau, m)Q_m(m, n_0). \]  

(12)

Here, the transition rate is obtained in the limit \( \Delta t \to 0 \). In matrix notation, the backward master equation is written as \( \partial_t p(t|\tau, n_0) = Q_p(t|\tau, n_0) \). In terms of \( \Delta_t \), it assumes the form (see (9))

\[ \partial_t p(t|\tau, n_0) = \sum_m [p(t|\tau, m) - p(t|\tau, n_0)] \Delta_t, \]

(13)

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 \( p(t|\tau, n_0) = \delta_{n,n_0} \). Just as the backward Fokker–Planck equation, the backward master equation proves useful for the computation of mean extinction and first passage times (see [110, 215] and section 3.3). Furthermore, it follows from the backward master equation (12) that the (conditional) average \( \langle A \rangle(t|\tau, n_0) := \sum_m A(m)p(t|\tau, m, n_0) \) of an observable \( A \) fulfills an equation of just the same form, namely

\[ \partial_t \langle A \rangle(t|\tau, n_0) = \sum_m \langle A \rangle(t|\tau, m)Q_m(m, n_0). \]

(14)

The final condition of the equation is given by \( \langle A \rangle(t|\tau, n_0) = A(n_0) \). The validity of equation (14) is the reason why we later employ a ‘backward’ path integral to represent the average \( A \) (see 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 \( \partial_t p(t|\tau, n_0) = Qp(t|\tau, n_0) \) and the backward master equation \( \partial_t p(t|\tau, n_0) = Q_p(t|\tau, n_0) \) are solved by [216]

\[ p(t|\tau, n) = e^{Q(t-\tau)I} \]

(15)

(recall that \( p(t|\tau, n) \) is the matrix with elements \( p(t, n|\tau, n_0) \)). The Chapman–Kolmogorov equation (6) 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 [217]). 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 [217, 218]. Algorithms that can deal with very large state spaces are considered in [219, 220]. For time-dependent transition rates, the matrix exponential generalizes to a Magnus expansion [221, 222].

In the previous paragraph, we restricted the dynamics to a finite state space to ensure the existence of the matrix exponential in (15). 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 [223]. To see that, we define the left stochastic matrix \( P := I + \lambda I^{-1}Q \), with the parameter \( \lambda \) being larger than the above supremum. Writing \( e^{Q\Delta t} = e^{-\lambda \Delta t}e^{\lambda \Delta t I}Q \) with \( \Delta t = \tau - \tau_0 \), the matrix exponential can be evaluated in terms of the convergent Taylor series

\[ p(t|\tau, n) = \sum_{k=0}^{\infty} \frac{\lambda^{-k}}{k!}Q^k, \]

(16)

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 \( \lambda \) (the Poisson process acts as a ‘clock’ with sufficiently high ticking rate \( \lambda \)). Such a decomposition is called a uniformization or randomization and was first proposed by Jensen [224]. The series (16) can be evaluated via numerically stable algorithms and truncation errors can be bounded [225, 226]. 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 [224–229].

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 [230, 231]. Feller derived an integral recurrence formula [117, 230], 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\(^4\). More information on such processes is provided in [117, 216].

\(^4\) Just as a population whose growth is described by the deterministic equation \( \partial_n n = n^2 \) explodes after a finite time, so does a population whose growth is described by the master equation (11) with transition rate \( w(n, m) = \delta_{n,m}ym(m - 1) \). This transition rate models the elementary reaction \( 2A \to 3A \) as explained in section 1.5. An explosion also occurs for the rate \( w(n, m) = \delta_{n,m}m^r \).
The path summation: \[ \int_0^\infty \sum_n \int_{t_0}^{t_n} \mathcal{W}_n(t) p(t, n|t_0, n_0) \, dt = \mathcal{W}_n(t_0|n_0) w(n_0, n_0) \]

In a numerical implementation of the SSA, the state \( n_1 \) 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 \( \mathcal{P} := \{ n_j \} \) 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 \( J^{-1} \sum \mathcal{W}_n \). For example, \( \tau := \eta + \tau_0 \) is distributed by

\[
\begin{align*}
(\mathcal{W}_n \ast \mathcal{W}_n)(\tau) &= \int_0^\tau d\tau_0 \mathcal{W}_n(\tau - \tau_0) \mathcal{W}_n(\tau_0) \\
&= \frac{w(n_1)w(n_0) e^{-w(n_0)\tau} - e^{-w(n_1)\tau}}{w(n_1) - w(n_0)} \Theta(\tau).
\end{align*}
\]

The probability that the system still resides in state \( n_j = n \) at time \( t \) is determined by the ‘survival probability’ \( S_n(t - t_j) = e^{-w(n)\tau}/\Theta(t - t_j) \). After putting all of these pieces together, we arrive at the following path summation representation of the conditional probability distribution:

\[
p(t, n|t_0, n_0) = \sum_{j=0}^\infty \sum_{J_{j-1} = 0}^{J-1} p_{j}(\mathcal{P}_j) \mathcal{W}_n(\tau_0) \mathcal{W}_n(\tau_0) \]

where \( \mathcal{W}_n(\tau) \) 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 \( p_{j}(\mathcal{P}_j) \). By an appropriate choice of integration variables, the probability \( p_{j}(\mathcal{P}_j) \) can also be written as

\[
p_{j}(\mathcal{P}_j) = \left( \prod_{j=0}^{J-1} \int_0^{\tau_{j-1}} d\tau_{j-1} \frac{w(n_{j-1}+1, n_{j-1})}{w(n_j)} \mathcal{W}_n(\tau_{j-1}) \right) \cdot S_n(\tau - (\tau_{j-1} + \cdots + \tau_0)).
\]

The survival probability \( S_n \) 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 [237]. In appendix B, we formally derive the path summation representation (19) from the Laplace transform of the forward master equation (11).
process). Moreover, care has to be taken when the analytic expressions are evaluated numerically because they involve pairwise differences of the exit rate \( w(n) \) (see (18)). 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 (20) can be approximated efficiently (for example, in terms of a Gamma distribution or by analytically determining the Laplace transform of (20), followed by a saddle-point approximation [246] 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 [240], 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 [242], 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 [168]. Efficient variations of the above ‘direct’ version of the SSA are, for example, described in [78, 247–249]. Algorithms for the fast simulation of biochemical networks or processes with spatial degrees of freedom are implemented in the simulation packages [250–257].

The evaluation of the average \( \langle A \rangle = \sum_n A(n)p(t,n) \) 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 \( \tau \)-leaping [78, 249, 258–267]. Another challenge is posed by the evaluation of processes with time-dependent transition rates [268–270].

For completeness, let us mention yet another numerical approach to the (forward) master equation. Since the master equation (10) constitutes a set of coupled linear first-order ODEs, it can of course be treated as such and be integrated numerically. 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 [271]). 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 \( p(t|t_0) = e^{0t_{t-t_0}} \) of the conditional probability distribution, nor its path summation representation (19) 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

\[
p(t,n|t_0,n_0) = \langle n_0 | \int_{t_0}^t e^{-\mathcal{S}(t)} | n \rangle,
\]

and in terms of the ‘backward’ path integral

\[
p(t,n|t_0,n_0) = \langle n_0 | \int_{0}^t e^{-\mathcal{S}(t)} | n \rangle.
\]

The meaning of the integral signs and of the bras \( \langle n \rangle \) and kets \( | n \rangle \) will become clear over the course of this review. 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 [103, 272] and the system-size expansion [98, 273], as well as the derivation of moment equations. Information on these methods can be found in classic textbooks [98, 110] and in a recent review [274]. 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 [275–283]. 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 [284]. In sections 4.4 and 5.3, 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) Fokker–Planck equation [16–19].

1.5. History of stochastic path integrals

The oldest path integral, both in the theory of stochastic processes and beyond [107, 285–287], is presumably Wiener’s integral for Brownian motion [80, 81]. The path integrals we consider here were devised somewhat later, namely in the 1970s and 80s: first for the Fokker–Planck (or Langevin) equation [16–19] and soon after for the master equation. For the master equation, the theoretical basis of these ‘stochastic’ path integrals was developed by Doi [6, 7]. He first expressed the creation and annihilation of molecules in a chemical reaction by the corresponding operators for the quantum harmonic oscillator [288] (modulo a normalization factor), introducing the concept of the Fock space [20] into non-equilibrium theory. Furthermore, he employed coherent states [22–24] to express averaged observables. Similar formalisms as the one of Doi were independently developed by Zel’dovich and Ovchinnikov [8], as well as by Grassberger and Scheunert [10]. Introductions to the Fock space approach, which are in part complementary to our review, are, for example, provided in [70, 91, 289–291]. The review of Mattis and Glasser [289] provides a chronological list of contributions to the field. These contributions include Rose’s renormalized kinetic theory [9], Mikhailov’s path integral [11, 12, 14], which is based on Zel’dovich’s and Ovchinnikov’s work, and Goldenfeld’s extension of the Fock space algebra to polymer crystal growth [13]. Furthermore, Peliti reviewed the Fock space approach and provided derivations of path integral representations of averaged observables and of the probability generating function [15]. 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 (22) and the novel backward representation (23). 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. $|n\rangle := \frac{x^n}{n!}$, the representation is obtained by summing (23) over an observable $A(n)$). Let us note that even though we adopt some of the notation of quantum theory, our review 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 [292] (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 $k A \rightarrow l A$ 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 $\gamma_{\ell}$, which we allow to depend on time. Thus, the rate at which the chemical reaction $k A \rightarrow l A$ induces a transition from state $m$ to state $n$ is $w_i(n,m) = \gamma_{\ell}(m)(m)_k \delta_{n,m-k,l}$. 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 (11) results in the ‘chemical’ master equation

$$\partial_\tau p(\tau, n|0_0, n_0) = \gamma_{\ell}[(n-l+k)\delta_{\ell}(\tau, n-l+k) - (n)\delta_{\ell}(\tau, n)]$$ \hspace{1cm} (24)$$

Microphysical arguments for its applicability to chemical reactions can be found in [77]. According to the chemical master equation, the mean particle number $\langle n \rangle = \sum_n p(\tau, n)$ obeys the equation $\partial_\tau \langle n \rangle = \gamma_{\ell}(l-k)\langle n \rangle$. 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

$$\partial_\tau \bar{n} = \gamma_{\ell}(l-k)\bar{n}$$ \hspace{1cm} (25)$$

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

Path integral representations of the chemical master equation (24) 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 representations are typically derived with the help of ‘creation and annihilation operators’ fulfilling a ‘bosonic’ commutation relation (see section 2.2.2). ‘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 fulfill an anti-commutation relation [26–31]. However, van Wijland recently showed that the use of operators fulfilling the bosonic...
commutation relation is also possible [90]. These approaches are considered in section 2.2.4.

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 technique in the review of Täuber, Howard, and Vollmayr-Lee [69] and in the book of Täuber [70]. Introductory texts are also provided by Cardy’s (lecture) notes [91, 293]. Roughly speaking, path integral representations of the chemical master equation (24) have been used to assess how a macroscopic law of mass action changes due to fluctuations, both below [96, 130, 132, 294–305] and above the (upper) critical dimension [295, 306, 307] or non-perturbative [306, 307, 317–319] 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 [67, 70, 320]. Moreover, path integral representations of master equations have recently been employed in such diverse contexts as the study of neural networks [321–323], of ecological populations [156, 157, 162, 163, 324, 325], and of the differentiation of stem-cells [326].

1.6. Résumé

Continuous-time Markov processes with discontinuous sample paths describe a broad range of phenomena, ranging from traffic jams on highways [186] and on cytoskeletal filaments [113–115, 190] to novel forms of condensation in bosonic systems [178, 191, 192]. In the introduction, we laid out the mathematical theory of these processes and derived the fundamental equations of their evolution: the forward and the backward master equations. Whereas the forward master equation (10) evolves a conditional probability distribution forward in time, the backward master equation (12) evolves the distribution backward in time. In the following main part of this review, 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 Walczak, Mugler, and Wiggins [34] and the Poisson representation of Gardiner and Chaturvedi [32, 33].

2. The probability generating function

The following two sections 2 and 3 are devoted to mapping the forward and backward master equations (10) and (12) 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

\[
g(\tau; q|l_0, n_0) := \sum_{n \geq 0} q^n p(\tau, n|l_0, n_0),
\]

(26)

at least when the corresponding transition rate describes a simple chemical reaction [122–128, 327]. The generating function effectively replaces the discrete variable \( n \) by the continuous variable \( q \). The absolute convergence of the sum in (26) is ensured (at least) for \(|q| \leq 1\). The generating function ‘generates’ probabilities in the sense that

\[
p(\tau, n|l_0, n_0) = \left. \frac{1}{n!} \frac{\partial^n g(\tau; q|l_0, n_0)}{\partial q^n} \right|_{q=0}.
\]

(27)

This inverse transformation from \( g \) to \( p \) involves the application of the (real) linear functional \( L_\sigma[f] := \frac{1}{n!} \frac{\partial^n f(\sigma)}{\partial q^n} \), which maps a (real-valued) function \( f \) to a (real) number. Moreover, it fulfills \( L_\sigma[f + \alpha g] = L_\sigma[f] + \alpha L_\sigma[g] \) 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 (26) 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 \( \varnothing \rightleftharpoons A \) in which molecules of type \( A \) form at rate \( \gamma > 0 \) and degrade at per capita rate \( \mu > 0 \). According to the chemical master equation (24), the probability of observing \( n \in \mathbb{N}_0 \) such molecules obeys the equation

\[
\partial_t p(\tau, n|l_0, n_0) = \gamma [p(\tau, n-1|\varnothing) - p(\tau, n|\varnothing)] + \mu [p(\tau, n+1|A) - np(\tau, n|\varnothing)],
\]

(28)

with initial condition \( p(0, n|l_0, n_0) = \delta_{n_0,n} \). This master equation respects the fact that the number of molecules cannot become negative through the reaction \( \varnothing \rightarrow A \). By differentiating the probability generating function \( g(\tau; q|l_0, n_0) \) in (26) with respect to the current time \( \tau \), one finds that it obeys the flow equation

\[
\partial_\tau g = (q - 1)(\gamma - \mu \partial_q)g.
\]

(29)

Its time evolution starts out from \( g(0,0|q|l_0, n_0) = q^{n_0} \). Instead of solving the flow equation right away, let us first simplify it by changing the basis function \( q^\mu \) of the generating function (26). As a first step, we change it to \((q + 1)^\alpha\), 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 \( e^{-\frac{\tau}{2}(q+1)} \) and arrive at the simplified flow equation

\[
\partial_\tau g = -\mu q \partial_q g.
\]

(30)

The generating function obeying this equation reads
\[
g(\tau; q|\cdot) = \sum_{n \in \mathbb{N}_0} (q + 1)^n e^{\frac{1}{2}(q+1)p(\tau, n|\cdot)}.
\] (31)

As before, the dots inside the functions’ arguments abbreviate the initial parameters \(t_0\) and \(n_0\).

The simplified flow equation (30) 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
\[
|n\rangle_q := (q + 1)^n e^{-\frac{1}{2}(q+1)}
\] (32)
and the corresponding generalized probability generating function (31) as
\[
|g(\tau|\cdot)\rangle_q := \sum_{n \in \mathbb{N}_0} |n\rangle_q p(\tau, n|\cdot).
\] (33)

In quantum mechanics, an object written as \(| \cdot \rangle\) is called a ‘ket’, a notation that was originally introduced by Dirac [21]. In the above two expressions, the kets simply represent ordinary linear algebra. The equation is solved by
\[
|g(\tau|\cdot)\rangle_q = e^{-\frac{1}{2}p(\tau, t_0|\cdot)}|n\rangle_q |n\rangle_{t_0|\cdot} = |n\rangle_q |n\rangle_{t_0|\cdot} = \sum_{|m| \in \mathbb{N}_0} |m\rangle |m\rangle_{t_0|\cdot}.
\] (34)

Thus, the ‘basis functionals’ in \(\{ |m| \}_{m \in \mathbb{N}_0}\) are orthogonal to the basis functions in \(\{ |n| \}_{n \in \mathbb{N}_0}\). The orthogonality condition can be used to recover the conditional probability distribution from (33) via
\[
p(\tau, n|\cdot) = \langle n|g(\tau|\cdot)|.\rangle
\] (37)

Besides being orthogonal to one another, the kets (32) and bras (35) fulfill the completeness relation \(\sum_{m} |m\rangle_{t_0|\cdot} f = f(q)\) with respect to analytic functions \(f\) (note that \(|f|\) as defined in (35) is just a real number and does not depend on \(q\)). Above, we mentioned that we later introduce alternative basis kets, including the Fourier basis function \(|n\rangle_q = e^{iq}\) for the solution of a random walk (with \(n \in \mathbb{Z}\), and the unit column vector \(|n\rangle = \mathbf{e}_n\) 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 \(|m,f| = \int_{-\pi}^{\pi} \frac{df}{2\pi} e^{-imf(q)} (m \in \mathbb{Z})\) and by complementing the unit column vector with the unit row vector \(|m| = \mathbf{e}_m^T (m \in \mathbb{N}_0)\).

In our present example and in most of this review, however, the generating function (33) represents a Taylor expansion around \(g(\tau|t_0|\cdot)\). The completeness condition \(\sum_{m} |n\rangle_{t_0|\cdot} = \mathbb{I}\) involves the (infinitely large) unit matrix \(\mathbb{I}\).

Thanks to the new basis function (32), the simplified flow equation (34) can be easily solved by separation of variables. Making the ansatz \(|g| = f(\tau)h(q)\), one obtains an equation whose two sides depend either on \(f\) or on \(h\) but not on both. The equation is solved by \(f(\tau) = e^{-k\mu(\tau-t_0)}\) and \(h(q) = q^k\), with \(k\) being a non-negative parameter. The non-negativity of \(k\) ensures the finiteness of the initial condition \(|g(t_0|t_0|n_0)| = |n_0\rangle_q\) 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 \(\langle k| := \frac{1}{\sqrt{\mathcal{N}}} \partial_x f(q)|q=0\rangle\) and \(|k\rangle_q := q^k\). These bras and kets are again orthogonal to one another in the sense of \(\langle k|l\rangle_q = \delta_{kl}\) and they also fulfill a completeness relation \(\sum_{k} |k\rangle_q \langle k| = \mathbb{I}\). Using the auxiliary bras and kets, the solution of the flow equation (34) can be written as
\[
|g(\tau|\cdot)| = \sum_{k \in \mathbb{N}_0} |k\rangle_q e^{-k\mu(\tau-t_0)} \langle k|n_0\rangle_q.
\] (38)

We wrote the expansion coefficient in this solution as \(\langle k|n_0\rangle\) to respect the initial condition \(|g(t_0|\cdot)| = |n_0\rangle_q\). The conditional probability distribution can be recovered from the generating function (38) via the inverse transformation (37) as
\[
p(\tau, n|t_0|\cdot) = \sum_{k \in \mathbb{N}_0} \langle n|k\rangle_q e^{-k\mu(\tau-t_0)} \langle k|n_0\rangle_q.
\] (39)

The coefficients \(\langle n|k\rangle\) and \(\langle k|n_0\rangle\) can be computed recursively as explained in [328]. Here we are interested in the asymptotic limit \(\tau \to \infty\) of the distribution (39) for which only the \(k = 0\) mode survives. Therefore, the distribution converges to the stationary Poisson distribution.

\(^{5}\mathrm{As\ before,\ sums\ whose\ range\ is\ not\ specified\ cover\ the\ whole\ state\ space.}\)
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 review. In section 4.2, the reaction \( \varnothing \iff A \) 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 \( \tau \in [t_0, t] \)):

\begin{align}
\sum_n p(\tau, n|t_0, n_0), \quad \sum_n p(t, n|\tau, n_0)|n_0), \quad \sum_n p(n|\tau, t_0, n_0), \quad \sum_n p(t, n|\tau, n_0)|n_0) .
\end{align}

Apparently, the series (41) coincides with the generalized probability generating function (33). In the next section 2.1, we formulate general conditions under which this function obeys a linear PDE. The remaining series (42)–(44) may not be as familiar. We call the series (42) a ‘marginalized distribution’. 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’ (43). 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 [32, 33]. The potential use of the series (44) 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 (42) and by the generating functional (43) in terms of a ‘backward’ path integral. Afterwards, in section 5, the flow equations obeyed by the generating function (41) and by the series expansion (44) are solved in terms of a ‘forward’ path integral. Inverse transformations, such as (37), 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 (10) can be cast into a linear PDE obeyed by the generalized probability generating function

\begin{align}
|g(\tau|t_0, n_0)|_{\tau} := \sum_n |n \rangle \langle n| p(\tau, n|t_0, n_0).
\end{align}

The basis function \( |n \rangle \langle n| \) is a function of the variable \( q \) and possibly of the time variable \( \tau \). 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 \( (n \in N) \), the local copy numbers of the molecule on a lattice \( (n \equiv \{ n_i \in N \}_{i \in Z}) \), or the copy numbers of multiple kinds of molecules \( (n \equiv (n^A, n^B, n^C) \in N^3) \). For the multivariate configurations, the basis function is typically decomposed into a product \( |n^A \rangle |n^B \rangle \cdots |n^C \rangle \) 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 4.5.2. Besides, we also consider a system of excluding particles in section 2.2.4. There, the (local) number of particles \( n \) is restricted to the values 0 and 1. Note that the sum in (45) extends over the whole state space.

The definition of the generating function (45) assumes the existence of a set \( \{|n] \} \) of basis functions for every time \( \tau \in [t_0, t] \). In addition, we assume that there exists a set \( \{ |m \rangle \} \) 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 \( \tau \), they act on the kets as

\begin{align}
\langle m| n \rangle = \delta_{m,n}, \quad (O)
\end{align}

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

\begin{align}
\sum_n |n], \langle n| f = f, \quad (C)
\end{align}

where \( f \) represents an appropriate test function. The completeness condition implies that the function \( f \) can be decomposed in the basis functions \( |n] \) 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

\begin{align}
p(\tau, n|t_0, n_0) = \langle n| g(\tau|t_0, n_0)). \quad (48)
\end{align}

Before deriving the flow equation obeyed by the generating function, let us note that the (O)thogonality condition
differs slightly from the corresponding conditions used in most other texts on stochastic path integrals (see, for example, [13, 15] or the ‘exclusive scalar product’ in [10]). 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 2.2.2 and 3.2.2 for the study of chemical reactions. Its inclusion would be rather unusual, however, for the Fourier basis introduced in sections 2.2.1 and 3.2.1. 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 [32, 33]. We discuss this connection in section 3.5.

To derive the flow equation obeyed by the generating function \( \langle \rangle \rangle \rangle \rangle () () \rangle () \rangle () \rangle () \rangle (), we differentiate its definition (45) with respect to the time variable \( t \). 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 (10). In matrix notation, this equation reads \( \partial_\tau p(\tau|t_0) = \mathcal{Q}_\tau p(\tau|t_0) \). Eventually, one finds that

\[
\partial_\tau |g\rangle = \sum_n \left( \partial_\tau |n\rangle + \sum_m |Q_m, n\rangle \right) p(\tau, n|\rangle \).
\]

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)orthogonality condition, one could also define this operator in a ‘constructive’ way as

\[
\mathcal{E}_\tau := \sum_n (\partial_\tau |n\rangle)(n|\rangle). \tag{51}
\]

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 \( \partial_q \) (later, in section 2.2.4 we also encounter a case in which it constitutes a power series with infinitely high powers of \( \partial_q \). For now, the pre-factors of 1, \( \partial_q \), \( \partial_q^2 \), 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 \( \partial_q \) 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 \( \mathcal{Q}_\tau \) (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 \( w_{\tau}(m, m) = \sum_n w_{\tau}(n, m) \) from a state \( m \) (with \( w_{\tau}(m, m) = 0 \), see (9)). We encode the information stored in \( \mathcal{Q}_\tau \) by a second differential operator called \( \mathcal{Q}_\tau(q, \partial_q) \). This operator should fulfil, for all values of \( n \),

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

In analogy with the transition (rate) matrix \( \mathcal{Q}_\tau \), 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 \( \partial_q \). In section 5, it will be assumed that \( \mathcal{Q}_\tau(q, \partial_q) \) can be expanded in powers of both \( q \) and \( \partial_q \). In a constructive approach, one could also define the operator as

\[
\mathcal{Q}_\tau := \sum_{m,n} |m\rangle Q_{\tau}(m, n|n\rangle. \tag{53}
\]

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 \( |n_q\rangle = e^{i n_q} \) (with \( n \in \mathbb{Z} \)), which we complemented with the functional \( \langle n|f = \int_{-\pi}^{\pi} dq e^{-i n q} f(q) \) 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)onstructive and (O)therogonal basis, it follows from (49) that the generalized generating function \( |g(\tau|t_0, n_0)\rangle \) obeys the flow equation

\[
\partial_\tau |g\rangle = (\mathcal{E}_\tau + \mathcal{Q}_\tau)|g\rangle =: \mathcal{\tilde{Q}}_\tau|g\rangle. \tag{54}
\]

Its initial condition reads \( |g(t_0|t_0, n_0)\rangle = |n_0\rangle \), with the possibly time-dependent basis function \( |n\rangle \) 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 \( \mathcal{\tilde{Q}}_\tau \) and \( \mathcal{Q}_\tau \) then agree because \( \mathcal{E}_\tau \) is zero. Therefore, we refer to both \( \mathcal{\tilde{Q}}_\tau \) 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 \( \mathcal{Q}_\tau \)-matrix and the \( \mathcal{Q}_\tau \)-operator, it is insightful to consider the unit column vectors \( |n\rangle := \mathcal{e}_n \) and the unit row vectors \( \langle n| := \mathcal{e}_n^\dagger \) (with \( m, n \in \mathbb{N}_0 \)). For these vectors, the right hand side of the transition operator (53) simply constitutes a representation of the \( \mathcal{Q}_\tau \)-matrix. Hence, \( \mathcal{Q}_\tau \) is not a differential operator in this case but coincides with the \( \mathcal{Q}_\tau \)-matrix. This observation does not come as a surprise because we already noted that the generating function (45) represents the vector \( p(\tau|t_0, n_0) \) of all probabilities in this case. Moreover, the corresponding flow equation (54) does not constitute a linear PDE but a vector representation of the forward master equation (10).

\[6\] Later, in our derivation of the forward path integral in section 5.1, we employ the finite difference approximation

\[
\lim_{\Delta t \to 0} (|g(\tau + \Delta t)| - |g(\tau)|)/\Delta t = \lim_{\Delta t \to 0} \mathcal{\tilde{Q}}_\tau(\tau)|g(\tau)|.
\]

This scheme con涅ves with the derivation of the forward master equation (10).
Following Doi, the transition operator $\tilde{Q}_\tau$ could be called a ‘time evolution’ or ‘Liouville’ operator [6, 7], or, following Zel’dovich and Ovchinnikov, a ‘Hamiltonian’ [8]. The latter name is due to the formal resemblance of the flow equation (54) to the Schrödinger equation in quantum mechanics [92] (this resemblance holds for any linear PDE that is first order in $\partial_\tau$). The name ‘Hamiltonian’ has gained in popularity throughout the recent years, possibly because the path integrals derived later in this review share many formal similarities with the path integrals employed in quantum mechanics [107, 285–287].

Nevertheless, let us point out that $\tilde{Q}_\tau$ is generally not Hermitian and that the generating function $g$ does not represent a wave function and also not a probability (unless one chooses the unit vectors $|n\rangle = \hat{e}_n$ and $\langle n| = \hat{e}_n^\dagger$ as basis). In this review, we stick to the name transition (rate) operator for $\tilde{Q}_\tau$ (and $Q_\tau$) to emphasize its connection to the transition (rate) matrix $Q_\tau$.

### 2.2. Bases for particular stochastic processes

In the previous section, we formulated four conditions under which the master equation (10) can be cast into the linear PDE (54) obeyed by the generalized generating function (45). In the following, we illustrate how these conditions can be met for various stochastic processes.

#### 2.2.1. 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 $\tau$ obeys the master equation

$$\partial_\tau p(\tau, n|t_0, n_0) = r_\tau [p(\tau, n - 1|\cdot) - p(\tau, n|\cdot)] + l_\tau [p(\tau, n + 1|\cdot) - p(\tau, n|\cdot)],$$

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 (54). But for this purpose, we first require an (O)rhogonal and (C)omplete basis, as well as a basis (E)volution operator $\mathcal{E}_\tau$, and a transition operator $\tilde{Q}_\tau$, (condition Q).

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

$$|n\rangle := e^{i \sigma q} \text{ and } \langle n| := \int_{2\pi} -\pi \frac{dq}{2\pi} e^{-i \sigma q} f(q),$$

with $n \in \mathbb{Z}$ and test function $f$. For this basis, the generalized generating function $|g\rangle = \sum_n e^{i \sigma q} p(\tau, n|\cdot) = \langle \sigma | q \rangle$ 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’

$$\mathbb{III}(q) = \sum_{n \in \mathbb{Z}} \delta_q = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{i \sigma q}.$$  \hspace{1cm} (57)

It thereby follows that

$$\sum_{n \in \mathbb{Z}} |n\rangle \langle n| f = \int_{-\pi}^{\pi} dq' \left(\mathbb{III}(q - q') f(q') = f(q) \right).$$  \hspace{1cm} (58)

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 $Q_\tau$. Its defining condition (Q) requires knowledge of the transition matrix $Q_\tau$ whose elements

$$Q_\tau(m, n) = r(\delta_{m,n+1} - \delta_{m,n}) + l(\delta_{m,n-1} - \delta_{m,n})$$

can be inferred by comparing the master equation (55) with its general form (10). The condition (Q) therefore reads

$$Q_\tau(n) = r_\tau (|n + 1\rangle - |n\rangle) + l_\tau (|n - 1\rangle - |n\rangle).$$

One can construct a transition operator with this property with the help of the functions $c(q) := e^{i \sigma q}$ and $a(q) := e^{-i \sigma q}$.

The function $c$ shifts the basis function $|n\rangle = e^{i \sigma q}$ 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 (60) can be defined as

$$Q_\tau(q, \partial_q) := r([c(q) - 1] + l[a(q) - 1]).$$

This operator can also be inferred from its constructive definition (53) by making use of the Dirac comb.

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

$$\partial_\tau |g\rangle = [r(e^{i \sigma q} - 1) + l(e^{-i \sigma q} - 1)]|g\rangle$$

for $t_0 \leq \tau \leq t$ with the initial value $|g(t_0|t_0, n_0\rangle) = |n_0\rangle = e^{i \sigma q_0}$. The flow equation is readily solved for $|g\rangle$ by

$$\exp\left(e^{-i \sigma q} - 1\right) \int_{t_0}^{t_2} ds r(s) + (e^{-i \sigma q} - 1) \int_{t_0}^{t_2} ds l(s)|n_0\rangle.$$}

The conditional probability distribution can be recovered from this generating function by performing the inverse Fourier transformation (48), 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 [329]. We outline the derivation of this distribution in appendix C. The distribution’s mean is $\mu = n_0 + \int_{t_0}^{t_2} ds (r(s) - l(s))$ and its variance $\sigma^2 = \int_{t_0}^{t_2} ds (r(s) + l(s))$. 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 $r$ and $l$ agree, the Skellam distribution tends to a Gaussian for large times.

#### 2.2.2. 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 \( k_i, l_j \in \mathbb{N}_0 \)). 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 \( L := \mathbb{Z}^d \). Upon labelling particles on different lattice sites by their positions, the hopping of a molecule of type \( A_1 \) from lattice site \( i \in L \) to lattice site \( j \in L \) could be regarded as the chemical reaction \( A_1^{(n)} \rightarrow A_1^{(p)} \). We consider the ‘chemical’ master equation associated to such hopping processes in section 4.5.2.

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 \( kA \rightarrow lA \) \((k, l \in \mathbb{N}_0)\). Since the forward master equation (10) 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 \( 2A \rightarrow \varnothing \) and the linear growth process \( A \rightarrow (1 + m)A \) \([96, 132, 317, 330, 331]\). 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 \([96, 132]\), according to non-perturbative calculations in up to six dimensions \([318]\)). The decomposition of a process into elementary reactions of the form \( kA \rightarrow lA \) is also possible in the contexts of growing polymer crystals \([13]\), aggregation phenomena \([294]\), and predator-prey ecosystems \([163]\). As explained in section 1.5, the transition rate \( w_{\tau} (m, n) = \gamma_{\tau} \delta_{m,n-k} (n)_k \) of the reaction \( kA \rightarrow lA \) 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 \( \gamma_{\tau} \) (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 \( n \in \mathbb{N}_0 \) and we require basis functions \( |n\rangle \) and basis functionals \( \langle n| \) only for such values.

Before specifying the (C)omplete and (O)rthonormal basis as well as the basis (E)volution operator \( \hat{E}_{\tau} \), let us first specify an appropriate transition operator \( Q_\tau \). Its corresponding condition (Q) depends on the transition matrix \( Q_\tau \), whose elements \( Q_{\tau}(m, n) = \gamma_{\tau} (n)_k (\delta_{m,n-k} - \delta_{m,n}) \) can be inferred from the rate \( w_{\tau} (m, n) \) with the help of the relation (9). Consequently, the condition (Q) reads

\[
Q_{\tau} |n\rangle = \gamma_{\tau} (n)_k (|n - k + l\rangle - |n\rangle).
\]

This condition is met by the transition operator

\[
Q_{\tau}(c, a) := \gamma_{\tau} (c^d - c^a) a^d,
\]

provided that there exist a ‘creation’ operator \( c \) and an ‘annihilation’ operator \( a \) acting as

\[
c|n\rangle = |n + 1\rangle \quad \text{and} \quad a|n\rangle = n|n - 1\rangle.
\]

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 (65) and (66) that the two operators fulfil the commutation relation

\[
[a_\tau, c_\tau] := a_\tau c_\tau - c_\tau a_\tau = 1
\]

at a fixed time \( \tau \). 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 \( \tau \) of the creation and annihilation operators. Together with the orthogonality condition, the commutation relation implies that the operators act on the basis functionals as

\[
\langle n|c = \langle n - 1| \quad \text{and} \quad \langle n|a = (n + 1)(n + 1].
\]

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) \([22, 332]\). 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 \( \omega \)). 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 (65) and (66) 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. \( c = c^\dagger \) and \( a = a^\dagger \)). Consequently, the relations corresponding to (68) and (69) read \( \langle n|a^\dagger = \sqrt{n}|n - 1\rangle \) and \( \langle n|a = \sqrt{n + 1}|n + 1\rangle \). Nevertheless, the commutation relation (67) 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 \([Q, P] = i\hbar \), which follows from \( P \) being the generator of spatial displacements in state space \([332]\)). 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 \([333]\).

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 (65) could be interpreted as adding a molecule to a system, and the corresponding annihilation operator (66) as removing a molecule. The commutation relation (67) 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 $k A \rightarrow I A$ (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 $c(n) = n + 1$ and $a(n) = n - 1$ commute).

Whether creation and annihilation operators with the properties (65) and (66) exist depends on the choice of the basis functions. For the study of chemical reactions, a useful choice proves to be the basis function

$$|n\rangle_q := (\zeta \bar{q} + \bar{q})^n e^{-\bar{q}(\bar{q} + \bar{q})}. \quad (70)$$

Here, $\bar{q}(\tau)$ and $\bar{x}(\tau)$ are arbitrary, possibly time-dependent functions and $\zeta = 0$ 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 $\zeta := 1$ but typically we choose $\zeta := 1$. For the latter choice, the basis function (70) simplifies to

$$|n\rangle_q = (q + \bar{q})^n e^{-\bar{q}(\bar{q} + \bar{q})}. \quad (71)$$

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

The two functions $\bar{q}$ and $\bar{x}$ may prove helpful in simplifying the flow equation obeyed by the generating function. For example, we chose $\bar{q} := 1$ and $\bar{x} := \gamma/\mu$ in the introduction to section 2 to simplify the flow equation of the reaction $\varnothing \rightarrow A$ (with growth rate coefficient $\gamma$ and decay rate coefficient $\mu$. Later, in section 7, $\bar{q}$ and $\bar{x}$ will act as the stationary paths of a path integral with $q$ and an auxiliary variable $x$ being deviations from them. If both $\bar{q}$ and $\bar{x}$ are chosen as zero, the basis function (71) simplifies to the basis function

$$|n\rangle_q = q^n \quad (72)$$

of the ordinary probability generating function.

For the general basis function (70), creation and annihilation operators with the properties (65) and (66) can be defined as

$$c(q, \partial_q) := \zeta q + \bar{q} \quad (73)$$

$$a(q, \partial_q) := \partial_q \bar{q} + \bar{x}. \quad (74)$$

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 (64) reads $Q_{\tau}(q, \partial_q) = \gamma_\tau(q - q^2)\partial_q^2$, resulting in the flow equation

$$\partial_\tau |g\rangle = \gamma_\tau(q - q^2)\partial_q^2 |g\rangle. \quad (75)$$

Thus far, we have not specified the basis functions. These functional can be defined using the annihilation operator (74). In particular, we complement the basis function (70) with the functionals

$$\langle n|f := \frac{d^n}{n!}f(q) \bigg|_{q = -\bar{q}/\zeta}, \quad (76)$$

where $f$ represents a test function. The (O)thogonality of the basis $\{|m\rangle_n\}_{m,n;\tau \in \mathbb{N}_0}$ follows directly from the action of the annihilation operator on the basis function $|n\rangle_q$, and from the vanishing of this function at $q = -\bar{q}/\zeta$, except for $n = 0$. The (C)ompleteness of the basis is also readily established.

The only piece still missing is the basis (E)volution operator $E_\tau$ to encode the time-dependence of the basis function (70). Differentiation of this function with respect to time and use of the creation and annihilation operators (73) and (74) show that this operator is given by

$$E_\tau(c, a) := (\partial_\tau q)(a - \bar{x}) - (\partial_\tau \bar{x})c. \quad (77)$$

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 $\bar{q}$ and $\bar{x}$. This simplification is, however, bought by making the basis functions time-dependent. In particular, we consider the simple growth process $\varnothing \rightarrow A$ for a time-independent growth rate coefficient $\gamma$. 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 $\partial_\tau |g\rangle = \gamma_\tau(g - 1)|g\rangle$ for this process. This flow equation can be simplified by redefining the basis function of the generating function as $|n\rangle_{\tau,\gamma} := (q + 1)e^{-\bar{q}(\bar{q} + 1)}$, with $\bar{x}$ solving the rate equation $\partial_\tau \bar{x} = \gamma \bar{x}$. Hence, the basis function depends explicitly on time through its dependence on $\bar{x}(\tau) = \bar{x}(t_0) + \gamma(\tau - t_0)$. Upon combining the transition operator (64) with the basis evolution operator (77), one finds that the generating function now obeys the flow equation $\partial_\tau |g\rangle = -\gamma |g\rangle$. The equation is readily solved by $|g(\tau|t_0, n_0)\rangle = e^{-\gamma(\tau-t_0)}|n_0\rangle|n_0\rangle$. The conditional probability distribution can be recovered via the inverse transformation (48) as

$$p(\tau, n|t_0, n_0) = e^{-\gamma(\tau-t_0)}|n_0\rangle|n_0\rangle. \quad (78)$$

The coefficient $|n_0\rangle|n_0\rangle$ can be evaluated by determining how the functional $|n_\gamma\rangle$ acts at time $t_0$. Using $a_\gamma - \bar{x}(\gamma) = a_0 - \bar{x}(t_0)$ and the binomial theorem, one can rewrite the action of this functional as

$$\langle n|_\gamma f = \sum_{m=0}^{n} \frac{\gamma (\gamma - t_0)^{n-m}}{(n-m)!} (m|_0 f). \quad (79)$$

By the (O)thogonality condition, the conditional probability distribution (78) therefore evaluates to the shifted Poisson distribution

$$p(\tau, n|t_0, n_0) = \frac{e^{-\gamma(\tau-t_0)}(\gamma(\tau-t_0))^{n-n_0}}{(n-n_0)!} \Theta_{n-n_0}, \quad (80)$$

where $\Theta_n := 1$ for $n > 0$ and $\Theta_0 := 0$ otherwise. The validity of this solution can be verified by solving the corresponding flow equation of the ordinary generating function.

7 Use $\langle \partial_q + Q_{\tau}^* f(q)\rangle_{q = -\bar{q}/\zeta} = Q_{\tau}^* e^{W(\partial_\tau + \bar{q}/\zeta ^2)}(f(q)|_{q = -\bar{q}/\zeta})$, and the analyticity of $e^{W(\partial_\tau + \bar{q}/\zeta ^2)}(f(q)|_{q = -\bar{q}/\zeta})$ for an analytic test function $f$.
In our above approach, we first established the form of the transition operator because the basis function (70) is not the only possible choice. Ohkubo recently proposed the use of orthogonal polynomials for this purpose [334]. In analogy to the eigenfunctions of the quantum harmonic oscillator [22, 332], one can, for example, choose the basis function as the Hermite polynomial

\[ |n \rangle := H_n(q_0) = (-1)^n e^{q_0^2/2} \frac{d^n}{dq^n} e^{-q^2/2}, \]  

with \( n \in \mathbb{N}_0 \). Hermite polynomials constitute an Appell sequence, i.e. they fulfill \( \partial_q H_n(q) = n H_{n-1}(q) \) (18.9.27 in [335]). With \( a := \partial_q \), this property coincides with the defining relation \( a|n\rangle = n|n-1\rangle \) of the annihilation operator in (66).

Furthermore, Hermite polynomials obey the recurrence relation \( H_{n+1}(q) = q H_n(q) - n H_{n-1}(q) \) (18.9.1 in [335]). Combined with the Appell property, \( c := q - \partial_q \) therefore fulfills the defining relation \( c|n\rangle = |n+1\rangle \) of a creation operator in (65). After complementing the basis function (81) with the functional

\[ \langle m|f := \frac{1}{m!} \int_{-\infty}^{\infty} dq \frac{e^{-q^2/2}}{\sqrt{2\pi}} H_m(q) f(q), \]  

the (O)rthogonality and (C)ompleteness of the basis are also established (18.3 and 18.18.6 in [335]). Thus, the chemical master equation (24) can be transformed into a flow equation obeyed by a generating function based on Hermite polynomials. To our knowledge, however, no stochastic process model is solved in terms of the Schrödinger equation in (92). In the following, we sketch the mathematical basis of an alternative approach, which employs the unit vectors we sketch the mathematical basis of an alternative approach, which employs the unit vectors

\[ \partial_q \langle n|\rangle = |n-1\rangle = \langle n-1|q \quad \text{and} \quad |n\rangle = \langle n|q. \]  

The orthogonality of the basis allows us to rewrite the elements of the transition matrix of the reaction \( k A \rightarrow IA \) as

\[ Q(m,n) = \langle m|/(\gamma (n-k)|m-k+1\rangle - |n\rangle). \]  

As before, \( p(\tau|t_0) \) represents the matrix of the conditional probabilities \( p(\tau, n|t_0) \), 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 \( k A \rightarrow IA \) with time-independent rate coefficients. The elements of the transition matrix associated to this reaction read

\[ Q(m,n) = \gamma(n) \delta_{m,n-k+1} - \delta_{m,n}. \]  

Since the particle numbers \( n \) and \( t_0 \) 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 \( p(\tau|t_0) = e^{(r-t_0)I} \) (in the sense of the expansion (16), 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 \( p(\tau|t_0) = e^{(r-t_0)I} \) 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. 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 \( |n\rangle := \delta_n \) and their orthogonal unit row vectors \( \langle n| := \delta_n^\dagger \) as basis (with \( n \in \mathbb{N}_0 \)). Individual probabilities can therefore be inferred from the above solution as

\[ p(\tau, n|t_0, n_0) = \langle n|e^{(r-t_0)I} - |n_0\rangle. \]  

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

The orthogonality of the basis allows us to rewrite the elements of the transition rate matrix of the reaction \( k A \rightarrow IA \) as

\[ Q(m,n) = \langle m|/(\gamma (n-k)|m-k+1\rangle - |n\rangle). \]  

A matrix with these elements can be written as

\[ Q = \gamma(c^l - c^k)a^k, \]  

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 seen that these matrices fulfill \( c|n\rangle = |n+1\rangle \) and \( a|n\rangle = |n-1\rangle \) with respect to the basis vectors. These relations coincide with our previous relations (65) and (66). Besides, the matrices also fulfill \( \langle n|c = \langle n-1| \) and \( \langle n|a = \langle n+1| \), as well as the commutation relation \( [a, c] = I \). 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 (85) has the same form as the transition operator (64). 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.

2.2.4. 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 [113–115]. In its simplest form, such a model describes the movement of mutually excluding motors as a hopping process on a one-dimensional lattice $\mathbb{L} \subseteq \mathbb{Z}$, 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 [26], adsorption processes [337, 338], and directed percolation [29]. In the following, we show how the generating function approach from section 2.1 and its complementary approach from section 2.2.3 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 [339, 340]. 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 $\mu$, and the rate of the reverse transition as $\gamma$. For simplicity, we assume these rate coefficients to be time-independent. Consequently, the master equation of the telegraph process reads

$$\partial_t P = \begin{pmatrix} -\mu & \gamma \\ \mu & -\gamma \end{pmatrix} P,$$  

(86)

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

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

(87)

One can solve the above master equation in various ways, for example, by evaluating the matrix exponential in terms of a series expansion of its exponential. The Kronecker operator acts on the basis functions as $\delta_{n,m}(n) = \delta_{n,m}|n\rangle$, and thus it can be used to replace the Kronecker deltas in (88). One thereby arrives at the flow equation

$$\partial_t |g\rangle = \mathcal{Q}|g\rangle = [\mu(a - 1)\delta_{\cdot,1} + \gamma(c - 1)\delta_{\cdot,0}]|g\rangle.$$  

(90)

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 (90) 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 $\partial_q$, i.e. that all the $q$ are to the left of all the $\partial_q$ in every summand. This order can be achieved by the repeated use of $[\partial_q, q\mathcal{Q}(q) = f(q)$. More information on the procedure can be found in [90]. 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 [153].

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 [341, 342] for details). Grassmann variables commute with real and complex numbers (i.e. $[q, a] = 0$ for $a \in \mathbb{C}$), but they anti-commute with themselves and with other Grassmann variables (i.e. $[q, q] = 0$). Grassmann variables have, for example, proven useful in calculating the correlation functions of kinetic Ising models [343], even when the system is driven far from equilibrium [344].

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

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

(91)

The creation matrix acts as $c|0\rangle = |1\rangle$ and $c|1\rangle = 0$, and the annihilation matrix as $a|1\rangle = 0$ and $a|0\rangle = 0$ (zero vector). Consequently, the matrix product $ca$ fulfils $ca|0\rangle = 0$.
and \(c_{01} = 1\), and thus it serves the same purpose as the Kronecker delta \(\delta_{01}\) in (88). Analogously, the operator \(ac\) serves the same purpose as \(\delta_{00}\). The master equation (86) can therefore be written in a form resembling the flow equation (90), namely as
\[
\partial_t \rho = [\mu(a - 1)ca + \gamma(c - 1)ac] \rho.
\]
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 \(\sigma_0, \sigma_z, \text{and} \sigma_0\)). More information on spin-representations of master equations is provided in [66, 337, 338, 345–348]. A spin-representation has, for example, been employed in the analysis of reaction–diffusion master equations via the density matrix renormalization group [349, 350]. In their study of directed percolation of species. Based on the (coherent) eigenstates of the resulting 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. Further information on the Jordan–Wigner transformation, Grassmann path integrals, and alternative approaches can be found in [26–28, 30, 31, 351–353].

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.

2.3.1. 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 [34]. 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 \(\gamma_i(n_i)\) depends only on the copy number \(n_i\) of the species coming earlier along the cascade. Second, each of the two species degrades at a constant per-capita rate \(\mu_i\). With \(n := (n_1, n_2) \in \mathbb{N}_0^2\), the corresponding master equation reads
\[
\partial_t p(\tau, n) = \gamma_i(n_i - 1)p(\tau, n - \hat{e}_i) - \gamma_i(n_i)p(\tau, n) + \gamma_2(n_1)n_2p(\tau, n - \hat{e}_2) - p(\tau, n) + \mu_1(n_1 + 1)p(\tau + \hat{e}_i) - np(\tau, n) + \mu_2(n_2 + 1)p(\tau + \hat{e}_2) - np(\tau, n).
\]
Here, the unit vector \(\hat{e}_i\) points in the direction of the \(i\)th species. The master equation can be cast into a flow equation for the generating function \(g(\tau|c) = \sum_n p(\tau, n|c)|n\rangle\) 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 \(|n\rangle_1 := |n_1\rangle|n_2\rangle\) with
\[
|n\rangle_k := (q_k + 1)^{n_k} e^{-\frac{n_k}{q_k+1}}.
\]
The values of the auxiliary parameters \(\gamma_1\) and \(\gamma_2\) are only specified in a numerical implementation of the method and affect its stability. Information on how their values are chosen is provided in [34]. 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 \(\partial_t g = (Q_0 + Q_1)g\) with the transition operators
\[
Q_0 := -\sum_{i \in \{1, 2\}} \mu_i q_i \partial_{q_i} \text{ and } Q_1 := -\sum_{i \in \{1, 2\}} q_i (\gamma_i - \gamma_i).
\]
Here, the two new operators \(\gamma_i\) and \(\gamma_i\) are defined in terms of their actions \(\gamma_i(n_i) = \gamma_i(n_i)\) 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 \(\gamma_i(n_i)\).

The operator \(Q_0\) has the same form as the transition operator of the bi-directional reaction \(\varnothing = A\) from the introductory example. Without the perturbation \(Q_1\), the above flow equation could thus be solved by extending the previous ansatz (38) to two species. To accommodate \(Q_1\) as well, it proves useful to generalize that ansatz to
\[
|g(\tau)\rangle = \sum_{k \in \mathbb{N}_0^2} |k\rangle G_k(\tau),
\]
with yet to be determined expansion coefficients \(G_k(\tau)\). The auxiliary ket is defined as \(|k\rangle_0 := |k_1\rangle_0|k_2\rangle_0\) with \(|k_1\rangle_0 = q_{k_1}^{k_1}\) and is orthogonal to the bra \(\langle k| := \langle k_1|\langle k_2|\) with \(\langle k_1|f := \frac{1}{k_1!} q_{k_1}^{k_1} f(q_{k_1})\). These bras can be used to extract the expansion coefficient via \(G_k(\tau) = \langle k| g(\tau)\rangle\). Differentiation of this coefficient with respect to time and imposing stationarity eventually results in a recurrence relation for \(G_k\) 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 (48).
In [34], Walczak et al computed the steady-state distribution of the transcrip-
tional regulatory cascade by solving the recurrence relation for $G_t$ 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 $\sim 1000$ coupled ODEs does not pose a problem for modern integrators and the integration provides the full temporal dynamics of the process (see [271] 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.

### 2.3.2. 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 $\tau \to \infty$ (see figure 2 in [34]). 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 [84]. 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 [63, 64, 146, 171, 172, 177, 354–363]. The recent review of Assaf and Meerson provides an in-depth discussion of the applicability of the real- and momentum-space approaches [364].

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 $\gamma$, and are replenished in the linear reaction $A \to 2A$ with rate coefficient $\mu \gg \gamma$. According to a deterministic model of the combined processes with rate equation $\partial_{\tau}n = \gamma n - 2\mu n^2$, the particle number $n$ 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 $\bar{n}_\infty$ for a long but finite time (see figure 3). Asymptotically, all particles become trapped in the ‘absorbing’ state $n = 0$. Consequently, the conditional probability distribution converges to $p(x, n|\bar{n}_\infty, n_0) = \delta_{\emptyset,0}$. 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 [35]. The value of the pre-exponential factor was determined by Turner and Maale-Mansour using a recurrence relation [365], by Kessler and Shnerb using a real-space WKB approximation [357], and by Assaf and Meerson upon combining the generating function technique with Sturm-Liouville theory [366] (the latter method was developed in [36, 37]; see also [367]). Using this method, Assaf and Meerson also succeeded in computing the shape of the metastable distribution. Moreover, Assaf et al showed how a momentum-space WKB approximation can be used to determine mean extinction times for processes with time-modulated rate coefficients [38].

Upon rescaling time as $\gamma(t) = \tau$, the chemical master equation (24) of the combined processes $A \to 2A$ and $2A \to \emptyset$ translates into the flow equation

$$\partial_{\tau}g = \left[ (q^2 - q)\partial_q + \frac{1}{2\bar{n}_\infty}(1 - q^2)\partial^2_q \right]g$$

(98)

de the ordinary generating function $g(\tau, q) = \sum_{n}q^n p(\tau, n)$ (see sections 2.1 and 2.2.2). A WKB approximation of the flow equation can be performed by inserting the ansatz

$$g(\tau, q) = e^{S(\tau, q)}$$

(99)

with the ‘action’

$$S(\tau, q) = 2\bar{n}_\infty \sum_{k=0}^{\infty} \frac{S_0(\tau, q)}{\bar{n}_\infty^k}$$

(100)

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 (99) often includes a minus sign in front of the action, which we neglect for convenience. The pre-factor ‘2’ of the action (100) is also included just for convenience. Upon inserting the ansatz (99) into the flow equation (98), one obtains the equality

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

(101)
As this path proceeds along a zero-energy line, the action (100) evaluates to
\[
S \approx 2\hbar_\infty S_0 = 2\hbar_\infty \int_0^1 \frac{qdq}{1+q} = -2\hbar_\infty(1 - \ln 2).
\] (106)
Upon returning to the original time scale via \( \tau \to \gamma \tau \), the mean extinction time follows as
\[
t = A \gamma^{-1} e^{2\pi A(1 - \ln 2)}
\] (107)
with pre-exponential factor \( A \). The value of this factor has been determined in [357, 365, 366] and reads, in terms of our rate coefficients, \( A = 2\sqrt{\pi\hbar_\infty} \). More information on the above procedure can be found in [35, 38]. The method has also been applied to classic epidemiological models [369] and to a variant of the Verhulst logistic model [370] (exact results on this model with added immigration have been obtained in [371] using the generating function technique).

2.3.3. A variational method. The generating function’s flow equation can also be analysed using a variational method as proposed by Sasai and Wolynes [39, 372]. Their approach is based on a method that Eyink had previously developed (primarily) for Fokker–Planck equations [83]. Instead of dealing with flow equation \( \partial_t |g| = \tilde{Q}_s |g| \) of the generating function in its differential form (or in terms of its spin matrix representation section 2.2.4), the variational method involves a functional variation of the ‘effective action’ \( \Gamma = \int_{n_0} d\tau \langle \psi^k | (\partial_\tau - \tilde{Q}_s) |\psi^k \rangle \) with \( |\psi^k\rangle = |g\rangle \). To perform this variation, one requires an ansatz for the objects \( |\psi^l\rangle \) and \( |\psi^R\rangle \), being parametrized by the variables \( \{\alpha^l_i\}_{i=1,...,s} \) and \( \{\alpha^R_i\}_{i=1,...,s} \). The values of these parameters are determined by requiring that \( |\psi^R\rangle \) is an extremum of the action. Besides its application to networks of genetic switches [39], the variational method has been applied to signalling in enzymatic cascades in [372]. An extension of the method to multivariate processes is described in [373]. Whether or not the variational method provides useful information mainly depends on making the right ansatz for \( |\psi^l\rangle \) 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 (10) can be transformed into a linear partial differential equation, a ‘flow equation’, obeyed by the probability generating function
\[
|g(\tau|t_0, n_0)\rangle = \sum_n^N p(\tau, n|t_0, n_0). \tag{108}
\]
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 $Q_\tau$, encoding the actual dynamics of the process. The generating function (108) then obeys the flow equation

$$\partial_\tau |g\rangle = (\mathcal{E}_\tau + Q_\tau)|g\rangle = \tilde{Q}_\tau|g\rangle.$$  \hspace{1cm} (109)

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 [83] and of Sasai and Wolynes [39], WKB approximations and spectral formulations of Elgart and Kamenev and of Assaf and Meerson [35–37], and the spectral method of Walczak, Mugler, and Wiggins [34]. 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 (109) 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 [35]. 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 (12). 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 [32, 33].

3.1. Flow of the marginalized distribution

The generalized probability generating function (45) 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 $\tau$ variable. By summing this distribution over a set of basis functions $\{|n_0\rangle\}$, one can define the series

$$p(t, n|\tau, n_0) := \sum_{n_0} p(t, n|\tau, n_0)|n_0\rangle.$$  \hspace{1cm} (110)

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)thogonal basis, the conditional probability distribution can be recovered from (110) via

$$p(t, n|\tau, n_0) = \langle n_0| p(t, n|\tau)\rangle.$$  \hspace{1cm} (111)

We call the function $\langle p(t, n|\tau)\rangle$ a ‘marginalized distribution’ because it proves most useful when the summation in (110) 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’ $|n_0\rangle := e^{in_0\tau}$, which we make heavy use of this basis function in the study of chemical reactions in section 3.2.2. Since the definition of the marginalized distribution in (110) does not affect the random variable $n$, the marginalized distribution of course solves the forward master equation $\partial_\tau |p(t, n|\tau)\rangle = \sum_{\langle m|} Q_\tau\langle n, m| p(t, m|\tau)\rangle$ with the initial condition $\langle p(t, n|\tau)|n_0\rangle$. 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 $|n_0\rangle := e^{in_0\tau}$, which we consider in section 3.2.1.

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 $\partial_\tau |p(t|\tau)\rangle = p(t|\tau)Q_\tau$ for this purpose (recall that $p(t|\tau)$ is the matrix with elements $p(t, n|\tau, n_0)$). Upon differentiating the marginalized distribution (110) with respect to the time parameter $\tau$, one obtains the equation

$$\partial_\tau |p\rangle = \sum_{n_0} p(\tau, n_0) \left(-\partial_\tau |n_0\rangle + \sum_{m} Q_\tau^\dagger\langle m|Q_\tau\langle n_0, m\rangle\right).$$  \hspace{1cm} (112)

The rate $Q_\tau^\dagger\langle m, n_0| = Q_\tau\langle n_0, m\rangle$ represents an element of the transposed transition matrix $Q_\tau^\dagger$.

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 $Q_\tau^\dagger(x, \partial_x)$ is required to
encode the information stored in the transition matrix. This operator should be a power series in $\partial_t$ and fulfill, for all $n_0$,

$$Q^\dagger_t(n_0) = \sum_m |m\rangle Q_t^m(m,n_0) \langle n_0| \tag{Q^\dagger}.$$ 

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

$$Q^\dagger_t := \sum_{m,n_0} |m\rangle Q^m_t(m,n_0)(n_0). \tag{114}$$

We wrote these expressions in terms of the transposed transition matrix because for the unit column vectors $|m\rangle = e_m$ and the unit row vectors $|n\rangle = e_n^\dagger$, $Q^\dagger_t$ and $Q_t^\dagger$ coincide (with $m,n \in \mathbb{N}_0$). As we mostly consider bases whose kets and bras represent functions and functionals, we call $Q^\dagger_t$ the ‘adjoint’ transition operator in the following. Often, an operator $O'(x, \partial_x)$ 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 $g$:

$$\int dx \ [O'(x, \partial_x)f(x)]g(x) = \int dx f(x)[O(x, \partial_x)g(x)]. \tag{115}$$

However, whether an operator $Q_t$, complementing the above $Q^\dagger_t$, 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 $E_t$ and an adjoint transition operator $Q_t^\dagger$ are found for a particular process, it follows from the backward-time equation (112) that the marginalized distribution $|p(t, n|\tau)\rangle$ obeys the flow equation$^8$

$$\partial_{\tau} |p\rangle = (-E_t + Q^\dagger_t)|p\rangle =: \nabla^\dagger_p |p\rangle. \tag{116}$$

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

In section 4, we show how the flow equation (116) can be solved in terms of a ‘backward’ path integral. Provided that the basis function $|n\rangle$ is chosen as 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 (45). 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 $\langle A \rangle = \sum_n A(n) \langle p(t, n|\tau)\rangle$.

$^8$In the derivation of the backward path integral in section 4.1, we use the finite difference approximation

$$\lim_{\Delta \tau \to 0} (|p(t|\tau - \Delta \tau)\rangle - |p(t|\tau)\rangle)/\Delta \tau = \sum_{\Delta \tau} Q^\dagger_{t+\Delta \tau}|p(t|\tau)\rangle.$$ 

The discretization scheme conforms with the derivation of the backward master equation (12).

\[ \text{Note that this average also obeys the flow equation (116). In section 3.3, we demonstrate how the flow equation can be used to compute mean extinction times.} \]

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 $\tau$. The above flow equation can then be solved by 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 (110), let us reconsider the random walk from section 2.2.1 and the chemical reaction from section 2.2.2. 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 [32, 33] in section 3.5.

#### 3.2.1. Random walks

The following solution of the random walk largely parallels the previous derivation in section 2.2.1. In particular, we again use the orthogonal and complete Fourier basis $|n\rangle_\pi = e^{in\pi}$ and $|n\rangle_n = \int e^{ikx} f(x) \, dx$ with $n \in \mathbb{Z}$. Due to the time-independence of the basis, the (E)volution operator $E_t$ is zero. The condition $Q^\dagger_{n_0}(n_0) = r_n(|n_0 - 1| - |n_0\rangle) + l_n(|n_0 + 1| - |n_0\rangle)$, (117)

As before, we employ the operators $c(x) = e^{ix}$ and $a(x) = e^{-ix}$, which act on the basis functions as $c(n) = |n + 1|$ and $a(n) = |n - 1|$, respectively. Therefore, the adjoint transition operator with the above property can be defined as

$$Q^\dagger_n(x, \partial_x) := r_n(a(x) - 1) + l_n(c(x) - 1). \tag{118}$$

As this operator does not contain any derivatives, it is self-adjoint in the sense of (115). Due to a mismatch in signs, it is, however, not the adjoint of the previous operator $Q_\pi$ in (61). This mismatch could be corrected by redefining the above basis function as $|n\rangle_\pi := e^{-in\pi}$. Ignoring this circumstance, the flow equation of the marginalized distribution follows as

$$\partial_{\tau} |p\rangle = [r_n(e^{-ix} - 1) + l_n(e^{ix} - 1)] |p\rangle, \tag{119}$$

and is solved by

$$|p\rangle = \exp(e^{-ix} - 1) \int e^{ix} ds \, r_n(e^{ix} - 1) \int e^{-ix} ds \, l_n |n\rangle. \tag{120}$$

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 $x \rightarrow -q$, $\tau \rightarrow 0$ and $t \rightarrow \tau$). Eventually, one recovers a Skellam distribution as the solution of the process.
3.2.2. 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 \( k A \rightarrow I A \). 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 2.2.2, we require an (O)rthogonal and (C)omplete basis, as well as an (E)volution operator \( E \), and an adjoint transition operator \( Q^l \) (condition (Q1)) to specify the flow equation of the marginalized distribution.

As discussed in section 2.2.2, the elements of the transition rate matrix of the reaction \( k A \rightarrow I A \) with rate coefficient \( \gamma_x \) are given by \( Q_x(m,n) = \gamma_x(n)b_{m,n-k+1} - b_{m,n} \). The condition (Q1) on the adjoint transition operator therefore reads

\[
Q^l_x|n_0\rangle = \gamma_x((n_0 - l + k)b_{n_0}|n_0 - l + k\rangle - (n_0)b_{n_0}|n_0\rangle). \tag{121}
\]

This condition is met by

\[
Q^l_x(c,a) := \gamma_xc^\dagger(a^\dagger - a^2), \tag{122}
\]

provided that there exist operators \( c \) and \( a \) fulfilling

\[
c|n\rangle = (n + 1)|n + 1\rangle \quad \text{and} \quad a|n\rangle = |n - 1\rangle, \tag{123}
\]

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 (65) and (66). Similar relations also hold with respect to the basis functionals, namely \( \langle n|c = n(n - 1) \) and \( \langle n|a = (n + 1) \) (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 2.2.2 can be adapted to meet the present requirements. In particular, the relations (123) and (124) can be fulfilled by moving the factorial from the basis functional (76) to the basis function (70) so that

\[
|n\rangle_x := \frac{(|\chi + \tilde{x}e^{-\delta t}\rangle e^{-\theta\langle\chi + \tilde{x}\rangle}}{n!}, \tag{125}
\]

and

\[
\langle n|f := a^\dagger f(x)|x\rangle e^{-\theta\langle\chi + \tilde{x}\rangle}. \tag{126}
\]

Let us briefly note that we changed the argument of the basis function from \( q \) to \( x \) as compared to section 2.2.2 because both the generating function approach and the marginalized distribution approach thereby result in the same path integral representation of averaged observables (see section 6). Apart from this notational change, the basis evolution operator

\[
E_x(c,a) := (\partial_x\tilde{x})(a - \tilde{q}) - (\partial_x\tilde{q})c \tag{127}
\]

keeps the form it had in (77). The creation and annihilation operators also keep their previous forms in (73) and (74), i.e.,

\[
c(x, \partial_x) := \chi x + \tilde{x} \tag{128}
\]

and

\[
a(x, \partial_x) := \partial_x x + \tilde{q}. \tag{129}
\]

Despite their similar appearance, the operator \( Q^l_x \) in (122) is not the adjoint of the operator \( Q_x \) in (64). Nevertheless, the two operators fulfill \( Q_x(q, x) = Q^l_x(x, q) \) for scalar arguments. This relation is essentially the reason why we interchanged the letters \( x \) and \( q \) as compared to section 2.2.2. 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 \( \zeta = 0 \), \( x(\tau) \), and \( \tilde{q}(\tau) \) in the basis function (125) depends on the problem at hand. In section 7, \( x \) and \( \tilde{q} \) will act as ‘stationary’ or ‘extremal’ paths, with \( x \) and an auxiliary variable \( q \) being deviations from them. For \( \zeta := \tilde{q} := 0 \) and \( \tilde{x} := 0 \), the basis function instead simplifies to the Poisson distribution

\[
|n\rangle_x = \frac{x^n e^{-x}}{n!}. \tag{130}
\]

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 \( A \rightarrow \mathcal{D} \) with rate coefficient \( \mu_x \). For the above choice of \( \zeta, \tilde{x}, \text{and} \tilde{q} \), the creation operator (128) reads \( c = x \) and the annihilation operator (129) reads \( a = \partial_x + 1 \). With the transition operator (122), the flow equation (116) obeyed by the marginalized distribution \( |p(t, n)| \rangle_x \) follows as

\[
\partial_{\tau_x}|p\rangle_x = -\mu_x\partial_{\tau}|p\rangle_x. \tag{131}
\]

This equation is solved by the Poisson distribution

\[
|p(t, n)| \rangle_x = \frac{(\alpha_x x)^n e^{-\alpha_x x}}{n!} \tag{132}
\]

whose mean \( \alpha_x x \) decays proportionally to \( \alpha_x \): \( := -\int \mu_x \). To interpret this solution, let us recall that the sum in the definition of the marginalized distribution (110) does not affect the particle number \( n \). Therefore, upon relabelling the time parameters, the above solution (132) also solves the forward master equation of the process, namely

\[
\partial_{\tau_x}|p(\tau, n)| \rangle_x = \mu_x(n + 1)|p(\tau, n + 1)| \rangle_x - n|p(\tau, n)| \rangle_x. \tag{133}
\]

Unlike the conditional distribution \( p(\tau, n)|t_0, n_0 \rangle \), however, the marginalized distribution \( |p(\tau, n)| \rangle_x \) 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 \( p(\tau, n)|t_0, n_0 \rangle = \langle n_0| p(\tau, n)|t_0 \rangle \) with \( \langle n_0| f = (\partial_x + 1)^n f(x)|_{x = 0} \).

This transformation results in the Binomial distribution

\[
p(\tau, n)|t_0, n_0 \rangle = \binom{n_0}{n} (\alpha_x x)^n (1 - \alpha_x) x^{n - n} \tag{134}
\]

Both the mean value \( e^{-\int \mu_x \nu_0} \) and the variance \( (1 - e^{-\int \mu_x \nu_0}) e^{-\int \mu_x \nu_0} \) of this distribution decay exponentially for large times, provided that \( \mu_x > 0 \).
If at most two reactants and two products are involved in a reaction, the flow equation (116) of the marginalized distribution has the mathematical form of a backward Fokker–Planck equation. That is, for example, the case for the coagulation reaction \( 2A \rightarrow A \). For the Poisson basis function, the marginalized distribution \( |p(t, n|\tau)\rangle \) of this process obeys the flow equation

\[
\partial_\tau |p\rangle = \alpha(x) \partial_x |p\rangle + \frac{1}{2} \beta(x) \partial^2_x |p\rangle, \tag{135}
\]

with the drift coefficient \( \alpha(x) := -\mu_x x^2 \) and the diffusion coefficient \( \beta(x) := -2\mu_x x^2 \). Unlike the diffusion coefficient of the ‘true’ backward Fokker–Planck equation (2), this diffusion coefficient may be negative (e.g. for \( x \in \mathbb{R} \setminus \{0\} \)). The final condition \(|p(t, n|t)\rangle \rangle_0 = x^n e^{-x^2/n} \) and the Feynman–Kac formula (A.3) imply that the above flow equation is solved by

\[
|p(t, n|\tau)\rangle_\pi = \left( \frac{x(t) e^{-\tau(x)}}{n!} \right)_W. \tag{136}
\]

Here, \( x(s) \) obeys the Itô SDE

\[
dx(s) = \alpha(x(s)) ds + \sqrt{\beta(x(s))} DW(s), \tag{137}
\]

which evolves \( x(s) \) from \( x(\tau) = x \) to \( x(t) \). The symbol \( \langle \cdot \rangle_W \) represents an average over realizations of the Wiener process \( W \). Since the diffusion coefficient \( \beta(x) = -2\mu_x x^2 \) 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 (136) for the calculation of \(|p(t, n|\tau)\rangle\), although it may complicate numerical evaluations. Recently, Wiese attempted the evaluation of the average in (136) via the generation of sample paths [336].

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 (135) to enable its direct numerical integration (an analogous problem is encountered for the flow equation of the generating function, see [366]).

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 autocatalysis, 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 [110, 215].

In [374], Drummond et al recently showed how the calculation can be simplified using the Poisson representation of Gardiner and Chaturvedi [32, 33]. 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 \(|n_0\rangle = x^n e^{-x}/n!\), 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 \( k A \rightarrow I A \). In addition, the transition matrix \( Q^\dag = Q \) 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 \(|p(\tau, n|0\rangle\rangle_\pi\), provided that its basis function is chosen as \(|n_0\rangle = x^n e^{-x}/n!\) (see the definition (110) of the marginalized distribution). Since the Poisson basis function is independent of time and the process under consideration homogeneous in time, the marginalized distribution \(|p(\tau, n|0\rangle\rangle_\pi\) obeys the forward-time flow equation (see (116))

\[
\partial_\tau |p\rangle = Q^\dag |p\rangle. \tag{138}
\]

We define the probability of finding the system in an ‘active’ state with \( n > 0 \) particles at time \( \tau \geq 0 \) as

\[
\alpha(\tau; x) := \sum_{n=1}^\infty |p(\tau, n|0\rangle = 1 - |p(\tau, 0|0\rangle. \tag{139}
\]

Over time, this probability flows into the ‘absorbing’ state \( n = 0 \) at rate \( f(\tau, x) := -\partial_\tau \alpha(\tau; x) \). Since \( f(\tau, x) d\tau \) is the probability of becoming absorbed during the time interval \([\tau, \tau + \Delta t]\), one can define the mean extinction time as \( \langle \tau \rangle_\pi := \int_0^\infty d\tau f(\tau, x) \). An integration by parts transforms this average into

\[
\langle \tau \rangle_\pi = - \int_0^\infty d\tau \partial_\tau \alpha(\tau; x) = \int_0^\infty d\tau \alpha(\tau; x). \tag{140}
\]

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 (139) implies that it fulfills the same forward-time flow equation as the marginalized distribution \(|p(\tau, n|0\rangle\rangle_\pi\). Since \( \lim_{\tau \to \infty} \alpha(\tau; x) = 0 \), the mean extinction time (140) therefore obeys

\[
Q^\dag (x, \partial_x) \langle \tau \rangle_\pi = \alpha(0; x) = 1 - e^{-x}. \tag{141}
\]

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 \( A \rightarrow 2 \) with decay rate coefficient \( \mu \), the equation (141) for the mean extinction time reads

\[
\mu \partial_x \langle \tau \rangle_\pi = (1 - e^{-x})x. \tag{142}
\]
This equation implies that the mean extinction time $\mu(\tau)_0$ increases logarithmically with the particles’ mean initial distance $x$ from the absorbing state (see figure 5). The explicit solution of the equation is $\mu(\tau)_0 = \ln x + \gamma - \text{Ei}(-x)$ with Euler’s constant $\gamma$ and the exponential integral $\text{Ei}$ (6.2.6 in [335]; the exponential integral ensures that $\langle \tau \rangle_0 = 0$). The application of the functional $\langle n_0| f = \langle \partial\partial + 1 \rangle_0^m f(x) \rangle_{\text{oo}}$ 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 $\mu(\tau)_n$ is given by the harmonic number $H_{n_0} := \sum_{i=1}^{n_0} \frac{1}{i}$. One can also infer this result directly from the backward master equation using the methods discussed in [110, 215].

Drummond et al have extended the above computation to chemical reactions with at most two reactants and two products [374]. For these reactions, the flow equation (138) has the mathematical form of a (forward) Fokker–Planck equation.

### 3.4. Flow of the generating functional

Both the probability generating function (45) and the marginalized distribution (110) 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’

$$\langle g(\tau)|t_0, n_0 \rangle := \sum_n \langle n| p(\tau, n)|t_0, n_0 \rangle$$

(143)

by summing the conditional distribution over the set $\{|n\}$ of basis functionals. The definition of the generating functional is meant with respect to test functions that can be expanded in the basis functions $\{|n\}$. 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

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

(144)

In the next section, we show how this inverse transformation reduces to the Poisson representation of Gardiner and Chaturvedi [32, 33] upon choosing the basis function $|n\rangle$ as a Poisson distribution.

The derivation of the (functional) flow equation of $\langle g \rangle$ proceeds analogously to the derivations in sections 2.1 and 3.1. Differentiation of its definition (143) with respect to $\tau$ and use of the forward master equation $\partial_\tau p(\tau)|t_0 \rangle = Q_0 p(\tau)|t_0 \rangle$ result in

$$\partial_\tau \langle g \rangle = \sum_n \partial_\tau \langle n| + \sum_m Q_0^m \langle n, m| \rangle \langle p(\tau)|n\rangle.$$

(145)

This equation can be cast into a linear PDE by using the basis (E)volution operator $\mathcal{E}_\tau$ and the adjoint transition operator $Q_0^\dagger$ (condition (Q)). In particular, since the Kronecker delta $\langle n|m \rangle = \delta_{nm}$ is independent of time, it holds that $\langle n|\partial_\tau|m \rangle = \langle -\partial_\tau(n)|m \rangle$. Consequently, the basis evolution operator in condition (E) fulfills

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

(146)

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 (Q) acts on basis functionals as

$$\langle n|Q_0^\dagger = \sum_m Q_0^m \langle n, m|.$$  

(147)

Both of the above expressions hold for all values of $n$. If the two differential operators $\mathcal{E}_\tau$ and $Q_0^\dagger$ exist, the generating functional $\langle g(\tau)|t_0, n_0 \rangle$ obeys the functional flow equation

$$\partial_\tau \langle g \rangle = \langle g(-\mathcal{E}_\tau + Q_0^\dagger) \rangle \langle g|Q_0^\dagger \rangle,$$

(148)

with initial condition $\langle g(t_0)|t_0, n_0 \rangle = \langle n_0 \rangle$. Note that this flow equation employs the same operator as the flow equation (116) of the marginalized distribution. Both equations can be used to derive the ‘backward’ path integral representation considered in section 4.

As a side note, let us remark that the flow equation $\partial_\tau \langle g \rangle = \langle g(\mathcal{E}_\tau + Q_0) \rangle \langle g|\mathcal{E}_\tau \rangle$ of the generating function in (54) also admits a functional counterpart. In particular, the series

$$\langle p(t, n)|\rangle := \sum_{n_0} \langle p(t, n)|t_0, n_0 \rangle|n_0\rangle$$

(149)

obeys the flow equation

$$\partial_\tau \langle p \rangle = \langle p|\mathcal{E}_\tau + Q_0 \rangle \langle p|,$$

(150)

with final value $\langle p(t, n)|t \rangle = \langle n \rangle$. The corresponding inverse transformation reads $p(t, n|t_0) = \langle p(t, n)|t_0 \rangle$. 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. Further uses of the series (149) remain to be explored.

### 3.5. The Poisson representation

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

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{poisson_distribution.png}
\caption{Mean extinction time $\mu(\tau)$ of the linear decay process $A \rightarrow \emptyset$ with decay rate coefficient $\mu$. The blue line represents the mean extinction time if the number of particles is initially Poisson distributed with mean $x \in \mathbb{R}_{\geq 0}$ ($\mu(\tau)_0 = \ln x + \gamma - \text{Ei}(-x)$). The red circles represent the mean extinction time if the initial number of particles is set to $n_0 \in \mathbb{N}_0$ ($\mu(\tau)_0 = H_{n_0}$).
}
\end{figure}
The insertion of the Poisson basis function \( |n\rangle = \frac{e^x}{n!} \) into the inverse transformation (144) results in

\[
p(\tau, n|t_0, n_0) = \int_{-\infty}^{\infty} dx \, g(\tau, x|t_0, n_0) \frac{e^{x}e^{-x}}{n!}. \tag{152}
\]

A representation of the probability distribution of this form is called a ‘Poisson representation’ and was first proposed by Gardiner and Chaturvedi [32, 33]. Since the integration in (152) proceeds along the real line, the above representation is referred to as a ‘real’ Poisson representation [375]. 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 \( g(t_0, x|t_0, n_0) = \delta(x)(\partial_+ + 1)^n \). By the definition

\[
\int_{-\infty}^{\infty} dx \, (\partial_+ \delta(x)) f(x) = \int_{-\infty}^{\infty} dx \, \delta(x)(-\partial_-) f(x) \tag{153}
\]

distributional derivatives (1.16.12 in [335], \( j \in \mathbb{N} \)), this kernel can also be written as \( ((1 - \partial_-)^n \delta(x)) \). The integral kernel \( \delta(x - \omega_0) \) instead results for a Poisson distribution with mean \( \omega_0 \). In [32], 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 [376] to deduce a stochastic path integral representation. Droz and McKane [377] 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 [110, 375] as well as the ‘gauge’ Poisson representation [378]. The former two representations as well as the real representation are discussed in the book of Gardiner [110]. The complex Poisson representation is obtained by continuing the Poisson basis function in (152) into the complex domain and performing the integration around a closed path \( C \) around 0 (once in counter-clockwise direction). Upon using Cauchy’s differentiation formula to redefine the basis functional as

\[
\langle n | f := \partial_+ e^f(x) | x = 0 = \int_C dx \, \frac{e^x}{2\pi i x^{n+1}} f(x), \tag{154}
\]

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 \( g(t_0, x|t_0, n_0) = \frac{e^x e^{-x}}{n!} \). As the interpretation of this kernel is also not straightforward, we refrain from calling it a ‘quasi-probability distribution’ [32].

Let us exemplify the flow equation obeyed by the integral kernel for the reaction \( k A \rightarrow 2A \) with rate coefficient \( \gamma \). The flow equation can be inferred from the corresponding flow equation (148) of the generating functional and the kernel’s definition in (151). Since we employ the Poisson basis function, the results from section 3.2.2 imply that the adjoint transition operator (122) reads

\[
Q^T(n) = \gamma e^{x}(\partial_+ + 1)^{2} - (\partial_+ + 1)^{2}. \tag{155}
\]

Consequently, one finds that the integral kernel \( g(\tau, x|\cdot) \) obeys the flow equation

\[
\partial_\tau g = Q(\tau(x, \partial_\tau) g \tag{156}
\]

\[
= \gamma.(1 - \partial_\tau)^{2} - (1 - \partial_\tau)^{2} x^{2} g. \tag{157}
\]

To arrive at this equation, we performed repeated integrations by parts while ignoring any potential boundary terms (see the definition of the adjoint operator in (115)). The importance of boundary terms is discussed in [378].

Thus far, most studies employing the Poisson representation have focused on networks of bimolecular reactions \( \sum j_l A_j \rightarrow \sum k_j A_j \) with \( \sum j_k < 2 \) and \( \sum j_l < 2 \). For these networks, the flow equation (156) 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 [32], 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 \( 2A \Rightarrow A + A \) in [379]. 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 [380] (see section 2.3.2). Moreover, a gauge Poisson representation was recently employed in a study of the coagulation reaction \( 2A \rightarrow A \) [381].

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 (45). In the present section, we complemented this flow equation by a backward-time flow equation obeyed by the marginalized distribution (110) and by a functional flow equation obeyed by the probability generating functional (143). Whereas the marginalized distribution

\[
|p(t, n|\tau) = \sum_n p(t, n|\tau, n_0)|n_0\rangle \tag{158}
\]

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

\[
|g(\tau|t_0, n_0) = \sum_n |p(\tau, n|t_0, n_0)\rangle \tag{159}
\]

was defined as the sum of the distribution over a set of basis functionals. In section 3.2, we introduced (O)thogonal and (C)omplete basis functions and functionals for the study of different stochastic processes, including the Poisson basis for the study of chemical reactions \( |m\rangle = \frac{e^x}{n!} \) and \( \langle n| f = (\partial_+ + 1)^{m} f(x) \mid x = 0 \). Provided that there also exist a
basis (E)volution operator $E_x$, and an adjoint transition operator $Q^\dagger_{\tau}$ (condition (Q3)), the marginalized distribution obeys the backward-time flow equation

$$\partial_{\tau} p(t,|\tau|p) = (-E_x + Q^\dagger_{\tau}) p(t,|\tau|p)$$

(160)

with final condition $p(t,|\tau|p) = |\eta|$. Moreover, the probability generating functional (143) then obeys the functional flow equation (148). The inverse transformation $p(\tau, t, n|\tau_0, n_0) = g(\tau|t, n_0)|n|$ of this functional generalizes the Poisson representation of Gardiner and Chaturvedi [32, 33] as shown in section 3.5. In section 3.3, we showed how the flow equation (160) 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 (41)–(44). 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 (41) and by the series (44) will lead us to the ‘forward’ path integral representation

$$p(t, n|t_0, n_0) = \langle n_0 | e^{-S} | n \rangle_{t_0},$$

(161)

and the flow equations obeyed by the marginalized distribution (42) and by the generating functional (43) to the ‘backward’ path integral representation

$$p(t, n|t_0, n_0) = \langle n_0 | e^{-S^\dagger} | n \rangle_{t_0}$$

(162)

The meanings of the integral signs $f^{(1)}_{\tau_0}$ and $f^{(-1)}_{\tau_0}$ 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 $\langle n |$. Counter-intuitively, however, this procedure requires us to solve a stochastic differential equation proceeding backward in time (see sections 5.2 and 5.3). Analogously, the backward path integral representation propagates the basis function $|n\rangle$ back to time $t_0$, where one then applies the functional $\langle n_0 |$. This procedure requires us to solve an ordinary, forward-time Itô SDE (see section 4.3). The name ‘adjoint’ path integral representation may also be used for the backward representation. As we will see below, its ‘action’ $S^\dagger$ involves the adjoint transition operator $Q^\dagger_{\tau}$.

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 [382], 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 4.5.2–4.5.4.

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 $\varnothing \rightleftharpoons A$ (section 4.2), the pair generating process $\varnothing \rightarrow 2A$ (section 4.5.1), and a process with diffusion and linear decay (sections 4.5.2–4.5.4). The forward path integral representation is derived in section 5 and is exemplified in deriving the generating function of linear processes $A \rightarrow I A$ with $I \geq 0$. For the linear growth process $A \rightarrow 2A$, 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 5.3 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 [16], de Dominicis [17], Janssen [18, 19], and Bausch, Janssen, and Wagner [19]. Moreover, we recover the Feynman–Kac formula from appendix A, an Onsager–Machlup representation [79], and Wiener’s path integral for Brownian motion [80, 81].

Before starting out with the derivations of the two path integral representations (161) and (162), let us note that these representations only apply if their underlying transition operators $Q^\dagger_{\tau}$ or $Q^\dagger_{\tau}$ can be written as power series in their arguments $(x, x')$ for the backward path integral and $q$ and $\partial_q$ 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 $\partial_x$ (all the $q$ to the left of all the $\partial_q$). This order can be established by the repeated use of the commutation relation $[\partial_x, f](x) = f(x)$ (or, more directly, by invoking $\langle \partial_x f(x) = \sum_{m=0}^n \sum_{m=0}^{\infty} \frac{x^{m+1}}{m+1} \frac{m!}{m!} \frac{\partial^m f(x)}{\partial x^m}$ (or $(x, x') f(x) = \sum_{m=0}^n \sum_{m=0}^{\infty} \frac{\partial^m f(x)}{\partial x^m}$, with the curly braces representing Stirling numbers of the second kind; see section 26.8 in [335]). The transition operator $Q^\dagger_{\tau} = \gamma_c e^{(c' - c)q}$ in (64) and the adjoint transition operator $Q^\dagger_{\tau} = \gamma_c q^\dagger (c' - c^\dagger)$ in (122) of the chemical reaction $A \rightarrow I A$ are already in their
normal-ordered forms (with the creation and annihilation operators in (73) and (74), or (128) and (129), respectively).

4.1. Derivation

We first derive the backward path integral representation (162) 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 \( \partial_{\tau} \{ p(\tau) \} = \mathcal{Q}_t \{ p(\tau) \} \) in (116) obeyed by the marginalized distribution. Its final condition reads \( p(t, n|t) = |n\rangle \).

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 next step of the derivation, we insert the integral \( \int_{t_0}^t dx^j \) in the above expression and pull \( \partial_t \) to the right of the exponential. Making use of the final condition \( \partial_t \{ p(\tau) \} = \mathcal{Q}_t \{ p(\tau) \} \) in (116) obeys the following discrete-time path integral representation:

\[
\langle p(t, n|t) \rangle_{n_0} = \int_{t_0}^t dx^j \mathcal{L}_n \{ x^j \} |n\rangle_{n_0}, \tag{163}
\]

where \( \mathcal{L}_n \{ x^j \} = \mathcal{Q}_t \{ x^j \} + O((\Delta t)^2) \),\(^9\) alternately, the following derivation could be performed by solving the flow equation (148) obeyed by the generating functional over the time interval \([t_0, t] \) as (see figure 6)

\[
\langle g(t|t_0, n_0) \rangle = \langle g(t_{N-1}|t_0, n_0) \rangle |\mathcal{L}_n^{[1]}(x^j_{N-1}, t_{N-1}, \partial_{x_{N-1}})\rangle. \tag{164}
\]

Here, the generating functional \( g(\cdot) \) acts on functions \( f(x_{N-1}, \cdot) \), meets the initial condition \( g(t_0|t_0, n_0) = \langle n_0 \rangle \), and is transformed back into the conditional probability distribution via

\[
p(t, n|t_0, n_0) = \langle g(t_0|t_0, n_0) | n \rangle. \tag{165}
\]

Both of the above approaches result in the same path integral representation. In the following, we use the equation (163) for this purpose.

As the next step of the derivation, we insert the integral representation of a Dirac delta between the generator \( \mathcal{L}_n \) and the marginalized distribution \( | p \rangle \) in (163), turning the right-hand side of this expression into

\[
\int dx^j \frac{dx^j}{2\pi} e^{-i\phi(x^j - x^j_{N-1})} |p(t, n|t_0)\rangle_{n_0}, \tag{166}
\]

The integrations over \( x_j \) and \( q_j \) are both performed along the real line from \(-\infty \) to \(+\infty \). If the adjoint transition operator \( \mathcal{Q}_t^\dagger \), and thus also \( \mathcal{L}_n^\dagger \), are normal-ordered power series, we can replace \( \partial_{x_{N-1}} \) by \( i\partial_{q_{N-1}} \) in the above expression and pull \( \mathcal{L}_n^\dagger \) to the right of the exponential. Making use of the final condition \( \partial_t \{ p(\tau) \} = \mathcal{Q}_t \{ p(\tau) \} \) in (116) obeys the above steps can be repeated until (165) reads

\[
\int dx^j \frac{dx^j}{2\pi} e^{-i\phi(x^j - x^j_{N-1})} |\mathcal{L}_n^{[1]}(x^j_{N-1}, t_{N-1}, \partial_{x_{N-1}})\rangle \langle n_0|, \tag{167}
\]

To proceed, we now replace \( \mathcal{L}_n^\dagger = 1 + \mathcal{Q}_t^\dagger \Delta t + O((\Delta t)^2) \) by the exponential \( e^{\mathcal{Q}_t^\dagger \Delta t} \) (for brevity, we drop the correction term in the following). Note that the exponential \( e^{\mathcal{Q}_t^\dagger \Delta t} \) does not involve differential operators because those were all replaced by the new variables \( q_j \). 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:

\[
\langle p(t, n|t) \rangle_{n_0} = \int dx^j \frac{dx^j}{2\pi} e^{-S_N} |n\rangle_{n_0} \tag{168}
\]

where \( S_N := \sum_{j=1}^N \left( q_j^2 - x_j^2 \right) \Delta t - \mathcal{L}_n^{[1]}(x^j_{N-1}, \partial_{x_{N-1}}) \).\(^{10}\) The final condition \( \langle p(t, n|t) \rangle = \langle n| n \rangle \) is trivially fulfilled because only \( N = 0 \) time slices fit between \( t \) and \( t \). The object \( S_N \) is called an ’action’. The exponential factor \( e^{-S_N} \) weighs the contribution of every path \( (q, x) \) 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 \( x(t_0 + j\Delta t) := x_j \) and \( q(t_0 + j\Delta t) := q_j \).

Further comments on the discretization scheme are provided in section 4.4. Combined with the inverse transformation (111), we thus arrive at the following backward path integral representation of the conditional probability distribution:

\[
p(t, n|t_0, n_0) = \langle n_0| |p(t, n|t_0)\rangle_{x_{N_{st}}}, \tag{170}
\]

with \( |p(t, n|t_0)\rangle_{x_{N_{st}}} = \int dx^j \frac{dx^j}{2\pi} e^{-S} |n\rangle_{n_{st}} \).\(^{11}\)
and \( S^t := \int_0^t \, \mathrm{d}t \, [iq \partial_x x - \mathcal{Q}_x(x, iq)]. \) \hspace{1cm} (172)

Here we included \( x(t_0) \) as an argument of the functional \( \langle n_0(x(t_0)) \rangle \) to express that this functional acts on \( x(\tau) \) only for \( \tau = t_0 \). Moreover, we defined \( F^{(t_0)} := \lim_{N \to \infty} \int_0^N \, \mathrm{d}t \) to indicate that the path integral involves integrations over \( x(\tau) \) and \( q(\tau) \), 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 (171) provides the marginalized distribution \( p(t, n|t_0) \). Second, this marginalized distribution is mapped to the conditional probability distribution \( p(t, n|t_0, n_0) \) by the action of the functional \( \langle n_0 \rangle \).

### 4.2. Simple growth and linear decay

Let us demonstrate the above two steps for the bi-directional reaction \( \varnothing \rightleftharpoons A \). The stationary distribution of this process was already derived in this introduction to section 2. Both the growth rate coefficient \( \gamma \) and the decay rate coefficient \( \mu \) shall be time-independent, but this assumption is easily relaxed. As the first step, the backward path integral (170) requires us to choose an appropriate basis. The time-independent Poisson basis with the basis function \( |n\rangle_x = \frac{x^ne^{-x}}{n!} \) proves to be convenient for this purpose (see section 3.2.2). The adjoint transition operator follows from (122), (128), and (129) as \( \mathcal{Q}_x^\dagger(x, \partial_x) = (\gamma - \mu x) \partial_x \). Insertion of this operator into the action (172) results in

\[
S^t = \int_0^t \, \mathrm{d}t \, jq \partial_x x - (\gamma - \mu x) \] \hspace{1cm} (173)

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 \( |p(t, n|t_0)\rangle_{(t_0)} \) with \( x(t_0) \equiv x_0 \) thereby follows as the product of Dirac deltas:

\[
\left( \prod_{j=1}^N \, \mathrm{d}x_j \, \delta(x_j - x_{j-1} - (\gamma - \mu x_{j-1})Dt) \right) |n\rangle_{x_0}. \]

The function \( |n\rangle_{x_0} \) 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

\[
|p(t, n|t_0)\rangle_{(t_0)} = |n\rangle_{x(t)} = \frac{x(t)^n e^{-x(t)}}{n!}, \]

with \( x(\tau) \) solving \( \partial_x x = \gamma - \mu x \). This equation coincides with the rate equation of the process \( \varnothing \rightleftharpoons A \). The solution \( x(\tau) = e^{-\mu(t-t_0)}(x(t_0) - \frac{\gamma}{\mu}) + \frac{\gamma}{\mu} \) 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 \( \gamma/\mu \). In the next section and in section 4.5, we generalize these results to more general processes.

The marginalized distribution (174) does not only solve the master equation of the reaction \( \varnothing \rightleftharpoons A \), 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 \( (n)(t) = x(t) \), while higher order moments can determined via \( (\mu^n)(t) = \sum_{j=0}^n \frac{1}{j!} j^x(t)^j \). The curly braces denote a Stirling number of the second kind (see section 26.8 in [335]).

The conditional probability distribution can be calculated from the marginalized distribution by applying the functional \( |n_0\rangle f = (\partial_{t_0} + \frac{\gamma}{\mu})f(x(t_0)) |(x(t_0) = 0) \) to the latter (see (126) and (129)). In the limit of a vanishing decay rate \( \mu \to 0 \), for which \( x(t) = x(t_0) + \gamma (t - t_0) \), we thereby recover the shifted Poisson distribution (80), i.e.

\[
p(t, n|t_0, n_0) = e^{-(t-t_0)(\gamma(t-t_0))}e^{-n_0} \Theta_{n_0-n_0}. \]

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 (134).

### 4.3. Feynman–Kac formula for jump processes

We now show how the backward path integral representation (171) of the marginalized distribution can be expressed in terms of an average over the paths of an Ito stochastic differential equation (SDE). The resulting expression bears similarities with the Feynman–Kac formula (3), especially when the adjoint transition operator \( \mathcal{Q}_x^\dagger(x, \partial_x) \) of the stochastic process under consideration is quadratic in \( \partial_x \). 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 4.5.4. The procedure outlined below serves as a general starting point for the exact or approximate evaluation of the backward path integral (171).

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

\[
\mathcal{Q}_x^\dagger(x, \partial_x) = \alpha_x(x) \partial_x + \frac{1}{2} \beta_x(x) \partial_x^2 + \mathcal{P}_x(x, \partial_x). \]

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

As the first step, let us rewrite the backward path integral representation (171) of the marginalized distribution as

\[
|p(t, n|t_0)\rangle_{(t_0)} = \int_{\mathcal{D}^N} \, \frac{\mathrm{d}x_0 \cdots \mathrm{d}x_N}{2\pi} e^{-ik_{N,0} x_0} Z_0 |(n_0)\rangle_{x_0}, \]

where \( Z_{p=0,x=0} \) represents a value of the \((Q,X)\)-generating functional.
\[ Z_{Q,X} := \int_{t_0}^{t_1} e^{i q_x x - Сd + \int_0^t d\tau \left[ Q(\tau) x(\tau) + X(\tau) i q(\tau) \right]}. \]  

(178)

Hence, we have singled out the integrations over \( x \) and \( q \) before performing the continuous-time limit (see (168) and (171)). We call \( Z_{Q,X} \) a generating functional because a functional differentiation of \( Z_{Q,X} \) with respect to \( Q(\tau) \) generates a factor \( x(\tau) \), and a functional differentiation of \( Z_{Q,X} \) with respect to \( X(\tau) \) generates a factor \( i q(\tau) \).

Given the action (172) with the adjoint transition operator (176), the above properties of the \((Q,X)\)-generating functional can now be used to rewrite this function as (see appendix D)

\[ Z_{Q,X} = e^{i q_x x - Сd + \int_0^t d\tau \left[ Q(\tau) x(\tau) + X(\tau) i q(\tau) \right]} Z^0_{Q,X} \]  

(179)

with \( Z^0_{Q,X} := \left\langle e^{i \int_0^t d\tau Q(\tau) x(\tau)} \right\rangle_W \) .

(180)

Here, \( \langle \cdot \rangle_W \) 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

\[ dx(\tau) = [\alpha_x(x) + X(\tau)]d\tau + \sqrt{\beta_x(x)} dW(\tau). \]  

(181)

The temporal evolution of \( x(\tau) \) starts out from \( x(t_0) \), which is determined by the argument of the marginalized distribution (177). In section 4.5.4, we demonstrate how the marginalized distribution can be evaluated in terms of a perturbation expansion of the \((Q,X)\)-generating functional (179) for a process of diffusing particles that are also decaying. In order to perform the perturbation expansion, let us already notice that the value \( x(\tau) \) of the path at time \( \tau \) depends on the ‘source’ \( X(\tau') \) only for times \( \tau' < \tau \), and that \( \frac{1}{\sqrt{\beta_x(x)}} \int_0^\tau d\tau' Q(\tau') f(\tau') = f(\tau) \) holds for all \( \tau \in (t_0, t) \). Moreover, let us note that the \((Q,X)\)-generating functional depends on \( q_x \) and \( x(t) \) but not on \( x_N \). These properties follow from the derivation of the representation (179), which we outline in appendix D.

An important special case of the above representation is constituted by processes whose perturbation operator \( P^i_x \) is zero. For \( Q = X = 0 \), the generating functional (179) then simplifies to \( Z_{0,0} = e^{i q_x(n)} \) and the marginalized distribution (177) follows in terms of the Feynman–Kac like formula

\[ |p(t,n|t_0))_{(n)} = \left\langle (|n|_{(n)})_{(n)} \right\rangle_W \].

(182)

Here, \( x(\tau) \) solves the Itô SDE

\[ dx(\tau) = \alpha_x(x)d\tau + \sqrt{\beta_x(x)} dW(\tau) .\]  

(183)

with the initial value \( x(t_0) \). For the Poisson basis function \( |n\rangle = \frac{e^{i q_x x}}{\sqrt{n!}} \), the above representation of the marginalized distribution coincides with our earlier result (136), which we encountered in the discussion of the coagulation reaction \( 2A \rightarrow A \). For the bi-directional reaction \( \varnothing \rightarrow A \) from the previous section with adjoint transition operator \( \tilde{Q}^T(x, \partial_x) = (\gamma - \mu_x) \partial_x \), the Itô SDE (183) simplifies to the deterministic rate equation \( \partial_x x = \gamma - \mu_x \). Thus, no averaging is required to evaluate the marginalized distribution (182) and one recovers our earlier solution (174).

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 (176) includes a vector-valued drift coefficient \( \alpha_x(x) \) and a diffusion matrix \( \beta_x(x) \). Moreover, the derivation of the corresponding generating functional (179) requires that there exists a matrix \( \sqrt{\beta_r} := \gamma, \gamma^T = \beta_r \). If the diffusion matrix \( \beta_r \) is positive-semidefinite, its positive-semidefinite and symmetric square root \( \sqrt{\beta_r} \) can be determined via diagonalization [111].

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 \( w_m(n, m) \) 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, \( \kappa_r(\Delta n, n) := w_m(n + \Delta n, n) \) denotes the rate at which the state \( n \) is left via jumps of size \( \Delta n \) (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.

4.4.1. Processes with continuous state spaces. We consider a process whose state is characterized by a continuous variable \( x \in \mathbb{R}_0 \). 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 \( \int_{-\infty}^{\infty} dx p(\tau, x|\cdot) = 1 \) and obey the master equation

\[ \partial_t p(\tau, x|\cdot) = \int_{-\infty}^{\infty} d\Delta x \left[ \kappa_r(\Delta x, x - \Delta x) p(\tau, x - \Delta x|\cdot) \right. \]

\[ \left. - \kappa_r(\Delta x, x) p(\tau, x|\cdot) \right] \]  

(184)

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 (11). Provided that the product \( \kappa_r(\Delta x, x) p(\tau, x|\cdot) \) is analytic in \( x \), one can perform a Taylor expansion of the above master equation to obtain the (forward) Kramers–Moyal expansion [1, 103, 272]

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

(185)

The ‘jump moments’ \( \mathcal{M}_m(\cdot)(x) \) are defined as
\[ \mathcal{M}_{\mathcal{E}}^{\mathcal{O}}(x) := \int_{\mathbb{R}} d\Delta x (\Delta x)^m \kappa_r(\Delta x, x). \]  
(186)

By Pawula’s theorem [104, 383, 384], 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 \( \mathcal{M}_{\mathcal{E}}^{(1)}(x) \) and its non-negative diffusion coefficient by \( \mathcal{M}_{\mathcal{E}}^{(2)}(x) \). 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 state space. For a jump process, the Kramers–Moyal expansion cannot stop. Nevertheless, a truncation of the Kramers–Moyal expansion cannot stop. Nevertheless, a truncation of the Kramers–Moyal expansion cannot stop. Nevertheless, a truncation of the Kramers–Moyal expansion cannot stop. Nevertheless, a truncation of the Kramers–Moyal expansion cannot stop.

The corresponding adjoint transition operator evaluates to

\[ \mathcal{M}_{\mathcal{E}}^{\mathcal{O}}(x) = \gamma_i (l - k)^{\mathcal{O}} x \Theta(x - k). \]  
(193)

The corresponding adjoint transition operator evaluates to

\[ \mathcal{Q}_r(x, i, q) = \gamma_i (x e^{-ij} + (e^{-iq})^j) \Theta(x - k). \]  
(194)

Path integrals with such an operator have been noted in [386–388], but their potential use remains to be fully explored. Curiously, upon ignoring the step function, the operator (194) has the same structure as the transition operator (122) upon identifying \( x^{-ij} \) with the creation operator \( c \), and \( e^{ij} \) with the annihilation operator \( a \) (Coe-Hopf transformation [386–390]). Note, however, that the stochastic processes associated to the two transition operators differ from each other (discrete versus continuous state space; different transition rates). The connection between the two associated path integrals (171) and (190) should be further explored.

4.4.3. Path integral representation of the backward Fokker–Planck equation. With the drift coefficient \( \alpha_r(x) := \mathcal{M}_{\mathcal{E}}^{(1)}(x) \) and the non-negative diffusion coefficient \( \beta_r(x) := \mathcal{M}_{\mathcal{E}}^{(2)}(x) \), a truncation of the adjoint transition operator (189) at its second summand reads

\[ \mathcal{Q}_r(x, \partial_{x_0} + \frac{1}{2} \beta_r(x_0) \partial_{x_0}^2. \]  
(195)

Thus, the corresponding Kramers–Moyal expansion (188) recovers the backward Fokker–Planck equation (2). Since the action (191) evaluates to

\[ S^i := \int_{t_0}^{t_f} d\tau \left[ i\tau \partial_{x}(x(\tau) - \mathcal{Q}_r(x, \partial_{x_0} + \frac{1}{2} \beta_r(x_0) \partial_{x_0}^2. \right. \]  
(196)

the corresponding path integral (190) 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 [16], de Dominicis [17], Janssen [18, 19], and Bausch, Janssen, and Wagner [19].
The application of this path integral to stochastic processes is, for example, discussed in the book of Täuber [70].

The transition operator (195) of the backward Fokker–Planck equation has the same form as the transition operator (176) in section 4.3, but it does not involve a perturbation operator \( \mathcal{P}_b \). Therefore, one can follow the steps in that section to evaluate the path integral (190) in terms of an average over the paths of an Itô SDE. This procedure shows that the backward Fokker–Planck equation is solved by

\[
p(t, x|t_0, x_0) = \left\langle \delta(x - x(t)) \right\rangle_{t_0},
\]

(197)

with \( x(\tau) \) solving the Itô SDE

\[
dx(\tau) = \alpha_x(x(\tau))d\tau + \sqrt{\beta_x(x(\tau))}dW(\tau)
\]

(198)

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).

4.4.4. 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 [79] begins apparent upon the completion of a square (as in a Hubbard-Stratonovich transformation [391, 392]). For this purpose, the diffusion coefficient \( \beta_x \) in the transition operator (195) must not only be non-negative but positive-definite. One can then complete the square in the variable \( q \) and perform the path integration over this variable. Returning to the discrete-time approximation (169) of the action, one obtains the following representation of the conditional probability distribution \( p(t, x|t_0, x_0) \) in terms of convolutions of Gaussian distributions:

\[
\lim_{N \to \infty} \left( \prod_{j=1}^{N-1} \int_{\mathbb{R}} dx_{j+1} \mathcal{G}_{\mu_j, \sigma_j}(x_{j+1} - x_j) \right) \delta(x - x_N).
\]

(199)

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

\[
\mathcal{G}_{\mu, \sigma}(x) := \frac{e^{-\frac{(x - \mu)^2}{2 \sigma^2}}}{\sqrt{2\pi\sigma^2}}.
\]

(200)

As before, \( \Delta t = (t - t_0)/N \) denotes the time intervals between \( t_0 \leq t_1 \leq \cdots \leq t_N = t \).

Upon identifying \( x(t_0 + j \Delta t) \) with \( x_j \), the representation (199) can be rewritten as

\[
p(t, x|t_0, x_0) = \int_{t_0}^{t} e^{-S} \delta(x - x(t)) \bigg|_{x(t_0) = x_0}.
\]

(201)

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 \( e^{-S} \) with the action

\[
S^+ := \lim_{N \to \infty} \sum_{j=0}^{N-1} \frac{(x_{j+1} - x_j - \alpha(x_j) \Delta t)^2}{2\beta(x_j) \Delta t}.
\]

(202)

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

\[
S^+ = \int_{t_0}^{t} dt \left( \frac{\partial_x x - \alpha(x) x^2}{2\beta(x)} \right).
\]

(203)

The integrand of this action is called an Onsager–Machlup function [79, 393] and the representation (201) 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

\[
\lim_{N \to \infty} \left( \prod_{j=0}^{N-1} \int_{\mathbb{R}} dx_{j+1} \frac{dx_{j+1}}{2\pi \beta(x_j)^{1/2} \Delta t} \right) \delta(\xi - x_{j+1})
\]

(204)

Mathematically rigorous formulations of the Onsager–Machlup representation have been attempted in [394–396]. For the other path integrals discussed in this review, such attempts have not yet been made to our knowledge.

4.4.5. 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 (see section 4.3). Alternative discretization schemes have been proposed as well and are, for example, employed in stochastic thermodynamics [377]. 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 \( D := \beta(x) \) is constant. A general discretization scheme—called the \( \alpha \)-scheme but here we use a \( \kappa \)—consists of shifting the argument \( x_t \) of the drift coefficient in the action (202) to \( \bar{x}_t := \kappa x_{t+1} + (1 - \kappa) x_t \) by writing

\[
x_{t+1} = \bar{x}_t - \kappa (x_{t+1} - x_t)
\]

(205)

(with \( \kappa \in [0, 1] \)). Afterwards, the drift coefficient is expanded in powers of \( x_{t+1} - x_t \), which is of order \( \sqrt{\Delta t} \) along relevant paths. Following [397], it suffices to keep only the first two terms of the expansion of \( \alpha(x) \) so that

\[
x_{t+1} - x_t = \alpha(x_t) \Delta t + \kappa \alpha'(x_t) \Delta t
\]

(206)

As the next step, the integration variables \( x_1, \ldots, x_N \) are transformed according to

\[
1 + \kappa \alpha'(x_t) \Delta t (x_{t+1} - x_t) \rightarrow x_{t+1} - x_t.
\]

(207)

The Jacobian of this transformation vanishes everywhere except on its diagonal and sub-diagonal. Its determinant therefore contributes an additional factor to the path integral and requires a redefinition of the action (202) as

\[
S_t^+ := \lim_{N \to \infty} \sum_{j=0}^{N-1} \frac{(x_{j+1} - x_j - \alpha(x_j) \Delta t)^2}{2D \Delta t} + \kappa \alpha'(x_t) \Delta t.
\]

One may abbreviate this limit by

\[
S^+ = \int_{t_0}^{t} dt \left( \frac{\partial_x x - \alpha(x) x^2}{2D} + \kappa \alpha'(x) \right).
\]

(208)
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 [70]). 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 $\kappa = \frac{1}{\tau}$ is, for example, employed in Seifert’s review on stochastic thermodynamics [237]. For a recent, more thorough discussion of the above discretization schemes, as well as of conflicting approaches, we refer the reader to [398] (the above action is discussed in the appendices).

4.4.6. Wiener’s path integral. Before returning to processes with a discrete state space, let us briefly note how the Onsager–Machlup representation (210) relates to Wiener’s path integral of Brownian motion [80, 81], as it is discussed in [399]. 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 $\tau \gamma^t d^2 k x$ and for the coagulation process $x(t)x(t)$ $2 x(t)$, as nei-

x = −∼ $d = 1$. If that is not the case, this representation is exact because the $D t t \tau \gamma_{s}$ solving the Itô SDE. For the pair genera-

$tion process follows as [see section 3.2.2]

\begin{align*}
\tilde{\mathcal{Q}}_\tau(x, \partial_t) &= Q(x, a) = \gamma e^{a} (a^l - a^k) \\
&= [\gamma(l - k)x^2] \partial_t + [\gamma_l(l - 1) - k(l - 1)]x^3 \frac{\partial_x^2}{2} \\
&+ \mathcal{P}^t_\tau(x, \partial_t).
\end{align*}

\null(209)

Here we employed the creation operator $c = x$ and the anni-
hilation operator $a = \partial_t + 1$, and we performed a series expansion with respect to $\partial_t$. Terms of third and higher order in $\partial_t$ were shoved into the perturbation operator $\mathcal{P}^t_\tau(x, \partial_t)$. In the following, we approximate the marginalized distribution of the reaction $k A \rightarrow A$ by first dropping both the diffusion coefficient and the perturbation operator from (210). 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 $\emptyset \rightarrow 2A$, this representation is exact because the perturbation operator associated to this process is zero (see section 4.5.1). Later, in section 4.5.4, we solve a process with a non-vanishing perturbation operator $\mathcal{P}^t_\tau$.

As a first approximation of the reaction $k A \rightarrow A$, let us drop all the terms of the adjoint transition operator (210) except for the drift coefficient $\alpha_\tau(x) = \gamma_x (l - k)x^4$. The SDE (181) then simplifies to the deterministic rate equation $\partial_t x = \gamma_x (l - k)x^4$ of the reaction. Its solution $x(\tau)$ acts as the mean of the marginalized distribution, which, according to (177)–(180), is again given by the Poisson distribution $|p(t, n, k_0)\rangle_n(t, \tau) = |n\rangle_n(t, \tau)$ in (174).

Going one step further, one may keep the diffusion coeffi-
cient in (210). The marginalized distribution then reads

$$
|p(t, n, k_0)\rangle_n(t, \tau) = \left\langle \frac{x(t)^x e^{-\tau x}}{n!} \right\rangle_w 
$$

with $x(\tau)$ solving the Itô SDE

$$
\frac{dx}{\gamma_x (l - k)x^4 dt} + \sqrt{\gamma_x (l(l - 1) - k(k - 1))x^6} \, dW(\tau).
$$

(211)

Hence, the Poisson distribution in (211) is averaged over all possible sample paths of the SDE (212), whose evolution starts out from the initial value $x(t_0)$ (see (136)). The above expression has also recently been derived by Wiese [336], although based on the forward path integral that we will discuss in section 5. Apparently, the expression under the square root in (212) is non-negative only if $k \in \{0, 1\}$ or if $l \geq 2$. If that is not the case, $x(\tau)$ strays off into the complex domain. SDEs with imaginary noise (or the corre-
spounding Langevin equations) have been studied in several recent articles, most notably for the binary annihilation reaction $2A \rightarrow \emptyset$ and for the coagulation process $2A \rightarrow A$ [336, 400, 401]. The numerical evaluation of such SDEs, however, often encountered severe convergence problems [336, 379]. The appearance of imaginary noise has been linked to the anti-correlation of particles in spatial systems [400], but as (212) 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 (211) becomes narrower than a Poisson distribution [336].

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 (171). 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 $k A \rightarrow A$ with rate coefficient $\gamma_x$. Using the Poisson basis function $|n\rangle_n = \frac{e^{-\gamma_x n} \gamma_x^{|n|}}{|n|!}$, the adjoint transition operator of this reaction can be written as (see section 3.2)

$$
\tilde{\mathcal{Q}}_\tau(x, \partial_t) = Q(x, a) = \gamma e^{a} (a^l - a^k)
$$

\null(209)

$$
= [\gamma_x(l - k)x^2] \partial_t + [\gamma_l(l - 1) - k(l - 1)]x^3 \frac{\partial_x^2}{2} \\
+ \mathcal{P}^t_\tau(x, \partial_t).
$$

\null(210)

Here we employed the creation operator $c = x$ and the anni-
hilation operator $a = \partial_t + 1$, and we performed a series expansion with respect to $\partial_t$. Terms of third and higher order in $\partial_t$ were shoved into the perturbation operator $\mathcal{P}^t_\tau(x, \partial_t)$. In the following, we approximate the marginalized distribution of the reaction $k A \rightarrow A$ by first dropping both the diffusion coefficient and the perturbation operator from (210). 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 $\emptyset \rightarrow 2A$ with growth rate coefficient $\gamma_x$, the drift and the (squared) diffusion coefficients agree: $\alpha_\tau = \beta_\tau = 2\gamma_x$. As neither of them depends on $x(\tau)$, the Itô SDE is readily solved by

$$
x(t) = x(t_0) + \int_{t_0}^t d\tau 2\gamma_x + \int_{t_0}^\tau dW(\tau) \sqrt{2\gamma_x}.
$$

After introducing the (rather daunting) parameter

$$
\eta_k := 
$$

\null(213)
Figure 7. Evolution of the marginalized distribution $|p(t, n|t_0)|$ in (214), which solves the pair generation process $\emptyset \rightarrow 2A$ (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 $\gamma_i = 1/(1 + \sqrt{r})$.

one can employ the binomial theorem to rewrite the right-hand side of the marginalized distribution (211) as

$$e^{-\gamma t_0} \sum_{k=0}^{N} \binom{N}{k} \int_{t_0}^{t} \eta^k \eta^{N-k}.$$ 

If the rate coefficient $\gamma$ is independent of time, one can show that the parameter $\eta$ coincides with the $k$th moment of a Gaussian distribution with zero mean and unit variance (i.e. $\eta_0 = 0$ 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). The marginalized distribution $|p(t, n|t_0)|_{n/2}$ therefore follows as

$$e^{-\gamma t_0} \sum_{k=0}^{n/2} \binom{n/2}{k} \int_{t_0}^{t} \eta^k \eta^{n-k/2}.$$ 

Here, $[n/2]$ represents the integral part of $n/2$. The marginalized distribution (214) 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 $\emptyset \rightarrow 2A$ 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 $\gamma_i$ depends on time. The evolution of the distribution is shown in figure 7 for the rate coefficient $\gamma_i = 1/(1 + \sqrt{r})$. 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.

4.5.2. 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 $2A \rightarrow \emptyset$ 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 $t^{-1/2}$ [130, 294, 402]. 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 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. [403] 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 [130–132, 294, 306, 404]. For a broader discussion of systems exhibiting a transition into an absorbing state see [68]. 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 $L$. 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 4.5.4, the particles will also be allowed to decay.

The configuration of particles on the network may be represented by the vector $n \in \mathbb{N}_0^{|L|}$, with $|L|$ being the total number of network nodes. The configuration changes whenever a particle hops from some node $i \in L$ to a neighbouring node $j \in \mathbb{N}_0 \subset L$. The probability $p(t, n|t_0, n_0)$ of finding the system in configuration $n$ then obeys the master equation

$$\partial_t p(t, n|\cdot) = \sum_{i \in L} \sum_{j \in \mathbb{N}_0} \left[ (n_i + 1) p(t, n + \hat{e}_i - \hat{e}_j|\cdot) - n_i p(t, n|\cdot) \right].$$ 

Here, $\hat{e}_i$ represents a unit vector that points in direction $i \in L$. Moreover, $\varepsilon_{r, i}$ 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 (24). It may therefore be cast into a linear PDE by extending the operators and basis functions from section 3.2.2 to multiple variables. In particular, we extend the Poisson basis function to

$$n_j := \prod_{i \in L} x_i^{n_i} e^{-x_i} / n_i!$$ 

and employ the creation and annihilation operators

$$c := x \quad \text{and} \quad a := \nabla + 1.$$ 

According to the flow equation (116), the marginalized distribution evolves via $\partial_{\tau} |p(t, n|\tau)| = \tilde{Q}_\tau |p(t, n|\tau)|$ with the adjoint transition operator
Here we introduced the discrete Laplace operator
\[ \Delta f_i := \sum_{j \in \mathbb{N}} (f_j - f_i), \]
which acts both on \( \varepsilon \tau \) 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
\[
| p(t, n| t_0) |_{x(t_0)} = \prod_{i \in \mathbb{L}} \frac{x_i(t_0) e^{-x_i(t_0)}}{n_i!}.
\]
Its mean \( x(t) \) solves the discrete diffusion equation
\[
\partial_t x_i = \sum_{j \in \mathbb{N}} (\varepsilon \tau x_j - \varepsilon \tau x_i) = \Delta x_i,
\]
with the initial condition \( x(t_0) \). Let us note that the marginalized distribution solves the master equation \((215)\), but with the initial number of particles being Poisson distributed locally.

The conditional distribution \( p(t, n| t_0, n_0) \) follows from the marginalized distribution \((219)\) by applying the functional
\[
\langle n_0 |_{x(t_0)} f \rangle = \left[ \prod_{i \in \mathbb{L}} (\partial_x |_{x(t_0)} + 1) n_0 \right] f(x(t_0)) \bigg|_{x(t_0) = 0}
\]
to it (see \((126)\) with \((129)\)). The evaluation requires a prior solution of the discrete diffusion equation \((220)\). This equation can be written in matrix form as \( \partial_t x = M x \) with \( M_{i,j} = (A_0 - |N_j| \delta_{i,j}) \varepsilon \tau \). Here, \( A \) represents the symmetric adjacency matrix of the network and \( |N| \) 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 \( x(\tau) = G(\tau|t_0) x(t_0) \) with the ‘propagator’ \( G \) solving \( \partial_t G(\tau|t_0) = M \tau G(\tau|t_0) \). We write the elements of the propagator as \( G(\tau, i|t_0, j) \). Its flow starts out from \( G(t_0|t_0) = 1 \). For later purposes, let us note the time-reversal property in terms of the matrix inverse \( G(\tau|t) = G(\tau|t) \) and also the conservation law \( \Gamma G(\tau|t_0) = \Gamma \).

It proves insightful to consider the master equation \((215)\) of the multi-particle hopping process for the random walk of a single particle on the one-dimensional lattice \( \mathbb{Z} \) with \( \varepsilon \tau = l = l \). The presence of only a single particle can be enforced by choosing \( n_0 \) as \( n_0 = 1 \) for all \( j \neq k \). By following the above steps, one eventually finds that the master equation \((215)\) is solved by the conditional probability distribution
\[
p(t, n| t_0, n_0) = \sum_{i \in \mathbb{L}} \left[ \prod_{j \neq 1} \delta_{n_0, j} G(t, i|t_0, k),
\]
with the propagator \( G \) solving the master equation \((55)\) of the simple random walk. Hence, the propagator evaluates to a Skellam distribution.

4.5.3. Diffusion in continuous space. To make the transition to a field theory, one may specify the network as the \( d \)-dimensional lattice \( \mathbb{L} = (\mathbb{Z})^d \) with the lattice spacing \( l > 0 \) going to zero. In order to take this limit, we define the variable
\[
x(\tau, r) := x_0(\tau),
\]
the rescaled Laplace operator \( \Delta := \frac{\Delta}{l^2} \), and the rescaled diffusion coefficient \( D_r(r) := \frac{D_r}{l^2} \). Assuming that these definitions can be continued to \( r \in \mathbb{R}^d \), the discrete diffusion equation \((220)\) becomes a PDE for the ‘field’ \( x(\tau, r) \) in the limit \( l \to 0 \), namely \( \partial_\tau x(\tau, r) = \Delta D_r(r) x(\tau, r) \). Here, \( \Delta \) represents the ordinary Laplace operator\(^{10}\). The solution of the PDE acts as the mean of the multivariate Poisson distribution \((219)\) whose extension to \( r \in \mathbb{R} \) is, however, not quite obvious. If the diffusion coefficient is homogeneous in space, the PDE is solved by
\[
x(\tau, r) = \int d^n r_0 G(\tau, r|t_0, r_0) x(r_0, r_0),
\]
with the Gaussian kernel
\[
G(\tau, r|\tau', r') = \frac{e^{-(r-r')^2/4} \int d^n r_0}{\sqrt{4\pi} \int d^n r_0}.
\]
The action \((172)\) can also be extended into continuous space. For that purpose, one may rescale the second path integral variable as \( q(\tau, r) := l^{-d} q_r(\tau) \) so that in the limit \( l \to 0 \),
\[
S_1 = \int d^n r_0 \int d^n r_0 \int \left[ d^n r_0 \int d^n r_0 \int i q_r(\partial_\tau x - \Delta \varepsilon \tau x_r) \right] \to \int d^n r_0 \int d^n r_0 \int d^n r_0 \int d^n r_0 \int i q_r(\partial_\tau x - \Delta D_r x).
\]
The above rescaling of \( x \) and \( q \) is not unique and depends on the problem at hand. Instead of dividing \( q_r \) by the volume factor \( l^d \), this factor is sometimes employed to cast \( x_r \) and the particle number \( n_r \) into densities \((130)\). 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 Regge field theory \([70, 405-407]\) like form \((96)\). The critical behaviour of these random walks falls into the universality class of directed percolation (DP) \((96, 132)\). Information on this universality class can be found in the book \((68)\), as well as in the original articles of Janssen and Grassberger, which established the extensive scope of the DP class \((408, 409)\).

4.5.4. Diffusion and decay. Let us exemplify how one can accommodate a non-vanishing perturbation operator \( P^l \) in the adjoint transition operator \((176)\). 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 \((218)\) shall specify the drift coefficient \( \alpha_\tau(x) := \frac{D_r(x)}{l^2} \varepsilon \tau \). For the sake of brevity, we assume that the
\[
\sum_{r' \in \mathbb{N}} \frac{f_r - f_r}{l^2} \to \sum_{i=1}^d \partial_i f(r) = \Delta f(r).
\]

\(^{10}\) With \( \mathbb{N} \) denoting the neighbouring lattice site of \( r \in (\mathbb{Z})^d \), the ordinary Laplace operator follows as the limit of a finite-difference approximation, i.e. as
corresponding decay rate coefficient $\mu_\tau$ 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 (131),

$$\mathcal{P}_t^\dagger(x, \nabla) = -\mu_\tau \sum_{i \in \mathbb{Z}} x_i \partial_i. \quad (228)$$

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

The first step of the derivation is to solve the differential equation (181) for $x$, which now reads $\partial_t x_i = \Delta \xi_{\tau,x_i} + X(\tau)$. The homogeneous solution of this equation is given by $x_i(\tau) := G(\tau|x_0) x_i(t_0)$, with $G$ being the propagator from the end of section 4.5.2. The solution of the full equation can be written as

$$x(\tau) = x_0(\tau) + \int_{t_0}^{\tau} d\tau' G(\tau|\tau') X(\tau'). \quad (229)$$

As the next step, the evaluation of the marginalized distribution (177) requires us to compute the $(Q,X)$-generating functional (179) for $Q = X = 0$. By performing a series expansion of its leading exponential, this function can be written as

$$Z_{0,0} = \sum_{k=0}^{\infty} \sum_{l=0}^{k} \frac{1}{l!} \left( i q_N \cdot \delta \frac{\delta}{\delta Q(t)} \right)^l \cdot \int_{t_0}^{\tau} d\eta_0 \mathcal{P}_t^+ \mathcal{P}_t^* \int_{t_0}^{\tau_0} d\eta \mathcal{P}_t^+ \mathcal{P}_t^* Z_{0,0}^0 \bigg|_{Q=X=0}, \quad (230)$$

with $\mathcal{P}_t^+ = \frac{\delta}{\delta x(\tau)} (-\mu_\tau) \frac{\delta}{\delta Q(t)}$. To evaluate $Z_{0,0}$, it helps to note that $\mathcal{P}_t^+ Z_{0,0} = \int_{t_0}^{\tau} d\tau' Q \cdot x$, from which it follows that for $t_0 \leq \tau \leq t$:

$$\delta \ln Z_{0,0}^0 \bigg|_{Q=X=0} = x_0(\tau) \quad \text{and} \quad (231)$$

$$\delta^2 \ln Z_{0,0}^0 \bigg|_{Q,X} \bigg|_{Q=X=0} = G(\tau, i|\tau', j) \Theta(\tau - \tau'). \quad (232)$$

The Heaviside step function is used with $\Theta(0) := 0$ to take into account that $x_0(\tau)$ depends on $X(\tau')$ only for $\tau' < \tau$ (see appendix D). All other derivatives of $\ln Z_{0,0}^0$, as well as itself, vanish for $Q = X = 0$.

Upon inserting the perturbation (228) into (230), one observes that every summand of the expansion can be written in terms of a combination of the derivatives (231) and (232), and a terminal factor $i q_N$. For example, the summand with $k = 2$ and $l = 1$ reads

$$i q_N \cdot \frac{\delta}{\delta Q(t)} \int_{t_0}^{\tau} d\eta \frac{\delta}{\delta X(\eta)} (-\mu_\tau_\eta) \frac{\delta}{\delta Q(\eta)} Z_{0,0}^0 \bigg|_{Q=X=0}, \quad (233)$$

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

$$\int_{t_0}^{\tau} d\eta G(\tau|\eta)(-\mu_\eta) x_0(\eta). \quad (234)$$

One can represent this expression graphically by a Feynman diagram according to the following rules. First, every diagram ends in a sink $\longrightarrow$, which contributes the factor $i q_N$.

The incoming line, or 'leg', of the sink represents the derivative $\frac{\delta}{\delta x(\tau)}$. This leg may either be left dangling, resulting in a factor $x_0(\tau)$ according to (231), or it may be 'contracted' with the outgoing line of a vertex $\longrightarrow$. The two legs of this vertex reflect the two derivatives in the perturbation $\mathcal{P}_t^+ = \frac{\delta}{\delta x(\tau)} (-\mu_\tau) \frac{\delta}{\delta Q(t)}$. According to (232), 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 (234) can therefore be represented graphically by

$$\int_{t_0}^{\tau} d\tau_1 G(\tau|\tau_1) x_0(\tau_1). \quad (235)$$

Note that dangling outgoing lines are not permitted because the derivative $\delta \ln Z_{0,0}^0(\tau)$ vanishes for $Q = X = 0$.

For more complex, non-linear processes, the Feynman diagrams may contain multiple kinds of vertices, each representing an individual summand of $\mathcal{P}_t^+$. If there exist vertices with more than two legs, the diagrams may exhibit internal 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 $\ln Z_{0,0}$ is given by the sum of only the connected diagrams [411], i.e. by

$$\ln Z_{0,0} = \rightarrow + \rightarrow + \rightarrow + \cdots \quad (236)$$

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

$$\ln Z_{0,0} = \sum_{k=0}^{\infty} i q_N \cdot \int_{t_0}^{\tau} d\eta_0 G(\tau|\eta)(-\mu_\eta_0) \cdots \int_{t_0}^{\tau_0} d\eta G(\tau_0|\eta_0)(-\mu_\eta_0) \quad (237)$$
After inserting this expression into the marginalized distribution (177), one recovers the multivariate Poisson distribution (219). Its mean, however, has now acquired the pre-factor \(e^{-\int_0^t d\tau \mu_n} \), reflecting the decay of the particles.

### 4.6. Résumé

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

\[
|p(t, n|t_0)\rangle = \int \langle n_0| e^{-S(t)} |n_0\rangle
\]

of the marginalized distribution (see section 3)

\[
|p(t, n|t_0)\rangle = \sum_{n_0} \langle n_0|p(tt, n_0, n_0)\rangle |n_0\rangle.
\]

When the basis function \(|n\rangle\) is chosen as a probability distribution (e.g. \(|n\rangle = e^{-n^2/2n^2}\) for all \(n \in \mathbb{N}\)), 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 (10) for the initial condition \(|p(t_0, n|t_0)\rangle = |n\rangle\) and it transforms into the conditional probability distribution as \(|p(t, n|\tau, n_0)\rangle = |n_0\rangle |p(t, n|\tau)\rangle\). In section 4.3, we showed how the backward path integral (239) 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 decay. We hope that the backward path integral (239) 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 [69, 412]. In section 6, we show how the formal representation readily follows from the backward path integral (239) upon summing the marginalized distribution over an observables \(A(t)\). The corresponding representation is commonly applied in the study of diffusion-limited reactions (e.g. [69, 96, 130–132]), 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 (162). 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 \(2A \rightarrow \nothing\). In the present section, however, we shift our focus to the forward path integral (161). 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 (54) obeying the generating function or from the flow equation (150) obeyed by the series (149) (see 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 [35]. As in section 4.1, we first split the time interval \([t_0, t]\) into \(N\) pieces \(t_0 \leq t_1 \leq \cdots \leq t_N := t\) of length \(\Delta t\). Over a sufficiently small interval \(\Delta t\), the flow equation (54) is then solved by

\[
|g(t|t_0, n_0)\rangle_{t_0} = \mathcal{L}_{t_0} \cdots \mathcal{L}_{t_N} |g(t|t_N, n_0)\rangle_{t_N}.
\]

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

\[
\mathcal{L}_{t_N} |q_{N\uparrow \downarrow} \rangle \times \int_{\mathbb{R}^N} \frac{d q_{N\downarrow}}{2\pi} e^{-i q_{N\downarrow} \cdot \Delta t} e^{-i q_{N\downarrow} \cdot \Delta q_{N\downarrow}} |g(t|t_N, n_0)\rangle_{q_{N\downarrow}}.
\]

Assuming that the transition operator \(\mathcal{Q}_t\), and therefore also \(\mathcal{L}_t\), are normal-ordered with respect to \(q\) and \(\partial_q\), we may replace \(\partial_{\partial_q}\) by \(i \partial_{\partial_q}\) and interchange \(\mathcal{L}_{t_N} \) with the exponential. This procedure is repeated \(N\) times before invoking the exponentialization \(\mathcal{L}_t = \exp(\mathcal{Q}_t \Delta t)\). Using the inverse transformation (48)
and the initial condition \(|g(t_0|t_0,n_0)⟩ = |n_0⟩\), one obtains the discrete-time path integral representation

\[
p(t,n|t_0,n_0) = \langle n|_{t_0,n_0}|g(t|t_0,n_0)⟩_{q_k}\]  

with \(|g(t|t_0,n_0)⟩_{q_k}\) given by \(\int_{q_0}^{N-1} e^{-S_k} |n_0⟩_{q_{0:k}}\) and

\[
S_N := \sum_{j=0}^{N-1} \Delta t \left( \bar{q}_j - q_j + q_j^f \right) - \bar{q}_j(q_j^f + i \bar{x}_j). \tag{245}
\]

Here we again employed the abbreviation (167), i.e.

\[
\int_{q_k}^l \frac{d q_k d x_k}{2\pi}. \tag{246}
\]

Moreover, the initial condition \(p(t_0,n|t_0,n_0) = \delta_{n_0}\) is again trivially fulfilled for \(N = 0\). Upon taking the continuous-time limit \(N → ∞\), so that \(\Delta t → 0\), the forward path integral representation of the master equation follows as

\[
p(t,n|t_0,n_0) = \langle n|_{t_0,q(t)}|g(t|t_0,n_0)⟩_{q(t)}\]  

with \(|g(t|t_0,n_0)⟩_{q(t)}\) given by \(\int_{q_0}^{t} e^{-S} |n_0⟩_{t_0,q(t_0)}\) and

\[
S := \int_{t_0}^t d τ \left[ i x τ ∂_x q - \bar{q}_τ(q, x) \right]. \tag{249}
\]

The limit \(\int_{t_0}^t \lim_{N→∞} \int_{0}^{N-1} \Delta t \) 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 (248) can, for example, be used to derive the generating function of the generic linear process \(A → l A\) with rate coefficient \(μ_l\) and \(l ∈ \mathbb{N}_0\). For that purpose, we choose the basis function as \(|n⟩_{t_0,q} = q^n\) so that \(|q⟩\) coincides with the ordinary generating function. Since the basis function does not depend on time, (64) alone specifies the transition operator

\[
\bar{Q}_τ(q, δ_ρ) = μ_l(q^l - q) δ_ρ \tag{250}
\]

of the flow equation \(\partial_τ |g(τ)⟩\) = \(\bar{Q}_τ(q, δ_ρ)|g(τ)⟩\). Consequently, the action

\[
S = \int_{q_0}^t d τ i x ∂_x q - μ_l(q^l - q) \tag{251}
\]

is linear in \(i x\). To evaluate the path integral (248), it helps to reconsider the discrete-time approximation

\[
S_N := \sum_{j=0}^{N-1} i x(q_j - q_j + μ_l(q_j^l + q_j + 1 - q_j + 1)) \tag{252}
\]

of the action. Upon integrating over all the \(q_j\)-variables and taking the limit \(Δ t → 0\), one obtains the generating function

\[
|g(t|t_0,n_0)⟩_{q(t_0)} = q(t_0)^n_0 \tag{253}
\]

with \(q(τ)\) solving \(∂_τ q = μ(q^l - q)\). The unique real solution of this equation with final value \(q(t)\) reads

\[
q(τ) = \frac{q(t)}{q(t)^{l-1} + e^{τ(l-1)} (1 - q(t)^{l-1})} \tag{254}
\]

For the linear decay process, or Yule–Furry, process \(A → 2A\) (\(l = 2\)), the inverse transformation (247) casts the generating function into

\[
p(t,n|t_0,n_0) = \left( \frac{1}{n!} \int_0^t d τ μ_2 \right)^n_0 \left[ 1 - e^{-τ μ_2} \right]^{n_0 - n}, \tag{255}
\]

for \(n > 0\) and into \(δ_{n,n_0}\) for \(n = 0\). The above distribution has the form of a negative Binomial distribution with probability of success \(e^{-τ μ_2}\), number of failures \(n - n_0\), and number of successes \(n_0\). The mean value \(e^{τ μ_2} n_0\) of the marginalized distribution (256) grows exponentially with time as long as \(μ_2 > 0\), and so does its variance \((e^{τ μ_2} - 1) e^{2 τ μ_2} - 1\).

For the linear decay process \(A → ∅\) (\(l = 0\)), the inverse transformation in (255) instead recovers the Binomial distribution (134).

5.3. Intermezzo: the forward Kramers–Moyal expansion

The above procedure can be generalized to processes whose transition operator is of the form

\[
\bar{Q}_τ(q, δ_ρ) = α_τ(q) δ_ρ + \frac{1}{2} β_τ(q) δ^2_ρ + P_τ(q, δ_ρ). \tag{257}
\]

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 (184) of a continuous process with a continuous state space can be written in terms of the (forward) Kramers–Moyal expansion

\[
∂_τ p(τ, q|·) = \sum_{m=1}^{∞} \frac{(-1)^m}{m!} \partial^m q [M^{(m)}(q) p(τ, q|·)] \tag{258}
\]

with initial condition \(p(t, q|t_0,q_0) = δ(q - q_0)\). 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

\[
∂_τ p(τ, q|·) = -\partial^2 q [M^2(q) p] + \frac{1}{2} \partial^2 q [M^2(q) p]. \tag{259}
\]

To apply the procedure from the previous section to this equation, it needs to be brought into the form \(∂_τ p(τ, q|·) = \bar{Q}_τ(q, δ_ρ)p(τ, q|·)\) with a normal-ordered
transition operator \(\tilde{\mathcal{A}}(q, \partial_q)\). It is readily established that this operator can be expressed in the form of (257) with the coefficients
\[
\alpha_r(q) := -\mathcal{M}_r^{(1)} + \partial_q \mathcal{M}_r^{(2)}, \quad \beta_r(q) := \mathcal{M}_r^{(2)}, \quad \text{and} \quad \mathcal{P}_r(q) := -\partial_q \mathcal{M}_r^{(1)} + \frac{1}{2} \partial^2_q \mathcal{M}_r^{(2)}.
\]
(260)
(261)
(262)

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
\[
p(t, q|t_0, q_0) = \int_{[t_0]} e^{-S} \delta(q_0 - q(t_0)) \bigg|_{q(t)=q} (263)
\]
with \(S := \int_{t_0}^t d\tau \left[ i x \partial_q - \tilde{\mathcal{A}}(q, i x) \right] \). (264)

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
\[
p(t, q|t_0, q_0) = \left\langle \left\langle e^{\int_{t_0}^t d\tau \mathcal{P}_r(q(\tau))} \delta(q_0 - q(t_0)) \right\rangle \right\rangle_w (265)
\]
with \(q(\tau)\) solving the backward-time SDE
\[
-dq(\tau) = \alpha_r(q(\tau)) d\tau + \sqrt{\beta_r(q(\tau))} dW(\tau). (266)
\]
The time evolution of this SDE starts out from the final value \(q(t) = q\). In a discrete-time approximation, the SDE reads
\[
g_j - g_{j+1} = \alpha_r(q_{j+1}) \Delta t + \sqrt{\beta_r(q_{j+1})} \Delta W_j. (267)
\]
The increments \(\Delta W_j\) are Gaussian distributed with mean 0 and variance \(\Delta t\). Let us note that the two path integral representations (190) and (263) of the conditional probability distribution belong to an infinite class of representations [42, 104, 415] (see also section 4.4.5). 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 (263) and of the representation (265), let us consider a process with linear drift and no diffusion, i.e. a process with jump moments \(\mathcal{M}_r^{(1)} = \mu q\) and \(\mathcal{M}_r^{(2)} = 0\). For this process, the representation (265) evaluates to
\[
p(t, q|t_0, q_0) = e^{-\mu(t-t_0)} \delta(q_0 - q) e^{-\mu(t-t_0)} = \delta(q - q_0) e^{\mu(t-t_0)}. (268)
\]
(269)

Thus, the leading exponential in (265) 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 (265) is the pure diffusion process with \(\mathcal{M}_r^{(1)} = 0\) and \(\mathcal{M}_r^{(2)} = D\). The derivation is performed most easily in the discrete-time approximation and results in the Wiener path integral
\[
p(t, q|t_0, q_0) = \lim_{N \to \infty} \left( \prod_{j=1}^{N} \int_{\mathbb{R}} d\tilde{q}_{j-1} \right) \left( \prod_{j=1}^{N} \mathcal{G}_{0, D, W} (\tilde{q}_{j-1} - \tilde{q}_j) \right) \delta(q_0 - q_0). (270)
\]
with \(\tilde{q}_N := q\) and Gaussian distribution \(\mathcal{G}_{0, D, W}\) (see (200)). An evaluation of the convolutions of Gaussian distributions shows that the process is solved by \(\mathcal{G}_{0, D, W} (q - q_0)\). Further uses of the path integral representation (263) of the (forward) Fokker–Planck equation remain to be explored.

5.4. Résumé

Here we derived the forward path integral representation
\[
|g(t|t_0, n_0)| = \int_{[t_0]} e^{-\alpha t} |n_0| (271)
\]
of the probability generating function \(|g(t|t_0, n_0)| = \sum_n n_0 p(t, n|t_0, n_0)\). 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 [35] 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 (271) 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 5.3, 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 $k A \rightarrow l A$ 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. The derived path integral has been applied in various contexts, particularly in the analysis of decay laws of diffusion-limited reactions and in the identification of universality classes [69, 70].

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 $p(t, n) = \sum_{n_0} p(t, n|t_0, n_0)p(t_0, n_0)$. Here we make use of the initial (single-time) distribution

$$p(t_0, n_0) = \frac{x(t_0)^n e^{-x(t_0)}}{n_0!}. \tag{272}$$

Consequently, the average value of an observable $A(n)$ at time $t$ evaluates to

$$\langle A \rangle_{\lambda(t_0)} = \sum_{n=0}^{\infty} A(n) p(t, n). \tag{273}$$

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 $|p(t, n|t_0)|$ in (110) for the Poisson basis function $|n_0\rangle_{\lambda(t_0)} = \frac{x(t_0)^{n_0} e^{-x(t_0)}}{n_0!}$, the above average can equivalently be written as

$$\langle A \rangle_{\lambda(t_0)} = \sum_{n=0}^{\infty} A(n) p(t, n|t_0))_{\lambda(t_0)}. \tag{274}$$

The path integral representation of $\langle A \rangle_{\lambda(t_0)}$ then follows directly from the backward path integral representation (171) of the marginalized distribution as

$$\langle A \rangle_{\lambda(t_0)} = \int_{\Omega} e^{-S} \langle A \rangle_{S(t)} \tag{275}$$

with $\langle A \rangle_{S(t)} := \sum_{n=0}^{\infty} x^n e^{-x} A(n). \tag{276}$

With the adjoint transition operator $Q^\dagger(x, a) = \gamma e^{k(a^2 - a^2)}$ of the reaction $k A \rightarrow l A$ (see section 3.2.2), the action $S^+$ in (172) reads

$$S^+ = \int_{t_0}^{t_1} dt \int_{\Lambda} [iq\partial, x - Q^\dagger(x, iq + 1)]. \tag{277}$$

The ‘+1’ in the transition operator is called the ‘Doi-shift’ [91]. 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 (110) as $|n_0\rangle_{\lambda(t_0)} = \frac{x(t_0)^n e^{-x(t_0)}}{n_0!}$, turning the average (274) into

$$\langle A \rangle_{\lambda(t_0)} = e^{-x(t_0)} \sum_{n=0}^{\infty} A(n) p(t, n|t_0))_{\lambda(t_0)}. \tag{278}$$

Upon rewriting the marginalized distribution in terms of the backward path integral (171), the action (277) acquires the addition summand $x(t_0) - x(t)$ and now involves the unshifted transition operator $Q^\dagger(x, iq)$. The average over a Poisson distribution in (276) 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 $\langle n \rangle_x = x$. This relation generalizes to factorial moments of order $k \in \mathbb{N}$ for which $\langle n_k \rangle_x = x^k$ (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 $n^k = \sum_{i=0}^{k} \frac{k!}{i!} (n)_i$, where the curly braces represent a Stirling number of the second kind (see section 26.8 in [335]). An extension of the path integral representation (275) to multi-time averages of the form

$$\sum_{n, n_0} A(n_2, n_3)p(n_2, n_3; n_1|t_0, t_0) \tag{279}$$

remains open (with $\tau_2 > \tau_1 > t_0$). The results of Elderfield [376] may prove helpful for this purpose.

In a slightly rewritten form, the Doi-shifted path integral (275) was, for example, employed by Lee in his study of the diffusion-controlled annihilation reaction $k A \rightarrow \emptyset$ with $k \geq 2$ [130, 416]. He found that below the critical dimension $d_c = 2(k - 1)$, the particle density asymptotically decays as $n \sim A_k t^{-d_c/2}$ with a universal amplitude $A_k$. At the critical dimension, the particle density instead obeys $n \sim A_k (\ln t)^{1/k-1}$. Neglecting the diffusion of particles, the above action (277) of the reaction $k A \rightarrow \emptyset$ reads

$$S^+ = \int_{t_0}^{t_1} dt \int_{\Lambda} [iq\partial, x + \gamma e^{k} \sum_{j=1}^{k} \{j\}] \langle iq \rangle^j. \tag{280}$$

Besides using different names for integration variables ($iq \rightarrow \psi$ and $x \rightarrow \psi$), the action employed by Lee involves an additional boundary term. This term can be introduced by rewriting the path integral representation (275) as

$$\langle A \rangle_{\lambda_0} = \int_{t_0}^{t_1} dt e^{-S} \langle A \rangle_{S(t)} \tag{281}$$

This path integral involves integrations over the variables $q(t)$ and $x(t_0)$, and the mean of the initial Poisson distribution is denoted by $\lambda_0$. Consequently, one needs to add the boundary term $iq(t_0)\langle x(t_0) - \lambda_0 \rangle$ to the action (280) to equate $x(t_0)$ with $\lambda_0$. The factor $iq(t_0)\langle x(t_0) \rangle$ of the new term is, however, often dropped eventually [130, 416]. The convention of Lee is also commonly used, for example, by Täuber [70]. The unshifted version of the path integral with transition operator $Q^\dagger(x, iq)$ is recovered via $iq + 1 \rightarrow iq$. 
For completeness, let us note that the path integral representation (275) can also be derived from the forward path integral (244), 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 (45) as \( \langle n \rangle := (q + 1)^n \) (insert \( \zeta := i, q := 1 \) and \( \xi := 0 \) into (70)). As the first step, the forward path integral (248) is summed over an initial Poisson distribution with mean \( x(0) \). The average of the factorial moment is then obtained via \( \langle n \rangle_{A(t_0)} = \partial_{x_{ik}}^n [g(t_0; x(t_0))]_{x_{ik}=0} \). To recover the action (277), one may note that the operator \( Q_\gamma(c, a) \) in (64) and the operator \( Q_\gamma^\dagger(c, a) \) in (122) fulfill \( Q_\gamma^\dagger(x, iq + 1, x) = Q_\gamma^\dagger(x, iq + 1) \) 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 (275) from the generating function is, for example, included in the article of Dickman and Vidigal [412] (see their equations (106) and (108)).

### 6.2. Intermezzo: alternative derivation based on coherent states

We noted previously in section 2.2.3 that the path integral representation (275) of the average

\[
\langle A \rangle_{A(t_0)} = \sum_{n=0}^{\infty} \frac{A(n)}{n!} \langle t_0 | n_0 \rangle \langle n_0 | e^{-iA(t_0)} | n_0 \rangle \tag{282}
\]

can be derived without first casting the master equation into a linear PDE [91, 130, 289, 336, 416]. The following section outlines this derivation for the process \( k A \to IA \) with time-independent rate coefficient \( \gamma \). Moreover, we assume \( A(n) \) to be analytic in \( n \).

The alternative derivation of the path integral representation (275) starts out from the exponential solution of the master equation, i.e., from (83)

\[
\rho(t, n | t_0, n_0) = \langle n | e^{\Omega(t-t_0)} | n_0 \rangle. \tag{283}
\]

The bras are chosen as the unit row vectors \( |n\rangle = \hat{e}_n^\dagger \) and the kets as the unit column vectors \( |\bar{n}\rangle = \hat{e}_n \). As before, the transition matrix of the reaction \( k A \to IA \) with rate coefficient \( \gamma \) reads (see (85))

\[
Q(c, a) = \gamma(c^\dagger - c)^k. \tag{284}
\]

The creation matrix fulfills \( c|n\rangle = |n + 1\rangle \) and \( \langle n | c = |n - 1\rangle \), and the annihilation matrix \( a|n\rangle = |n\rangle - |n - 1\rangle \) and \( \langle n | a = \langle n + 1 | n - 1\rangle \). Thus, the basis column vectors can be generated incrementally via \( |n\rangle = c^n |0\rangle \) and the basis row vectors via \( |n\rangle = (0|e^{n\gamma/c})^\dagger \).

Since the bra \( |n\rangle \) is a left eigenvector of the number matrix \( N := ca \) with eigenvalue \( n \), one may write \( A(n)|n\rangle = \langle n | A(N) |n\rangle \) for an analytic observable \( A \). After inserting the exponential solution (283) into the averaged observable (282), an evaluation of the sums therefore results in

\[
\langle A \rangle_{x_0} = \langle 0 | e^{A(c)\Omega(c^\dagger - c)^k} e^{A(c^\dagger - c)^k} | 0 \rangle, \tag{285}
\]

Here we employed the (infinitely-large) unit matrix \( \mathbb{I} \). Moreover, we changed the variable \( x(t_0) \) to \( x_0 \) in anticipation of a discrete-time approximation.

Following the lecture notes of Cardy [91], 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 \([a, c] = \mathbb{I}\). First, it follows from this relation that \([a, c^n] = n c^{n-1}\mathbb{I}\) holds for all \( n \in \mathbb{N} \), and more generally that \([a, [a, \ldots, [a, c^n]]] = (n) c^{n-1}\mathbb{I}\) holds for nested commutators with \( j \leq n \) annihilation matrices. nestled commutators of higher order vanish. The Hadamard lemma [417] can therefore be employed to write (with \( z \in \mathbb{C} \) and \( n \in \mathbb{N} \))

\[
e^{z^n} c^n = \left( c^n + [a, c^n]z + \frac{1}{2} [a, [a, c^n]]z^2 + \ldots \right) e^{z^n} \tag{286}
\]

\[
e^{z^n} c^n = \left( c^n + nc^{n-1} \mathbb{I} + \frac{n}{2} c^{n-2} \mathbb{I} + \ldots \right) e^{z^n} = (c^n + z^n) e^{z^n}. \tag{287}
\]

This expression generalizes to the following shift operations for an analytic function \( f \):

\[
e^{z^n} f(c) = f(c + z\mathbb{I}) e^{z^n} \tag{288}
\]

\[
f(a) e^{z^n} = f(c + z\mathbb{I}) e^{z^n}. \tag{289}
\]

Shifting of the first exponential in (285) to the right therefore results in

\[
\langle A \rangle_{x_0} = \langle 0 | A((c + 1) a) e^{Q(c + 1, a)\Delta t} e^{\gamma(c^\dagger - c)^k} | 0 \rangle. \tag{290}
\]

To proceed, we split the time interval \([t_0, t]\) into \( N \) pieces of length \( \Delta t = (t - t_0) / N \). The Trotter formula [418] can then be used to express the right hand side of (290) in terms of the limit

\[
\lim_{N \to \infty} \langle 0 | A((c + 1) a) [1 + Q(c + 1, a)\Delta t]^N e^{\gamma(c^\dagger - c)^k} | 0 \rangle. \tag{291}
\]

This expression is now rewritten by inserting \( N \) of the identity matrices [15]

\[
1 = \sum_{m=0}^{\infty} \frac{1}{m!} \left( \int_{\mathbb{R}} dx \left( \partial_x^m x^m \right) \int_{\mathbb{R}} \frac{d\alpha}{2\pi e^{-i\alpha x}} \right) |m\rangle \langle n| \tag{292}
\]

\[
= \int_{\mathbb{R}} \frac{d\alpha}{2\pi} e^{(-\alpha - 1)x} \langle m | e^{\alpha x} | n \rangle. \tag{293}
\]

We obtained the expression in the second line by performing integrations by parts while neglecting any potential boundary terms (note that the derivations in sections 4 and 5 did not involve integrations by parts). Moreover, we introduced the right eigenvector \( |z\rangle := e^{i(c^\dagger - c)} |0\rangle \) of the annihilation matrix \( (a|z\rangle = z |z\rangle \) with \( z \in \mathbb{C} \), and the left eigenvector \( \langle z^*| := \langle 0 | e^{z^n} \) of the creation matrix \( \langle z^* | c = z \langle z^* | \). These vectors are commonly referred to as ‘coherent states’. The eigenvector conditions can be easily verified by rewriting the vectors as
\[
|z\rangle = \sum_{m=0}^{\infty} \frac{z^m e^{-z}}{m!}|m\rangle \quad \text{and} \quad \langle z'| = \sum_{n=0}^{\infty} z^n|n\rangle. \tag{294}
\]

Insertion of the identity matrices into (291) results in
\[
\langle A \rangle_{x_0} = \lim_{N \to \infty} \int_1^N \langle 0|\mathcal{A}(c + 1)a) e^{\mathcal{S}[c]}|0\rangle \\
\times \prod_{j=1}^N [e^{-i g^{-1} \xi - a}[1 + \mathcal{Q}(c + 1, a) \Delta t] e^{i \xi - x}]|0\rangle. \tag{295}
\]

This expression can be simplified with the help of the shift operations in (288) and (289). Since the \( \mathcal{Q}\)-matrix (284) is a normal-ordered polynomial in \( c \) and \( a \) (i.e. all the \( c \) are to the left of all the \( a \)), and both \( \langle 0|c \rangle \) and \( \langle a|0\rangle \) vanish, the above expression evaluates to
\[
\langle A \rangle_{x_0} = \lim_{N \to \infty} \int_1^N \langle 0|\mathcal{A}(c + 1)(a + x \mathbb{1})|0\rangle \\
\times \prod_{j=1}^N [e^{-i g^{-1} \xi - x}|1 + \mathcal{Q}(i q_j + 1, x_j - 1) \Delta t]. \tag{296}
\]

The factor \( \langle 0|\mathcal{A}(c + 1)(a + x \mathbb{1})|0\rangle \) could be evaluated by normal-ordering the observable with respect to \( c \) and \( a \) before employing \( \langle 0|c = 0 \text{ and } a|0\rangle = 0 \) again. The resulting object is sometimes called a ‘normal-ordered observable’ [336]. In the following, we show that this object agrees with the average over a Poisson distribution in (276). The proof of this assertion is based on the observation that \( f_j(x) := \langle 0|(c + 1)(a + x \mathbb{1})|0\rangle \) fulfills the following defining relation of Touchard polynomials (see [419, 420] for information on these polynomials):
\[
f_{j+1}(x) = x \sum_{i=0}^{j} \binom{j}{i} f_i(x) \quad \text{with} \quad f_0(x) = 1. \tag{297}
\]

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 \( \langle n^j \rangle |x \rangle \) as defined in (276), our above assertion holds true.

The path integral representation (275) of averaged observables is recovered as the continuous-time limit of (296) (upon rewriting \( 1 + \mathcal{Q} \Delta t \) as an exponential). The action (277) is also recovered because the transition matrix \( \mathcal{Q}(c,a) \) in (284) and the adjoint transition operator \( \mathcal{Q}_+^j(c,a) \) in (122) fulfill \( \mathcal{Q}(i q + 1, x) = \mathcal{Q}(x, i q + 1) \) for scalar arguments.

6.3. Perturbation expansions

The path integral representation (275) 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 [15, 70] or a non-perturbative analysis [410] 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 (176), so that
\[
\mathcal{Q}_+^j(x, \partial_t + 1) = \mathcal{A}_t(x) \partial_t + \mathcal{B}_t(x) \frac{\partial^2}{2} + \mathcal{P}_+^j(x, \partial_t). \tag{298}
\]

By following the steps in appendix D, the path integral representation of a factorial moment can be written as
\[
\langle \langle n \rangle \rangle_{x(t_0)} = \frac{\delta^j}{\delta \mathcal{Q}(t)} \left. e^{\int_{t_0}^{t_f} dt' \mathcal{P}_+^i(x(t), \partial x') \frac{\delta}{\delta \mathcal{Q}(x(t'))} \mathcal{Z}_{Q,X} \right|_{Q=x=x=0}, \tag{299}
\]
where \( Z_{Q,X}^{0,0} \) represents the \( (Q, X)\)-generating functional (180). Moreover, \( x(t) \) solves the Itô SDE
\[
dx(t) = [\alpha_t(x) + X(x(t)) \partial x + \beta_t(x) \partial W(t)] dt + \sqrt{\beta_t(x)} \partial W(t) \tag{300}
\]
with initial condition \( x(t_0) \). If both the diffusion and perturbation parts of the transition operator (298) vanish, the factorial moment at time \( t \) obeys the statistics of a Poisson distribution, i.e. \( \langle \langle n \rangle \rangle = \langle x(t) \rangle \). In the general case, the representation (299) may be evaluated perturbatively by expanding its exponential as [15]
\[
e^{\int_{t_0}^{t_f} dt' \mathcal{P}_+^i} = \sum_{m=0}^{\infty} \int_{t_0}^{t_f} dt_m \mathcal{P}_+^{i_1} \cdots \int_{t_0}^{t_2} dt \mathcal{P}_+^{i_2} \cdots \int_{t_0}^{t_1} dt \mathcal{P}_+^{i_m} \tag{301}
\]
Note that this expression involves only a single sum and not a double sum as the expansion (230) did.

6.4. Coagulation

Let us exemplify the perturbation expansion of (299) for particles that diffuse and coagulate in the reaction \( 2A \rightarrow A \). This reaction exhibits the same asymptotic particle decay as the binary annihilation reaction \( 2A \rightarrow \emptyset \) and thus belongs to the same universality class [294, 421]. Only a pre-factor differs between the perturbation operators of the two reactions (see below). A full renormalization group analysis of general annihilation reactions \( k A \rightarrow l A \) with \( l < k \) was performed by Lee [130, 416] (see also [69]). 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 \( 2A \rightarrow A \) and the annihilation reaction \( 2A \rightarrow \emptyset \) 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 \( \emptyset \rightarrow A \). The fluctuations around the ensuing non-trivial steady state have been explored with the help of path integrals in [298, 308].

In the following, we consider particles that diffuse on an arbitrary network \( L \) and that coagulate in the reaction \( 2A \rightarrow A \). Hence, all results from section 4.5.4 apply here as well, except that the transition operator of the coagulation reaction now acts as a perturbation. With the local coagulation rate coefficient \( \mu_{t,i} \) at lattice site \( i \in L \), this perturbation reads (see (122))
\[
\mathcal{P}_+^i(x, i q) = \sum_{i \in L} \mu_{t,i} x_i^2 \left[ (i q + 1) - (i q + 1)^2 \right] \tag{302}
\]
\[
= \sum_{i \in L} (-\mu_{t,i}) x_i^2 (i q) \left[ (i q + 1) - (i q + 1)^2 \right]. \tag{303}
\]
For the binary annihilation reaction \( 2A \rightarrow \emptyset \), 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 4.5.4, summands of the expansion of (299) can be represented by Feynman diagrams. For the mean local particle number \( \langle n_i \rangle \), each of the diagrams is composed of a sink with one incoming leg, and possibly of the vertices

\[
\begin{array}{c}
\mu_{\tau,i} \\
\text{and}
\end{array}
\]

\quad (304)

According to the functional derivative (232), 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 \( \partial_x \), which represents the homogeneous solution of \( \partial_t x_i = \Delta^\varepsilon x_i + X_i(\tau) \) (see (229) and (231)).

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

\[
\begin{array}{c}
\mu_{\tau,i} \\
\text{and}
\end{array}
\]

\quad (305)

This loop is part of the \( m = 2 \) summand of (301). Its mathematical expression is obtained by tracing the diagram from right to left and reads

\[
\int_{t_0}^t dt_1 \sum_{j \in \mathbb{L}} G(t, i|t_1, j)(-\mu_{\tau,j}) + \int_{t_0}^t dt_1 \sum_{j \in \mathbb{L}} 2G(t_1, j|\tau, k)(-\mu_{\tau,k})|\lambda_{h, \xi}(\tau)|^2.
\]

(306)

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 \( \partial_t Q(t) \), is not associated to an additional pre-factor (unlike in section 4.5.4).

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 (299), one can define the ‘tree-level average’

\[
\bar{n}_i(t) = \langle n_i(t) \rangle + \int_{t_0}^t dt_1 \sum_{j \in \mathbb{L}} G(t, i|t_1, j)(-\mu_{\tau,j})|\lambda_{h, \xi}(\tau)|^2
\]

\quad (307)

The corresponding mathematical expression reads

\[
\bar{n}_i(t) = \langle n_i(t) \rangle + \int_{t_0}^t dt_1 \sum_{j \in \mathbb{L}} G(t, i|t_1, j)(-\mu_{\tau,j})|\lambda_{h, \xi}(\tau)|^2
\]

\quad (307)

Differentiation of this equation with respect to \( t \) confirms the validity of the rate equation (309) (recall the definition of the propagator in section 4.5.2).

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 [69, 70, 130, 293, 294, 302, 304, 305, 316, 400], of branching and annihilating random walks and percolation processes [69, 70, 96, 132, 133, 301, 303, 422, 423], and of elementary multi-species reactions [131, 295–297, 299, 300, 309–311, 313, 424, 425]. 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 \( k A \rightarrow l A \). Provided that the number of particles \( n \) in the system is initially Poisson
distributed with mean \( x(0) \), we showed that the average of an observable \( A(n) \) can be represented by the path integral

\[
\langle A \rangle_{x(t_0)} = \int_{t_0} dy e^{-S(y)} \langle A \rangle_{x(t)}
\]  

(311)

with \( \langle A \rangle = \sum_{n=0}^{\infty} \frac{x^n e^{-x}}{n!} A(n) \).  

(312)

We derived this representation in section 6.1 by summing the backward path integral representation (171) 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 [70,96,130–133]. Its unshifted version is obtained for a redefined basis function. Unlike path integral representations encountered in the literature, our representation (311) 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 (312) of \( A(n) \) over a Poisson distribution (see section 6.2).

As shown in section 6.3, the path integral representation (311) 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 \( 2A \rightarrow A \). 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 [69,70]. Recently, non-perturbative renormalization group techniques have been developed for the evaluation of stochastic path integrals [317,410]. These techniques have proved particularly useful in studying branching and annihilating random walks [317,318] and annihilation processes [306,307].

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 [11].

More recently, Elgart and Kamenev extended the method for the study of rare event probabilities [35] and for the classification of phase transitions in reaction–diffusion models with a single type of particles [320]. These studies are effectively based on the forward path integral representation (244) 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 (168) 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 (24) and employ the basis functions that we introduced in sections 2.2.2 and 3.2.2. Moreover, we assume the number of particles to be initially Poisson distributed with mean \( x(0) \). 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

\[
|p(t,n|t_0)\rangle_{x(t_0)} = \sum_{n_0=0}^{\infty} p(t,n|t_0,n_0) \frac{x(t_0)^{n_0} e^{-x(t_0)}}{n_0!}.
\]

(313)

For this purpose, we employ the forward path integral (244) to formulate an alternative representation of the ordinary probability generating function

\[
|g(t|t_0,n_0)\rangle_{q(t_0)} := \sum_{n=0}^{\infty} q(t)^n p(t,n|t_0,n_0).
\]

(314)

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’ \( \bar{q}(t) \), i.e. \( q(t) = \bar{q}(t) + \Delta q(t) \). The path \( \bar{q}(t) \) and an auxiliary path \( \bar{x}(\tau) \) 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 [426]. In the next section, we apply the method to the binary annihilation process \( 2A \rightarrow \varnothing \).

In order to implement the above steps, we define the basis function

\[
|n\rangle_{r,\Delta q(\tau)} := (\zeta \Delta q(\tau) + \bar{q}(\tau))^{n} e^{-\int(\zeta \Delta q(\tau) + \bar{q}(\tau))}.
\]

(315)

for yet to be specified paths \( \bar{q}(\tau) \) and \( \bar{x}(\tau) \), and a free parameter \( \zeta \). The basis function has the same form as the basis function (70) but with its second argument being written as \( \Delta q(\tau) \). Provided that the path \( \bar{q}(\tau) \) fulfils the final condition \( \bar{q}(t) = q(t) \), one can trivially rewrite the generating function (314) in terms of the above basis function as

\[
|g\rangle_{q(t_0)} = e^{\bar{q}(t)\frac{d}{dt}\left(0\right)} \left\{ \sum_{n=0}^{\infty} |n\rangle_{r,\Delta q(t)} p(t,n) \right\}_{\Delta q(t_0)=0}.
\]

(316)

The term in brackets has the form of the generalized generating function (45), and thus it can be rewritten in terms of the forward path integral (248) 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
\[ |g(t_0)|_{\tilde{q}(t_0)} = e^{i\tilde{q}(t_0)\tilde{q}(t_0) - \int_{t_0}^{t_1} e^{-S} |n_0\rangle_{t_0} \Delta q(t_0)]_{\Delta q(t_0) = 0}. \]  

(317)

In analogy with (246), the integral sign is defined as the continuous-time limit

\[ \int_{t_0}^{t_1} = \lim_{N \to \infty} \prod_{j=0}^{N-1} \int_{2\pi} 2\pi e^{\Delta q_j}(t_j). \]  

(318)

The action (249) inside the generating function (317) reads

\[ S = \int_{t_0}^{t_1} |\Delta x(\partial_{\tau} - \Delta q) - \tilde{Q}(\Delta q, i\Delta x)|. \]  

(319)

As emphasized above, this 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 (314), it proves convenient to work in terms of the marginalized distribution (313) is recovered from the redefined generating function via

\[ \langle g(t_0), x(t_0) \rangle_{t_1} = \frac{1}{n!} \langle \partial_{\tilde{q}^n(\Delta q)} g(t_0), x(t_0) \rangle_{t_1} \bigg|_{\Delta q(t_0) = 0}. \]  

(321)

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 \( \tilde{x}(t) = x(t) \). We specify these paths in such a way that the action (319) becomes free of terms that are linear in the deviations \( \Delta q \) and \( \Delta x \). For this purpose, let us recall the definitions of the creation operator \( c = \zeta \Delta q + \tilde{q} \) and of the annihilation operator \( a = \partial_{\Delta q} + \tilde{x} \) from section 2.2.2, as well as the definition of the transition operator

\[ \tilde{Q}(\Delta q, \Delta x) = Q(c, a) + E_x(c, a). \]  

(322)

The basis evolution operator is specified in (77) as

\[ E_x(c, a) = (\partial_{\tilde{x}}) (a - \tilde{x}) - (\partial_{\tilde{x}}) \tilde{x}. \]  

(323)

Upon performing a Taylor expansion of the operator \( \tilde{Q}(\Delta q, i\Delta x) \) in the action (319) with respect to the deviations \( \Delta q \) and \( \Delta x \), one observes that terms that are linear in the deviations vanish from the action if the paths \( \tilde{x}(\tau) \) and \( \tilde{q}(\tau) \) fulfill \( 11 \)

\[ \partial_{\tilde{x}} \tilde{x} = \frac{\partial Q(\tilde{q}, \tilde{x})}{\partial \tilde{q}} \]  

with \( \tilde{x}(t_0) = x(t_0) \) and

\[ \partial_{\tilde{x}} \tilde{x} = \frac{\partial Q(\tilde{q}, \tilde{x})}{\partial \tilde{x}} \]  

with \( \tilde{q}(t) = q(t) \).

(324)

(325)

These equations resemble Hamilton’s equations from classical mechanics. Just as in classical mechanics, \( Q(\tilde{q}, \tilde{x}) \) is conserved along solutions of the equations if it does not depend on time itself. This property follows from \( \frac{d}{dt} Q(\tilde{q}, \tilde{x}) = \partial_{\tilde{q}} Q(\tilde{q}, \tilde{x}) + \partial_{\tilde{x}} Q(\tilde{q}, \tilde{x}) \), with \( \frac{d}{dt} \) being the total time derivative. Since the ordinary generating function obeys the flow equation \( \partial_{\tilde{q}} Q(\tilde{q}, \tilde{x}) = Q(\tilde{q}, \tilde{x}) \), the conservation of total probability requires that \( Q(\tilde{q}, \tilde{x}) = 0 \). This condition is, for example, fulfilled by the transition operator \( Q(x, \partial_x) = \gamma_x (\tilde{q}_x - q_x) \partial_x \), of the generic reaction \( k A \rightarrow I A \) (see (64)). For the final value \( q(t) = 1 \), Hamilton’s equations (324) and (325) are solved by \( \tilde{q}(\tau) = 1 \) with \( \tilde{x}(\tau) \) solving the rate equation \( \partial_\tau \tilde{x} = \gamma_x (\tilde{q}_x - q_x) \) of the process. Elgart and Kamenev analysed the topology of \( Q(\tilde{q}, \tilde{x}) = 0 \) lines of reaction–diffusion models with a single type of particles to classify the phase transitions of these models \( [320] \) (they use the label \( p \) instead of \( \tilde{q} \), and \( q \) instead of \( \tilde{x} \)).

Provided that Hamilton’s equations (324) and (325) are fulfilled, the action (319) evaluates to

\[ S = \tilde{x}(\tau) q(\tau) - x(t_0) + \tilde{S} + \Delta S \]  

with the definitions

\[ \tilde{S} := x(t_0)(1 - \tilde{q}(t_0)) + \int_{t_0}^{t_1} \partial_{\tilde{x}} q_\tau \tilde{x} \]  

(327)

and

\[ \Delta S := \int_{t_0}^{t_1} (\partial_{\Delta x} \partial_{\tau} \Delta q - \Delta Q_{\tau}). \]  

(328)

The transition operator \( \Delta Q \), absorbs all terms of the Taylor expansion of \( Q_{\tau} \) that are of second or higher order in the deviations. Combined with (320), the above action results in the following path integral representation of the generating function:

\[ |g(t_0), x(t_0)\rangle_{Q(\tilde{q}, \tilde{x})} = e^{-\tilde{S}} \int_{t_0}^{t_1} e^{-\Delta S} \bigg|_{\Delta q(t_0) = 0}. \]  

(329)

Although this path integral representation may seem daunting, our primary interest lies only in its leading-order approximation \( |g| \approx e^{-\tilde{S}} \). This approximation is exact if \( \Delta Q \) vanishes. That is, for example, the case for the simple growth process \( \tilde{S} \rightarrow A \). For the binary annihilation reaction \( 2A \rightarrow \tilde{S} \), the leading-order approximation was evaluated by Elgart and Kamenev up to a pre-exponential factor \( [35] \). The pre-exponential factor was later approximated by Assaf and Meerson for large times \( [36] \). In the next section, we evaluate the leading-order approximation of the binary annihilation reaction and evaluate its pre-exponential factor for arbitrary times.
The first term of this equation was neglected in previous studies [35, 36] (for large times \( t \), \( \bar{q}(t_0) \approx 1 \)). Using the conservation of \((\bar{q}^2 - 1)x^2\) once again, the action in the leading-order approximation \( |\psi(q(t))e^{-\Delta S(q(t))} | \) can be written as

\[
S(q(t)) = [1 - \bar{q}(t_0)x(t_0) + \frac{1 - \bar{q}(t_0)^2}{2x_0^2(t)}].
\]

(335)

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 \(|\psi(q(t))e^{-\Delta S(q(t))} |\) respects the normalization of the underlying probability distribution because

\[
\sum_{n=0}^{\infty} p(t, n|t_0, n_0) = |\psi(q(t))_0 = 1. \]

(336)

Here we used that for \( q(t) = 1 \), Hamilton’s equation (331) is solved by \( \bar{q}(\tau) = 1 \), implying that \( S(q(t)) = 1 = 0 \).

The probability distribution follows from the generating function via the inverse transformation \(|p(t, n|t_0)_{q(t_0)} = \frac{1}{n!} |\psi_n(q(t))e^{\Delta S(q(t))} |_{q(t_0) = 0} \). This transformation 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

\[
|p(t, n|t_0)_{q(t_0)} = \frac{1}{2\pi i} \oint C \frac{dq(t)}{q(t)} e^{-\Delta S(q(t)) - n \ln q(t)}. \]

(337)

Here, \( C \) represents a closed path around zero in the complex domain and is integrated over once in counter-clockwise direction.

The contour integral in (337) can be evaluated in a saddle-point approximation as

\[
|p(t, n|t_0)_{q(t_0)} \approx \frac{q_n e^{-\Delta S(q_n)}}{2\pi(n - q_n^2 S''(q_n))}. \]

(338)

The saddle-point \( q_n \) is found by solving the equation \( n/q_n = -S'(q_n) \). By differentiating the implicit solution (334) with respect to \( q(t) \), the saddle-point condition can be rewritten as

\[
\frac{n}{q_n} = x(t_0) \sqrt{1 - \bar{q}(t_0)^2} \sqrt{1 - \bar{q}_n^2}. \]

(339)

A closed equation for \( q_n \) is obtained by combining this equation with the implicit solution (334). The resulting equation is solved by \( q_n = 1 \) for \( n = x(t) \), 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 \( q_n \) has been obtained, the value \( \bar{q}(t_0) \) can be inferred from (339).

One piece is still missing for the numeric evaluation of the probability distribution (338), namely its denominator. It evaluates to

72. Binary annihilation

Let us demonstrate the use of the representation (329) for the binary annihilation reaction \( 2A \rightarrow \phi \) with rate coefficient \( \mu_r = 1 \). The transition operator \( Q_{\tau} = \mu_r (1 - \bar{q})^2 \) of the binary annihilation process vanishes for \( x = 0 \) and \( \bar{q} = 1 \) (orange dotted lines). The green disk represents the fixed point \( (\bar{q}, x) = (1, 0) \) 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).

\[
Q_{\tau} = \mu_r (1 - \bar{q})^2 \]

(320)

A phase portrait of Hamilton’s equations (321) and (322) read

\[
\partial_q \bar{q} = 2 \mu_r (1 - \bar{q}^2) \bar{x} \quad \text{with} \quad \dot{x}(t_0) = x(t_0) \quad \text{and} \quad \partial_x \bar{x} = -2 \mu_r \bar{x}^2.
\]

(330)

A phase portrait of these equations is shown in figure 9.

\[
\text{Equation (331) is solved by} \quad \bar{q} = 1, \quad \text{for which the previous equation simplifies to the rate equation} \quad \partial_q \bar{q} = -2 \mu_r \bar{x}^2 \quad \text{of the process}. \quad \text{This rate equation is solved by}
\]

\[
\bar{x}(t) = x(t) \sqrt{1 + x(t)/\bar{x}_0(t)}
\]

(332)

with the asymptotic limit \( \bar{x}_0(t) = (2 \int_{t_0}^{\infty} \mu_r(t) \tau^{-1} \). Since Hamilton’s equations conserve \((\bar{q}^2 - 1)x^2\), one can rewrite the equation (331) as

\[
\partial_q \bar{q} = -2 \mu_r x(t_0) \sqrt{1 - \bar{q}(t_0)^2} \sqrt{1 - \bar{q}^2}.
\]

(333)

Here we assume \( \bar{q}(t) < 1 \) but the derivation can also be performed for \( \bar{q}(t) > 1 \) (these inequalities are preserved along the flow). The above equation only allows for an implicit solution that provides \( \bar{q}(t_0) \) for a given \( q(t) \), namely

\[
\arccos \bar{q}(t_0) + \frac{x(t_0)}{\bar{x}_0(t)} \sqrt{1 - \bar{q}(t_0)^2} = \arccos q(t).
\]

(334)
The resulting path \( \tilde{x}() \rightarrow \tilde{x}(t) \rightarrow \tilde{x}(T) \) is required to evaluate \( \int_{\tilde{x}(0)}^{\tilde{x}(T)} \tilde{Q}(\tilde{x}, \tilde{q}) \). 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 \( 2A \rightarrow \emptyset \). On the downside, the leading order of the “backward” approach does not respect the normalization of \( p(t, n|t_0, n_0) \). The normalization of the distribution has to be implemented by hand or by evaluating higher order terms.

To represent the marginalized distribution (343) by a path integral, we require that for a given particle number \( n \) and for a given initial mean \( \bar{x}(t_0) \), there exist functions \( \tilde{x}(\tau) \) and \( \tilde{q}(\tau) \) fulfilling

\[
\partial_{\tau} \tilde{x} = \frac{\partial \tilde{Q}(\tilde{x}, \tilde{q})}{\partial \tilde{q}} \quad \text{with} \quad \tilde{x}(t_0) = x(t_0)
\]

\[
\partial_{\tau} \tilde{q} = -\frac{\partial \tilde{Q}(\tilde{x}, \tilde{q})}{\partial \tilde{x}} \quad \text{with} \quad \tilde{q}(t) = \frac{n}{\tilde{x}(t)}
\]

Since the transition operators \( \tilde{Q}_A \) in (122) and \( \tilde{Q}_S \) in (64) are connected via \( \tilde{Q}^+_A(\tilde{x}, \tilde{q}) = Q_S(\tilde{q}, \tilde{x}) \), the above equations differ...
from the previous equations (324) and (325) only in the final condition on \( \bar{q} \). As shown below, the marginalized distribution (343) can then be expressed as
\[
| p(t, n|t_0)_{\Delta(n)} \rangle = \frac{\bar{x}(t \rangle e^{-S^\dagger}}{n!} \int_{\Delta(n)=0}^{|t|} e^{-\Delta S} \bigg| \frac{\bar{x}(t \rangle e^{-S^\dagger}}{n!} \bigg|_{\Delta(n)=0} (346)
\]
with \( S^\dagger := x(t_0) + \int_{t_0}^t dt \{ q \partial_x \bar{x} - Q^\dagger_{1}(x, q) \} \) (347)
and \( \Delta S^\dagger := n \left( \frac{\Delta x(t)}{\bar{x}(t)} - \ln \left( 1 + \frac{\Delta x(t)}{\bar{x}(t)} \right) \right) \)
\[+ \int_{t_0}^t dt \{ i \Delta q \partial_x \Delta x - \Delta Q^\dagger_{1} \} \] (348)
The variables \( \Delta x \) and \( \Delta q \) again represent deviations from the stationary paths \( \bar{x} \) and \( \bar{q} \). The transition operator \( \Delta Q^\dagger_{1} \) encompasses all the terms of an expansion of \( Q^\dagger_{1}(\Delta x + \bar{x}, \bar{q}) \) 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 \( \Delta x \) and \( \Delta q \).

The derivation of the above representation proceeds analogously to section 7.1. It begins by using \( \bar{x}(t_0) = x(t_0) \) and the basis function \( | n \rangle_{\Delta \lambda} = \frac{1}{n!} (\Delta x + \bar{x})^n e^{-\bar{q}(\Delta x + \bar{x})} \) from (125) to rewrite the right-hand side of the marginalized distribution (343) as
\[
\epsilon(\bar{q}(t_0) - x(t_0)) \sum_{n=0}^{\infty} p(t, n|t_0, n_0)|n_0, \Delta(n)\rangle \bigg|_{\Delta(n)=0} \bigg|_{\Delta(n)=0} \quad (349)
\]
The sum in this expression can be represented by the backward path integral (168), turning the expression into
\[
\epsilon(\bar{q}(t_0) - x(t_0)) \int_{t_0}^t dt e^{-S^\dagger} | n_0, \Delta(n)\rangle \bigg|_{\Delta(n)=0} \quad (350)
\]
Recalling the definition of \( \mathcal{E}_x \) in (127), the transition operator
\[
\tilde{Q}_{1}(\Delta x, i \Delta q) := Q^\dagger_{1}(\Delta x + \bar{x}, \bar{q}) - \mathcal{E}_x(\Delta x + \bar{x}, \bar{q}) \] (351)
can now be expanded in the deviations. We thereby obtain the action
\[
S^\dagger = \bar{q}(t_0)\bar{x}(t_0) - \bar{q}(t)(\bar{x}(t)) + \int_{t_0}^t dt \{ \bar{q} \partial_x \bar{x} - Q^\dagger_{1}(\bar{x}, \bar{q}) \}
\[+ \int_{t_0}^t dt \{ i \Delta q \partial_x \Delta x - \Delta Q^\dagger_{1} \} \] (352)
The path integral representation (346) follows upon inserting this action into (350) and employing the final condition \( \bar{q}(t) = n\bar{x}(t) \).

7.4. Binary annihilation
To complement the approximation of the binary annihilation reaction \( 2A \rightarrow \emptyset \) with rate coefficient \( \mu_r \) in section 7.2, we now perform an approximation of the process using the backward path integral representation (346) of the marginalized distribution. We first perform the approximation in leading order, which amounts to the evaluation of the pre-factor of (346). For the simple growth process \( \bar{q} \rightarrow A \) and for the linear decay process \( A \rightarrow \emptyset \), the pre-factor provides exact solutions.

7.4.1. Leading order. The evaluation of the leading-order approximation
\[
| p(t, n|t_0)_{\Delta(n)} \rangle \approx \frac{\bar{x}(t \rangle e^{-S^\dagger}}{n!} \] (353)
proceeds very much like the derivation in section 7.2. Using the adjoint transition operator \( Q^\dagger_{1}(\bar{c}_r, \bar{a}_r) = \mu_r \bar{c}_r^2 (1 - \bar{a}_r^2) \) from (122), one finds that Hamilton’s equations
\[
\partial_x \bar{x} = -2\mu_r \bar{q} \bar{x}^2 \quad \text{with } \bar{x}(t_0) = x(t_0) \] (354)
\[
\partial_r \bar{q} = 2\mu_r (1 - \bar{q}^2) \bar{x} \quad \text{with } \bar{q}(t) = n\bar{x}(t) \] (355)
agree with (330) and (331), apart from the final condition on \( \bar{q} \). The conservation of \( (\bar{q}^2 - 1)\bar{x}^2 \) and the asymptotic deterministic limit \( \bar{x}_\infty(t) = (2 \int_{t_0}^t dt \mu_r)^{-1} \) can be used to rewrite the action (347) as
\[
\tilde{S}^\dagger = n + \left( 1 - \frac{\bar{x}(t)^2 - n^2}{x(t)^2} \right) x(t_0) + \frac{\bar{x}(t)^2 - n^2}{2\bar{x}_\infty(t)} \] (356)
In addition, the conservation law can be used to infer the flow equation
\[
\partial_t \bar{q} = -2\mu_r \sqrt{1 - \bar{q}^2} \sqrt{\bar{x}(t)^2 - n^2} \] (357)
Here we assume \( \bar{q}(t) < 1 \) but the derivation can also be performed for \( \bar{q}(t) > 1 \). The equation is implicitly solved by
\[
\arccos \sqrt{1 - \bar{q}(t)^2 - n^2} = \arccos \frac{n}{\bar{x}(t)} \] (358)
For a given initial mean \( x(t_0) \), this equation can be solved numerically for \( \bar{x}(t) \), which is then inserted into (353). 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.

7.4.2. Beyond leading order. Instead of normalizing the function (353) by hand, it can be normalized by evaluating the path integral (346). In the following, we perform this evaluation, but only after restricting the action \( \Delta S^\dagger \) in (348) to terms that are of second order in the deviations \( \Delta x \) and \( \Delta q \), i.e. to
\[
\Delta S^\dagger \approx -\frac{(\Delta x(t))^2}{2 \bar{x}(t)^2} + \int_{t_0}^t dt (i \Delta q \partial_x \Delta x - \Delta Q^\dagger_{1}) \] (359)
The corresponding transition operator
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 upper row show the marginalized probability distribution \( p(t,n|0) \) at times (a) \( t = 0.001 \), (b) \( t = 0.1 \), and (c) \( t = 1 \) on a linear scale. The figures in the lower row 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 \( \mu_r = 1 \). 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 (353) was corrected by the factor \( 1/\sqrt{\det A} \) (see section 7.4.2). The numerical solution of the fixed point equation (358) failed for large times (missing circles in (f)).

\[
\Delta Q^I_t \approx i \Delta q \Delta x \alpha_r - \frac{(\Delta q)^2}{2} \beta_r + \frac{(\Delta x)^2}{2} c_r
\]

is obtained as explained in section 7.3 (with \( \zeta := i \)). Its coefficients read

\[
\alpha_r := -4 \mu_r \bar{x} q, \quad \beta_r := 2 \mu_r \bar{x}^2, \quad c_r := 2 \mu_r (q^2 - 1).
\]

The path integral can now be evaluated in one of the following two ways.

On the one hand, the transition operator \( \Delta Q^I_t \) 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 (346) can be expressed in terms of the following average over a Wiener process \( W \):

\[
\sum_{n} e^{-\Delta S^I(t_0)} \int_{(t_0)} e^{\Delta S^I(t) - \frac{(\Delta q(t))^2}{2} \beta_r} \int_{(t_0)} e^{\Delta S^I(t) - \frac{(\Delta x(t))^2}{2} c_r} \text{d}W(t).
\]

Here, \( \Delta x(t) \) solves the Itô SDE

\[
d\Delta x = \alpha_r \Delta x \text{d}t + \sqrt{\beta_r} \text{d}W(t),
\]

with initial value \( \Delta x(t_0) = 0 \). Unfortunately, the computation of a sufficient number of sample paths to approximate the average (362) was found to be rather slow.

As an alternative to the above way, one can discretize the action \( \Delta S^I \) and cast it into the quadratic form \( \Delta S^I = \frac{1}{2} \xi^T A \xi \) with \( \zeta := \{ \Delta q_j, \Delta x_j \}_{j=1}^N \). The matrix \( A \) is symmetric and tridiagonal. If \( A \) is also positive-definite, the path integral (346) evaluates to \( 1/\sqrt{\det A} \). Curiously, we found that this factor even matches the average (362) when \( A \) is not positive-definite. Therefore, we used the factor \( 1/\sqrt{\det A} \) to correct the approximation (353) 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 (358) for a fixed point \( \bar{x}(t) \).

\section{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 (314) as an intermediate step, the expansion of the backward path integral representation provides the marginalized distribution (343). 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 \( 2A \rightarrow \emptyset \). Our approximation based on the forward path integral amends an earlier computation of Elgart and Kamenev [35]. The computation requires a saddle-point approximation of Cauchy’s
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 [138, 139, 142], the step of a molecular motor along a cytoskeletal filament [113–115], or the flipping of a spin [134–137]. The stochastic evolution of such processes is commonly modelled in terms of a master equation. 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 work, 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 [98, 110] and in the review [274].

The stochastic simulation algorithm (section 1.4)

The SSA [75–78] and its variations [78, 247–249] 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 [138, 139, 142] and extensions of it have been implemented in various simulation packages [250–257]. A fast but only approximate alternative to the SSA for the simulation of processes evolving on multiple time scales is τ-leaping [78, 249, 258–267]. Algorithms for the simulation of processes with time-dependent transition rates are discussed in [268, 269]. 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’ (19). Various alternatives to the SSA have been proposed for the numerical evaluation of this representation, mostly based on its Laplace transform [238–243]. 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 \( \partial_{\tau} p(\tau|\tau_0) = Qp(\tau|\tau_0) \) is solved by the matrix exponential \( p(\tau|\tau_0) = e^{Q\tau} \). Here, \( Q \) represents the transition matrix of a process and \( p(\tau|\tau_0) \) 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. \( \sup_{(m, m')} |Q(m, m)| < \infty \)), the conditional probability distribution can be represented by the ‘uniformization’ formula (16).

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 [219, 220]).

Flow equations (sections 2 and 3)

For the chemical master equation (24) of the reaction \( A \rightarrow \emptyset \) with rate coefficient \( \gamma_r \), it is readily shown that the probability generating function

\[
|g(\tau|\tau_0, n_0)\rangle = \sum_n \langle n|p(\tau, n|\tau_0, n_0)
\]

with basis function \( |n\rangle = a^n \) obeys the linear partial differential equation, or ‘flow equation’

\[
\partial_{\tau}|g\rangle = \gamma_r (q^\dagger - q^\gamma) \partial_{q^\dagger}|g\rangle.
\]

Analogously, the marginalized distribution

\[
|p(t, n|\tau_0)\rangle = \sum_{n_0} p(t, n|\tau, n_0)|n_0\rangle
\]

with Poisson basis function \( |n_0\rangle = \frac{x^{n_0}\exp(-x)}{n_0!} \) obeys the backward-time flow equation

\[
\partial_{\tau} |p\rangle = \gamma_r x^{\dagger}[(\partial_x + 1)^\dagger - (\partial_x + 1)^\dagger]|p\rangle.
\]

After one has solved one of these equations, the conditional probability distribution can be recovered via the inverse transformations \( p(\tau, n|\tau_0, n_0) = \langle n|g(\tau|\tau_0, n_0)\rangle \) 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
\[
\langle g(\tau | t_0, n_0) \rangle = \sum_n p(\tau, n | t_0, n_0).
\] (368)

For the Poisson basis function \( p(\tau, n | t_0, n_0) = e^{-\tau \sum_{n=0}^{\infty} \frac{n!}{n!} A(n) } \), the inverse transformation
\[
p(\tau, n | t_0, n_0) = \langle g(\tau | t_0, n_0) \rangle n!
\]
of this functional recovers the Poisson representation of Gardiner and Chaturvedi [32, 33]. The Poisson representation can, for example, be used for the computation of mean extinction times [374], but we found the marginalized distribution to be more convenient for this purpose (see 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 [39, 83], spectral formulations and WKB approximations [34, 36–38, 328, 364, 366, 367, 370, 427]. 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 review (information on these approximations can be found in [63, 64, 146, 171, 172, 177, 354–360, 362, 363, 364]). 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 [366] for a discussion of ‘lacking’ boundary conditions in a study of the branching-annihilation reaction \( A \to 2A \) and \( 2A \to 0 \)).

Recently, we have been made aware of novel ways of treating the generating function’s flow equation based on duality relations [428–432]. 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 [14, 69, 70, 96, 130–133, 162, 293–319, 321, 323–325, 400, 422–425]. A classical example of such a phase transition is the transition between an active and an absorbing state of a system [68]. 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
\[
\langle A \rangle_{x(t_0)} = \int_{x(t)} e^{-\mathcal{S}} \langle A \rangle_{x(t_0)}.
\] (369)

with
\[
\langle A \rangle_{x(t_0)} := \sum_{n=0}^{\infty} \frac{x^n e^{-x}}{n!} A(n).
\] (370)

The above path integral can be derived both from the novel backward path integral representation
\[
p(t, n | t_0, n_0) = \langle n_0 \rangle \int_{t_0}^{t} e^{-S^{\prime}} | n \rangle,
\] (371)
of the conditional probability distribution (see section 4), and from the forward path integral representation
\[
p(t, n | t_0, n_0) = \langle n \rangle \int_{t_0}^{t} e^{-S} | n_0 \rangle
\] (372)
(seen section 5). The means of the integral signs and of the actions \( S \) and \( S^{\prime} \) 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 [69, 70], information on non-perturbative techniques in [410]. 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 Kramer–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 [16–19]. 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 (371) and (372) of the conditional probability distribution, or use the representation (369) 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 (372) can be approximated by expanding its action around ‘stationary paths’ [35]. The paths obey Hamilton’s equations of the form
\[
\partial_t \tilde{x} = \frac{\partial Q_\tau}{\partial \tilde{q}} \quad \text{and} \quad \partial_t \tilde{q} = -\frac{\partial Q_\tau}{\partial \tilde{x}},
\] (373)
with the transition operator \( Q_\tau \) 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 (371). 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 (373) as has been proposed in [320].

We hope that our review 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.

Acknowledgments

For critical reading of a pre-submission manuscript and for providing valuable feedback, we 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, we would like to thank Johannes Nefele, Isabella Krämer, and Cornelius Weig for helpful discussions during the preparation of the manuscript. This research was financially supported by the German Excellence Initiative via the program ‘NanoSystems Initiative Munich’ (NIM).

Appendix 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 [110]). In particular, we assume that for \( \tau \in [t_0, t] \), a function \( u(\tau, x) \) obeys the linear PDE

$$
\partial_{\tau} u(\tau, x) = \alpha(\tau, x) \partial_x u + \frac{1}{2} \beta(\tau, x) \partial_x^2 u \quad \text{(A.1)}
$$

with final value \( u(t, x) = G(x) \). \( \text{(A.2)} \)

The function \( \alpha \) is called a drift coefficient and \( \beta \) a diffusion coefficient. According to the Feynman–Kac formula, the above PDE is solved by

$$
u(\tau, x) = \langle \langle G(X(t)) \rangle \rangle \quad \text{(A.3)}
$$

with \( \langle \langle \rangle \rangle \) 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)

$$
\text{d}X(s) = \alpha_s(X(s)) \text{d}s + \sqrt{\beta_s(X(s))} \text{d}W(s) \quad \text{(A.4)}
$$

with initial value \( X(\tau) = x \). \( \text{(A.5)} \)

If the initial value \( x \) of the SDE is chosen as a real number, the drift coefficient \( \alpha_s \) is a real function, and the diffusion coefficient \( \beta_s \) 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 \( \sqrt{\beta_s} := \gamma_s \) fulfilling \( \gamma_s \gamma_s^T = \beta_s \). If the matrix \( \beta_s \) is symmetric and positive-semidefinite, one may choose \( \gamma_s \) as its unique symmetric and positive-semidefinite square root [111].

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 (A.3) then implies that \( \langle \langle \delta(x - X(\tau)) \rangle \rangle \) 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 (A.3), we assume that \( X(s) \) satisfies the SDE (A.4). By Itô’s Lemma and the PDE (A.1), it then holds that

$$
\text{d}u(s, X(s)) = \frac{\partial u(s, X(s))}{\partial X(s)} \sqrt{\beta_s(X(s))} \text{d}W(s). \quad \text{(A.6)}
$$

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 \( \langle \langle \text{d}W \rangle \rangle_W = 0 \), it follows that \( \langle \langle u(t, X(t)) \rangle \rangle_W = \langle \langle u(t, X(\tau)) \rangle \rangle_W \). This expression coincides with the Feynman–Kac formula (A.3) because the initial condition (A.5) implies that \( u(\tau, X(\tau)) = u(\tau, x) \) does not depend on the Wiener process, and because the final condition (A.2) implies that \( u(t, X(t)) = G(X(t)) \).

Appendix B. Proof of the path summation representation in section 1.4

In the following, we prove that the path summation representation (19) solves the forward master equation (11) 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

$$
\partial_s p(\tau | 0, n_0) = (w - d) p(\tau | \cdot) \quad \text{(B.1)}
$$

with the probability vector \( p(\tau | t_0, n_0) \). Note that the matrix notation assumes some mapping between the state space of \( n \) and an index set \( I \subset \mathbb{N} \). Following [239], the Laplace transform

$$
L_s f(s) := \int_0^\infty \text{d}\tau e^{-st} f(\tau) \quad \text{(B.2)}
$$

of the above master equation is given by

$$
\text{s} L_s p(s | \cdot) - \hat{e}_{n_0} = (w - d) L_s p(s | \cdot) \quad \text{(B.3)}
$$

Here, \( p(0 | 0, n_0) = \hat{e}_{n_0} \) 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 \( \mathbb{I} \), the above expression can be solved for \( L_s p(s | \cdot) \) and be rewritten as

$$
L_s p(s | \cdot) = [\mathbb{I} - (s \mathbb{I} + d)^{-1} w]^{-1} (s \mathbb{I} + d)^{-1} \hat{e}_{n_0} \quad \text{(B.4)}
$$

The value of the real part of \( s \) is determined by the parameter \( \epsilon \) in the inverse Laplace transformation

$$
\hat{f}(t) = \frac{1}{2\pi i} \int_{\epsilon - i\infty}^{\epsilon + i\infty} \text{d}s e^{st} L_s f(s). \quad \text{(B.5)}
$$

We assume that \( \epsilon \) can be chosen so large that the first factor in the solution (B.4) can be rewritten as a geometric series, i.e. as
Appendix C. Solution of the random walk in sections 2.2.1 and 3.2.1

We here provide the conditional probability distribution \( p(\tau, n|t_0, n_0) \) solving the random walk from sections 2.2.1 and 3.2.1. In the first of these sections, we showed that this distribution is obtained by applying the functional \( \langle n| f = \int_0^\tau dt \exp(-i\omega t) f(q) \rangle \) to the generating function \( g(\tau|t_0, n_0) \), which we derived as

\[
\exp\left(e^{i\omega} - 1 \int_{t_0}^{\tau} ds \, r_s + (e^{-i\omega} - 1) \int_{t_0}^{\tau} ds \, l_s + in_0\omega \right).
\]

Multiple series expansions show that this expression can be rewritten as \( e^{-\int_{t_0}^{\tau} ds \, (r_s + l_s)} \) times

\[
\sum_{k=0}^\infty \sum_{m=0}^{k} \left( \int_{t_0}^{\tau} ds \, r_s \right)^m \left( \int_{t_0}^{\tau} ds \, l_s \right)^{k-m} \frac{m!}{(k-m)!} e^{i2n_0k+nl_0}\omega.
\]

We ignore the constant pre-factor \( e^{-\int_{t_0}^{\tau} ds \, (r_s + l_s)} \) for now and apply the functional \( \langle n| \) to this expression. After carefully noting the restrictions imposed by Kromericker delta's, the resulting expression reads

\[
\sum_{k=0}^\infty \left( \int_{t_0}^{\tau} ds \, r_s \cdot \int_{t_0}^{\tau} ds \, l_s \right)^k \frac{k!}{(k+n-n_0)!}.
\]

This sum can be expressed by a modified Bessel function of the first kind (10.25.2 in [335]), resulting in

\[
\left( \int_{t_0}^{\tau} ds \, r_s \right) \frac{\alpha-n_0}{\tau} \left( I_{n_0-n_0} \left( 2 \left( \int_{t_0}^{\tau} ds \, r_s \cdot \int_{t_0}^{\tau} ds \, l_s \right)^{1/2} \right) \right).
\]

Multiplied with \( e^{-\int_{t_0}^{\tau} ds \, (r_s + l_s)} \), this expression corresponds to a Skellam distribution with mean \( \mu = n_0 + \int_{t_0}^{\tau} ds \, (r_s + l_s) \) and variance \( \sigma^2 = \int_{t_0}^{\tau} ds \, (r_s + l_s) \).

Appendix D. Details on the 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 (171) of the marginalized distribution in terms of the \((Q,X)\)-generating functional (179). Upon comparing the discrete-time approximation of the marginalized distribution (168) with its representation in (177), one observes that the corresponding \((Q,X)\)-generating functional should read

\[
Z_{QX,N} := \int_1^{N-1} e^{\sum \omega_j n_j - S_N^X e^{\omega_N}}
\]

with

\[
Z_N := \sum_{j=1}^{N} \Delta t Q_j x_{j-1} + \sum_{j=1}^{N-1} \Delta t X_j - i\eta_j.
\]

A differentiation of (D.1) with respect to \( \Delta t Q_j \) generates a factor \( x_{j-1} \), a differentiation with respect to \( \Delta t X_j \) a factor \( i\eta_j \). Upon recalling the definition of the action \( S_X^X \) in (169), the first exponential in (D.1) can be rewritten via

\[
i\eta_j x_{j-1} - S_X^{(N)}
\]

\[
= - \sum_{j=1}^{N-1} i\eta_j \left( x_{j-1} - x_{j-1} + \alpha_{j-1}(x_{j-1}-\Delta t) \right)
\]

\[
+ \sum_{j=1}^{N-1} \Delta t \beta_{j-1}(x_{j-1}-\Delta t) + \sum_{j=1}^{N-1} \Delta t \beta_{j-1}(x_{j-1} - i\eta_j)
\]

\[
+ i\eta_j \delta x_{j-1}.
\]

The equality holds up to corrections of \( O(\Delta t) \) (because of the sums, all remaining terms are of \( O(1) \)). Note that the right hand side of (D.4) is independent of \( x_N \). One can linearize the terms that are quadratic in \( \eta_j \) by the completion of a square. In particular, we write

\[
\exp \left( \sum_{j=1}^{N-1} \Delta t \beta_{j-1}(x_{j-1} - \Delta t) \right)
\]

\[
= \exp \left( \sum_{j=1}^{N-1} i\eta_j \left( \beta_{j-1}(x_{j-1} - \Delta t) \right) \right)
\]

\[
= \exp \left( \sum_{j=1}^{N-1} \Delta t \beta_{j-1}(x_{j-1} - \Delta t) \right)
\]

with the average being defined as

\[
\langle \langle \cdot \rangle \rangle_W := \left( \int_1^{N-1} d\Delta t W_j f_{0,\Delta t}(\Delta t W_j) \right)
\]

The average employs the Gaussian distribution

\[
\rho_{0,\Delta t}(\Delta W_j) = \frac{e^{-(\Delta W_j)^2/(2\Delta t)}}{\sqrt{2\pi\Delta t}}
\]

with zero mean and variance \( \Delta t \) (see (200)). The sequence \( \Delta W_1, \ldots, \Delta W_N \) can be interpreted as the steps of a discretized Wiener process. To proceed with the derivation, we now move the perturbation operator \( P^* \) to the front of the \((Q,X)\)-generating functional (D.1) by writing
Upon combining all of the above steps, one finds that

\[
Z_{Q,X,N}^{0} := \left( \prod_{j=1}^{N} \int_{R} dx_j \delta(b(x_j) - x_{j-1} - \{\alpha_{j-1}, x_{j-1} + X_{j-1}\} \Delta t + \sqrt{\beta_{j-1}(x_{j-1}) \Delta W_j}) \right) e^ {\sum_{j=1}^{N-1} \Delta t \left( \frac{1}{2} \frac{\partial^2}{\partial x_j^2} + \frac{1}{2} \frac{\partial^2}{\partial x_j^2} \right) } Z_{Q,X,N} \tag{D.9}
\]

with the definition

\[
\sum_{j=1}^{N-1} \Delta t \left( \frac{1}{2} \frac{\partial^2}{\partial x_j^2} + \frac{1}{2} \frac{\partial^2}{\partial x_j^2} \right) \] 0

\[e^ {\sum_{j=1}^{N-1} \Delta t \left( \frac{1}{2} \frac{\partial^2}{\partial x_j^2} + \frac{1}{2} \frac{\partial^2}{\partial x_j^2} \right) } Z_{Q,X,N} \tag{D.8}
\]

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 (D.9) 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 (177) with generating functional (179) and Itô SDE (181).

References

[1] Einstein A 1905 Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen Ann. Phys. 322 549–60
[2] Fokker A D 1914 Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld Ann. Phys. 348 810–20
[3] Planck M 1917 Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie Sitz.-ber. Preuß. Akad. Wiss. 24 324–41
[4] Kolmogoroff A 1931 Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung Math. Ann. 104 415–58
[5] Nordsieck A, Lamb W E Jr and Uhlenbeck G E 1940 On the theory of cosmic-ray showers I: the Furry model and the fluctuation problem Physica 7 344–60
[6] Doi M 1976 Second quantization representation for classical many-particle system J. Phys. A: Math. Gen. 9 1465–77
[7] Doi M 1976 Stochastic theory of diffusion-controlled reaction J. Phys. A: Math. Gen. 9 1479–95
[8] Ze’l’ dovich Ya B and Ovchinnikov A A 1978 The mass action law and the kinetics of chemical reactions with allowance for thermodynamic fluctuations of the density Sov. Phys.— JETP 47 829–34
[9] Rose H A 1979 Renormalized kinetic theory of nonequilibrium many-particle classical systems J. Stat. Phys. 20 415–47
[10] Grassberger P and Scheunert M 1980 Fock-space methods for identical classical objects Fortschr. Phys. 28 547–78
[11] Mikhailov A S 1981 Path integrals in chemical kinetics I Phys. Lett. A 85 214–6
[12] Mikhailov A S 1981 Path integrals in chemical kinetics II Phys. Lett. A 85 427–9
[13] Goldenfeld N 1984 Kinetics of a model for nucleation-controlled polymer crystal growth J. Phys. A: Math. Gen. 17 2807–21
[14] Mikhailov A S and Yashin V V 1985 Quantum-field methods in the theory of diffusion-controlled reactions J. Stat. Phys. 38 347–59
[15] Peltii L 1985 Path integral approach to birth-death processes on a lattice J. Phys. France 46 1469–83
[16] Martin P C, Siggia E D and Rose H A 1973 Statistical dynamics of classical systems Phys. Rev. A 8 423–37
[17] de Dominicis C 1976 Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques J. Phys. Colloq. 37 C1–247–C1–253
[18] Janssen H-K 1976 On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties Z. Phys. B 23 377–80
[19] Bausch R, Janssen H-K and Wagner H 1976 Renormalized field theory of critical dynamics Z. Phys. B 24 113–27
[20] Fock V 1932 Konfigurationsraum und zweite Quantelung Z. Phys. 75 622–47
[21] Dirac P A M 1939 A new notation for quantum mechanics Math. Proc. Camb. Phil. Soc. 35 416–8
[22] Schrödinger E 1926 Der stetige Übergang von der Mikro- zur Makromechanik Naturwissenschaften 14 664–6
[23] Sudarshan E C G 1963 Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams Phys. Rev. Lett. 10 277–9
[24] Glauber R J 1963 Coherent and incoherent states of the radiation field Phys. Rev. 131 2766–88
[25] Wick G C 1950 The evaluation of the collision matrix Phys. Rev. 80 268–72
[26] Sandow S and Trimper S 1993 Aggregation processes in a master-equation approach Europhys. Lett. 21 799–804
[27] Pätzold F, Sandow S and Trimper S 1994 Diffusion and correlation in a coherent representation Z. Phys. B 95 357–62
[28] Bares P-A and Mobilia M 1999 Diffusion-limited reactions of hard-core particles in one dimension Phys. Rev. E 59 1996–2009
[29] Brunel V, Oerding K and van Wijland F 2000 Fermionic field theory for directed percolation in (1 + 1)-dimensions J. Phys. A: Math. Gen. 33 1085–97
[30] Schulz M and Reineker P 2005 Exact substitute processes for diffusion-reaction systems with local complete exclusion J. Phys. A: Math. Gen. 38 429–68
[31] Silva É M, Muzy P T and Santana A E 2008 Fock space for fermion-like lattices and the linear Glauber model Physica A 387 5101–9
[32] Gardiner C W and Chaturvedi S 1977 The Poisson representation. I. A new technique for chemical master equations J. Stat. Phys. 17 429–68
[33] Chaturvedi S and Gardiner C W 1978 The Poisson representation. II. Two-time correlation functions J. Stat. Phys. 18 501–22
[34] Walczak A M, Mugler A and Wiggins C H 2009 A stochastic spectral analysis of transcriptional regulatory cascades Proc. Natl Acad. Sci. USA 106 6529–34
[35] Elgart V and Kamenev A 2004 Rare event statistics in reaction–diffusion systems Phys. Rev. E 70 041106
[36] Assaf M and Meerson B 2006 Spectral formulation and WKB approximation for rare-event statistics in reaction systems Phys. Rev. E 74 041115
[37] Assaf M and Meerson B 2006 Spectral theory of metastability and extinction in birth-death systems Phys. Rev. Lett. 97 200602
[38] Assaf M, Kamenev A and Meerson B 2008 Population extinction in a time-modulated environment Phys. Rev. E 78 041123
[39] Sasai M and Wolynes P G 2003 Stochastic gene expression as a many-body problem Proc. Natl Acad. Sci. USA 100 2374–9
[40] Edwards S F 1964 The statistical dynamics of homogeneous turbulence J. Fluid Mech. 18 293–73
[41] Sreenivasan K R and Eting G L 2005 Sam Edwards and the turbulence theory Stealing the Gold ed P Goldbart et al (Oxford: Oxford University Press) pp u–5
[42] Haken H 1976 Generalized Onsager–Machlup function and classes of path integral solutions of the Fokker–Planck equation and the master equation Z. Phys. B 24 321–6
[43] Wódkiewicz K 1981 Functional representation of a non-Markovian probability distribution in statistical mechanics Phys. Lett. 84 56–8
[44] Pesquera L, Rodriguez M A and Santos E 1983 Path integrals for non-Markovian processes Phys. Lett. A 94 287–9
[45] Fox R F 1986 Functional-calculus approach to stochastic differential equations Phys. Rev. A 33 467–76
[46] Fox R F 1986 Uniform convergence to an effective Fokker–Planck equation for weakly colored noise Phys. Rev. A 34 4525–7
[47] Luciani J F and Verga A D 1987 Functional integral approach to bistability in the presence of correlated noise Europhys. Lett. 4 255–61
[48] Förster A and Mikhaïlov A S 1988 Optimal fluctuations leading to transitions in bistable systems Phys. Lett. A 126 459–62
[49] Luciani J F and Verga A D 1988 Bistability driven by correlated noise: Functional integral treatment J. Stat. Phys. 56 567–97
[50] Bray A J and McKane A J 1989 Instanton calculation of the escape rate for activation over a potential barrier driven by colored noise Phys. Rev. Lett. 62 493–6
[51] Hänggi P 1989 Path integral solutions for non-Markovian processes Z. Phys. B 75 275–81
[52] McKane A J 1989 Noise-induced escape rate over a potential barrier: results for general noise Phys. Rev. A 40 4050–3
[53] McKane A J, Luckock H C and Bray A J 1990 Path integrals and non-Markov processes. I. General formalism Phys. Rev. A 41 644–56
[54] Bray A J, McKane A J and Newman T J 1990 Path integrals and non-Markov processes. II. Escape rates and stationary distributions in the weak-noise limit Phys. Rev. A 41 657–67
[55] Luckock H C and McKane A J 1990 Path integrals and non-Markov processes. III. Calculation of the escape-rate prefactor in the weak-noise limit Phys. Rev. A 42 1982–96
[56] Venkatesh T G and Pataki L M 1993 Effective Fokker–Planck equation: Path-integral formalism Phys. Rev. E 48 2402–12
[57] Einchcomb S J B and McKane A J 1996 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 M Millonas (New York: Springer) pp 139–54
[58] Mahanta C and Venkatesh T G 2000 Damped stochastic system driven by colored noise: analytical solution by a path integral approach Phys. Rev. E 62 1509–20
[59] Wiegel F W 1986 Introduction to Path-Integral Methods in Physics and Polymer Science (Singapore: World Scientific)
[60] Doi M and Edwards S F 1988 The Theory of Polymer Dynamics (International Series of Monographs on Physics vol 73) (Oxford: Clarendon)
[61] Vilgis T A 2000 Polymer theory: path integrals and scaling Phys. Rep. 336 167–254
[62] Kleinert H 2009 Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets 5th edn (New Jersey: World Scientific) ch 15
[63] Bressloff P C and Newby J M 2014 Path integrals and large deviations in stochastic hybrid systems Phys. Rev. E 89 042701
[64] Bressloff P C 2014 Stochastic Processes in Cell Biology (Interdisciplinary Applied Mathematics vol 41) (Cham: Springer) ch 10
[65] Bressloff P C 2015 Path-integral methods for analyzing the effects of fluctuations in stochastic hybrid neural networks J. Math. Neurosci. 5 1–33
[66] Hinrichsen H 2000 Non-equilibrium critical phenomena and phase transitions into absorbing states Adv. Phys. 49 815–958
[67] Ödor G 2004 Universality classes in nonequilibrium lattice systems Rev. Mod. Phys. 76 663–724
[68] Henkel M, Hinrichsen H and Lübeck S 2008 Non-Equilibrium Phase Transitions: Volume I: Absorbing Phase Transitions (Theoretical and Mathematical Physics) (Dordrecht: Springer)
[69] Täuber U C, Howard M and Vollmayr-Lee B P 2005 Applications of field-theoretic renormalization group methods to reaction–diffusion problems J. Phys. A: Math. Gen. 38 R79–131
[70] Täuber U C 2014 Critical Dynamics: a Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior (Cambridge: Cambridge University Press)
[71] Zwanzig R 1961 Memory effects in irreversible thermodynamics Phys. Rev. 124 983–92
[72] Zwanzig R 1964 On the identity of three generalized master equations Physica 30 1109–23
[73] Mori H 1965 Transport, collective motion and Brownian motion Prog. Theor. Phys. 33 423–55
[74] Oppenheim I and Shuler K E 1965 Master equations and Markov processes Phys. Rev. 138 B1007–11
[75] Gillespie D T 1976 A general method for numerically simulating the stochastic time evolution of coupled chemical reactions J. Comput. Phys. 22 403–34
[76] Gillespie D T 1977 Exact stochastic simulation of coupled chemical reactions J. Phys. Chem. 81 2340–61
[77] Gillespie D T 1992 A rigorous derivation of the chemical master equation Physica A 188 404–25
[78] Gillespie D T 2007 Stochastic simulation of chemical kinetics Annu. Rev. Phys. Chem. 58 35–55
[79] Onsager L and Machlup S 1953 Fluctuations and irreversible processes Phys. Rev. 91 1505–12
[80] Wiener N 1921 The average of an analytic functional Proc. Natl Acad. Sci. USA 7 253–60
[81] Wiener N 1921 The average of an analytic functional and the Brownian movement Proc. Natl Acad. Sci. USA 7 294–8
[82] Brown R 1828 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 Phil. Mag. Ser. 2 4 161–73
[83] Eyink G L 1996 Action principle in nonequilibrium statistical dynamics Phys. Rev. E 54 3419–35
[84] Bender C M and Orszag S A 1999 Advanced Mathematical Methods for Scientists and Engineers I (New York: Springer) ch 10
[85] Kac M 1949 On distributions of certain Wiener functionals Trans. Am. Math. Soc. 65 1–13
[86] Itô K 1944 Stochastic integral Proc. Imp. Acad. 20 519–24
[87] Itô K 1946 On a stochastic integral equation Proc. Japan Acad. 22 32–5
[88] Itô K 1950 Stochastic differential equations in a differentiable manifold Nagoya Math. J. 1 35–47
[89] Lemons D S 1997 Paul Langevin’s 1908 paper ‘On the theory of Brownian motion’ [1908 ‘Sur la théorie du mouvement brownien,’ C. R. Acad. Sci. 146 530–3] Am. J. Phys. 65 1079–81
[90] van Wijland F 2001 Field theory for reaction–diffusion processes with hard-core particles Phys. Rev. E 63 022101
[91] Cardy J 2008 Reaction–diffusion processes Non-Equilibrium Statistical Mechanics and Turbulence (London Mathematical Society Lecture Note Series ed S Nazarenko and O V Zaboronski (Cambridge: Cambridge University Press) pp 108–61
[92] Schrödinger E 1926 An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* **28** 1049–70

[93] Heisenberg W 1925 Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Z. Phys.* **33** 879–93

[94] Born M and Jordan P 1925 Zur Quantenmechanik. Z. Phys. **34** 858–88

[95] Born M, Heisenberg W and Jordan P 1926 Zur Quantenmechanik. II. Z. Phys. **35** 557–615

[96] Cardy J L and Täuber U C 1998 Field theory of branching and annihilating random walks. *J. Stat. Phys.* **90** 1–56

[97] Sinervo B and Lively C M 1996 The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380** 240–3

[98] van Kampen N G 2007 *Stochastic Processes in Physics and Chemistry* 3rd edn (Amsterdam: Elsevier)

[99] Huang R, Chavez I, Taute K M, Lüdic B, Jeney S, Raizen M G and Florin E L 2011 Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid. *Nat. Phys.* **7** 576–80

[100] Franosch T, Grimm M, Belushkin M, Mor F M, Foffi G, von Smoluchowski M 1906 Zur kinetischen Theorie der Brownischen Molekularbewegung und der Suspensionen. *Ann. Phys.* **326** 756–80

[101] Klein O 1921 Zur statistischen Theorie der Suspensionen und Lösungen. *Ark. Mat. Astr. Fys.* **16** 1–51 (PhD Thesis)

[102] Kramers H A 1940 Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7** 284–304

[103] Risken H 1996 *The Fokker–Planck Equation—Methods of Solution and Applications* (Springer Series in Synergetics vol 18) 2nd edn (Berlin: Springer)

[104] Kimura M 1957 Some problems of stochastic-processes in genetics. *Ann. Math. Stat.* **28** 882–901

[105] de Vladar H P and Barton N H 2011 The contribution of statistical physics to evolutionary biology. *Trends Ecol. Evol.* **26** 424–32

[106] Feynman R P 2005 The principle of least action in quantum mechanics. *Feynman’s Thesis: a New Approach to Quantum Theory* ed L M Brown (Singapore: World Scientific) pp 1–69

[107] Kloeden P E and Platen E 1992 *Numerical Solution of Stochastic Differential Equations* (Applications of Mathematics—Stochastic Modelling and Applied Probability vol 23) (Berlin: Springer)

[108] Majumdar S N 2005 Brownian functionals in physics and computer science. *Curr. Sci.* **89** 2076–92

[109] Gardiner C 2009 *Stochastic Methods: a Handbook for the Natural and Social Sciences* (Springer Series in Synergetics vol 13) 4th edn (Berlin: Springer)

[110] Harville D A 1997 *Matrix Algebra From a Statistician’s Perspective* (New York: Springer) ch 21.9

[111] Volpe G and Wehr J 2016 Effective drifts in dynamical systems with multiplicative noise: a review of recent progress. *Rep. Prog. Phys.* **79** 053901

[112] Lipowsky R, Klumpp S and Nieuwenhuizen T M 2001 Random walks of cytoskeletal motors in open and closed compartments. *Phys. Rev. Lett.* **87** 101801

[113] Parmeggiani A, Franosch T and Frey E 2003 Phase coexistence in driven one-dimensional transport. *Phys. Rev. Lett.* **90** 086601

[114] Parmeggiani A, Franosch T and Frey E 2004 Totally asymmetric simple exclusion process with Langmuir kinetics. *Phys. Rev. E* **70** 046107

[115] Howard J 1997 Molecular motors: structural adaptations to cellular functions. *Nature* **389** 561–7

[116] Anderson W J 1991 *Continuous-Time Markov Chains: an Applications-Oriented Approach* (Springer Series in Statistics: Probability and its Applications) (New York: Springer)

[117] Kolmogoroff A 1936 Zur Theorie der Markoffschen Ketten. *Math. Ann.* **112** 155–60

[118] Furry W H 1937 On fluctuation phenomena in the passage of high energy electrons through lead. *Phys. Rev.* **52** 569–81

[119] Feller W 1939 Die Grundlagen der Volterraschen Theorie des Kämpfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheor.* **5** 11–40

[120] Delbrück M 1940 Statistical fluctuations in autocatalytic reactions. *J. Chem. Phys.* **8** 120–4

[121] Singer K 1953 Application of the theory of stochastic processes to the study of irreproducible chemical reactions and nucleation processes. *J. R. Stat. Soc. B* **15** 92–106

[122] Bartholomay A 1958 Stochastic models for chemical reactions: I. Theory of the unimolecular reaction process. *B. Math. Biophys.* **20** 175–90

[123] Krieger I M and Gans P J 1960 First-order stochastic processes. *J. Chem. Phys.* **32** 247–50

[124] Ishida K 1964 Stochastic model for bimolecular reaction. *J. Chem. Phys.* **41** 2472–8

[125] McQuarrie D A 1963 Kinetics of small systems. I. *J. Chem. Phys.* **38** 433–6

[126] McQuarrie D A, Jachimowski C J and Russell M E 1964 Kinetics of small systems. II. *J. Chem. Phys.* **40** 2914–21

[127] McQuarrie D A 1967 Stochastic approach to chemical kinetics. *J. Appl. Probab.* **4** 413–78

[128] Nicolis G and Prigogine I 1977 *Self-Organization in Nonequilibrium Systems: from Dissipative Structures to Order Through Fluctuations* (New York: Wiley) ch 10

[129] Lee B P 1994 Renormalization group calculation for the reaction $A + \emptyset \rightarrow 2$ *J. Phys. A: Math. Gen.* **27** 2633–52

[130] Lee B P and Cardy J 1995 Renormalization group study of the $A + B \rightarrow \emptyset$ diffusion-limited reaction. *J. Stat. Phys.* **80** 971–1007

[131] Cardy J and Täuber U C 1996 Theory of branching and annihilating random walks. *Phys. Rev. Lett.* **77** 4780–3

[132] Janssen H K and Täuber U C 2005 The field theory approach to percolation processes. *Ann. Phys.* **315** 147–92

[133] Glauber R J 1963 Time-dependent statistics of the Ising model. *J. Math. Phys.* **4** 294–307

[134] Kawasaki K 1966 Diffusion constants near the critical point for time-dependent Ising models. I. *Phys. Rev.* **145** 224–30

[135] Kawasaki K 1966 Diffusion constants near the critical point for time-dependent Ising models. II. *Phys. Rev.* **148** 375–81

[136] Kawasaki K 1966 Diffusion constants near the critical point for time-dependent Ising models. III. Self-diffusion constant. *Phys. Rev.* **150** 285–90

[137] Elowitz M B and Leibler S 2000 A synthetic oscillatory network of transcriptional regulators. *Nature* **403** 335–8

[138] Thattai M and van Oudenaarden A 2001 Intrinsinc noise in gene regulatory networks. *Proc. Natl Acad. Sci. USA* **98** 8614–9

[139] Rao C V, Wolf D M and Arkin A P 2002 Control, exploitation and tolerance of intracellular noise. *Nature* **420** 237–7

[140] Karlebach G and Shamir R 2008 Modelling and analysis of gene regulatory networks. *J. Math. Biol.* **56** 20053901

[141] Shahrezaei V and Swain P S 2008 Analytical distributions for stochastic disease dynamics. *Proc. Natl Acad. Sci. USA* **105** 17256–61

[142] Tsimring L S 2014 Noise in biology. *Rep. Prog. Phys.* **77** 026601

[143] Bailey N T J 1990 A simple stochastic epidemic. *Biometrika* **77** 193–202

[144] Keeling M J and Ross J V 2008 On methods for studying stochastic disease dynamics. *J. R. Soc. Interface* **5** 171–81
[146] Black A J and McKane A J 2011 WKB calculation of an epidemic outbreak distribution J. Stat. Mech. P12006
[147] Rock K, Brand S, Moir J and Keeling M J 2014 Dynamics of infectious diseases Rep. Prog. Phys. 77 026602
[148] Clayton E, Doupe D P, Klein A M, Winton D J, Simons B D and Jones P H 2007 A single type of progenitor cell maintains normal epidermis Nature 446 185–9
[149] Chou T 2007 Peeling and sliding in nucleosome repositioning Phys. Rev. Lett. 99 058105
[150] Volkov I, Banavar J R, Hubbell S P and Maritan A 2003 Neutral theory and relative species abundance in ecology Nature 424 1035–7
[151] McKane A J and Newman T J 2004 Stochastic models in population biology and their deterministic analogs Phys. Rev. E 70 041902
[152] McKane A J and Newman T J 2005 Predator-prey cycles from resonant amplification of demographic stochasticity Phys. Rev. Lett. 94 218102
[153] Mobilia M, Georgiev I T and Täuber U C 2007 Phase transitions and spatio-temporal fluctuations in stochastic lattice Lotka–Volterra models J. Stat. Phys. 128 447–83
[154] Reichenbach T, Mobilia M and Frey E 2006 Coexistence versus extinction in the stochastic cyclic Lotka–Volterra model Phys. Rev. E 74 051907
[155] Volkov I, Banavar J R, Hubbell S P and Maritan A 2007 Patterns of relative species abundance in rainforests and coral reefs Nature 450 45–9
[156] Butler T and Reynolds D 2009 Predator-prey quasicycles from a path-integral formalism Phys. Rev. E 79 032901
[157] Butler T and Goldenfeld N 2009 Robust ecological pattern formation induced by demographic noise Phys. Rev. E 80 030902
[158] Park J-M, Muñoz E and Deem M W 2010 Quasispecies theory for finite populations Phys. Rev. E 81 011902
[159] Noble A E, Hastings A A and Fagan W F 2011 Multivariate Moran process with Lotka–Volterra phenomenology Phys. Rev. Lett. 107 228101
[160] Dobrinevski A and Frey E 2012 Extinction in neutrally stable stochastic Lotka–Volterra models Phys. Rev. E 85 051903
[161] Black A J and McKane A J 2012 Stochastic formulation of ecological models and their applications Trends Ecol. Evol. 27 337–45
[162] Täuber U C 2012 Population oscillations in spatial stochastic Lotka–Volterra models: a field-theoretic perturbational analysis J. Phys. A: Math. Theor. 45 405002
[163] Shih H-Y and Goldenfeld N 2014 Path-integral calculation for the emergence of rapid evolution from demographic stochasticity Phys. Rev. E 90 050702
[164] Cremer J, Melbinger A and Frey E 2012 Growth dynamics and the evolution of cooperation in microbial populations Sci. Rep. 2 281
[165] Reiter M, Rulands S and Frey E 2014 Range expansion of heterogeneous populations Phys. Rev. Lett. 112 148103
[166] Weber M F, Poxleitner G, Hebisch E, Frey E and Opitz M 2014 Chemical warfare and survival strategies in bacterial range expansions J. R. Soc. Interface 11 20140172
[167] Wienand K, Lechner M, Becker F, Jung H and Frey E 2015 Non-selective evolution of growing populations PLoS One 10 e0134300
[168] Reichenbach T, Mobilia M and Frey E 2007 Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games Nature 448 1046–9
[169] Reichenbach T, Mobilia M and Frey E 2007 Noise and correlations in a spatial population model with cyclic competition Phys. Rev. Lett. 99 238105
[170] Reichenbach T, Mobilia M and Frey E 2008 Self-organization of mobile populations in cyclic competition J. Theor. Biol. 254 366–83
[171] Mobilia M and Assaf M 2010 Fixation in evolutionary games under non-vanishing selection Europhys. Lett. 91 10002
[172] Assaf M and Mobilia M 2010 Large fluctuations and fixation in evolutionary games J. Stat. Mech. P09009
[173] Frey E 2010 Evolutionary game theory: theoretical concepts and applications to microbial communities Physica A 389 4265–98
[174] Melbinger A, Cremer J and Frey E 2010 Evolutionary game theory in growing populations Phys. Rev. Lett. 105 178101
[175] Cremer J, Melbinger A and Frey E 2011 Evolutionary and population dynamics: a coupled approach Phys. Rev. E 84 051921
[176] Taulusen A, Claussnitz J C and Hauert C 2012 Stochastic differential equations for evolutionary dynamics with demographic noise and mutations Phys. Rev. E 85 041901
[177] Black A J, Taulusen A and Galla T 2012 Mixing times in evolutionary game dynamics Phys. Rev. Lett. 109 028101
[178] Knebel J, Weber M F, Krüger T and Frey E 2015 Evolutionary games of condensates in coupled birth-death processes Nat. Commun. 6 6977
[179] Melbinger A, Cremer J and Frey E 2015 The emergence of cooperation from a single mutant during microbial life cycles J. R. Soc. Interface 12 20150171
[180] Krug J and Spohn H 1991 Kinetic roughening of growing surfaces Solids Far from Equilibrium (Collection Aea-Saclay: Monographs and Texts in Statistical Physics) ed C Godrèche (Cambridge: Cambridge University Press)
[181] Weidlich W 1991 Physics and social science—the approach of synergetics Phys. Rep. 204 1–163
[182] Weidlich W and Braun M 1992 The Master equation approach to nonlinear economics J. Evol. Econ. 2 233–65
[183] Lux T 1995 Herd behaviour, bubbles and crashes Econ. J. 105 881–96
[184] Castellano C, Fortunato S and Loreto V 2009 Statistical physics of social dynamics Rev. Mod. Phys. 81 591–646
[185] Gross D, Shortle J F, Thompson J M and Harris C M 2008 Fundamentals of Queueing Theory 4th edn (Hoboken, NJ: Wiley)
[186] Nagel K and Schreckenberg M 1992 A cellular automaton model for freeway traffic J. Phys. I France 2 2221–9
[187] Helbing D 2001 Traffic and related self-driven many-particle systems Rev. Mod. Phys. 73 1067–141
[188] Reichenbach T, Franosch T and Frey E 2006 Exclusion processes with internal states Phys. Rev. Lett. 97 050603
[189] Kriecherbauer T and Krug J 2010 A pedestrian’s view on interacting particle systems, KPZ universality and random matrices J. Phys. A: Math. Theor. 43 403001
[190] Chou T, Mallick K and Zia R K P 2011 Non-equilibrium statistical mechanics: from a paradigmatic model to biological transport Rep. Prog. Phys. 74 116601
[191] Vorberg D, Wustmann W, Ketzermerick R and Eckardt A 2013 Generalized Bose–Einstein condensation into multiple states in driven-dissipative systems Phys. Rev. Lett. 111 240405
[192] Vorberg D, Wustmann W, Schomerus H, Ketzermerick R and Eckardt A 2012 Bose–Einstein condensation into multiple states of ideal bosonic and fermionic quantum gases Phys. Rev. E 92 062119
[193] Evans M R and Waclaw B 2014 Condensation in stochastic processes Nat. Commun. 6 10002
[194] Grosskinsky S, Redig F and Vafayi K 2010 Correlation inequalities for interacting particle systems with duality J. Stat. Phys. 141 242–63
[195] Giardinà C, Redig F and Vafayi K 2010 Correlation inequalities for interacting particle systems with duality J. Stat. Phys. 141 242–63
[295] Howard M and Cardy J 1995 Fluctuation effects and multiscaling of the reaction–diffusion front for $A + B \rightarrow \emptyset$
J. Phys. A: Math. Gen. 28 3599–621

[296] Howard M 1996 Fluctuation kinetics in a multispecies reaction–diffusion system J. Phys. A: Math. Gen. 29 3437–60

[297] Oerding K 1996 The $A + B \rightarrow \emptyset$ annihilation reaction in a quenched random velocity field J. Phys. A: Math. Gen. 29 7051–65

[298] Rey P-A and Droz M 1997 A renormalization group study of a class of reaction–diffusion models, with particles input J. Phys. A: Math. Gen. 30 1101–14

[299] Sasamoto T, Mori S and Wadati M 1997 Universal properties of the $na + nb \rightarrow \emptyset$ diffusion-limited reaction Physica A 247 357–78

[300] van Wijland F, Oerding K and Hilhorst H J 1998 Wilson renormalization of a reaction–diffusion process Physica A 251 179–201

[301] Täuber U C, Howard M J and Hinrichsen H 1998 Multicritical behavior in coupled directed percolation processes Phys. Rev. Lett. 80 2165–8

[302] Hinrichsen H and Howard M 1999 A model for anomalous directed percolation Eur. Phys. J. B 7 635–43

[303] Goldschmidt Y Y, Hinrichsen H, Howard M and Täuber U C 1999 Nonequilibrium critical behavior in unidirectionally coupled stochastic processes Phys. Rev. E 59 6381–408

[304] Hnatic M and Honkonen J 2000 Velocity-fluctuation-induced anomalous kinetics of the $A + A \rightarrow \emptyset$ reaction Phys. Rev. E 61 3904–11

[305] Vernon D C 2003 Long range hops and the pair annihilation reaction $A + A \rightarrow \emptyset$: renormalization group and simulation Phys. Rev. E 68 041103

[306] Winkler A A and Frey E 2012 Validity of the law of mass action in three-dimensional coagulation processes Phys. Rev. Lett. 108 108301

[307] Hommighausen I, Winkler A A and Frey E 2013 Fluctuation effects in the pair-annihilation process with Lévy dynamics Phys. Rev. E 88 012111

[308] Droz M and Sasvari L 1993 Renormalization-group approach to simple reaction–diffusion phenomena Phys. Rev. E 48 R2341–6

[309] Leppänen P and Cardy J 1994 Scaling of reaction zones in the $A + B \rightarrow 0$ diffusion-limited reaction Phys. Rev. E 50 R3287–90

[310] Howard M J and Barkema G T 1996 Shear flows and segregation in the reaction $A + B \rightarrow \emptyset$ Phys. Rev. E 53 5949–56

[311] Konkoli Z, Johansson H and Lee B P 1999 Fluctuation effects in steric reaction–diffusion systems Phys. Rev. E 59 R3787–90

[312] Dickman R and Vidigal R 2002 Path-integral representation for a stochastic sandpile J. Phys. A: Math. Gen. 35 7269–85

[313] Hilhorst H J, Deloubrière O, Wasenberger M J and Täuber U C 2004 Segregation in diffusion-limited multispecies pair annihilation J. Phys. A: Math. Gen. 37 7063–93

[314] Janssen H K, van Wijland F, Deloubrière O and Täuber U C 2004 Pair contact process with diffusion: failure of master equation field theory Phys. Rev. E 70 056114

[315] Whitelam S, Berthier L and Garrahan J P 2005 Renormalization group study of a kinetically constrained model for strong glasses Phys. Rev. E 71 026128

[316] Hnatic M, Honkonen J and Lučivjanský T 2013 Two-loop calculation of anomalous kinetics of the reaction $A + A \rightarrow \emptyset$ in randomly stirred fluid Eur. Phys. J. B 86 214

[317] Canet L, Delamotte B, Deloubrière O and Wschebor N 2004 Nonperturbative renormalization-group study of reaction–diffusion processes Phys. Rev. Lett. 92 195703

[318] Canet L, Chaté H and Delamotte B 2004 Quantitative phase diagrams of branching and annihilating random walks Phys. Rev. Lett. 92 255703

[319] Canet L, Chaté H, Delamotte B, Dornic I and Muñoz M A 2005 Nonperturbative fixed point in a nonequilibrium phase transition Phys. Rev. Lett. 95 100601

[320] Elgart V and Kamenev A 2006 Classification of phase transitions in reaction–diffusion models Phys. Rev. E 74 041101

[321] Biuße M A and Cowan J D 2007 Field-theoretic approach to fluctuation effects in neural networks Phys. Rev. E 75 051919

[322] Biuße M A and Cowan J D 2009 Statistical mechanics of the neocortex Prog. Biophys. Mol. Biol. 99 53–86

[323] Biuße M A and Chow C C 2013 Beyond mean field theory: statistical field theory for neural networks J. Stat. Mech. P03003

[324] Bettelheim E, Agam O and Shnerb N M 2001 ‘Quantum phase transitions’ in classical nonequilibrium processes Physica E 9 600–8

[325] Täuber U C 2011 Stochastic population oscillations in spatial prey-predator models J. Phys.: Conf. Ser. 319 012019

[326] Zhang B and Wolynes P G 2014 Stem cell differentiation as a many-body problem Proc. Natl Acad. Sci. USA 111 10185–90

[327] Feller W 1949 On the theory of stochastic processes, with particular reference to applications Proc. [First] Berkeley Symp. on Mathematical Statistics and Probability (Berkeley, CA: University of California Press) pp 403–32

[328] Walczak A M, Mugler A and Wiggins C H 2012 Analytic methods for modeling stochastic regulatory networks Computational Modeling of Signaling Networks (Methods in Molecular Biology vol 880) ed X Liu and M D Betterton (New York: Humana Press) pp 273–322

[329] Skellam J G 1946 The frequency distribution of the difference between two Poisson variates belonging to different populations J. R. Stat. Soc. 109 296

[330] Takayasu H and Tretyakov A Yu 1992 Extinction, survival and dynamical phase transition of branching annihilating random walk Phys. Rev. Lett. 68 3060–3

[331] Jensen I 1993 Critical behavior of branching annihilating random walks with an odd number of offsprings Phys. Rev. E 47 R1–4

[332] Ballentine L E 1998 Quantum Mechanics: a Modern Development (Singapore: World Scientific) ch 6.2

[333] Peskin M E and Schroeder D V 1995 An Introduction To Quantum Field Theory (Boulder, CO: Westview Press)

[334] Oikubo J 2012 One-parameter extension of the Doi–Peliti formalism and its relation with orthogonal polynomials Phys. Rev. E 86 042102

[335] Olver F W J, Lozier D W, Boisvert R F and Clark C W (ed) (Berkeley, CA: University of California Press) pp 403–32

[336] Wiese K J 2016 Coherent-state path integral versus coarse-grained effective stochastic equation of motion: from reaction diffusion to stochastic sandpiles Phys. Rev. E 93 042117

[337] Stinchcombe R B, Gryenberg M D and Barma M 1993 Diffusive dynamics of deposition-evaporation systems, jamming and broken symmetries in related quantum-spin models Phys. Rev. E 47 4018–36

[338] Gryenberg M D, Newman T J and Stinchcombe R B 1994 Exact-solutions for stochastic adsorption-desorption models and catalytic surface processes Phys. Rev. E 50 3729–31
[339] Goldstein S 1951 On diffusion by discontinuous movements and on the telegraph equation *Q. J. Mech. Appl. Math.* **4** 129–56

[340] Kac M 1974 A stochastic model related to the telegrapher’s equation *Rocky Mt. J. Math.* **4** 497–510 (reprinting of an article published in 1956)

[341] Berezin F A 1966 *The Method of Second Quantization* (Orlando, FL: Academic) ch 1.3

[342] Combes M and Robert D 2012 *Coherent States and Applications in Mathematical Physics* (Theoretical and Mathematical Physics) (Dordrecht: Springer) ch 11.2

[343] Alev M A 2000 On the description of the Glauber dynamics of one dimensional disordered Ising chain *Physica A* **277** 261–73

[344] Mobilia M, Zia R K P and Schmittmann B 2004 Complete exclusive particle systems *J. Stat. Phys.* **7642** E 5–10 (reprinting of an equation 41505001 from equilibrium into equilibrium in one-dimensional transport systems)

[345] Aliev M A 2000 On the description of the Glauber dynamics of one dimensional disordered Ising chain *Physica A* **277** 261–73

[346] Schütz G and Sandow S 1994 Non-abelian symmetries of stochastic processes: derivation of correlation functions for random-vertex models and disordered-interacting-particle systems *Phys. Rev. E* **49** 2726–41

[347] Schütz G M 1995 Reaction–diffusion processes of hard-core particles *J. Stat. Phys.* **79** 243–64

[348] Henkel M, Orlandini E and Santos J 1997 Reaction–diffusion processes from equivalent integrable quantum chains *Ann. Phys.* **259** 163–231

[349] Carlon E, Henkel M and Schollwöck U 1999 Density matrix renormalization group and reaction–diffusion processes *Eur. Phys. J. B* **12** 99–114

[350] Carlon E, Henkel M and Schollwöck U 2001 Critical properties of the reaction–diffusion model 2A → 3A, 2A → 0 *Phys. Rev. E* **63** 036104

[351] Park S-C, Kim D and Park J-M 2000 Path-integral formulation of stochastic processes for the exclusive particle systems *Phys. Rev. E* **62** 7642–5

[352] Park S-C and Park J-M 2005 Generating function, path integral representation and equivalence for stochastic exclusive particle systems *Phys. Rev. E* **71** 026113

[353] Taillier J, Kurchan J and Lecomte V 2008 Mapping out-of-equilibrium into equilibrium in one-dimensional transport models *J. Phys. A: Math. Theor.* **41** 505001

[354] Kubo R, Matsu K and Kitahara K 1973 Fluctuation and relaxation of macrovariables *J. Stat. Phys.* **9** 51–96

[355] Gang H 1987 Stationary solution of master equations in the large-system-size limit *Phys. Rev. A* **36** 5782–90

[356] Dykman M I, Mori E, Ross J and Hunt P M 1994 Large fluctuations and optimal paths in chemical kinetics *J. Chem. Phys.* **100** 5735–50

[357] Kessler D A and Shnerb N M 2007 Extinction rates for fluctuation-induced metastabilities: a real-space WKB approach *J. Stat. Phys.* **127** 861–86

[358] Meerson B and Sasorov P V 2008 Noise-driven unlimited population growth *Phys. Rev. E* **78** 060103

[359] Escudero C and Kamenev A 2009 Switching rates of multistep reactions *Phys. Rev. E* **79** 041149

[360] Assaf M and Meerson B 2011 Extinction rates of established spatial populations *Phys. Rev. E* **81** 021116

[361] Ovaskainen O and Meerson B 2010 Stochastic models of population extinction *Trends Ecol. Evol.* **25** 643–52

[362] Meerson B and Sasorov P V 2011 Extinction rates of established spatial populations *Phys. Rev. E* **83** 011129

[363] Smith N R and Meerson B 2016 Extinction of oscillating populations *Phys. Rev. E* **93** 032109

[364] Assaf M and Meerson B 2016 WKB theory of large deviations in stochastic populations (arXiv:1612.01470)

[365] Turner J W and Malek-Mansour M 1978 On the absorbing zero boundary problem in birth and death processes *Physica A* **93** 517–25

[366] Assaf M and Meerson B 2007 Spectral theory of metastability and extinction in a branching-annihilation reaction *Phys. Rev. E* **75** 031122

[367] Assaf M, Meerson B and Sasorov P V 2010 Large fluctuations in stochastic population dynamics: momentum-space calculations *J. Stat. Mech.* **P07018

[368] Evans L C 2010 *Partial Differential Equations* (Graduate Studies in Mathematics vol 19) 2nd edn (Providence, RI: American Mathematical Society)

[369] Schwartz I B, Forgoston E, Bianco S and Shaw L B 2011 Converging towards the optimal path to extinction *J. R. Soc. Interface* **8** 1699–707

[370] Assaf M, Kamenev A and Meerson B 2009 Population extinction risk in the aftermath of a catastrophic event *Phys. Rev. E* **79** 011127

[371] Meerson B and Ovaskainen O 2013 Immigration-extinction dynamics of stochastic populations *Phys. Rev. E* **88** 012124

[372] Lan Y, Wolynes P G and Papania G A 2006 A variational approach to the stochastic aspects of cellular signal transduction *J. Chem. Phys.* **125** 124106

[373] Okhobo J 2008 Approximation scheme for master equations: A variational approach to multivariate case *J. Chem. Phys.* **129** 044108

[374] Drummond P D, Vaughan T G and Drummond A J 2010 Extinction times in autocatalytic systems *J. Phys. Chem. A* **114** 10481–91

[375] Drummond P D, Gardiner C W and Walls D F 1981 Quasiprobability methods for nonlinear chemical and optical systems *Phys. Rev. A* **24** 914–26

[376] Elderfield D 1985 Exact macroscopic dynamics in non-equilibrium chemical-systems *J. Phys. A: Math. Gen.* **18** 2049–59

[377] Droz M and McKane A 1994 Equivalence between Poisson representation and Fock space formalism for birth-death processes *J. Phys. A: Math. Gen.* **27** L467–74

[378] Drummond P D 2004 Gauge poisson representations for birth/death master equations *Eur. Phys. J. B* **38** 617–34

[379] Deloubrière R, Frachebourg L, Hilhorst H J and Kitahara K 2002 Imaginary noise and parity conservation in the reaction A + A → 0 *Physica A* **308** 135–47

[380] Petroyan K G and Hu C-K 2014 Nonequilibrium Lypunov function and a fluctuation relation for stochastic systems: Poisson-representation approach *Phys. Rev. E* **89** 042132

[381] Burnett J and Ford I J 2015 Coagulation kinetics beyond mean field theory using an optimised Poisson representation *J. Chem. Phys.* **142** 036101

[382] Swanson M S 1992 *Path Integrals and Quantum Processes* (San Diego, CA: Academic) ch 5.1

[383] Pawula R F 1967 Generalizations and extensions of the Fukker–Planck–Kolmogorov equations *IEEE Trans. Inf. Theory* **13** 33–41

[384] Pawula R F 1967 Approximation of the linear Boltzmann equation by the Fukker–Planck equation *Phys. Rev. E* **162** 186–8

[385] Thomas P, Fleck C, Grima R and Popovic N 2014 System size expansion using Feynman rules and diagrams *J. Phys. A: Math. Theor.* **47** 455007

[386] Andreanov A, Biroli G, Bouchaud J-P and Lefèvre A 2006 Field theories and exact stochastic equations for interacting particle systems *Phys. Rev. E* **74** 030101(R)

[387] Lefèvre A and Biroli G 2007 Dynamics of interacting particle systems: stochastic process and field theory *J. Stat. Mech.* **P07024**
[388] Itakura K, Ohkubo J and Sasa S-I 2010 Two Langevin equations in the Doi–Peliti formalism J. Phys. A: Math. Theor. 43 125001

[389] Janssen H K, Täuber U C and Frey E 1999 Exact results for the Kardar-Parisi-Zhang equation with spatially correlated noise Eur. Phys. J. B 9 491–511

[390] Frey E, Täuber U C and Janssen H K 1999 Scaling regimes and critical dimensions in the Kardar–Parisi–Zhang problem Europhys. Lett. 47 14–20

[391] Stratonovich R L 1957 On a method of calculating quantum distribution functions Sov. Phys.—Dokl. 2 416

[392] Hubbard J 1959 Calculation of partition functions Phys. Rev. Lett. 3 77–8

[393] Horsthemke W and Bach A 1975 Onsager–Machlup function for one dimensional nonlinear diffusion processes Z. Phys. B 22 189–92

[394] Dürr D and Bach A 1978 The Onsager–Machlup function as Lagrangian for the most probable path of a diffusion process Commun. Math. Phys. 60 153–70

[395] Ito H 1978 Probabilistic construction of Langrangean of diffusion process and its application Prog. Theor. Phys. 59 725–41

[396] Takahashi Y and Watanabe S 1981 The probability functionals (Onsager–Machlup functions) of diffusion processes Stochastic Integrals (Lecture Notes in Mathematics vol 851) ed D Williams (Berlin: Springer) pp 433–63

[397] Zinn-Justin J 2002 Quantum Field Theory and Critical Phenomena (International Series of Monographs on Physics vol 113) 4th edn (Oxford: Oxford University Press) ch 4.6

[398] Tang Y, Yuan R and Ao P 2014 Summing over trajectories of stochastic dynamics with multiplicative noise J. Chem. Phys. 141 044125

[399] Chaichian M and Demichev A 2001 Path Integrals in Physics: volume I: Stochastic Processes and Quantum Mechanics (Series in Mathematical and Computational Physics vol 1) (Bristol: IOP) ch 1

[400] Howard M J and Täuber U C 1997 ‘Real’ versus ‘imaginary’ noise in diffusion-limited reactions J. Phys. A: Math. Gen. 30 7721–31

[401] Hochberg D, Zorzano M-P and Morán F 2006 Complex noise in diffusion-limited reactions of replicating and competing species Phys. Rev. E 70 066109

[402] Torney D C and McConnell H M 1983 Diffusion-limited reactions in one dimension J. Phys. Chem. 87 1941–51

[403] Redner S 2001 A Guide to First-Passage Processes (Cambridge: Cambridge University Press)

[404] Jensen I 1994 Critical exponents for branching annihilating random walks with an even number of offspring Phys. Rev. E 50 3623–33

[405] Moshe M 1978 Recent developments in Reggeon field theory Phys. Rep. 37 255–345

[406] Frey E, Täuber U C and Schwabl F 1994 Crossover from self-similar to self-affine structures in percolation Europhys. Lett. 26 413–8

[407] Frey E, Täuber U C and Schwabl F 1994 Crossover from isotropic to directed percolation Phys. Rev. E 49 5058–72

[408] Janssen H K 1981 On the nonequilibrium phase transition in reaction–diffusion systems with an absorbing stationary state Z. Phys. B 42 151–4

[409] Grassberger P 1982 On phase transitions in Schlogl’s second model Z. Phys. B 47 365–74

[410] Canet L, Chaté H and Delamotte B 2011 General framework of the non-perturbative renormalization group for non-equilibrium steady states J. Phys. A: Math. Theor. 44 495001

[411] Zinn-Justin J 2007 Phase Transitions and Renormalization Group (Oxford Graduate Texts) (Oxford: Oxford University Press)

[412] Dickman R and Vidigal R 2003 Path integrals and perturbation theory for stochastic processes Braz. J. Phys. 33 73–93

[413] Yule G U 1925 A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. Phil. Trans. R. Soc. B 213 21–87

[414] DeGroot M H and Schervish M J 2012 Probability and Statistics 4th edn (Boston: Pearson) ch 5.5

[415] Wissel C 1979 Manifolds of equivalent path integral solutions of the Fokker–Planck equation Z. Phys. B 35 185–91

[416] Lee B P 1994 Critical behavior in non-equilibrium systems PhD Thesis University of California, Santa Barbara

[417] Mansour T and Schork M 2015 Commutation Relations, Normal Ordering and Stirling Numbers (Discrete Mathematics and Its Applications) (Boca Raton, FL: Chapman and Hall) p 410

[418] Trotter H F 1959 On the product of semi-groups of operators Proc. Am. Math. Soc. 10 545–51

[419] Touchard J 1939 Sur les cycles des substitutions Acta Math. 70 243–97

[420] Peccati G and Taqqu M S 2011 Wiener Chaos: Moments, Cumulants and Diagrams: a Survey with Computer Implementation (Bocconi & Springer Series) vol 1 (Berlin: Springer) ch 3.3

[421] Kang K and Redner S 1984 Fluctuation effects in Smoluchowski reaction-kinetics Phys. Rev. A 30 2833–6

[422] Benítez F and Wschebor N 2012 Branching-rate expansion around annihilating random walks Phys. Rev. E 86 010104

[423] Benítez F and Wschebor N 2013 Branching and annihilating random walks: exact results at low branching rate Phys. Rev. E 87 052132

[424] Rey P-A and Cardy J 1999 Asymptotic form of the approach to equilibrium in reversible recombination reactions J. Phys. A: Math. Gen. 32 1585–603

[425] Konkoli Z and Johannesson H 2000 Two-species reaction–diffusion system with equal diffusion constants: anomalous density decay at large times Phys. Rev. E 62 3276–80

[426] Blelloch N and Handelsman R A 1986 Asymptotic Expansions of Integrals (New York: Dover) ch 6.1

[427] Mugler A, Walczak A M and Wiggins C H 2009 Spectral radius in diffusion-limited reactions E. Phys. Rev. A 30 010104

[428] Ohkubo J 2010 Duality in interacting particle systems and boson representation J. Stat. Phys. 139 454–65

[429] Ohkubo J 2010 Solving partial differential equation via stochastic process Unconventional Computation vol 6079, ed C S Calude et al (Berlin: Springer) pp 105–14

[430] Ohkubo J 2013 Extended duality relations between birth–death processes and partial differential equations J. Phys. A: Math. Theor. 46 375004

[431] Ohkubo J 2015 Duality-based calculations for transition probabilities in stochastic chemical reactions 95 023304

[432] Benítez F, Duclut C, Chaté H, Delamotte B, Dornic I and Muñoz M A 2016 Langevin equations for reaction–diffusion processes Phys. Rev. Lett. 117 100601
Markus F Weber received his PhD in Physics at the Ludwig-Maximilians-Universität München in 2016. He conducted his doctoral research in the group of Erwin Frey. The present review arose out of this work. He is interested in the mathematical theory of stochastic processes and in the mathematical modelling of biological systems. His recent articles focus on bacterial range expansions, evolutionary game theory, and on a non-equilibrium condensation phenomenon of bosons.

Erwin Frey received his PhD in Physics at the Technische Universität München in 1989 and was Postdoctoral Fellow at Harvard University. Since 2001, he is Professor of Theoretical Physics at the Ludwig-Maximilians-Universität München. He is interested in how the interplay between stochastic fluctuations, interactions, and geometry shapes system-level properties and their functional characteristics. A central goal of his research is to identify and characterize universal action principles in biological systems.