THE UNIVERSITY *of* EDINBURGH

Exact steady states of minimal models of nonequilibrium statistical mechanics

Ivan Lobaskin

Doctor of Philosophy
The University of Edinburgh
2023

Abstract

Systems out of equilibrium with their environment are ubiquitous in nature. Of particular relevance to biological applications are models in which each microscopic component spontaneously generates its own motion. Known collectively as active matter, such models are natural effective descriptions of many biological systems, from subcellular motors to flocks of birds. One would like to understand such phenomena using the tools of statistical mechanics, yet the inherent nonequilibrium setting means that the most powerful classical results of that field cannot be applied. This circumstance has fuelled interest in exactly solvable models of active matter. The aim in studying such models is twofold. Firstly, as exactly solvable model are often minimal, it makes them good candidates as generic coarse-grained descriptions of real-world processes. Secondly, even if the model in question does not correspond directly to some situation realizable in experiment, its exact solution may suggest some general principles, which could also apply to more complex phenomena.

A typical tool to investigate the properties of a large system is to study the behaviour of a probe particle placed in such an environment. In this context, cases of interest are both an active particle in a passive environment or an active particle in an active environment. One model that has attracted much attention in this regard is the asymmetric simple exclusion process (ASEP), which is a prototypical minimal model of driven diffusive transport. In this thesis, I consider two variations of the ASEP on a ring geometry. The first is a system of symmetrically diffusing particles with one totally asymmetric (driven) defect particle. The second is a system of partially asymmetric particles, with one defect that may overtake the other particles. I analyze the steady states of these systems using two exact methods: the matrix product ansatz, and, for the second model the Bethe ansatz. This allows me to derive the exact density profiles and mean currents for these models, and, for the second model, the diffusion

constant. Moreover, I use the Yang-Baxter formalism to study the general class of two-species partially asymmetric processes with overtaking. This allows me to determine conditions under which such models can be solved using the Bethe ansatz.

Lay summary

Statistical mechanics is a field of physics that deals with large, complicated systems that consist of many small components. Many systems in biology fit this description at different levels, from sub-cellular motors to cells to organisms.

Traditionally, the approach taken by physicists to study statistical mechanics has been to assume that the individual components behave like little balls that obey Newton's laws of motion. The rules for each individual particle are simple, but the challenge is to figure out how they combine to produce effects on a large scale. Although this approach works well for gases and crystals, it has some major issues when applied to biological systems. One of the core reasons for this is that Newton's laws work the same way forwards and backwards in time. If one were to watch a movie of two billiard balls colliding and flying apart, it would be hard to say whether the movie was played forwards or backwards. However, the world of biology is full of processes in which there is an obvious temporal direction. Unfortunately, the physics of such irreversible processes is relatively undeveloped. We are not yet at the stage where we can simulate a lion on our computer at the cellular level. To get to that point, we must first understand much simpler processes.

In this thesis I study two very simple irreversible models. They consist of hard balls hopping on a one-dimensional line with a preference to hop in one direction. Although this might seem much too simple for real biological systems, it is a relatively good description of certain microscopic processes happening inside our cells every second. I study these models using advanced mathematical techniques, which gives some insight about how their motion at the microscopic scale manifest in complex large scale behaviour.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Parts of this work have been published in [1–4].

(*Ivan Lobaskin, 2023*)

Acknowledgements

I would like to thank my supervisor, Martin R Evans, and our collaborator, Kirone Mallick, for a very enjoyable and educational PhD experience. Their guidance has been instrumental in forming my views on science and its practice.

I am also grateful to my parents and my Edinburgh physics PhD peers, without whose support in all matters academic and non-academic, I could not have completed this work. I especially appreciated this during the period of lockdown due to the outbreak of the COVID-19 epidemic, which persisted through nigh on half of my time as a PhD student.

I acknowledge studentship funding from EPSRC under Grant No. EP/R513209/1.

Contents

List of Figures

List of Tables

Chapter 1

Background

1.1 Introduction

One of the most remarkable achievements of classical physics was the formulation of a general theory of equilibrium thermodynamics. Indeed, Einstein famously said of it, "It is the only physical theory of universal content, which I am convinced, that within the framework of applicability of its basic concepts will never be overthrown" [5]. Yet, despite much effort over the last decades, a similar general theory for systems far from equilibrium is still lacking.

A particularly important class of nonequilibrium processes are ones which are driven by some internal or external source and may thus be sustained in nonequilibrium steady states [6]. Such processes are abundant in the natural world, particularly in the realm of biology. However, to describe nonequilibrium steady states poses a formidable challenge. Even calculating something as fundamental as the probabilities of different states may be nontrivial, as one cannot utilise the standard Gibbs-Boltzmann measure of equilibrium statistical mechanics.

One of the most fruitful avenues of research in addressing these issues has been the study of minimal, low-dimensional models of nonequilibrium processes [7]. These models are often inspired by the motion of micro-organisms or biological molecular motors. Although the full mechanical and chemical context of such processes may be very complicated, much insight can be gained by taking a coarse-grained view. The moving object is represented as a point particle that

performs random hopping motion according to some simple dynamical rules. The philosophy behind this approach is that this extremely simplified model may be able to capture the qualitatively most important features of the motion of the process, and yet be simple enough to allow one to derive exact mathematical results, without having to rely on a pre-existing framework, such as the Gibbs-Boltzmann formalism.

Perhaps the most studied model of this type is the totally asymmetric simple exclusion process (TASEP) [7], which has been called the "Ising model of nonequilibrium statistical mechanics". It was initially inspired by the motion of ribosomes during RNA synthesis but has since then been used to model a wide range of processes, including DNA transcription and traffic flow. It consists of particles hopping randomly on a lattice in one direction with a hard-core exclusion interaction. Its intrinsic simplicity has lead to a multitude of exact results being obtained in the mathematical and physical literature. Furthermore, many generalisations have been considered, such as processes with partially asymmetric hopping and processes with multiple particle species.

In obtaining exact results pertaining to the steady-state properties of asymmetric exclusion processes, two powerful techniques have been especially successful. The first is the matrix product formalism [8]. This approach allows one to find the exact steady-state measures of many variants of asymmetric exclusion processes. It involves representing the state of the process as a product of non-commuting matrices that satisfy certain relations. If such matrices are found, this approach in principle allows one to compute any spatial correlation function.

The second is the Bethe ansatz [9]. This is in essence a technique for diagonalising certain linear operators. It was first introduced to study quantum spin chains but has since then found application in many other problems, including the time evolution of some asymmetric exclusion processes. A particularly useful corollary of this is that the Bethe ansatz may be used to calculate the exact current statistics in certain exclusion processes in the long time limit [10].

The results presented in this thesis will involve the application of the matrix product formalism and the Bethe ansatz to certain interesting special cases of asymmetric exclusion processes.

While the one-species TASEP, partially asymmetric SEP (PASEP) and symmetric SEP (SSEP) are all reasonably well understood, exact results for multi-species problems have mostly been restricted to certain special cases. As a simple limiting

case, one can consider a process with a single particle of one species. It may then be interpreted as a defect particle. Studying processes with defect particles thus allows one to build an understanding of the behaviour of multi-species processes.

Defect particles also allow one to study the response of the process they are embedded in to external perturbations. One may essentially view the defect particle as a controlled probe that is inserted into the underlying process. It is then interesting to see how the presence of the defect will affect the system at a macroscopic level. The case of a defect in a TASEP is relatively well understood [11, 12]. In this thesis we will instead investigate how driven defects affect the steady states of symmetric and partially asymmetric processes.

An important issue that arises with exact methods, such as the Bethe ansatz and the matrix product formalism, is the extent of their applicability. It is known in the literature that for exact methods to work for certain problems, a particular choice of parameters may be required. Thus even if a one-species model is solvable with a given method, the addition of a defect particle may make it intractable. In the context of the matrix product formalism, the solvability of two-species models has been studied in the literature. Yet for the Bethe ansatz, there is not a clear answer to which two-species models are solvable. The second main goal of this thesis is to address this issue.

In chapter 3, we study a single asymmetric particle in a bath of symmetrically diffusing particles using the matrix product formalism. This is an interesting setup to explore the question of the minimal ingredients required to drive a system out of equilibrium.

A natural generalisation one might consider is a system of partially asymmetric particles with a single defect particle that hops with different rates and may overtake other particles. Again, it is interesting to investigate whether a single particle can have macroscopic effects on the steady state (in this case nonequilibrium) of the system. In chapter 4, we use the matrix product formalism to analyse a special case of this general problem, in which the defect particle overtakes the other particles with its usual hopping rates.

In chapter 5, we take a general view of two-species partially asymmetric exclusion processes with overtaking and address the question of what conditions on the model parameters are required to ensure that they may be treated using the Bethe ansatz.

Finally, in chapter 6, we return to the model introduced in 4 and analyse it using the Bethe ansatz.

In total, this thesis investigates the effects of driven defect particles on the steady states of passive and driven lattice gases by means of exact mathematical methods. It also addresses the limits of applicability of the matrix product formalism and the Bethe ansatz in stochastic processes. This adds to the catalogue of exact results for nonequilibrium steady states.

1.1.1 Chapter outline

In the rest of this chapter, we discuss some background material that is necessary for an understanding of the remainder of this thesis. We begin by briefly outlining the traditional approach to statistical mechanics in section 1.2, which deals with steady states of the equilibrium type. Although this approach will not be used in the remainder of this thesis, it serves as a point of reference for comparison and contrast with the methods used for nonequilibrium systems. In section 1.3, we introduce the mathematical setting in which we will work in this thesis, namely continuous time Markov processes. We explain why this is a relevant framework for biological applications and outline a few basic results, which will be important in the later chapters. In section 1.4, we give a brief introduction to large deviation theory. Although we will not require anything beyond the most basic results in this thesis, we will see that the language of this field is a natural one for statistical mechanics, both in and out of equilibrium. In section 1.5, we introduce the key model, which will serve as the starting point of the models considered in this thesis: the asymmetric simple exclusion process. We define the model precisely and analyse the mean-field theory of this model to give some intuition for its phenomenology.

1.2 Equilibrium statistical mechanics

We now provide a brief overview of the traditional approach to statistical mechanics. Although these methods will not be used in the remainder of this thesis, they are included for two main reasons. The first reason is that understanding the limitations of these methods is important to motivate the methods that will be used. The second reason is that a lot of our subsequent

calculations will be informed by an understanding of these traditional methods. Indeed, although these methods are not applicable *verbatim*, we will see that they contain certain specific features that will reappear in subsequent calculations. We will highlight these features as they come up.

1.2.1 The meaning of equilibrium

Classical statistical mechanics describes steady states of systems in equilibrium. By equilibrium, one is to understand the condition of detailed balance, *i.e.* the vanishing of net fluxes between any two states (see also section 1.3.8). This is actually a very strong restriction and it is the reason why classical statistical mechanics is not applicable to the systems we will study in this thesis. The assumption of equilibrium is typically justified on the grounds of microscopic reversibility. As the dynamics of, say, molecules in a gas, are assumed to be time-reversible (as is typical for systems governed by Hamiltonian dynamics), it can be shown that the resulting steady state must be of the equilibrium type (see section1.3.8).

This also tells us how one breaks equilibrium; namely by making dynamics irreversible at a microscopic level. This can be achieved at a global level, in which case the system is called driven, or at the level of individual particles, in which case the system is called active. However, we remark that this distinction is not always clear-cut.

1.2.2 Microcanonical ensemble

For now, we consider systems in equilibrium. In this case, one can use the classic Gibbs-Boltzmann formalism (see any standard textbook on statistical physics, for instance [13–15]).

We consider some system with a state space X. This could be, for instance, the set of spin configurations in an Ising-type model, the set of positions and momenta of a classical gas, or the Hilbert space of a quantum system. The individual states $x \in X$ are called microstates. On a macroscopic scale, one observes not individual microstates x, but rather some coarse observables $Y(x)$, where we take Y to be an operator on X. By calling Y "coarse", we mean that there are generally many states x that give the same value of the observable $Y(x)$. Systems that

are sufficiently large are typically found in states in which the values of all coarse observables that correspond to conserved quantities of the underlying dynamics have some constant mean value. We call such states macrostates, where a given macrostate is characterised by the values of all relevant macroscopic observables. We will use the notation X_y to denote the macrostate in which the mean value of some specified observable is $Y(x) = y$.

The simplest case to consider is a closed system or microcanonical ensemble. This is a set of conditions under which all macroscopic variables (*e.g.* total energy E, particle number N, volume V) are held fixed. Under these circumstances, as we are not able to distinguish any state from any other using the coarse variables, the probability of all microstates may be taken to be the same. Then the statistical weight of any macrostate X_y is simply $|X_y|$.

We note that this typically scales exponentially with particle number N. As an example, for a set of N spin-1/2 spins, the total number of available states is 2^N. This exponential scaling of probabilities with respect to a large parameter actually means that one can incorporate this ensemble in a broader framework called large deviation theory, which does not rely on an assumption of equilibrium (see section1.4).

The exponential scaling motivates us to define the celebrated Boltzmann entropy,

$$S_y = \log |X_y|, \tag{1.1}$$

so that $S \sim N$ (note that we have chosen units in which $k_B = 1$). The entropy is an easier object to handle than the probability directly. By simple statistics it then follows that the most likely macrostate is the one which maximises the entropy S_y. Moreover, the probabilities of all states other than the most likely one will be suppressed exponentially, which means that we can really take the maximal entropy state as the only state the system is found in. This is essentially the second law of thermodynamics.

1.2.3 Canonical ensemble

In most situations, one encounters systems that are not closed but in thermal contact with their environment. This means that at least energy may be exchanged between the system and its environment. The set of conditions

under which all macroscopic variables are held fixed but energy is allowed to be exchanged is called the canonical ensemble.

Then the equilibrium configuration is found by maximising the total entropy of the system and its environment while varying how energy is distributed between them. This may be expressed as,

$$\frac{\mathrm{d}}{\mathrm{d}E}[S_{sys}(E) + S_{env}(E_{tot} - E)]_{E=E_{eq}} = 0. \tag{1.2}$$

We may take the environment to be much larger than the system, in which case $E_{tot} \gg E$. Then the second term can be approximated as

$$\frac{\mathrm{d}}{\mathrm{d}E}[S_{env}(E_{tot} - E)]_{E=E_{eq}} \approx - \left.\frac{\mathrm{d}S_{env}(E)}{\mathrm{d}E}\right|_{E=E_{tot}} = -\beta. \tag{1.3}$$

This is evidently a property of the environment, ignorant of what system it is in thermal contact with. We may then express the condition for equilibrium in terms of only the system as a minimisation of a potential defined as follows,

$$\beta F = \beta E - S_{sys}, \tag{1.4}$$

where β is taken to be a parameter. Systems with the same value of β will be in thermal equilibrium. For systems with different values of β, energy will typically flow from the system with the lower value of β to the one with the higher, until equilibrium is reached. Thus β may be identified with inverse temperature $\beta = 1/T$.

The potential F is called the Helmholtz free energy. Free energy minimisation for open systems is the equivalent of the principle of entropy maximisation for closed systems. The relation (1.4) between S and F is known as a Legendre transform. We will see how Legendre transforms appear naturally in large deviation theory in section1.4.

To determine the probabilities of individual states in the canonical ensemble, we may again consider the system and environment together as a closed system. Then from (1.1), the probability of some state X_y is given by

$$P(X_y) \propto \mathrm{e}^{S_{sys}(E_y) + S_{env}(E_{tot} - E_y)}. \tag{1.5}$$

Again, the environment may be taken to be much larger than the system, so

$S_{env} \gg S_{sys}$. We then Taylor expand the environment entropy,

$$S_{env}(E_{tot} - E_y) = S_{env}(E_{tot}) - E_y\beta + \ldots, \tag{1.6}$$

where we have used the definition (1.3). The first term does not depend on the state Y, so it may be ignored. We thus obtain,

$$P(X_y) = \frac{\mathrm{e}^{-\beta E_y}}{Z}, \tag{1.7}$$

where Z is a normalisation called a partition function defined as

$$Z = \sum_{x \in X} \mathrm{e}^{-\beta E_x}. \tag{1.8}$$

The expression (1.7) is known as the canonical or Gibbs-Boltzmann measure and it is the foundation of equilibrium statistical mechanics. It can be shown that the free energy F and the partition function Z are related as

$$-\beta F = \log Z. \tag{1.9}$$

We will present an elegant proof of this in section1.4. Note that it is reminiscent of (1.1).

It is remarkable that the expression for probabilities of states (1.7) can be written down so easily for any system in equilibrium. This is especially striking when juxtaposed with the calculations necessary to derive probabilities in nonequilibrium steady states, which will be presented in the next chapter.

1.2.4 Multiple conserved quantities

The Gibbs-Boltzmann measure (1.7) has a particularly simple form thanks to the hidden assumption that energy is the only relevant conserved quantity in the underlying dynamics. It is relatively straightforward to generalise to a situation with more conserved quantities. Supposing we have several conserved quantities E_i, then we can write the simply by introducing a "temperature" (conjugate variable) β_i for each conserved quantity,

$$P(x) = \frac{\exp[-\sum_i \beta_i E_i(x)]}{Z}. \tag{1.10}$$

A common example of such a construction is the grand canonical ensemble. There, particle number is included as an explicitly conserved quantity and the relevant conjugate variable is $\beta\mu$, where μ is the chemical potential.

Typical systems considered in classical thermodynamics tend to involve few conserved quantities, such as energy E, particle number N and volume V. However, one also encounters systems in which the number of conserved quantities is very large, which changes the thermodynamics considerably. Such systems are described using so-called generalised Gibbs ensembles [16, 17]. The idea of generalised Gibbs ensembles is particularly important for integrable systems, as these have a number of conserved quantities that scale with the number of particles.

We mention generalised Gibbs ensembles here for reference, but they will not be used in this thesis. Although we will consider systems that are integrable in some sense, they will not possess any physically meaningful conserved quantities except particle number.

1.3 Continuous time Markov jump processes

We now turn from a consideration of steady states to dynamics. Specifically, in this section we give an introduction to continuous time Markov jump processes, which will be the mathematical setting in this thesis. As this is somewhat different from the formalisms used in other domains of physics, we begin by motivating why this is a relevant approach to take, particularly if one is interested in describing biological phenomena. As with equilibrium statistical mechanics, the field of Markov processes is quite old, so there is an abundance of introductory literature, for instance [18–20].

1.3.1 A motivating example

Consider the following example. To fulfil various biological functions, our cells are constantly producing proteins in a collection of processes known as gene expression. As part of this procedure, information has to be read off from a

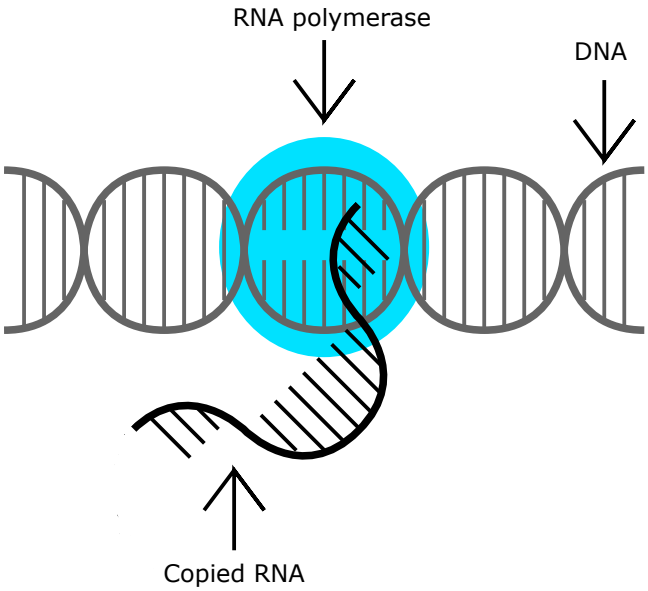

Figure 1.1 *Schematic illustration of DNA transcription.*

relevant piece of DNA to ensure the correct protein is produced [1]. This process is known as DNA transcription. It begins with a particle called an RNA polymerase (RNAP) attaching to a strand of DNA. The RNAP then begins to move along the strand, one base pair at a time, creating a copy of one half of the DNA. This is depicted in figure 1.1. The motion of the RNAP looks like a series of discrete steps of a fixed size [22]. Motion of a similar type has also proposed as a model for other molecular motors, such as kinesins and myosins [23–27].

This type of motion is fundamentally different from that of the constituents of a typical inanimate system, such as the molecules in a gas. It makes little sense to attempt to describe it using a Hamiltonian, since such fundamental concepts as momentum are ill-defined. Although one assumes that if one went down to a sufficiently small level, taking into accounts all the physical and chemical interactions, the underlying equations of motion would ultimately be time reversible and deterministic, such a description would provide little insight into the actual observed nature of this motion. Rather, what we observe is that the motion has the following properties: *(i)* it consists of discrete jumps, *(ii)* it occurs at random time intervals, *(iii)* it stops as soon as a jump has been performed (*i.e.* there is no inertia), *(iv)* the jumps can occur spontaneously,

[1] Indeed, this is stated in the so-called central dogma of molecular biology: "DNA makes RNA, RNA makes protein" [21].

without any external input.

Evidently, rather than using the framework of classical mechanics, it is better to model this as a stochastic process that jumps at random intervals between a set of discrete states. The simplest case assumption to make is that the times between jumps are exponentially distributed and that the probability of the subsequent state depends only on the current state (*i.e.* the system has no memory of past states). With these assumptions, this process falls into the class of continuous time Markov jump processes.

Remark on irreversibility We briefly comment on an apparent paradox that the astute reader may have noticed in the previous section. Namely, we claimed that the microscopic dynamics of the system in question are time-reversible, and yet the coarse-grained stochastic model generally is not. This is essentially the famous Loschmidt paradox, which may be stated as: "how can time-irreversible macroscopic laws be obtained from time-reversible microscopic ones?"

This is a very old and subtle question, which lies at the foundations of almost all statistical mechanics. Remarkably, there is still work being published to this day with new insights on this topic, for instance [28]. Here, we only provide a heuristic argument, which is sufficient for the purposes of this thesis.

The key point is that time-reversibility only holds for closed systems, *i.e.* situations in which external forces may be neglected. However, most systems are very far from this idealised scenario. Rather, they are constantly interacting with a highly complex environment. The environment is in fact generally so complex, that there is no hope of describing its exact microstate. Instead, one must resort to a statistical description in terms of some ensemble of states. Then the information of any effects the system has on the environment is lost, or in other words, entropy increases as a result of their interaction. Hence, in this framework, irreversibility of coarse-grained models emerges from our inability to precisely know the full microstate of the environment (really the whole universe) at any given time.

1.3.2 Types of Markov processes

Although they will not be examined in this thesis, we also mention other types of Markov processes, which are relevant for other situations. In total,

Markov processes are usually divided into four types, consisting of the possible combinations of continuous or discrete time steps, and a continuous or discrete state space. Discrete time steps can be relevant if the process evolves over some regular intervals of time, for instance if a new state is picked once a day, or if one is modelling a population that evolves in discrete generations.

A continuous state space is evidently natural for diffusion-like process. However, it can also be the case that a discrete description is better, even if the underlying state space is continuous. For instance the state space of a protein can be described by the possible values of the dihedral angles between each consecutive amino acid component. However, in practice, a protein spends much of its time in a small region of this space, corresponding to some folded or unfolded configuration, before transitioning quickly to another region, corresponding to a different configuration. It is thus better to take a coarse-grained description, in which each configuration is represented by a discrete state, with discontinuous jumps between these states [29].

1.3.3 Mathematical definition

We now define the continuous time Markov jump process more precisely. We denote the state of the system at as X_t. The easiest example is to think of this as the position of a diffusing particle. For more complicated cases, like multi-particle systems, it is to be interpreted as the configuration of the whole system. Markovian (memoryless) continuous time evolution means that the system remains in state X for a random time τ, which is distributed exponentially with some rate $r(X)$,

$$\tau \sim \mathrm{Exp}(r(X)), \tag{1.11}$$

before jumping to a new state X'. This can also be interpreted as while the system is in state X, in each infinitesimal time interval $\mathrm{d}t$, it has a probability of $r(X)\mathrm{d}t$ to leave that state. We denote the probability of choosing the state X' by $p(X, X')$. These are assumed to depend only on the current state, X, and not on the history of the evolution of the process.

The assumptions we have made are quite strong, as they presuppose that there is no temporal correlations in the times between jumps or the probability of jumping between certain states. The justification for using it is that it greatly

simplifies the mathematical formalism and often produces results that agree with experimental data. Thus it is assumed by default, unless there is a good reason not to.

A common physical justification for these assumptions is the separation in time scales between microscopic and mesoscopic degrees of freedom [19, 30]. Taking the classic example of a Brownian colloid in a fluid, the velocity of the colloid is much smaller than the velocity of the fluid molecules. The fluid molecules can therefore essentially be assumed to thermalise instantaneously around the colloid, so the states of the fluid the colloid interacts with are effectively taken randomly from a canonical measure.

1.3.4 Master equation

With this definition, the whole dynamics can be described by the set of rates at which the system transitions from a state X to another state X'. We denote this as

$$k(X, X') = r(X)p(X, X'). \tag{1.12}$$

Thus $k(X, X')$ may also be seen as the rate per unit time at which the system transitions from state X to X'. Note that this implies the key property

$$\sum_{X' \neq X} k(X, X') = r(X). \tag{1.13}$$

Then the evolution of the probability of being in state X at time t is given by the master equation

$$\frac{\mathrm{d}P_t(X)}{\mathrm{d}t} = \sum_{X' \neq X} [-k(X, X')P_t(X) + k(X', X)P_t(X')]. \tag{1.14}$$

The first term in the sum corresponds to the flux of probability out of X into X' and the second term corresponds to the flux into X from X'.

We can write (1.14) in vector form by combining the transition rates into a matrix,

$$\frac{\mathrm{d}P_t(X)}{\mathrm{d}t} = \sum_{X'} \mathcal{M}(X, X')P_t(X'), \tag{1.15}$$

where $\mathcal{M}$ is a matrix, defined as

$$\mathcal{M}(X, X') = \begin{cases} k(X', X), & X' \neq X \\ -r(X), & X' = X \end{cases}. \tag{1.16}$$

$\mathcal{M}$ is interchangeably called a transition rate matrix, a Markov operator [2] or Liouville operator. From (1.13), we have the fundamental property of transition rate matrices, namely that their columns sum to 0,

$$\sum_{X'} \mathcal{M}(X', X) = \sum_{X' \neq X} k(X, X') - r(X) = 0. \tag{1.17}$$

This implies the conservation of total probability. To see this, we sum the right hand side of (1.15) over all X to obtain,

$$\sum_{X} \sum_{X'} \mathcal{M}(X, X') P_t(X') = \sum_{X'} \left(\sum_{X} \mathcal{M}(X, X') \right) P_t(X') = 0. \tag{1.18}$$

We remark that (1.15) has mathematically the same form as the Schrödinger equation with the replacement $\mathcal{M} \to (\mathrm{i}\hbar)^{-1}\hat{H}$. The main difference is that the Schrödinger equation describes the evolution of complex amplitudes, while the master equation describes the evolution of probabilities. Thus, unitary time evolution in the Schrödinger equation instead becomes probability-preserving time evolution in the master equation.

1.3.5 Spectral analysis

We can arrange the probabilities of all states at time t in a single vector $\mathbf{P}_t$. Then the solution of the master equation (1.15) can be formally written as

$$\mathbf{P}_t = \mathrm{e}^{\mathcal{M}t}\mathbf{P}_0, \tag{1.19}$$

with some initial condition $\mathbf{P}_0$ and where the exponential is to be understood as the matrix exponential,

$$\mathrm{e}^{A} = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{1.20}$$

[2] This term is sometimes reserved for the discrete time case.

Using the eigendecomposition of $\mathcal{M}$, one can write this more explicitly as

$$\mathbf{P}_t = \sum_i \mathrm{e}^{\lambda_i t} \Psi_i^{(R)} (\Psi_i^{(L)} \cdot \mathbf{P}_0), \tag{1.21}$$

where λ_i, $\Psi_i^{(R)}$, $\Psi_i^{(L)}$ are the eigenvalues of $\mathcal{M}$ and the corresponding right and left eigenvectors. We remark that the eigenvectors $\Psi_i^{(R)}$, $\Psi_i^{(L)}$ are generally not vectors of probabilities themselves.

From (1.21), it is evident that the spectrum of $\mathcal{M}$ is key for understanding the dynamics of probabilities evolving according to the master equation. At long times, modes corresponding to the right eigenvectors $\Psi_i^{(R)}$ will grow, maintain a constant magnitude or decay, depending on whether $\Re(\lambda_i)$ is positive, 0 or negative respectively. Actually, it can be shown using the Perron-Frobenius theorem that for all eigenvalues of a transition rate matrix we have $\Re(\lambda_i) \leq 0$ [19]. Indeed, it is intuitively obvious from the interpretation of $\mathcal{M}$ as the generator of the evolution of a probability distribution, that as probability cannot grow indefinitely, there cannot be eigenvalues with positive real part. This means that all modes will generally decay with, perhaps, some transient oscillation, except those for which $\Re(\lambda_i) = 0$.

1.3.6 Steady states

The eigenvectors of the 0 eigenvalues evidently play an important role. It is easy to see that there is always at least one 0 eigenvalue: the columns of $\mathcal{M}$ add to 0, so the row vector with all entries equal to 1 is always a left eigenvector with eigenvalue 0.

From the master equation (1.15), we see that the right eigenvectors of the 0 eigenvalues represent stationary configurations of the system. They satisfy the so-called global balance condition,

$$\sum_{X' \neq X} [-k(X, X')P_\infty(X) + k(X', X)P_\infty(X')] = 0. \tag{1.22}$$

Although this is always true for steady states of Markov processes, it is usually too general to be useful. We will show in section1.3.8 that for equilibrium steady states, a much stronger condition called detailed balance is satisfied. Generally, even for nonequilibrium steady states, one generally looks for some sort of local

balance condition.

As the right eigenvectors of zero eigenvalues represent steady-state distributions, one must be able to write them as valid probability vectors (*i.e.* with only non-negative entries that sum to 1). Conversely, it can be shown that for all right eigenvectors with non-zero eigenvalues, the sum of their entries is 0. This is another manifestation of the conservation of total probability.

1.3.7 Ergodicity

In general, it is possible that the 0 eigenvalue is degenerate. For example, this is the case if there are multiple absorbing states (*i.e.* states out of which all rates are 0). An important special case, however, is when there is only one eigenvalue with 0 real part. Then any initial condition will eventually converge to the eigenvector of the 0 eigenvalue. Markov processes for which this is true are called ergodic.

A corollary is that for an ergodic process, if a trajectory is sampled over a long enough time, the empirical measure will eventually converge to the unique stationary eigenvector. This corresponds to the classic notion of ergodicity: averaging over a trajectory should produce the same result as averaging over an ensemble. However, we point out that the definition via Markov processes is more general, as the ergodic hypothesis comes out as a corollary.

1.3.8 Detailed balance

Given an ergodic Markov process, a natural question is whether its steady state is of the equilibrium or nonequilibrium type. The answer to this is that a steady state can be called equilibrium if and only if it satisfies detailed balance. This is defined as the condition that the net probability fluxes between any two states should vanish in the steady state. If we denote the steady state probabilities as P_∞, this is expressed mathematically as

$$k(X, X')P_\infty(X) = k(X', X)P_\infty(X'), \qquad \forall\, X, X'. \tag{1.23}$$

To see why this is equivalent to thermodynamic equilibrium, recall the Arrhenius equation for the activation rate of a thermodynamic process,

$$k = A\mathrm{e}^{-\beta\delta E}, \tag{1.24}$$

where k is the rate, A is a constant and δE is the activation energy or potential barrier separating the initial and final states. Now consider two states with energies E_1, E_2. We take $E_1 > E_2$ and let δE be the height of a barrier to be crossed going from state 1 to state 2. Then going from state 2 to state 1, the barrier heights is $E_1 - E_2 + \delta E$. According to Arrhenius's law, the ratio of rates of going in either direction will be given by,

$$\frac{k_{2\to1}}{k_{1\to2}} = \frac{A\mathrm{e}^{-\beta(E_1-E_2+\delta E)}}{A\mathrm{e}^{-\beta\delta E}} = \mathrm{e}^{-\beta(E_1-E_2)}. \tag{1.25}$$

At the same time, according to Gibbs-Boltzmann theory, the ratio of probabilities, given by the canonical measure, of finding the system in those states is

$$\frac{P(1)}{P(2)} = \frac{Z^{-1}\mathrm{e}^{-\beta E_1}}{Z^{-1}\mathrm{e}^{-\beta E_2}} = \mathrm{e}^{-\beta(E_1-E_2)}. \tag{1.26}$$

Hence we have $k_{2\to1}/k_{1\to2} = P(1)/P(2)$, which is the definition (1.23).

Kolmogorov's criterion

Detailed balance may alternatively be defined by the so-called Kolmogorov's criterion [30]. It states that for detailed balance to hold, the product of rates for any closed cycle $X_1 \to X_2 \to \cdots \to X_N \to X_1$ must equal the product of rates of the reverse cycle $X_1 \to X_N \to \cdots \to X_2 \to X_1$,

$$k(X_1, X_2)\ldots k(X_N, X_1) = k(X_1, X_N)\ldots k(X_2, X_1). \tag{1.27}$$

This definition has the advantage that it does not require the construction of the steady state and thus diagonalisation of $\mathcal{M}$, though we remark that enumerating all closed cycles in a large state space can be a harder task than simply calculating the steady state.

We can see that (1.23) implies Kolmogorov's criterion as follows,

$$\begin{aligned}\frac{k(X_1,X_2)\dots k(X_N,X_1)}{k(X_1,X_N)\dots k(X_2,X_1)} &= \frac{k(X_1,X_2)}{k(X_2,X_1)}\frac{k(X_2,X_3)}{k(X_3,X_2)}\cdots\frac{k(X_{N-1},X_N)}{k(X_N,X_{N-1})}\frac{k(X_N,X_1)}{k(X_1,X_N)} \\ &= \frac{P(X_2)}{P(X_1)}\frac{P(X_3)}{P(X_2)}\cdots\frac{P(X_N)}{P(X_{N-1})}\frac{P(X_1)}{P(X_N)} = 1. \qquad (1.28)\end{aligned}$$

The inverse implication can also be shown but it is omitted here as it is more involved.

The contrapositive of Kolmogorov's criterion states that if detailed balance is broken, there must be some cycles of states which carry net currents. Indeed, currents are used both as signatures and characteristic macroscopic descriptors of nonequilibrium steady states [30].

1.4 Large deviation theory

One of the key goals of nonequilibrium statistical mechanics is to understand which general laws or principles that are well-established in the equilibrium context can be extended to nonequilibrium. In pursuit of this goal, one of the most useful frameworks that has emerged is large deviation theory [31].

The basis of large deviation theory is the observation that many distributions found in statistical physics attain an exponential form in some asymptotic limit. Consider some random variable X_N, which depends on the parameter N, and we are interested in the large N limit. In statistical mechanics, the role of N is usually played by particle number or time. A typical case is for X_N to be an extensive variable, so that $\langle X_N \rangle \sim N$. For example, one would expect such a scaling if X_N is the sum of contributions from N components. We will assume this is the case for simplicity, though this is not necessary for the theory. Then in the large N limit, it is often found that the distribution of X_N takes the form,

$$P_N\left(\frac{X_N}{N} = x\right) \sim \mathrm{e}^{-NI(x)}, \qquad N \to \infty, \qquad (1.29)$$

where $I(x)$ is a function (that does not depend on the parameter N) that dictates the rate at which probabilities decay as N is made large. For this reason, it is called the rate function. More rigorously, $P_N(x)$ is said to satisfy a large deviation

principle with rate function $I(x)$ if the following limit exists:

$$\lim_{N\to\infty} -\frac{\log P_N(x)}{N} = I(x). \tag{1.30}$$

We see from (1.29) that as N becomes large, the probabilities of most values x will become exponentially suppressed and $P_N(x)$ will become dominated by the maxima of $I(x)$.

This can be seen as a generalisation of the law of large numbers and the central limit theorem. Essentially, the central limit theorem is what one obtains if one expands a rate function $I(x)$ as a Taylor series up to quadratic order, giving a Gaussian distribution. The rate function $I(x)$ provides information not only on how the the distribution $P_N(x)$ behaves close to its mean, but also how it behaves for large deviations from the mean $|X_N - \langle X_N\rangle| \sim O(N)$ [3].

Example: Poisson process

As a familiar example, we consider a Poisson process with rate α. This is defined as a sequence of events, where the time between one event and the next is a random variable taken from an exponential distribution with rate α. Suppose this process has been going on for some time t, up to which time there have been $Y(t)$ events. The probability that the number of events $Y(t)$ equals some value k is given by the Poisson distribution with parameter αt. We wish to investigate the behaviour of $P_t(Y(t)/t = v)$ in the limit $t \to \infty$. From the Poisson distribution, we have the exact form,

$$P_t\left(\frac{Y(t)}{t} = v\right) = \frac{(\alpha t)^{vt}}{(vt)!}\mathrm{e}^{-\alpha t} = \exp\left[-\alpha t + vt\log(\alpha t) - \log((vt)!)\right]. \tag{1.31}$$

Then using the Stirling approximation, $\log((vt)!) \approx (vt)\log(vt) - vt$, after a little simplification, we obtain

$$P_t\left(\frac{Y(t)}{t} = v\right) \approx \exp\left[-t(\alpha - v + v\log v - v\log\alpha)\right]. \tag{1.32}$$

[3] This is why this field is called *large* deviation theory.

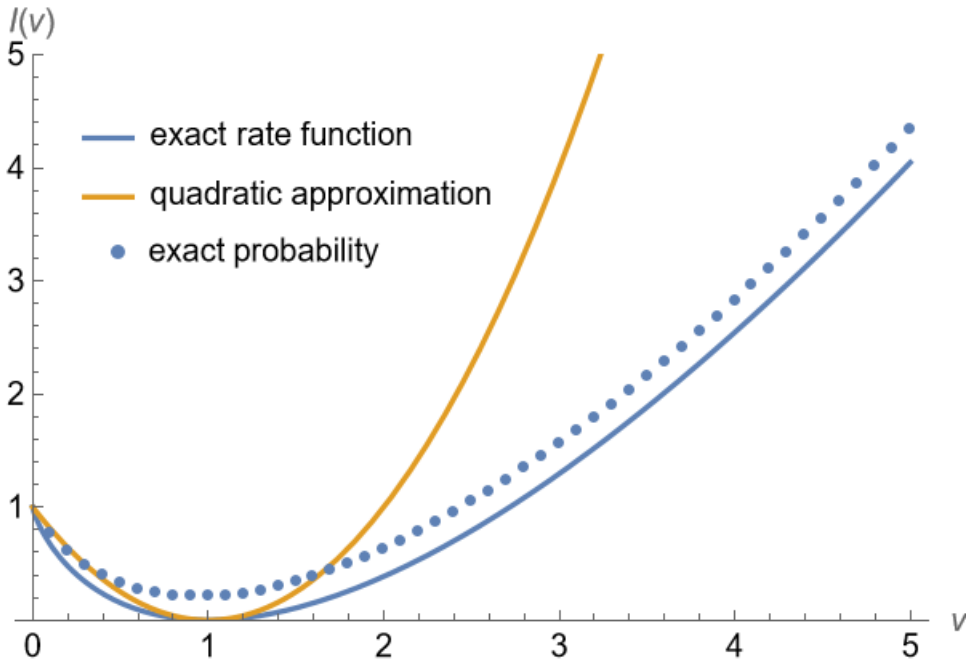

Figure 1.2 *Rate function for Poisson process* (1.33)*, for* $\alpha = 1$*. The quadratic approximation (i.e. the Taylor series expansion of the rate function up to second order around the minimum) and the exact values of* $-1/t \log P_t(v)$ *(for* $t = 10$*) are also shown. The quadratic approximation and the exact rate function agree close to the minimum (at* $v = 1$ *here), but for large deviations, the exact rate function gives a much better estimate. Note that there is a roughly constant difference between the rate function and the exact finite* t *values due to sub-exponentially scaling prefactors in the latter.*

Hence we see that the Poisson distribution satisfies a large deviation principle with rate function

$$I(v) = \alpha - v + v \log v - v \log \alpha. \tag{1.33}$$

This is illustrated in figure 1.2.

1.4.1 Scaled cumulant generating function

One of the most important results from large deviation theory is the relationship between the rate function and the scaled cumulant generating function. This is expressed mathematically in the Gärtner-Ellis theorem [32, 33]. The importance of this result comes from the fact that it provides a fairly general method for computing rate functions. Indeed, other than for certain simple cases, direct computation of the rate function, as was done in the previous section, is very

challenging.

We first consider the moment generating function $\langle e^{\gamma X_N}\rangle$, where γ is a variable conjugate to X_N. Note that by definition, we can calculate moments of X_N by taking formal derivatives of the moment generating function at $\gamma = 0$,

$$\langle X_N^n\rangle = \left.\frac{d^n\langle e^{\gamma X_N}\rangle}{d\gamma^n}\right|_{\gamma=0}. \tag{1.34}$$

Suppose X_N takes values in some domain Ω. Then using the exponential form (1.29), we have

$$\langle e^{\gamma X_N}\rangle \approx \int_\Omega e^{N[\gamma x - I(x)]}dx. \tag{1.35}$$

As N is assumed to be large, this integral will be dominated by the values close to the maximum of $\gamma x - I(x)$. We may thus use the saddle point approximation [34] to write,

$$\langle e^{\gamma X_N}\rangle \approx \max_{x\in\Omega}\{e^{N[\gamma x - I(x)]}\}. \tag{1.36}$$

Defining the scaled cumulant generating function of X_N,

$$\lambda_N(\gamma) = \frac{\log\langle e^{\gamma X_N}\rangle}{N}, \tag{1.37}$$

we see that it will satisfy the limit

$$\lambda(\gamma) = \lim_{N\to\infty}\lambda_N(\gamma) = \max_{x\in\Omega}\{\gamma x - I(x)\}. \tag{1.38}$$

Thus in the asymptotic limit $N \to \infty$, the rate function $I(x)$ and the scaled cumulant generating function $\lambda(\gamma)$ are related by a Legendre transform. The formal statement of this fact is given by the Gärtner-Ellis theorem. For practical purposes, as is sufficient from the point of view of physics, it asserts that if the scaled cumulant generating function $\lambda_N(\gamma)$ converges to some differentiable function $\lambda(\gamma)$ in the limit $N \to \infty$, then the underlying distribution satisfies a large deviation principle, with a rate function given by the (inverse) Legendre transform,

$$I(x) = \max_\gamma\{\gamma x - \lambda(\gamma)\}. \tag{1.39}$$

Therefore, one way to compute a rate function $I(x)$ is to first find the scaled

cumulant generating function $\lambda(\gamma)$ and perform a Legendre transform. The latter is often easier than calculating $I(x)$ directly.

Example: Skellam process

As an example to illustrate this approach, we consider the difference of two Poisson random variables. This is known as a Skellam process [35]. It arises naturally in diffusion on a lattice as the net displacement of a particle performing biased diffusion.

Specifically, consider a particle on an infinite lattice hopping to the right with rate α and to the left with rate β. Let $Y_R(t)$ and $Y_L(t)$ be the number of hops it has performed to the right and left respectively up to time t. Each of those are distributed according to Poisson distributions with means αt and βt respectively. However, now we are interested in the statistics of the net displacement, which is given by $Y(t) = Y_R(t) - Y_L(t)$.

One approach would be to write out the probability explicitly. The probability that $Y(t) = vt$ is given by the probability that $Y_R(t) = vt + k$ and $Y_L(t) = k$, where k can be any non-negative integer. This gives,

$$P_t\left(\frac{Y(t)}{t} = v\right) = \sum_{k=0}^{\infty} \frac{(\alpha t)^{k+vt}}{(k+vt)!}\mathrm{e}^{-\alpha t}\frac{(\beta t)^k}{k!}\mathrm{e}^{-\beta t}. \tag{1.40}$$

Then to derive the rate function, one would have to perform this summation and extract the asymptotics. The Skellam distribution can actually be written in terms of modified Bessel functions, but as we will show now, the Gärtner-Ellis theorem allows a much simpler approach.

To apply the Gärtner-Ellis theorem, we use the result that the cumulant generating function of two independent random variables is the sum of the cumulant generating functions of the variables. Then firstly, to obtain the cumulant generating function of a Poisson process, we use (1.33) and (1.38). From $I'(v_{\max}) = \gamma$, we find that the maximal value of v is given by,

$$v_{\max} = \alpha \mathrm{e}^{\gamma}. \tag{1.41}$$

Then the scaled cumulant generating function of Y_R is given by,

$$\lambda_R(\gamma) = \alpha(\mathrm{e}^{\gamma} - 1). \tag{1.42}$$

Similarly, we see that the scaled cumulant generating function of the variable $-Y_L$ is given by,

$$\lambda_L(\gamma) = \beta(\mathrm{e}^{-\gamma} - 1). \tag{1.43}$$

Then adding (1.42) and (1.43), we get that the scaled cumulant generating function of a Skellam process with rates α, β can be written as,

$$\lambda(\gamma) = (1 - \mathrm{e}^{-\gamma})(\alpha\mathrm{e}^{\gamma} - \beta). \tag{1.44}$$

Now to find the rate function, we need to perform a Legendre transform with respect to γ. The maximal value of γ, defined by $\lambda'(\gamma_{\max}) = v$, is given by

$$\mathrm{e}^{\gamma_{\max}} = \frac{v + \sqrt{v^2 + 4\alpha\beta}}{2\alpha}. \tag{1.45}$$

The equation for $\mathrm{e}^{\gamma_{\max}}$ is actually quadratic and contains another root but this can be discounted as it is negative. Using this, after a little algebra, it is found,

$$I(v) = v \log\left(\frac{v + \sqrt{v^2 + 4\alpha\beta}}{2\alpha}\right) + \alpha + \beta - \sqrt{v^2 + 4\alpha\beta}. \tag{1.46}$$

This is shown in figure 1.3. Thus we were able to obtain the rate function using just some basic algebra. In many cases, the Gärtner-Ellis theorem is the only tractable way to calculate the rate function.

1.4.2 Example: equilibrium statistical mechanics

We now show how equilibrium statistical mechanics can be expressed in the language of large deviation theory. Specifically, we will see that entropy plays the role of a rate function and the Legendre transform corresponds moving between different ensembles. A detailed overview is given in [36].

In a microcanonical ensembles, all microstates are assumed to be equiprobable. Then the probability of a macrostate y is proportional to the number of microstates which verify it, $|X_y|$. This can be written in terms of the Boltzmann

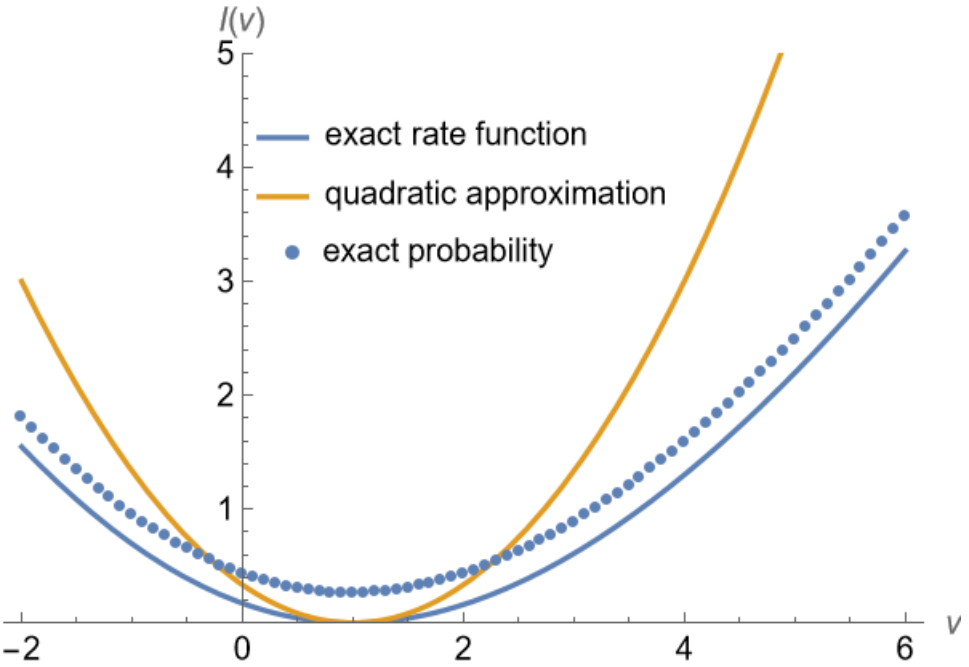

Figure 1.3 *Rate function for Skellam process* (1.46)*, for* $\alpha = 2$*,* $\beta = 1$*. The quadratic approximation (i.e. the Taylor series expansion of the rate function up to second order around the minimum) and the exact values of* $-1/t \log P_t(v)$ *(for* $t = 10$*) are also shown. The quadratic approximation and the exact rate function agree close to the minimum (at* $v = 1$ *here), but for large deviations, the exact rate function gives a much better estimate. Note that there is a roughly constant difference between the rate function and the exact finite* t *values due to sub-exponentially scaling prefactors in the latter.*

entropy S, (1.1), where we have chosen units in which $k_B = 1$,

$$P(y) \propto |X_y| = \mathrm{e}^{S(y)}. \tag{1.47}$$

Entropy is an extensive quantity, meaning it scales as $S \sim N$, where N is the number of particles in the system. Hence, in the thermodynamic limit, the microcanonical measure naturally follows a large deviation principle,

$$P(y) \sim \mathrm{e}^{Ns(y)}, \tag{1.48}$$

in which the rate function is the negative reduced entropy $I(y) = -S(y)/N$.

Now we may define the generating function of energy with some conjugate parameter $-\beta$,

$$Z = \langle \mathrm{e}^{-\beta E} \rangle = \int \mathrm{d}E \, \frac{|X_E|}{|X|} \mathrm{e}^{-\beta E}. \tag{1.49}$$

We have suggestively named it Z, as it can be identified with the canonical partition function. Then using the saddle point approximation, we obtain

$$N^{-1} \log Z \approx \min_{\epsilon} \{\beta\epsilon - s(\epsilon)\}, \tag{1.50}$$

where $\epsilon = E/N$ is the energy per particle. The expression on the right hand side can be identified with the reduced free energy, $\beta f = \beta F/N$, which proves the relation (1.9). We see that the reduced free energy and entropy are related by Legendre transforms,

$$\beta f(\beta) = \min_{\epsilon} \{\beta\epsilon - s(\epsilon)\}, \qquad s(\epsilon) = \min_{\beta} \{\beta\epsilon - \beta f(\beta)\}. \tag{1.51}$$

The min appears instead of the max, because entropy is defined with the opposite sign compared to the rate function.

Large deviation theory thus provides a different perspective on equilibrium statistical mechanics. Namely, the free energy plays the role of a scaled cumulant generating function and temperature enters merely as a variable conjugate to energy. This is a valuable insight as such relations are more general than the equilibrium setting and we will see similar structures appear in nonequilibrium problems.

1.4.3 Large deviation theory in nonequilibrium statistical physics

Importantly, large deviation theory has also been applied in nonequilibrium statistical mechanics. One often finds that measures follow a large deviation principle, with either particle number N or time t acting as the large parameter.

We have already seen that the displacement statistics of a particle undergoing biased diffusion satisfy a large deviation principle, where the large parameter is time. The same is true of the collective displacement of many interacting particles, which will be studied in chapter 6.

A widely studied example of large deviation-type relations in nonequilibrium systems are fluctuation relations. Consider some time-averaged observable $A_t = t^{-1}\int \mathrm{d}t' A(t')$. Then a fluctuation theorem is established if one can prove a relation of the form,

$$\frac{P(A_\tau = a)}{P(A_\tau = -a)} \sim \mathrm{e}^{tI(a)}, \tag{1.52}$$

for some rate function $I(a)$. Such results are particularly interesting from the point of view of nonequilibrium physics as they provide a direct way to analyse the degree of irreversibility of processes. Fluctuation theorems have been proved for such observables as entropy production rate [37, 38] and work [39]. See [40] for a general review.

The examples described above are actually quite simple applications of large deviation theory, as we are examining the statistics of a single scalar observable. However, one could in principle also examine the whole dynamic measure of a stochastic process. This requires to generalise the framework we have shown essentially to make it describe vectors or tensors of observables (such as the vector of densities on each lattice site in a lattice model), rather than a single observable. This is known as large deviation theory at level 2 and has found much application in the field of nonequilibrium systems.

For instance, it has been shown that trajectories of stochastic PDEs satisfy large deviation principles in the weak noise limit [41, 42]. Similarly, large deviation theory has been used to to study stochastic lattice gas models (these models will be discussed more in section1.5) [43–45].

Another application of large deviation theory for trajectories of nonequilibrium systems that has seen remarkable success is that of macroscopic fluctuation theory [46]. This will be discussed more in section1.5.6.

1.5 Asymmetric simple exclusion processes

We turn to a specific class of Markov jump processes that will be the focus of the remainder of this thesis. One of the most extensively studied models of driven diffusive motion is the one-dimensional asymmetric simple exclusion process (ASEP). This is a minimal model of single file diffusion in a one-dimensional or quasi-one-dimensional environment. Due to the minimalist nature of its definition, it naturally lends itself as a starting point for modelling a large variety of systems. It was first introduced to model the aforementioned hopping of various molecular motors (see section1.3.1) [47–49], but has since also been used as an effective model for colloids in a narrow channel [50, 51] and cars on a street [52–55]. Some general reviews are [7, 8]

1.5.1 Model definition

The system we wish to model consists of a number of components (RNAPs, colloids, cars) that are moving effectively in one dimension with a preferential direction of motion and a hardcore exclusion interaction. We will model these components as particles hopping on a lattice. To model the asymmetry of the hopping, we give the particles different hopping rates to the right and to the left. By rescaling time, we may set the right hopping rate to 1. The left hopping rate takes some value x. We will take $x < 1$ throughout this thesis without loss of generality.

To enforce hardcore exclusion, when a particle attempts a hop onto a site that is already occupied, this hop is rejected. Thus the maximum occupation of all sites is 1. This is illustrated in figure 1.4. We also introduce the following notation, which will be much used hereafter. We represent empty sites (holes) by 0 and particles by 1. Then we can represent the process we just described by

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \ . \tag{1.53}$$

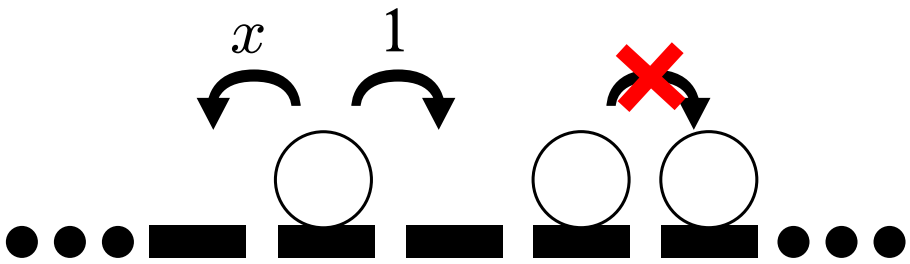

Figure 1.4 *Illustration of asymmetric simple exclusion process with hopping rate 1 to the right and x to the left.*

1.5.2 Boundary conditions

Next, we need to examine what types of boundary conditions we might wish to consider for modelling different physical scenarios. This makes a big difference both to the phenomenology one observes and to what kinds of methods may be used to obtain exact results.

Periodic boundary conditions

Perhaps the simplest boundary conditions to examine are periodic ones. One simply restricts the lattice to some finite number of sites L and equates site 1 with $L + 1$. In other words, the process is set on a ring. The evolution of this process is described by a master equation with a finite-dimensional transition rate matrix, which reaches a nonequilibrium steady state. Arguably periodic boundary conditions are not as relevant in nature as open boundary conditions. However, they help us to understand the behaviour of the process, as they can make the problem easier to solve exactly. All the problems we examine in this thesis will use periodic boundary conditions.

Infinite lattice

If we are modelling cars on a very long street, one might consider the process on an infinite lattice. This choice of boundary conditions has been used a lot in the literature, particularly on the mathematical side (for instance [56–59]). As the system is infinite, this process is described by an infinite transition rate matrix and does not reach a steady state, unless the initial conditions already constitute an invariant measure. Evidently then for this process, one needs to also make a judicious choice of initial conditions.

Although one might be tempted to interpret this as a limiting case of the periodic boundary conditions in the limit $L \to \infty$, the order of taking the system size and time to infinity makes a difference. Indeed, as we have already noted, the process on a ring typically reaches a steady state, whereas the process on the infinite lattice does not.

Open boundary conditions

Lastly, one might choose open boundary conditions. This means that the process is set on a finite lattice with L sites. Moreover, one imposes special rates on the first and last sites. Typically, one injects particle to site 1 with some rate a and removes particles from site L with some rate b. Indeed, this is the natural choice of boundary conditions for modelling RNAP hopping, as this takes place on a finite segment with some entry and exit rates.

This process is again described by a finite transition rate matrix and reaches a nonequilibrium steady state. However, the state space of this process is bigger than that of one on a periodic lattice of the same size, as open boundary conditions allow for the particle number to vary.

The open boundary problem has a long history in the physics and mathematics literature (for instance [60–64]). One important aspect to note is that the choice of entry and exit rates makes a big difference on whether the problem can be solved exactly using certain methods.

1.5.3 Two-species problems

A simple way to generalise the basic ASEP is to consider two different types of particles. The most general way to do this is to give the second species completely independent hopping rates p and q and to allow particles of different species to overtake each other with some rates α and β. This process can be represented as

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{q}{\overset{p}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 21 \,, \tag{1.54}$$

where 0 and 1 denote holes and particles of species 1, as before, and we have introduced 2 to denote particles of species 2. There has been much literature devoted to special cases of this process. We now describe a few of the most

studied ones.

Second-class particle

One of the earliest and most studied special cases is that of a "second-class" particle. This is obtained by setting $p = \alpha = 1$ and $q = \beta = x$, giving,

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{x}{\overset{1}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{x}{\overset{1}{\rightleftarrows}} 21 \, . \tag{1.55}$$

This means that the species 2 particles move with the same rates as species 1 particles. However, species 1 particles essentially do not differentiate between species 2 particles and holes. Because of this, they can be seen as having priority in the dynamics, which is why they are called "first-class", with the species 2 particles being "second-class".

This setup has been studied widely, particularly in the case when there is only one second-class particle [65–69]. Because the first-class particles do not distinguish between second-class particles and holes, their motion is unaffected. However, the motion of second-class particles depends on the local density of first-class particles. It turns out that this property means that second-class particles can be used to track the discontinuities in first-class particle density profiles. In fact, they were first introduced in the literature precisely for this purpose.

No overtaking

Another case that is appropriate for certain situations (such as cars on a street) is to forbid overtaking by setting $\alpha = \beta = 0$. We will examine such a model in chapter 3.

It is important to note that in this case, the sequence of particle species cannot change. Therefore this generally becomes a problem with quenched disorder.

More than two species

One may generalise these processes in the obvious way to an arbitrary number of species, although the calculations become more involved at each step. Nevertheless, there has been some very fruitful research, particularly on the

generalised second-class particle model. One considers N species, with each species k having priority over all species $k+1, k+2, \ldots, N$. This has been studied for instance in [70, 71]. Although this construct may seem somewhat artificial, it turns out to be particularly amenable to exact methods, which will be discussed in chapter 2.

Another choice one might make is to consider a fully disordered ASEP, where each particle i has distinct hopping rates p_i, q_i. This has been examined for instance in [72–74]. We will present some exact results on this problem in chapter 2.

1.5.4 Mean-field theory of single-species ASEP

We now analyse the single-species ASEP more closely, as it will serve as the basis of all the models we will consider in this thesis. Let $\eta_i(t)$ be a variable that is equal to 1 if site i is occupied and 0 if it is empty at time t. Then its average, $\langle \eta_i(t) \rangle$, can be interpreted as the probability that site i is occupied at time t or the mean density at site i at time t. We can write its evolution as,

$$\begin{aligned} \frac{\mathrm{d}\langle \eta_i \rangle}{\mathrm{d}t} &= -\langle \eta_i(1-\eta_{i+1}) \rangle - x\langle (1-\eta_{i-1})\eta_i \rangle + \langle \eta_{i-1}(1-\eta_i) \rangle \\ &\quad + x\langle (1-\eta_i)\eta_{i+1} \rangle, \end{aligned} \tag{1.56}$$

where the argument t is suppressed. We interpret this equation as follows. The first two terms on the right hand site are the probability loss terms. The site i can become empty only if it is currently occupied. Then it becomes empty with rate 1 if site $i+1$ is empty and with rate x if site $i-1$ is empty. Similarly, the other two terms represent the probability gain, which occurs with the corresponding rates if the site i is currently empty and either site $i+1$ or $i-1$ are occupied.

We note that to solve this equation of motion exactly, we first need to calculate the time evolution of the correlation functions $\langle \eta_{i-1}\eta_i \rangle$ and $\langle \eta_i \eta_{i+1} \rangle$. Writing down the equations of motion for those, we obtain equations involving three point correlations, and, continuing this pattern, a hierarchy of evolution equations up to correlations of all sites in the system [4]. Thus it seems that we have not reduced the complexity of the problem, as to solve it, we still have to solve the full global dynamics. Fortunately, as the ASEP is a very simple model, there exist powerful exact methods for analysing the problem, particularly in the steady state. These

[4] A similar structure occurs in kinetic theory, where it is known as the BBGKY hierarchy.

will be discussed in chapter 2. Alternatively, one might truncate the hierarchy at some point to obtain an approximate result. This is most commonly done at first order, yielding a simple mean-field theory. Specifically, this is obtained by setting $\langle \eta_{i-1}\eta_i \rangle = \langle \eta_{i-1} \rangle \langle \eta_i \rangle$.

The latter equation can be understood as a first order approximation scheme in which the covariance of the occupations of different sites is taken to be small,

$$\langle \eta_{i-1}\eta_i \rangle_c = \langle \eta_{i-1}\eta_i \rangle - \langle \eta_{i-1} \rangle \langle \eta_i \rangle \ll \langle \eta_{i-1} \rangle \langle \eta_i \rangle. \tag{1.57}$$

Thus, this is a good approximation if the fluctuations of neighbouring sites are effectively independent (though the fluctuations of a site still in general depend on the *mean* occupation of its neighbour).

The usual way to proceed is to take the continuum limit (an alternative approach, in which the lattice is kept discrete, is presented in appendix 1.A). We define $y = i/L$ and we wish to take the limit $L \to \infty$. The idea is to assume that the density field $\eta(y,t) = \langle \eta_{Ly} \rangle$ varies slowly with y, so we can approximate it by its Taylor series,

$$\langle \eta_{i\pm 1} \rangle \approx \eta(y,t) \pm L^{-1}\nabla\eta(y,t) + \frac{L^{-2}}{2}\nabla^2\eta(y,t) + O(L^{-3}). \tag{1.58}$$

Then to leading order in L^{-1}, (1.56) becomes,

$$\frac{\partial \eta(y,t)}{\partial t} = -L^{-1}(1-x)\nabla[\eta(y,t)(1-\eta(y,t))] + \frac{1}{2}L^{-2}(1+x)\nabla^2\eta(y,t). \tag{1.59}$$

This has the form of a conservation law,

$$\frac{\partial \eta(y,t)}{\partial t} = -\nabla \cdot J, \tag{1.60}$$

with the current given by

$$J(\eta) = L^{-1}(1-x)\eta(1-\eta) - \frac{1}{2}L^{-2}(1+x)\nabla\eta. \tag{1.61}$$

This is an important expression as the steady states will be characterised by a constant current. Recalling Kolmogorov's criterion (see section1.3.4), we see that a currentless ($J = 0$) steady state implies detailed balance, whereas nonequilibrium steady states are characterised by a non-vanishing current ($J \neq 0$).

We now have to separate three different cases, depending on the strength of the asymmetry, $1-x$. These are *(i)* the symmetric case, $x=1$, *(ii)* the (strongly) asymmetric case, $1-x=O(1)$, and *(iii)* the weakly asymmetric case, $|1-x|=O(L^{-1})$ [5].

In each case we must rescale time, *i.e.* set $t=L^{\alpha}\tau$ for some α, to ensure that the left- and right-hand sides of (1.59) are of the same order in the inverse system size expansion. In the cases considered here, we will use either the ballistic timescale, $\alpha=1$, or the diffusive timescale, $\alpha=2$.

Furthermore, we can obtain the steady-state density profiles by setting the time derivative to 0. Then integrating once with respect to y, we generate a constant of integration, which correspond to the mean steady-state current J (this can be seen from (1.60)).

No asymmetry: $x=1$. In the symmetric case, $x=1$, the first term in (1.59) vanishes. Then taking the diffusive timescale, $t=L^2\tau$, we have,

$$\frac{\partial\eta(y,\tau)}{\partial\tau}=\nabla^2\eta(y,\tau). \tag{1.62}$$

This is simply the diffusion equation with unit diffusion coefficient. It makes sense intuitively that as the particles are performing simple diffusion on a lattice, one obtains a diffusion equation on a large scale.

However, this should not be confused with the behaviour of individual particles. It can be shown, for instance, that the mean square displacement of a single tagged particle in this process grows as $\langle x^2\rangle\sim t^{1/2}$, rather than the standard Brownian behaviour $\langle x^2\rangle\sim t$ [75]. This is because the no-overtaking restriction slows down diffusion.

Note that if we simply chose to reinterpret collision events as overtaking events by swapping particle labels, this model would become a model of freely diffusing particles. The continuum description is ignorant of this choice, as it only models the mean density at a particular site and does not keep track of particle indexing. Therefore it makes sense that it gives the same theory for free and single-file diffusion.

[5] Confusingly, the term "weakly asymmetric" is sometimes used to describe the limit $|1-x|=O(L^{-1/2})$, which has also received attention in the literature. Our terminology here is consistent with for instance [45].

Setting $\partial\eta(y,\tau)/\partial\tau = 0$ and integrating once with respect to y, we readily obtain steady-state density profiles of the linear type,

$$\eta_{sym}(y) = \eta_0 - Jy, \tag{1.63}$$

where η_0 is a constant. This is the familiar linear temperature profile in a one-dimensional conductor with two ends being held at different fixed temperatures. Whether or not we have detailed balance, $J = 0$, depends on the boundary conditions.

Strong asymmetry: $1 - x = O(1)$. In the case of strong asymmetry, $1 - x = O(1)$, the second term in (1.59) is subdominant. Then we choose the ballistic timescale, $t = L\tau$, which gives,

$$\frac{\partial\eta(y,\tau)}{\partial\tau} = -(1-x)\nabla[\eta(y,\tau)(1-\eta(y,\tau))]. \tag{1.64}$$

This is the inviscid Burgers' equation. It is a prototypical equation for one-dimensional fluid flow.

We can interpret (1.64) as a collection of small patches with densities η, propagating with group velocities, v_g. These patches are known as kinematic waves. The group velocity v_g of a kinematic wave with density η is given by,

$$v_g(\eta) = \frac{\partial J(\eta)}{\partial\eta} = (1-x)(1-2\eta). \tag{1.65}$$

In particular, this means that low density waves travel faster than high density waves. Then if we start with an initial condition in which there is a higher density region ahead of some low density region, the waves from the low density region will catch up and begin to pile up behind the high density region. This phenomenon is known as shock formation. Formally, the PDE (1.64) breaks down at this point, as $\eta(y,\tau)$ develops a discontinuity at the location of the shock, and one needs to consider higher order spatial derivatives to resolve it. However, the original underlying particle model is always well-defined. Hence, exact solutions to the ASEP provide an avenue for studying the microscopic structures of shocks.

It is interesting to observe that at the mean-field level, there is no difference between the totally and partially asymmetric cases, as setting $x = 0$ does not change the structure of the equation (1.64). However, as we will see in chapter 2,

exact solutions of totally and partially asymmetric cases are remarkably different, with the partially asymmetric case typically being considerably more complex.

Setting $\partial\eta(y,\tau)/\partial\tau = 0$, we obtain uniform steady-state density profiles $\eta_a(y) = \eta_a$, with the current-density relation

$$J = (1-x)\eta_a(1-\eta_a). \tag{1.66}$$

Thus, one typically obtains uniform density profiles.

However, we note that as this is quadratic, one has two different densities with the same mean current,

$$\eta_a^{\pm} = \frac{1 \pm \sqrt{1 - 4J/(1-x)}}{2}. \tag{1.67}$$

Then (1.64) permits shocks in the steady state with a low density region η_a^- followed by a high density region η_a^+.

Weak asymmetry: $|1-x| = O(L^{-1})$. The final case we can consider is that of weak asymmetry. We let $x = 1 - L^{-1}\delta$, where $\delta = O(1)$. Then we choose the diffusive timescale, $t = L^2\tau$, in which case (1.59) becomes,

$$\frac{\partial\eta(y,\tau)}{\partial\tau} = -\delta\nabla[\eta(y,\tau)(1-\eta(y,\tau))] + \nabla^2\eta(y,\tau). \tag{1.68}$$

This is the viscous Burgers' equation. It has a form similar to (1.64), with an extra diffusive term. The diffusive term helps to regulate the shocks, which means that this equation does not develop discontinuities, unlike (1.64).

Physically, we can interpret this as follows. When a shock begins to form, there is a competition between convection, which favours shock formation, and diffusion, which smooths out non-uniform profiles. In the inviscid Burgers' equation (1.64), the convection acts on a much faster timescale than diffusion, so diffusion cannot stop the shock from forming. On the other hand, in the viscous case (1.68), convection has been slowed down to the same timescale as diffusion, so a balance between the two effects can be reached.

Setting $\partial\eta(y,\tau)/\partial\tau = 0$, we obtain,

$$\nabla\eta_{wa}(y) = \delta\eta_{wa}(y)(1-\eta_{wa}(y)) - J. \tag{1.69}$$

This is the most involved case. Firstly, similarly to the case with finite asymmetry, we define the two densities,

$$\eta_{wa}^{\pm} = \frac{1 \pm j}{2}, \tag{1.70}$$

where

$$j = \sqrt{1 - 4J/\delta}. \tag{1.71}$$

This allows us to write (1.69) as

$$\eta'_{wa}(y) = -\delta(\eta_{wa}(y) - \eta_{wa}^{+})(\eta_{wa}(y) - \eta_{wa}^{-}). \tag{1.72}$$

This is a simple first order ODE. Solving it using partial fractions, we obtain the implicit solution

$$\left| \frac{\eta_{wa}(y) - \eta_{wa}^{+}}{\eta_{wa}(y) - \eta_{wa}^{-}} \right| = \mathrm{e}^{-\delta j (y - y_0)}, \tag{1.73}$$

where y_0 is a constant coming from the integration. The constants J, y_0 must be fixed by the boundary conditions.

The solution behaves differently depending on whether $\eta_{wa}(y)$ is within the interval $[\eta_{wa}^{-}, \eta_{wa}^{+}]$ or outside it. From (1.72), we see that the profile $\eta_{wa}(y)$ cannot cross those densities, as its gradient becomes 0 when it approaches them.

Then if $\eta_{wa}^{-} < \eta_{wa}(y) < \eta_{wa}^{+}$, we obtain solutions of the form,

$$\eta_{wa}(y) = \frac{1}{2}\left[1 + j \tanh\left(\frac{\delta j}{2}(y - y_0)\right)\right]. \tag{1.74}$$

This is a continuous profile that interpolates between two densities η_{wa}^{-} at $y \to -\infty$ and η_{wa}^{+} at $y \to +\infty$, with the crossover centered at y_0 and a crossover length of $2/(\delta j)$. We see that in the strong asymmetry limit, $\delta \to \infty$, the crossover length becomes smaller, so the profile becomes sharper, approaching a discontinuous shock.

If $\eta_{wa}(y)$ is outside the interval $[\eta_{wa}^{-}, \eta_{wa}^{+}]$, we have profiles of the form

$$\eta_{wa}(y) = \frac{1}{2}\left[1 + j \coth\left(\frac{\delta j}{2}(y - y_0)\right)\right]. \tag{1.75}$$

This is a mostly uniform profile with a divergence at $y = y_0$. The boundary

conditions must ensure that y_0 is outside the physically relevant domain of y.

1.5.5 Truncations beyond mean-field theory

It is in principle also possible to keep two-point correlation functions and instead truncate the evolution equation hierarchy at some higher order. Such an approach has been employed for instance in [76, 77]. However, while standard mean-field theory has a clear physical interpretation as a sort of "molecular chaos" assumption, truncations at higher orders become progressively harder to justify on physical grounds and generally require more microscopic analysis to be validated. Moreover, as the complexity of the problem grows with each added equation, one must decide whether such a complication is worth the effort, particularly as the result will still not be exact.

1.5.6 Macroscopic fluctuation theory

To go beyond the mean-field theory, one general approach that has seen remarkable success is macroscopic fluctuation theory [6] (see [78] for the original work and [46] for a review). We briefly comment on it, although we do not go into detail as it will not be used in this thesis. Macroscopic fluctuation theory allows one to derive the large deviations (see section1.4) for the joint current-density distributions of certain processes on a hydrodynamic scale. From this, it is possible to compute fluctuations of macroscopic observables and also solve conditioned problems, like the optimal density profile to achieve a specific current. The equations obtained from macroscopic fluctuation theory are generally quite complicated and solving them is a nontrivial task. There has been recent progress on solving them for certain systems using the classic inverse scattering transform [79–81].

A key assumption in standard macroscopic fluctuation theory is the diffusive timescale $t = L^{-2}\tau$. This means that it can be used for the symmetric and weakly asymmetric case, but the cases of finite and total asymmetry are excluded. However there has been some very recent progress on developing a macroscopic fluctuation theory to systems with ballistic scaling [82].

[6] This choice of name unfortunately gives it the same acronym as mean-field theory, which can cause some confusion.

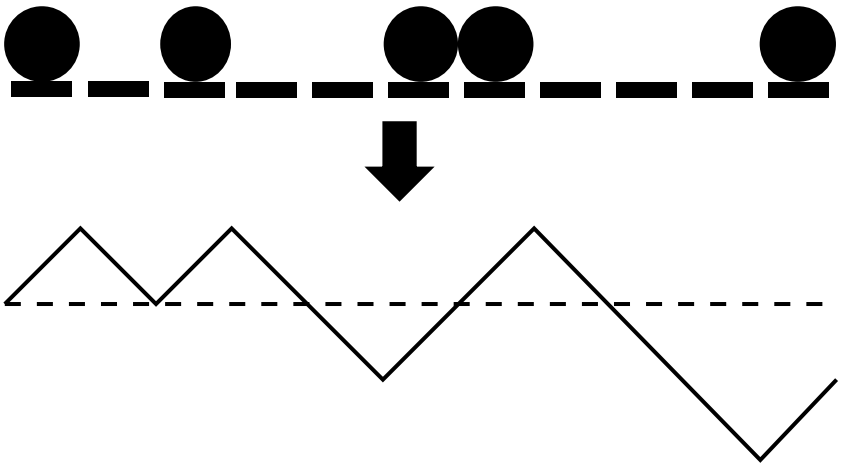

Figure 1.5 *Mapping from exclusion process to fluctuating surface. Particles correspond to a surface gradient of* $+1$ *and empty sites to a gradient of* -1*. The surface can alternatively be interpreted as a polymer in a random medium.*

1.5.7 Connections to other problems

Apart from its relevance to biological and other applications, the ASEP is also interesting due to its connection to numerous other problems in statistical physics. We mention some of the most notable ones.

Surface growth models

Consider the following mapping. Taking an ASEP in some configuration, we draw a line parallel to the lattice from left to right. The line is comprised of connected segments. When we have an empty site, the corresponding line segment is taken to be diagonally downwards, whereas when we have a site containing a particle, the corresponding line segment is taken to be diagonally upwards. This is shown in figure 1.5. Interpreting this line as a boundary between two phases of some substance, the dynamics of the ASEP generate the fluctuating dynamics of the interface.

This is described by the Edwards-Wilkinson equation (noisy diffusion equation) in the symmetric case ($x = 1$) and the KPZ equation in the asymmetric case [83, 84]. The KPZ equation is one of the most studied equations in statistical physics. It is a nonlinear, stochastic PDE, which makes it very difficult to analyse. The mapping to the ASEP provides one of the few avenues for obtaining exact results for it.

Directed polymers in random media

Consider carrying out the same mapping to a line as above. Rather than interpreting the line as a surface between two phases, we can interpret it as a polymer (see figure 1.5). The stochasticity is to be interpreted as a random medium, where each location has a different potential. The statistical physics of such a polymer in a random medium is one of the most widely studied problems in the field of disordered systems [85–87].

Spin chains

A more direct connection exists between ASEPs and certain quantum spin chains. This is due to the fact that the Hamiltonians of spin chains and the Markov operators of ASEPs are both naturally expressed as sums of tensor products of operators on a 2×2 state space [88, 89]. Markov operators can essentially be interpreted as Hamiltonians in imaginary time. Specifically, the symmetric exclusion process corresponds to an XXX spin chain [90], whereas the ASEP corresponds to an XXZ spin chain [91, 92]. The two problems are therefore in some sense equivalent.

However, considering both perspectives can provide insight that only one model might not. For instance, the ASEP picture naturally motivates questions such as the statistics of particle displacement, which has no obvious analogue in the spin chain picture. Moreover, many parameter choices in the ASEP result in non-Hermitian Hamiltonians.

1.6 Summary

We have introduced several fundamental topics, which will serve as the foundation of the material discussed in this thesis. We began with a brief overview of equilibrium statistical mechanics, which will not be used, but which serves as a guiding framework for how we will attempt to understand nonequilibrium systems. Then we discussed the framework of continuous time Markov jump processes and a particularly important such process, the ASEP. We also gave an introduction to large deviation theory, which serves as a unifying framework for equilibrium and nonequilibrium statistical mechanics and will be used in our

study of ASEPs.

1.6.1 Outline of remainder of thesis

The remainder of this thesis is structured as follows. In chapter 2, we explain three methods for obtaining exact results for the steady states of ASEPs, namely the mapping to the zero-range process (ZRP), the matrix product formalism and the Bethe ansatz. Chapters 3–6 then involve application of these methods to various interesting special cases of two-species ASEPs on a ring.

In chapter 3, we analyse a symmetric exclusion process with a single asymmetric defect particle. We obtain its exact steady-state measure using the ZRP mapping and the matrix product formalism.

In chapter 4, we study a partially asymmetric process with a single defect particle that may overtake the normal particles. We analyse the general case using mean-field theory and find an unusual matrix product ansatz that allows us to solve the steady state for a special choice of parameters.

In chapter 5, we consider a general class of two-species processes with overtaking and investigate which choices of hopping rates ensure that the model is solvable using the Bethe ansatz.

In chapter 6, we return to the model we studied in 4 and analyse it using the Bethe ansatz, which allows us to calculate the fluctuations of the current of normal particles.

1.A ASEP mean-field steady-state density profile via homographic maps

In examining the steady state in the ASEP, an interesting alternative to taking the continuum limit is to exploit the algebraic structure of the evolution equation (1.56). This approach has the benefit that it can be used even for small system sizes. However, the solution is more algebraically cumbersome and less physically intuitive than the solution obtained in the continuum limit, which means that it is rarely used, except for a few instances, such as [60, 93].

Setting the time derivative in (1.56) to 0, we obtain,

$$\langle \eta_i(1-\eta_{i+1})\rangle - x\langle (1-\eta_i)\eta_{i+1}\rangle = J = \text{const.} \tag{1.76}$$

This is simply the statement that the current across each bond $i \to (i+1)$ must be constant in the steady state. Then taking the mean-field approximation $\langle \eta_i\eta_{i+1}\rangle = \langle \eta_i\rangle\langle \eta_{i+1}\rangle$, we can reformulate this as a recursion relation for the mean densities,

$$\langle \eta_{i+1}\rangle = \frac{\langle \eta_i\rangle - J}{(1-x)\langle \eta_i\rangle + x}. \tag{1.77}$$

The structure of this recursion relation — a fraction of two linear functions — is called a homographic map. We can obtain an explicit solution for the density profile $\langle \eta_i\rangle$ as follows. First, we introduce the change of variable

$$\langle \eta_i\rangle = \frac{Vu_i + W}{Xu_i + Y}, \tag{1.78}$$

where V, W, X, Y are some constants, to be fixed. The inverse transform is given by,

$$u_i = -\frac{Y\langle \eta_i\rangle - W}{X\langle \eta_i\rangle - V}. \tag{1.79}$$

By a suitable choice of the constants V, W, X, Y, we can make the recursion for the transformed variables u_i take a simple form. Initially, we have,

$$\begin{aligned} u_{i+1} &= -\frac{Y\langle \eta_{i+1}\rangle - W}{X\langle \eta_{i+1}\rangle - V} && (1.80)\\ &= -\frac{(Y - W(1-x))\langle \eta_i\rangle - (YJ + Wx)}{(X - V(1-x))\langle \eta_i\rangle - (XJ + Vx)} && (1.81)\\ &= -\frac{Au_i - B}{Cu_i - D}, && (1.82) \end{aligned}$$

where

$$\begin{aligned} A &= V(Y - W(1-x)) - X(YJ + Wx) && (1.83a)\\ B &= W(Y - W(1-x)) - Y(YJ + Wx) && (1.83b)\\ C &= V(X - V(1-x)) - X(XJ + Vx) && (1.83c)\\ D &= W(X - V(1-x)) - Y(XJ + Vx). && (1.83d) \end{aligned}$$

We can now choose V, W, X, Y in such a way that $B = C = 0$. This can be achieved by requiring,

$$(1-x)WY - (1-x)W^2 - JY^2 = 0 \tag{1.84a}$$
$$(1-x)XV - (1-x)V^2 - JX^2 = 0. \tag{1.84b}$$

Then setting $X = Y = -1$, we obtain the following equations for V, W,

$$J/(1-x) + W + W^2 = 0 \tag{1.85a}$$
$$J/(1-x) + V + V^2 = 0. \tag{1.85b}$$

This is solved by,

$$V,\ W = Z_\pm = \frac{-1 \pm j}{2}, \tag{1.86}$$

where

$$j = \sqrt{1 - 4J/(1-x)}. \tag{1.87}$$

Note that, say for a PASEP with periodic boundaries, the current-density relation is,

$$J = (1-x)\rho(1-\rho), \tag{1.88}$$

where $\rho \in [0,1]$ is the mean density. This attains its minimal value of $J = 0$ at $\rho = 0$ and $\rho = 1$, and its maximal value of $J = (1-x)/4$ at $\rho = 1/2$. Thus we expect $j \in [0,1]$.

As we have chosen $X = Y$, choosing the same value $Z_\pm$ for V, W would make the transformation (1.78) singular. Instead, we choose $V = Z_+$, $W = Z_-$.

Then the recursion for u_i becomes simply a geometric sequence,

$$u_{i+1} = \frac{A}{D}u_i = \left(\frac{A}{D}\right)^i u_1. \tag{1.89}$$

With the choices $X = Y = -1$, $V = Z_+$, $W = Z_-$, we have

$$u_{i+1} = \left(\frac{j_-}{j_+}\right)^i u_1, \tag{1.90}$$

where

$$j_{\pm} = 1 + x \pm (1 - x)j. \tag{1.91}$$

From (1.90), the corresponding density profile $\langle \eta_i \rangle$ can then be obtained by substituting into (1.78), which gives

$$\begin{aligned} \langle \eta_{i+1} \rangle &= -\frac{Z_+ u_{i+1} + Z_-}{u_{i+1} + 1} && (1.92) \\ &= -\frac{Z_+ (j_-/j_+)^i u_1 + Z_-}{(j_-/j_+)^i u_1 + 1} && (1.93) \\ &= \frac{1}{2}\left[1 - j\frac{u_1 j_-^i - j_+^i}{u_1 j_-^i + j_+^i}\right]. && (1.94) \end{aligned}$$

Then $\langle \eta_1 \rangle$ and j have to be fixed using the boundary conditions.

It is an interesting mathematical exercise to derive the mean-field steady-state density profile in this way, without taking the continuum limit, and seemingly wholly divorced from any analysis of PDEs. However, the solution obtained this way is somewhat impractical, as the complexity of the expressions obtained blows up geometrically. In particular, if one is studying a system on a ring with a fixed total number of particles, the imposition of this global constraint results in a polynomial equation of very high degree for j. However, we note that this approach has the obvious advantage over the continuum limit if one is considering small system sizes.

This exercise contains an important general lesson regarding mean-field theory: although writing down a mean-field theory is much more straightforward than using exact methods, the resulting equations can still be quite complicated, or even intractable.

Chapter 2

Methods

2.1 Introduction

In this chapter, we introduce the approaches we will use to exactly analyse steady states and dynamical properties of simple exclusion processes (SEPs) (see section 1.5). We demonstrate how these methods work on simple cases. These demonstrations are based on established literature in the field, though the presentation has been modified in places. This is done partly to retrospectively simplify some of these calculations and partly to bring the notation in line with the rest of this thesis.

In section 2.2, we introduce another transport model: the zero-range process (ZRP). In section 2.2.2, we show how its steady state can be solved exactly using a product measure. Although this solution is interesting in its own right, the main point of interest this model presents in this thesis is that SEPs can be mapped to it, which is shown in section 2.2.3. This provides a simple method for finding exact steady states of SEPs.

In section 2.3, we introduce the matrix product formalism. This is a generalisation of (scalar) product measures. It is somewhat more formal in nature than the approach using the ZRP mapping, but makes the calculation of local observables more intuitive. We give a general proof of validity of matrix product measures for SEPs in section 2.3.2. In section 2.3.3, we focus on a particularly simple and important class of matrix product measures and present a way of checking whether a particular model falls within this class. In section 2.3.5, we give examples of

calculations of physical observables for SEPs using the matrix product formalism.

In section 2.4, we introduce the coordinate Bethe ansatz. This is a method to diagonalise certain linear operators, including the Markov operator that generates the time evolution of some SEPs. It is particularly well suited for studying the long-time statistics of particle motion. In section 2.4.1, we give an informal explanation of the method, to provide some intuition. In section 2.4.2, we demonstrate more precisely how this works on the simple case of a totally asymmetric SEP (TASEP) on a ring. In section 2.4.3, we modify the calculations from the previous section to allow for direct computation of the scaled cumulant generating function of the total particle displacement at long times. In section 2.4.4, we demonstrate how to extract explicit expressions from the formal solution obtained in the previous section.

2.2 Zero-range process

In this section, we introduce the zero-range process (ZRP). This is a model of particles hopping on discrete sites, similar to the SEP (see section 1.5). However, unlike the SEP, there is no limit on site occupancy. A remarkable property of ZRPs (particularly in the context of nonequilibrium models) is that there is a general exact solution for their steady states in the form of a product measure.

This section is laid out as follows. In section 2.2.2, we demonstrate the product measure solution for the steady state. In section 2.2.3, we introduce a mapping between ZRPs and SEPs, which will allow us to obtain exact solutions for the steady states of certain SEPs. Calculations in this section will be demonstrated on a heterogeneous, totally asymmetric ZRP on a ring, which will be shown to correspond to a disordered TASEP. This section mainly follows the reviews [94, 95].

2.2.1 Model definition

The ZRP is a mass transport model in which indistinguishable particles hop on a graph with no constraint on the number of particles that can occupy any one site. In general, one can consider this process on graphs with arbitrary topologies and boundary conditions. We will focus on the case of a one-dimensional chain with

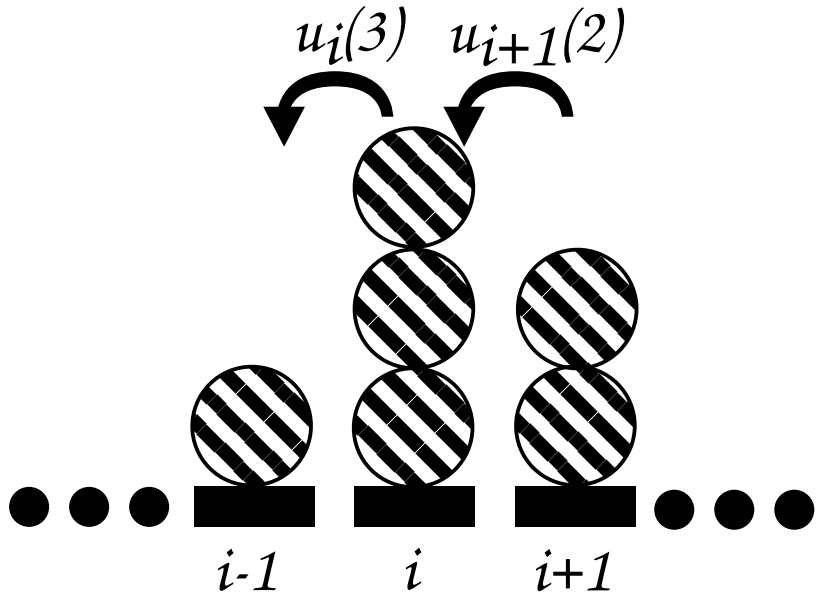

Figure 2.1 *Illustration of heterogeneous totally asymmetric zero-range process in one dimension.*

periodic boundary conditions (a ring), as this is the topology that is relevant for SEPs on a ring (see section 2.2.3). The zero-range condition signifies that the rate at which particles hop out of a site can only depend on the occupation number of that site. Hence, the particles can be said to have an interaction with zero range.

Consider an M site ring. We denote the site occupation numbers by n_i, where $i = 1, 2, \ldots, M$, is the site index. Then the zero-range condition means that the rate u_i at which particles leave site i is a function of only n_i, *i.e.* $u_i(n_1, \ldots, n_M) = u_i(n_i)$. An illustration is given in in Figure 2.1. Note that by labelling rates with the site index, we are allowing different hopping rates on different sites, which makes this a heterogeneous ZRP. We will see in section 2.2.3 that this corresponds to a multi-species SEP. The homogeneous case, $u_i(n) = u_j(n)$ for all i, j, has also been studied extensively, though it will not be relevant for the models considered in this thesis.

In general, we may allow particles from site i to jump to any other site j with some rate $u_{i,j}$, depending on the topology of the graph (where $\sum_j u_{i,j} = u_i$). For processes on a ring, we have $u_{i,j} = 0$ for all j except $j = i \pm 1$.

2.2.2 Product measure solution of the steady state

An important property of ZRPs is that their steady states are solved exactly by a product measure. We demonstrate this with a simple example.

Consider a ZRP on a one-dimensional periodic lattice with M sites and N particles. We take the hopping to be unidirectional, with the direction of hopping going to the left, *i.e.* $u_{i,j} = \delta_{j,i-1} u_i$. As will be seen in section 2.2.3, this choice of hopping direction gives the natural direction of hopping (to the right) in the corresponding ASEP.

A configuration is given by the occupation numbers of all sites, $\{n_1, n_2, \ldots, n_M\}$. Then the product measure structure of the steady state means that the probability of observing some configuration in the steady state can be written as a product of local weights, $f_i(n_i)$, each of which depend only on the state of one site,

$$P(\{n_1, \ldots, n_M\}) = Z_{M,N}^{-1} \prod_{i=1}^{M} f_i(n_i), \tag{2.1}$$

where $Z_{M,N}$ is a normalisation. Taking this as an ansatz, we now verify that this solution works. This will also allow us to fix the weights f_i in terms of the system parameters.

Recall that we have defined the rate at which particles hop from site i to $i-1$ as $u_i(n_i)$. In the steady state, we have the global balance condition, coming from the master equation (see section 1.3.4),

$$\begin{aligned} \sum_{i=1}^{M} \theta(n_i) u_i(n_i) P(\{\ldots, n_i, n_{i+1}, \ldots\}) = \\ \sum_{i=1}^{M} \theta(n_i) u_{i+1}(n_{i+1}) P(\{\ldots, n_i - 1, n_{i+1} + 1, \ldots\}), \end{aligned} \tag{2.2}$$

where we have used the Heaviside theta θ, defined as

$$\theta(n) = \begin{cases} 0, & n \leq 0 \\ 1, & n > 0 \end{cases}, \tag{2.3}$$

and the indices i are to be understood mod M. Now we plug in the ansatz (2.1).

To obtain a solution, we assume the local balance condition,

$$u_i(n_i)f_i(n_i)f_{i+1}(n_{i+1}) = u_{i+1}(n_{i+1})f_i(n_i - 1)f_{i+1}(n_{i+1} + 1). \tag{2.4}$$

Dividing by $f_i(n_i - 1)f_{i+1}(n_{i+1})$, we obtain

$$u_i(n_i)\frac{f_i(n_i)}{f_i(n_i - 1)} = u_{i+1}(n_{i+1})\frac{f_{i+1}(n_{i+1} + 1)}{f_{i+1}(n_{i+1})} = \text{const.} \tag{2.5}$$

Evidently this can be repeated for all i and all values of n_i. The weights $f_i(0)$ can be set to unity without loss of generality. This then implies that

$$f_i(n_i) = \prod_{k=1}^{n_i}[u_i(k)]^{-1}. \tag{2.6}$$

Therefore the full probability of a configuration is

$$P(\{n_1, \ldots, n_M\}) = Z_{M,N}^{-1}\prod_{i=1}^{M}\prod_{k=1}^{n_i}[u_i(k)]^{-1}, \tag{2.7}$$

where the normalisation $Z_{M,N}$ is given by,

$$Z_{M,N} = \sum_{\{n_1,\ldots,n_M\}}\prod_{i=1}^{M}\prod_{k=1}^{n_i}[u_i(k)]^{-1}\delta_{\sum_{i=1}^{M} n_i,N}. \tag{2.8}$$

This is an explicit expression for the normalisation, which is already a remarkable result, as this is a nonequilibrium model, so we could not write down a partition function *a priori*. However, it is unwieldy due to the global particle number constraint. We can deal with this by expressing the normalisation as a complex contour integral using the generating function trick (see section 2.B). This gives,

$$Z_{M,N} = \oint \frac{\mathrm{d}z}{2\pi\mathrm{i}} z^{-(N+1)}\prod_{i=1}^{M}\sum_{n_i=0}^{\infty}\frac{z^{n_i}}{\prod_{k=1}^{n_i} u_i(k)}, \tag{2.9}$$

where the contour of integration is a small circle around the origin in the complex plane. Although this expression may still seem intractable, in many cases the form of u_i is simple enough to allow the sum to be simplified.

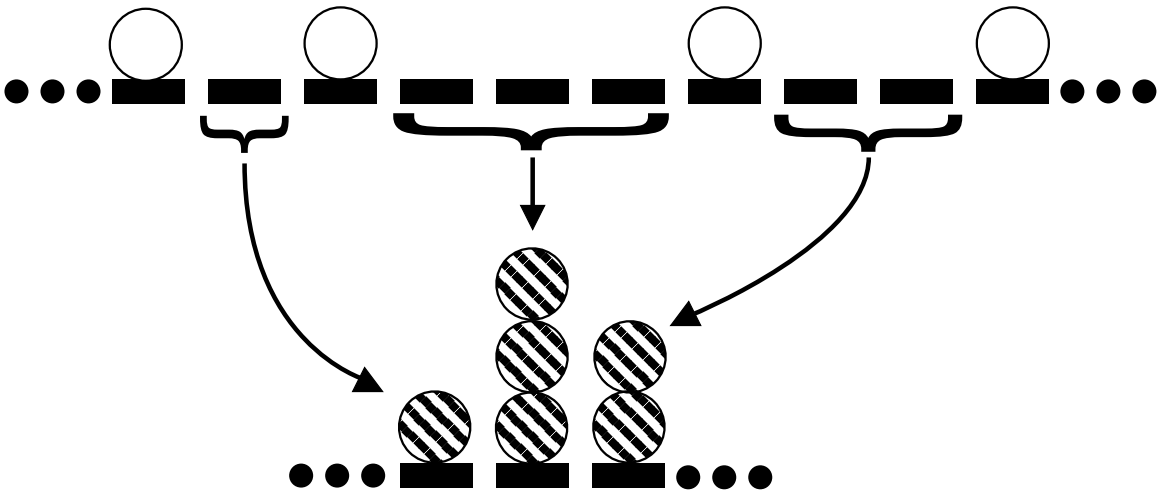

Figure 2.2 *Mapping between simple exclusion process (top) and zero-range process (bottom).*

2.2.3 Mapping to simple exclusion process

Although the ZRP is interesting as a mass transfer model in its own right, for the scope of this thesis, its main purpose is to allow us to construct exact steady-state measures for SEPs. This is done by mapping a given SEP to a corresponding ZRP. Then we can find the steady-state measure of the ZRP using the method outlined in the previous section and translate it to the corresponding SEP. We demonstrate the mapping for a disordered TASEP.

Consider a TASEP on a periodic lattice with L sites, M particles and $N = L - M$ empty sites. Here we examine the fully disordered case, with each hopping to the right with a distinct rate p_i. The fully disordered case has been considered for instance in [72, 74]. Note that this general case contains various interesting subcases, such as multi-species processes and processes with a defect particle.

We map this to a ZRP with M sites and N particles by associating the i-th SEP particle to the i-th ZRP site. We let the occupation number of the i-th ZRP site, n_i, be equal to the number of empty sites to the right of the i-th SEP particle. This is illustrated in Figure 2.2 When the i-th SEP particle hops forward, this transfers one empty site from being to the right of the i-th particle to being to the right of the $(i-1)$-th particle. This corresponds to a hop of a ZRP particle from the i-th to the $(i-1)$-th site. Thus the ZRP hopping rates are given by,

$$u_i(n_i) = p_i. \tag{2.10}$$

In particular, they are independent of the occupation number. We remark that, as the ZRP particles correspond to holes in the SEP, the currents in the two systems flow in opposite directions, which is why we defined the ZRP to flow to the left.

Steady-state measure

From (2.6) and (2.10), we see that weight functions then have a simple geometric form,

$$f_i(n_i) = p_i^{-n_i}. \tag{2.11}$$

Then from (2.9), we have the normalisation,

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(N+1)} \prod_{i=1}^{M} \sum_{n_i=0}^{\infty} \frac{z^{n_i}}{p_i^n} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(N+1)} \prod_{i=1}^{M} \frac{1}{1 - z/p_i}. \tag{2.12}$$

As the weights in this case have a very simple form, we can simplify this a bit further. We can evaluate the integral by using the result that a contour integral of a meromorphic function equals minus the sum of the residues of the poles outside the contour. Here, the poles outside the contour are at $z = p_i$ for $i = 1, \ldots, M$. Assuming for simplicity that $p_i \neq p_j$ for all $i \neq j$, we have,

$$\begin{aligned} Z_{L,M} &= (-)^{M+1} \sum_{i=1}^{M} \mathrm{Res}\left(z^{-(N+1)} \prod_{j=1}^{M} \frac{p_j}{z - p_j}, z = p_i \right) && (2.13) \\ &= (-)^{M+1} \sum_{i=1}^{M} p_i^{-(N+1)} \prod_{j \neq i} \frac{p_j}{p_i - p_j}. && (2.14) \end{aligned}$$

Thus we have obtained an expression for the exact steady-state measure of a fully heterogeneous TASEP as a finite sum, though we remark that this does not necessarily make the computation of all physical observables trivial. There is generally still the combinatorial challenge of enumerating the configurations that correspond to some particular value of that observable.

Mean velocity

One result that does follow easily is the mean velocity of the particles. Note that due to the fact that there is no overtaking, all particles have the same mean velocity, so it does not matter which particle we consider. Consider some particle k. Its steady-state velocity is given by,

$$\langle v_k \rangle = p_k P(n_k \geq 1). \tag{2.15}$$

To calculate the probability $P(n_k \geq 1)$, we need to sum the weights of all configurations in which the condition $n_k \geq 1$ is satisfied. This is given by an expression similar to (2.12) but with the lower limit of the sum over n_k being 1 (the lower limits of the sums of all n_i for $i \neq k$ remain 0). Then using the identity $\sum_{n=1}^{\infty} z^n = z/(1-z)$ and recalling $L = M + N$, we obtain,

$$\langle v_k \rangle = \frac{Z_{L-1,M}}{Z_{L,M}}. \tag{2.16}$$

As expected, the result is independent of which particle k we chose. We remark that the form of this expression resembles a derivative of $\log Z_{L,M}$. This suggests that for SEPs, currents may play a similar role to free energies in equilibrium systems.

We can obtain an exact finite-size result by plugging in (2.14). This, however, does not have a particularly insightful form. For a large system size ($L \to \infty$), we see from (2.14) that, as each term in the summation is weighted by $p_i^{-(N+1)}$, where $N \to \infty$, the normalisation $Z_{L,M}$ is dominated by the term with the smallest hopping rate,

$$p_* = \min_i p_i. \tag{2.17}$$

Then we see from (2.14) that in (2.16), the only difference between the dominant term in the numerator and denominator is a factor of p_*, which gives,

$$\langle v_k \rangle \approx p_*, \tag{2.18}$$

for $k = 1, \ldots, M$. The interpretation of this result is clear. As overtaking is not allowed, all particles eventually form a traffic jam behind the slowest particle and move at its speed.

Mean number of holes in front of a particle

Another result that follows easily is the mean number of holes in front of an arbitrary particle k, $\langle n_k \rangle$. This corresponds to the mean occupation of site k in the ZRP. We take $k = 1$ without loss of generality. The mean number of holes to the right of this particle is given by

$$\langle n_1 \rangle = \sum_{n=0}^{\infty} nP(n_1 = n). \tag{2.19}$$

Note that it somewhat simplifies calculations not to terminate the sum explicitly at $n = N$. We can do this safely with the convention $P(n_1 = n) = 0$ for $n > N$.

Then to obtain an explicit expression for the probability $P(n_1 = n)$, we need to assign the local weight $f_1(n) = p_1^{-n}$ (2.11) to site 1 for $n = 0, 1, 2, \ldots$. The remainder of the system may have any configuration with $N - n$ holes. We can enumerate these with a "partition function" similar to (2.12), but with $i = 1$ omitted in the product and the number of holes being $N - n$. In total, this gives,

$$\langle n_1 \rangle = Z_{L,M}^{-1} \sum_{n=0}^{\infty} n p_1^{-n} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(N+1-n)} \prod_{i=2}^{M} \frac{1}{1 - z/p_i}. \tag{2.20}$$

Using the variation on the geometric sum

$$\sum_{n=0}^{\infty} n x^n = \frac{x}{(1-x)^2}, \tag{2.21}$$

we can write

$$\langle n_1 \rangle = Z_{L,M}^{-1} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(N+1)} \frac{z/p_1}{(1 - z/p_1)^2} \prod_{i=2}^{M} \frac{1}{1 - z/p_i}. \tag{2.22}$$

We now evaluate this by switching to the residues outside the contour, similar to (2.14). The only difference is that now the pole at $z = p_1$ is of degree 2. We

obtain,

$$\langle n_1 \rangle = Z_{L,M}^{-1}(-)^M \sum_{i=1}^{M} \mathrm{Res}\left(z^{-N} \frac{p_1}{(z-p_1)^2} \prod_{j=2}^{M} \frac{p_j}{z-p_j}, z = p_i \right) \tag{2.23}$$

$$= Z_{L,M}^{-1}(-)^{M+1} \left[p_1^{-N} \left(\prod_{j=2}^{M} \frac{p_j}{p_1 - p_j} \right) \left(N + \sum_{k=2}^{M} \frac{p_1}{p_1 - p_k} \right) - \sum_{i=2}^{M} p_i^{-N} \frac{1}{p_i - p_1} \prod_{j \neq i} \frac{p_j}{p_i - p_j} \right]. \tag{2.24}$$

Now the asymptotic behaviour is very different depending on whether p_1 is the smallest rate, p_*, or not. If it is, then this expression and the normalisation $Z_{L,M}$ are dominated by the first term (with weight p_1^{-N}). Then approximating

$$Z_{L,M} \approx (-)^{M+1} p_1^{-(N+1)} \prod_{j=2}^{M} \frac{p_j}{p_1 - p_j} \tag{2.25a}$$

$$\langle n_1 \rangle \approx Z_{L,M}^{-1}(-)^{M+1} p_1^{-N} \left(\prod_{j=2}^{M} \frac{p_j}{p_1 - p_j} \right) \left(N + \sum_{k=2}^{M} \frac{p_1}{p_1 - p_k} \right), \tag{2.25b}$$

gives after some simplification

$$\langle n_1 \rangle \approx N - \sum_{k=2}^{M} \frac{1}{(p_k/p_1) - 1}. \tag{2.26}$$

This reinforces the picture we described in the previous section: if $i = 1$ is the slowest particle, the space in front of it is occupied by a macroscopic (*i.e.* $O(N)$) number of holes. In the ZRP picture, this means that site 1 contains a macroscopic number of particles, or in other words, the system undergoes condensation.

On the other hand, if p_1 is not the smallest weight, the sums are instead dominated by one of the other terms with the weight p_*. Using a similar approximation to the p_1 dominant case for $Z_{L,M}$ and $\langle n_1 \rangle$ gives,

$$\langle n_1 \rangle \approx \frac{1}{(p_1/p_*) - 1}, \tag{2.27}$$

which is a microscopic value ($\ll N$).

Condensation in ZRPs has been studied extensively (*e.g.* [72, 74, 96, 97]). We remark that condensation can also occur in a ZRP with no disorder, but

with occupation dependent hopping rates $u(n)$, though the mechanism by which condensation occurs is different in that case [94].

Two-point function

We now demonstrate how the solution of the SEP steady state using the ZRP mapping can be used to compute two-point functions in the SEP. Although this approach for computing two-point functions has been used in literature (for instance in [50]), we consider the matrix product approach better for this task, which will be discussed in section 2.3.5. Thus, this is included purely for the purpose of demonstration. In this subsection we consider the homogeneous case $p_i = p$ for all i for simplicity, though we will see that even this calculation is quite cumbersome using the ZRP approach.

We pick one particle in the SEP, say particle 1, and label the site on which it is currently located as site 1. Then we wish to compute the probability that site i is also occupied, *i.e.* $\langle \eta_1 \eta_i \rangle$. To obtain this, we need to sum the probabilities that any one of the particles $2, 3, \ldots, M$ is on site i. Note that particle k being on site i means that $\sum_{k'=1}^{k-1} n_{k'} = i - k$. Then for the remainder of the system, we must enumerate all possibilities that are consistent with $\sum_{k'=k}^{M} n_{k'} = N - (i-k)$. In total, this gives,

$$\begin{aligned} \langle \eta_1 \eta_i \rangle &= Z_{L,M}^{-1} \sum_{k=2}^{M} Z_{i-1,k-1} Z_{L-i+1,M-k+1} && (2.28) \\ &= Z_{L,M}^{-1} \sum_{k=2}^{M} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(i-k+1)} \left(\frac{1}{1-z/p} \right)^{k-1} \\ &\quad \times \oint \frac{\mathrm{d}w}{2\pi \mathrm{i}} w^{-(N-i+k+1)} \left(\frac{1}{1-w/p} \right)^{M-k+1} && (2.29) \end{aligned}$$

Note that we can safely extend the upper limit of the sum to ∞, as for $k > M$ the w integral has no simple pole at $w = 0$ and so vanishes. Then performing the k sum, we obtain, after some simplification,

$$\langle \eta_1 \eta_i \rangle = Z_{L,M}^{-1} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(i-1)} \oint \frac{\mathrm{d}w}{2\pi \mathrm{i}} w^{-(N-i+2)} \left(\frac{1}{1-w/p} \right)^{M-1} \frac{1}{w-z}. \quad (2.30)$$

We can now evaluate the z integral by taking the negative of the residue outside

the contour (at $z = w$). This gives,

$$\langle \eta_1 \eta_i \rangle = Z_{L,M}^{-1} \oint \frac{\mathrm{d}w}{2\pi \mathrm{i}} w^{-(N+1)} \left(\frac{1}{1 - w/p} \right)^{M-1} = \frac{Z_{L-1,M-1}}{Z_{L,M}}. \tag{2.31}$$

For the homogeneous case, the partition functions are readily evaluated using the negative binomial expansion [98] to give,

$$Z_{L,M} = \binom{L-1}{M-1} p^{-(L-M)}, \tag{2.32}$$

which leads to

$$\langle \eta_1 \eta_i \rangle = \frac{M-1}{L-1}. \tag{2.33}$$

This result was to be expected as the homogeneous TASEP has a uniform measure. Yet this calculation shows that to answer even such a simple question is quite cumbersome using the ZRP mapping: we had to evaluate a sum and two complex contour integrals.

Disordered PASEP

Extending the results from section 2.2.2, we now consider a disordered PASEP. This has been considered for instance in [72, 99]. We again have an L-site ring with M particles, but now particle i can hop to the right with rate p_i and to the left with rate q_i. This evidently maps to a partially asymmetric ZRP, so we have to generalise the solution from section 2.2.2. Note that there is a slight mismatch between indices in the SEP and ZRP. Particle i in the SEP hopping to the right corresponds to a particle hopping from site i to site $i-1$ in the ZRP, but particle i in the SEP hopping to the left corresponds to a particle hopping from site $i-1$ to site i in the ZRP. Therefore in the ZRP, site i has hopping rates p_i to the left, but q_{i+1} to the right.

The global balance condition (2.2) now becomes,

$$\sum_{i=1}^{M} \theta(n_i)(p_i + q_{i+1})P(\{\dots, n_{i-1}, n_i, n_{i+1}, \dots\}) = \sum_{i=1}^{M} \theta(n_i)[p_{i+1}P(\{\dots, n_{i-1}, n_i - 1, n_{i+1} + 1, \dots\}) + q_i P(\{\dots n_{i-1} + 1, n_i - 1, n_{i+1}, \dots\})]. \tag{2.34}$$

Making an ansatz of the product type, this yields the local balance condition

$$(p_i + q_{i+1})f_{i-1}(n_{i-1})f_i(n_i)f_{i+1}(n_{i+1}) = p_{i+1}f_{i-1}(n_{i-1})f_i(n_i - 1)f_{i+1}(n_{i+1} + 1) + q_i f_{i-1}(n_{i-1} + 1)f_i(n_i - 1)f_{i+1}(n_{i+1}). \tag{2.35}$$

Dividing by $f_{i-1}(n_{i-1})f_i(n_i - 1)f_{i+1}(n_{i+1})$ gives,

$$(p_i + q_{i+1})\frac{f_i(n_i)}{f_i(n_i - 1)} = p_{i+1}\frac{f_{i+1}(n_{i+1} + 1)}{f_{i+1}(n_{i+1})} + q_i\frac{f_{i-1}(n_{i-1} + 1)}{f_{i-1}(n_{i-1})}. \tag{2.36}$$

This is solved by a geometric ansatz [94],

$$f_i(n_i) = g_i^{n_i}, \tag{2.37}$$

where now g_i are not simply the rates but some other constants, which are to be determined by solving the equations

$$(p_i + q_{i+1})g_i = p_{i+1}g_{i+1} + q_i g_{i-1}. \tag{2.38}$$

To solve this, we first note that we can rearrange this equation to read,

$$p_{i+1}g_{i+1} - q_{i+1}g_i = p_i g_i - q_i g_{i-1} = \text{const.} \tag{2.39}$$

We can set this to 1 without loss of generality. Note that this implies that the velocity of particle k can be written in the same form as in the totally asymmetric case,

$$\langle v_k \rangle = \frac{(p_i g_i - q_i g_{i-1})Z_{L-1,M}}{Z_{L,M}} = \frac{Z_{L-1,M}}{Z_{L,M}}. \tag{2.40}$$

To solve for the weights g_i, we may proceed as follows. Iterating (2.39), we obtain,

$$\begin{aligned} g_i &= \frac{1}{p_i}(1 + q_i g_{i-1}) && (2.41)\\ g_i &= \frac{1}{p_i}\left(1 + \frac{q_i}{p_{i-1}}(1 + q_{i-1}g_{i-2})\right) && (2.42)\\ &\vdots \\ g_i &= \frac{1}{p_i}\left(1 + \frac{q_i}{p_{i-1}}\left(1 + \frac{q_{i-1}}{p_{i-2}}\left(1 + \ldots \frac{q_{i+2}}{p_{i+1}}(1 + q_{i+1}g_i)\right)\right)\right). && (2.43) \end{aligned}$$

The last term (containing g_i) is thus multiplied by all q_j and divided by all p_j. Bringing it to the other side of the equation and solving for g_i gives

$$g_i = \frac{1}{p_i}\frac{1 + \frac{q_i}{p_{i-1}}\left(1 + \frac{q_{i-1}}{p_{i-2}}(1 + \ldots)\right)}{1 - \frac{q_1 q_2 \ldots q_M}{p_1 p_2 \ldots p_M}}, \tag{2.44}$$

where the nested terms terminate at q_{i+2}/p_{i+1}. Using this, we can explicitly construct the steady-state measure for the disordered PASEP, which would have required much more work using the matrix product approach or the Bethe ansatz.

Yet these weights are very complicated and practically unusable for explicit calculations. In particular, they are significantly more complicated than the weights in the totally asymmetric case (2.11). There, the weights for each particle depended only on the hopping rate of that particle, p_i, whereas in the partially asymmetric case all weights depend on all hopping rates. We will see similar phenomena in the matrix product approach and the Bethe ansatz as well. We conclude that partially asymmetric systems are generally much harder to deal with exactly than totally asymmetric ones.

2.3 Matrix product states

We now introduce another method to obtain exact steady-state measures for exclusion processes: the matrix product ansatz. This approach has the advantage that it provides a naturally local description of the SEP steady state, which makes it easy to write down expressions for local observables, like local mean densities and two-point functions. It can also be used for open boundary problems and models with overtaking, where the ZRP mapping fails, although requiring a simple matrix product structure may impose some restrictions on

model parameters. The precise meaning of "simple structure" and the associated restrictions will be explained in section 2.3.3.

The idea of what would eventually become matrix product states has its roots in work on the ground state correlation functions in certain 1D quantum models [100, 101]. In the context of stochastic processes, the matrix product approach was first used in [61]. There, it was used to solve a one-species open boundary TASEP, though it has since then also been used for numerous other problems on a ring, including a TASEP on a ring [102], models with a "second-class" particle [68], models with a slow defect particle and overtaking [11], and multi-species problems [103]. We note that for one-species processes on a ring, the steady-state measure is always uniform, which makes the matrix product structure trivial [104], so we will consider two-species processes in examples in this section. This section mainly follows the review [8].

We first motivate the matrix product formalism and give the matrix product measure as an ansatz for the steady state of a two-species SEP. Then, in section 2.3.2, we give a proof of the validity of the matrix product ansatz, provided a certain condition holds. In section 2.3.3, we consider an important special case of this condition and derive a set of relations in terms of the model parameters that must be satisfied in order for a given SEP to have a solution that lies within this special case. Finally, in section 2.3.5, we give some examples of calculations of local observables for a SEP using the matrix product formalism.

2.3.1 Statement of the matrix product ansatz

Our motivation for using the matrix product approach is that we wish to obtain a description for the steady state of a SEP that is local. In general, the simplest local measures are of the (scalar) product type, as we saw in section 2.2.2. However, that solution was local in the ZRP picture, but became non-local in the SEP picture. We could make a scalar product measure ansatz for a SEP, but in most cases this is insufficient. It turns out that we can still obtain a measure of the product type, provided we allow the weights to be matrices, rather than scalars.

Let us consider a two-species exclusion process on a periodic lattice with L sites We denote the configuration of the system as $\{\tau_1, \ldots, \tau_L\}$, where we use the notation $\tau_i = 0$, 1, 2 for empty sites and particles of species 1 and 2 respectively.

Then we postulate that there exist matrices X_0, X_1, X_2 such that the weight of that configuration is given by

$$P(\{\tau_1, \ldots, \tau_L\}) = Z_{L,M}^{-1} \mathrm{tr}(X_{\tau_1} \ldots X_{\tau_L}), \tag{2.45}$$

where $Z_{L,M}$ is a normalisation. The use of the trace reflects the rotational symmetry implied by the periodic boundary conditions. For open boundary conditions, one would generally need to multiply the matrix product by some row and column vector from the left and right respectively.

2.3.2 Proof of validity

We present a general proof of validity of the matrix product ansatz for two-species exclusion processes on a ring, which was given in [105]. This proof relies on the condition (2.48), which is valid for most known solutions of the matrix product type, though we will see a case in which this condition does not seem to hold and the proof needs to be modified in chapter 4.

Let us denote rates of processes as $w_{\tau\tau'} = k(\tau\tau' \to \tau'\tau)$. The global balance condition (see section 1.3.4) reads

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} P(\{\tau_1, \ldots, \tau_L\}) = \sum_{i=0}^{L} \big[& -w_{\tau_i \tau_{i+1}} P(\{\ldots, \tau_i, \tau_{i+1}, \ldots\}) \\ & + w_{\tau_{i+1} \tau_i} P(\{\ldots, \tau_{i+1}, \tau_i, \ldots\}) \big] = 0, \end{aligned} \tag{2.46}$$

where the indices i are to be understood mod L. Plugging in the ansatz (2.45), we obtain,

$$\sum_{i=0}^{L} \left[w_{\tau_i \tau_{i+1}} \mathrm{tr}(\ldots X_{\tau_i} X_{\tau_{i+1}} \ldots) - w_{\tau_{i+1} \tau_i} \mathrm{tr}(\ldots X_{\tau_{i+1}} X_{\tau_i} \ldots) \right] = 0. \tag{2.47}$$

For the most general known proof of the validity of the matrix product ansatz, we require that there exist auxiliary matrices $\tilde{X}_\tau$, $\tilde{X}_{\tau'}$, such that the following relations hold for all τ, τ',

$$w_{\tau\tau'} X_\tau X_{\tau'} - w_{\tau'\tau} X_{\tau'} X_\tau = X_\tau \tilde{X}_{\tau'} - \tilde{X}_\tau X_{\tau'}. \tag{2.48}$$

From this, the proof is straightforward. The left-hand side of (2.47) becomes

$$\sum_{i=0}^{L} \Big[\mathrm{tr}(\dots X_{\tau_i} \tilde{X}_{\tau_{i+1}} \dots) - \mathrm{tr}(\dots \tilde{X}_{\tau_i} X_{\tau_{i+1}} \dots) \Big], \tag{2.49}$$

which vanishes in a telescoping fashion on carrying out the summation. This completes the proof.

2.3.3 Scalar matrix reduction relations

The requirement that there exist auxiliary matrices $\tilde{X}_\tau$ satisfying (2.48) allows in principle any two-species process. However, without further simplifications, this can be too cumbersome for explicit calculations.

A particularly useful special case, which turns out to be sufficient for many important examples, is when the auxiliary matrices $\tilde{X}_\tau$ are scalars, x_τ. This special case was analysed exhaustively in [106]. The relations (2.48) become

$$w_{\tau\tau'} X_\tau X_{\tau'} - w_{\tau'\tau} X_{\tau'} X_\tau = x_{\tau'} X_\tau - x_\tau X_{\tau'}. \tag{2.50}$$

These are called matrix reduction relations because they convert terms involving two matrices into terms with just one matrix. Using these relations, one can systematically reduce products of L matrices to smaller ones, which can be evaluated more easily. This is demonstrated in section 2.3.5. Such algebraic relations also allow one to find explicit representations for the matrices X_τ in certain cases [61, 63, 107, 108], although we will not use them in this thesis.

Now, given some two-species exclusion process, one would like to be able to tell quickly whether a scalar set of matrix product relations exists. There is a simple consistency check one can perform, without having to construct the matrices explicitly [106]. It is derived as follows.

A necessary condition for a set of matrices with scalars x_0, x_1, x_2 to form a consistent description of the steady state is that reductions carried out in any order give the same result. Then we could, for instance, take the state $\{021\}$ and using the relations (2.50) relate it to the state $\{120\}$ in two different ways: $\{021\} \to \{012\} \to \{102\} \to \{120\}$ and $\{021\} \to \{201\} \to \{210\} \to \{120\}$.

Equating the terms produced by these sequences gives,

$$\begin{aligned} x_1 w_{02}(w_{12} - w_{21} - w_{10} + w_{01}) X_0 X_2 + x_2 w_{01}(w_{12} - w_{21} + w_{20} - w_{02}) X_0 X_2 \\ + x_0 w_{21}(w_{20} - w_{02} - w_{10} + w_{01}) X_2 X_1 + x_1 x_2 (w_{12} - w_{21} + w_{20} - w_{10}) X_0 \\ + x_1 x_0 (w_{21} - w_{02}) X_2 + x_2 x_0 (w_{10} - w_{12} - w_{20} + w_{02}) X_1 = 0. \end{aligned} \tag{2.51}$$

Then requiring the prefactor of each term to vanish gives,

$$x_1 w_{02}(w_{12} - w_{21} - w_{10} + w_{01}) = 0, \tag{2.52a}$$

$$x_2 w_{01}(w_{12} - w_{21} + w_{20} - w_{02}) = 0, \tag{2.52b}$$

$$x_0 w_{21}(w_{20} - w_{02} - w_{10} + w_{01}) = 0, \tag{2.52c}$$

$$x_1 x_2 (w_{12} - w_{21} + w_{20} - w_{10}) = 0, \tag{2.52d}$$

$$x_1 x_0 (w_{21} - w_{02}) = 0, \tag{2.52e}$$

$$x_2 x_0 (w_{10} - w_{12} - w_{20} + w_{02}) = 0. \tag{2.52f}$$

This can be used to generate a full classification of two-species exclusion processes that can be solved using a matrix product with simple reduction relations. This classification is given in [106]. We do not reproduce it here as most models contained therein will not be discussed in this thesis. Instead, we demonstrate how one might use these restrictions on a simple example of two species with no overtaking.

Consider a system in which particles of species 1 and 2 are both totally asymmetric with hopping rates 1 and α, and no overtaking. We can represent this as follows,

$$10 \xrightarrow{1} 01 \quad ; \quad 20 \xrightarrow{\alpha} 02.$$

With our choice of rates, we have $w_{01} = w_{01} = w_{12} = w_{21} = 0$, $w_{10} = 1$ and $w_{20} = \alpha$. Then (2.52a), (2.52b), (2.52c) and (2.52e) are automatically satisfied. To satisfy (2.52d) and (2.52f), we must either set $x_2 = 0$ or $x_1 = x_0 = 0$. If we try the latter option, we discover that $X_1 X_0 = 0$. This cannot be true, as it would make most weights vanish. Instead, we set $x_2 = 0$. Actually, the absence of overtaking, also requires $x_1 = 0$, which can be seen by considering (2.50) with $\tau = 1$, $\tau' = 2$. Then x_0 remains a free parameters. It is convenient to choose $x_0 = 1$. Putting this into (2.50), we obtain

$$X_1 X_0 = X_1, \qquad X_2 X_0 = \alpha^{-1} X_2. \tag{2.53}$$

2.3.4 Normalisation

One advantage of the matrix product formulation is that it makes it intuitive to write down local correlation functions. For this, it is convenient to introduce the following matrix,

$$C(z_1, z_2) = X_0 + z_1 X_1 + z_2 X_2, \tag{2.54}$$

where z_1, z_2 are dummy variables, which can also be interpreted as the fugacities of particles of species 1 and 2 respectively. We can use this matrix to denote a site whose occupancy we wish to leave unrestricted, which makes it easy to enumerate system configurations. For instance, we can write a "nonequilibrium partition function" (normalisation), which enumerates the weights of all allowed configurations as follows,

$$Z_{L,M_1,M_2} = \{z_1^{M_1} z_2^{M_2}\} \mathrm{tr}(C(z_1, z_2)^L), \tag{2.55}$$

where M_1, M_2 are the number of particles of species 1 and 2 respectively, and the notation $\{z^k\}$ denotes that we must take the coefficient of z^k in the power series that follows. We can rewrite this as a complex contour integral using Cauchy's residue theorem, similar to the generating function trick 2.B,

$$Z_{L,M_1,M_2} = \oint \frac{\mathrm{d}z_1}{2\pi \mathrm{i}} \oint \frac{\mathrm{d}z_2}{2\pi \mathrm{i}} \frac{\mathrm{tr}(C(z_1, z_2)^L)}{z_1^{M_1+1} z_2^{M_2+1}}, \tag{2.56}$$

where the contours of integration are, again, small circles around $z_1 = 0$ and $z_2 = 0$ respectively. To proceed, one needs to simplify this using the matrix reduction relations. How specifically this works must be dealt with on a case-by-case basis, as different reduction relations will generally require very different approaches.

We give a demonstration on a simple case. Consider, again, the TASEP with a defect and no overtaking, as in section 2.3.3, but now with only a single particle of species 2 (*i.e.* $M_2 = 1$). Due to translational invariance of the steady state, we can fix the single species 2 particle to be on site 1 and set $z_2 = 0$ for the remainder of the system. In this reference frame, the position of the defect particle is always on site 1 and the remaining particles move on sites $2, \ldots, L$. When the defect hops, we keep its position as site 1 and instead move all other particles in the

opposite direction by one site. Then we have

$$Z_{L,M_1,1} = \oint \frac{\mathrm{d}z_1}{2\pi\mathrm{i}} \frac{\mathrm{tr}(X_2 C(z_1,0)^{L-1})}{z_1^{M_1+1}}. \tag{2.57}$$

Applying the reduction relations (2.53), we have,

$$X_2 C(z_1,0) = X_2(X_0 + z_1 X_1) = \alpha^{-1} X_2 + z_1 X_2 X_1. \tag{2.58}$$

This gives,

$$Z_{L,M_1,1} = \alpha^{-1} Z_{L-1,M_1,1} + \oint \frac{\mathrm{d}z_1}{2\pi\mathrm{i}} \frac{\mathrm{tr}(X_2 X_1 C(z_1,0)^{L-2})}{z_1^{M_1}}. \tag{2.59}$$

Now note that the reduction relations (2.53) imply

$$X_1 C(z_1,0)^{L-2} = (1 + z_1 X_1)^{L-2} X_1. \tag{2.60}$$

Then we evaluate the contour integral by expanding the binomial and reading off the coefficient of $z_1^{M_1-1}$. Hence, we obtain,

$$Z_{L,M_1,1} = \alpha^{-1} Z_{L-1,M_1,1} + \binom{L-2}{M_1-1} \mathrm{tr}(X_2 X_1^{M_1}). \tag{2.61}$$

We can now repeat this step to similarly express $Z_{L-1,M_1,1}$ in terms of $Z_{L-2,M_1,1}$. This can be continued until we reach $Z_{M_1-1,M_1,1}$, which vanishes, as all terms in the integrand have a factor of z^k with $k \leq -2$. All terms produced by this expansion will have a factor of $\mathrm{tr}(X_2 X_1^{M_1})$, which we may set to unity without loss of generality. This then allows us to obtain

$$Z_{L,M_1,1} = \sum_{j=0}^{L-M_1-1} \binom{L-2-j}{M_1-1} \alpha^{-j}. \tag{2.62}$$

2.3.5 Correlation functions

Now we can write n-point functions simply by putting matrices representing particles of the desired species on the desired location and Cs elsewhere. For instance, the two-point function $\langle \delta_{\tau_i,1} \delta_{\tau_{i+k},1} \rangle$ can be written as,

$$\langle \delta_{\tau_i,1} \delta_{\tau_{i+k},1} \rangle = \frac{\{z_1^{M_1-2} z_2^{M_2}\} \mathrm{tr}(X_1 C^{k-1} X_1 C^{L-k-1})}{\{z_1^{M_1} z_2^{M_2}\} \mathrm{tr}(C^L)}. \tag{2.63}$$

Although we have an exact formal solution, obtaining an explicit expression may not be trivial. Generally, if one has matrix reduction relations, (2.50), one will try to use them to re-express the numerator in a form similar to a normalisation and then obtain a finite-size expression for or extract the asymptotic behaviour of the normalisation.

Again, consider the TASEP with a single slow defect. We again use the reference frame in which the defect is always on site 1. The easiest two-point function to examine is the probability that site i to the right of the species 2 particle is occupied. This is given by,

$$\langle \eta_i \rangle_{L,M_1,1} = Z^{-1}_{L,M_1,1} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{\mathrm{tr}(X_2 C^{i-1} X_1 C^{L-1-i})}{z^{M_1}}. \tag{2.64}$$

Using a similar procedure to the calculation of $Z_{L,M_1,1}$, we can express this as,

$$\langle \eta_i \rangle_{L,M_1,1} = Z^{-1}_{L,M_1,1} \left[\binom{L-1-i}{M_1-1} \alpha^{-(i-1)} + \sum_{j=0}^{i-2} \binom{L-3-j}{M_1-2} \alpha^{-j} \right]. \tag{2.65}$$

It can actually be shown via asymptotic analysis that in the limit $L \to \infty$, with $\rho = M/L$ held fixed that there is a phase transition at $\alpha = (1-\rho)$. For α above this value, the density profile is uniform. For α below this value, the density profile becomes a shock with a region of density $(1-\alpha)$ behind the defect and a region of density 0 in front of it. This is illustrated in figure 2.3.

2.3.6 Fluxes

Another result one obtains immediately are net fluxes of a given process through a bond. Taking a process with scalar reduction relations, consider, for instance, the net rate of the process $10 \to 01$, Φ_{10}, at some location in the system. Then using the matrix product formulation, its expectation can be written as

$$\langle \Phi_{10} \rangle = \frac{w_{10}\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_1 X_0 C^{L-2}) - w_{01}\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_0 X_1 C^{L-2})}{Z_{L,M_1,M_2}}. \tag{2.66}$$

Using the reduction relations (2.50), we can write this as

$$\langle \Phi_{10} \rangle = \frac{x_0\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_1 C^{L-2}) - x_1\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_0 C^{L-2})}{Z_{L,M_1,M_2}}. \tag{2.67}$$

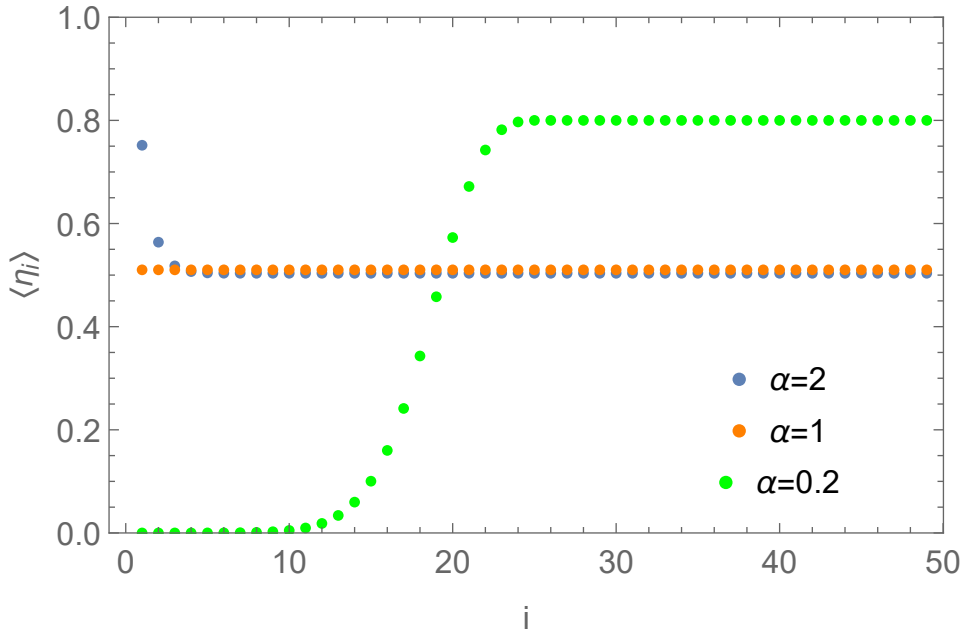

Figure 2.3 *Density profile as given by* (2.65)*. The parameters used were* $L = 50$*,* $M = 25$*. At this mean density,* $\rho = 0.5$*, in the infinite system size limit, the system experiences a phase transition at* $\alpha = 0.5$*. It can be seen that for* $\alpha > 0.5$*, the profile is mostly uniform, whereas for* $\alpha < 0.5$*, it becomes a shock.*

Now we note that the two terms in the numerator are related to the mean densities of 1s and 0s in systems with parameters $\{L-1, M_1, M_2\}$ and $\{L-1, M_1-1, M_2\}$ respectively,

$$\langle \delta_{\tau_i,1} \rangle_{L-1,M_1,M_2} = \frac{\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_1 C^{L-2})}{Z_{L-1,M_1,M_2}} = \frac{M_1}{L-1} \tag{2.68a}$$

$$\langle \delta_{\tau_i,0} \rangle_{L-1,M_1-1,M_2} = \frac{\{z_1^{M_1-1} z_2^{M_2}\} \mathrm{tr}(X_0 C^{L-2})}{Z_{L-1,M_1-1,M_2}} = \frac{L-M_1-M_2}{L-1}, \tag{2.68b}$$

where the rightmost expressions follow from the translational invariance of the system. This allows us to write

$$\langle \Phi_{10} \rangle = x_0 \frac{M_1}{L-1} \frac{Z_{L-1,M_1,M_2}}{Z_{L,M_1,M_2}} - x_1 \frac{L-M_1-M_2}{L-1} \frac{Z_{L-1,M_1-1,M_2}}{Z_{L,M_1,M_2}}. \tag{2.69}$$

Similar to section 2.2.3, we obtain expressions that resemble derivatives of the logarithm of a normalisation.

2.4 Bethe ansatz

We now present the last technique that we will use to perform exact calculations in steady states of asymmetric exclusion processes: the Bethe ansatz. At its core, this is a method for diagonalising certain linear operators. In principle, the Bethe ansatz allows one to calculate the steady-state measure. However, we will see that the nature of the solution obtained is not particularly convenient for this. Instead, through a generating function construction, it will allow us to calculate the scaled cumulant generating functions of particle displacement, with time serving as the large parameter. We also mention that the Bethe ansatz has been used in another way in the context of the ASEP, namely to calculate the gap between the eigenvalues with the largest and second-largest real part of the Markov operator. This has been done for instance for the TASEP on a ring [91], the PASEP on a ring [92], the TASEP with open boundaries [109, 110] and the PASEP with open boundaries [111]. However, in this thesis we will only use it to investigate the large deviations of particle displacement statistics. This section mainly follows [10].

The Bethe ansatz was first introduced to exactly diagonalise the quantum Heisenberg spin chain Hamiltonian [9]. It has since been realised that the same ansatz also works for many other linear operators, including the Hamiltonians of one-dimensional quantum gases [112, 113], transfer matrices of two-dimensional classical equilibrium lattice models [114, 115] and the Markov operators of exclusion processes [90, 91, 116]. The mathematical reason this approach works in the latter case is that Markov operators of exclusion processes can be written as tensor products of local operators that represent the rates of processes involving two sites. These can be expressed in terms of Pauli matrices, which are used to express Hamiltonians of quantum spin chains. Thus the two problems are algebraically very similar.

2.4.1 Heuristic explanation

An intuitive, physical way to think about the Bethe ansatz is that it describes systems in which particles in some sense behave mostly as free particles, with two-body collisions that change the internal state of the system but preserve the values of momenta. Suppose we are considering a system with M particles and

denote the state of the system by their positions, $\{x_1, \ldots, x_M\}$. We may then write an ansatz for an eigenstate of the time evolution operator consisting of plane waves for each particle,

$$\psi(x_1, \ldots, x_M) \sim \prod_{i=1}^{M} z_i^{x_i}, \tag{2.70}$$

where we have used the notation $z_i = \mathrm{e}^{\mathrm{i}k_i}$. The variables z_i are called Bethe roots and k_i Bethe momenta. The latter are more commonly used in the quantum context, but the former are more common in the exclusion process context, so we will use that notation in this thesis.

According to our assumption, after scattering, the only change that can occur is that the momenta k_i become permuted among the particles and there is some phase change. Thus for a general state, we include the possibility of all permutations of k_i with some unknown amplitudes,

$$\psi(x_1, \ldots, x_M) = \sum_{\sigma} A_\sigma \prod_{i=1}^{M} z_{\sigma(i)}^{x_i}, \tag{2.71}$$

where the notation σ indicates all permutations of the integers $i = 1, \ldots, M$. This is the basic form of the Bethe ansatz. If the Bethe ansatz diagonalises the linear operator, $\mathcal{M}$, one is able to express its eigenvalues as purely a function of the Bethe roots,

$$\mathcal{M}\psi(x_1, \ldots, x_M) = \lambda(z_1, \ldots, z_M)\psi(x_1, \ldots, x_M). \tag{2.72}$$

The Bethe ansatz method proceeds as follows. By examining the effect of the operator $\mathcal{M}$ on a vector of the form (2.71), one is able to obtain a set of equations, called Bethe equations, for the Bethe roots z_i. Then, solving the Bethe equations, one finds the Bethe roots and thus the spectrum of $\mathcal{M}$.

We remark that even once this is obtained, the form of the eigenvectors themselves, as given by (2.71), is quite complicated, as it involves a sum over $M!$ permutations.

2.4.2 Bethe ansatz solution for TASEP

We now demonstrate how the Bethe ansatz works in practice by deriving the Bethe equations for a simple case: the totally asymmetric simple exclusion process (TASEP) on a ring. This solution was first fully presented in [91] (though the applicability of the Bethe ansatz on this model was already pointed out in [90, 116]).

We consider an L-site ring with M particles hopping with unit rate. As before, we denote the state of the system by the positions of the particles, $\{x_1, \ldots, x_M\}$. Then we can write the master equation (see section 1.3.4) in the following form,

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t(x_1, \ldots, x_M) = \sum_{i=1}^{M} [-P_t(\ldots x_i \ldots) + P_t(\ldots x_i - 1 \ldots)], \quad (2.73a)$$

$$P_t(1, x_2, \ldots, x_M) = P_t(x_2, \ldots, x_M, L+1). \quad (2.73b)$$

The first equation contains the evolution of the system in configurations in which no two particles are on adjacent sites. This can be thought of as the time evolution of a system of free particles. It must be supplemented with a restriction that enforces the exclusion rule. How exactly this manifests will be shown shortly. The second equation imposes periodic boundary conditions.

Assuming that a state vector ψ of the Bethe ansatz form (2.71) is an eigenstate of the master equation (2.72) for some eigenvalue λ, the master equation becomes,

$$\lambda = -M + \sum_{i=1}^{M} z_i^{-1}, \quad (2.74a)$$

$$A_{\ldots ji \ldots} = -\frac{1 - z_j}{1 - z_i} A_{\ldots ij \ldots}, \qquad i, j = 1, \ldots, M \quad (2.74b)$$

$$A_{i \ldots j} = A_{\ldots ji} z_i^L, \qquad i = 1, \ldots, M. \quad (2.74c)$$

The first equation expresses the eigenvalue in terms of the Bethe roots. Importantly, note that it is given solely in terms of the Bethe roots and does not include the amplitudes. It may be easily derived by assuming that the configuration of the system is such that no two particles are on adjacent sites.

The second equation essentially ensures that this expression is also valid for configurations in which there are two particles on adjacent sites, given that the exclusion rule must be obeyed. It is derived as follows. In order for (2.74a) to be

valid in configurations in which, for instance $x_{i+1} = x_i + 1$, we must have for all pairs of indices j, k,

$$\begin{aligned}(-2 + z_j^{-1} + z_k^{-1})[A_{\dots jk\dots} z_j^{x_i} z_k^{x_i+1} + A_{\dots kj\dots} z_k^{x_i} z_j^{x_i+1}] \\ = -[A_{\dots jk\dots} z_j^{x_i} z_k^{x_i+1} + A_{\dots kj\dots} z_k^{x_i} z_j^{x_i+1}] \\ +[A_{\dots jk\dots} z_j^{x_i-1} z_k^{x_i+1} + A_{\dots kj\dots} z_k^{x_i-1} z_j^{x_i+1}].\end{aligned} \tag{2.75}$$

Simplifying and rearranging this yields (2.74b). These relations effectively incorporate the exclusion interaction into the ansatz and are therefore sometimes called collision relations.

An alternative, "quick and dirty" way to derive these relations is to say that one requires the ansatz to formally satisfy the eigenvalue problem,

$$\lambda(z_1, \dots, z_M)\psi(x_1, \dots, x_M) = \sum_{i=1}^{M} [-\psi(\dots x_i \dots) + \psi(\dots x_i - 1 \dots)], \tag{2.76}$$

in all configurations, even when two particles are on adjacent sites. For this to be consistent with the simple exclusion rule, one requires the cancellation of all terms corresponding to "forbidden" processes, for instance for $x_{i+1} = x_i + 1$,

$$-\psi(\dots x_i, x_i + 1 \dots) + \psi(\dots x_i, x_i \dots) = 0, \tag{2.77}$$

although, of course, the terms involved in this equation are not physically meaningful. Putting in the ansatz (2.71), and requiring for the cancellation to hold for any pair of indices, one obtains (2.74b).

The third equation incorporates the periodic boundary conditions.

As the eigenvalue is expressed solely in terms of the Bethe roots, we seek to combine these equations in such a way that the amplitudes are eliminated and we are left with a self-consistent set of equations for the Bethe roots. Applying (2.74b) to the right-hand side of (2.74c) $M - 1$ times to permute the index i to the first position, we obtain

$$A_{i\dots j} = (-)^{M-1} z_i^L \left(\prod_{k \neq i} \frac{1 - z_k}{1 - z_i} \right) A_{i\dots j}. \tag{2.78}$$

Cancelling the amplitudes on either side and inserting the harmless term $-(1 -$

$z_i)/(1-z_i) = -1$, we obtain the Bethe equation,

$$(1-z_i)^M = Cz_i^L, \tag{2.79}$$

where C is a constant,

$$C = (-)^M \prod_{j=1}^{M}(1-z_j). \tag{2.80}$$

This is a polynomial equation for z_i of degree L. One needs to solve it and choose M roots, which can be plugged into (2.74a) to obtain $\binom{L}{M}$ eigenvalues. This is reassuring, as that is the dimension of the state space of this system.

Note that the Bethe equation (2.79) contains the constant C, which in turn depends on the Bethe roots through (2.80). It can be fixed using the global periodicity condition, which comes from the fact that moving the whole system around the ring once should not change the state. To express this algebraically, one can apply (2.74c) M times (once for each index) and cancel the amplitudes to obtain,

$$z_1^L \dots z_M^L = 1. \tag{2.81}$$

Thus the product of all Bethe roots must equal some L-th root of unity, $\mathrm{e}^{2\pi \mathrm{i} k/L}$, for $k = 0, 1, 2, \dots, L-1$. The choice of root of unity is analogous to choosing the total momentum of a quantum system (recall that $z_i = \mathrm{e}^{\mathrm{i}k_i}$). Choosing different roots of unity will give different eigenvalues. In particular, for the "ground state" eigenvalue, one must choose the "zero momentum" root, 1.

2.4.3 Bethe ansatz solution for large deviation function of TASEP current

In principle, the Bethe ansatz may allow for a full computation of all eigenvalues and eigenvectors of a linear operator (although there are some known cases in which the Bethe ansatz does not allow one to construct all eigenvectors, for instance [117]). However, in practice it can be very difficult to extract explicit expressions for physical observables, even if the Bethe equations are given. In the context of exclusion processes, an approach that has proved very fruitful is to consider a deformation of the Markov operator. One is able to obtain the

long time particle displacement statistics by evaluating just one eigenvalue. This approach was first used for the TASEP in [10]. It has since then also been used for other versions of ASEPs, including TASEPs with slow defect particles and overtaking [12, 118], PASEPs on a ring [119, 120] and PASEPs with open boundaries [121].

This construction relies on the fact that the particle displacement at long times satisfies a large deviation principle (see section 1.4) with time acting as the large parameter. The result is an expression for the scaled cumulant generating function of particle displacement. As this is related to the rate function by a Legendre transform (see section 1.5.4), it is commonly referred to as the rate function in this context. We demonstrate this approach on the TASEP on a ring.

First, recall the master equation, which we can write as,

$$\frac{\mathrm{d}P_t(\mathcal{C})}{\mathrm{d}t} = \sum_{\mathcal{C}'} \mathcal{M}(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}'), \tag{2.82}$$

where we have used the shorthand $\mathcal{C}$ to denote the configuration of the system.

Let $Y(t)$ be a variable that counts the total number of jumps of all particles up to time t. To track its evolution, we write down the master equation for the joint probability distribution of $Y(t)$ with the configuration of the system, $P_t(\mathcal{C}, Y)$. To obtain the evolution of this joint probability distribution, it is convenient to split the Markov operator into parts as follows,

$$\mathcal{M} = \mathcal{M}_0 + \mathcal{M}_1, \tag{2.83}$$

where $\mathcal{M}_1(\mathcal{C},\mathcal{C}') = k(\mathcal{C}',\mathcal{C})$ and $\mathcal{M}_0(\mathcal{C},\mathcal{C}') = -\delta_{\mathcal{C},\mathcal{C}'}\sum_{\mathcal{C}''\neq\mathcal{C}} k(\mathcal{C},\mathcal{C}'')$. To clarify the purpose of this partition, we note that subspaces of states with the same value of Y are closed under application of $\mathcal{M}_0$, whereas $\mathcal{M}_1$ connects the subspace of states with $Y-1$ to that of Y. Then we can write the evolution equation of the joint probability distribution as

$$\frac{\mathrm{d}P_t(\mathcal{C},Y)}{\mathrm{d}t} = \sum_{\mathcal{C}'}[\mathcal{M}_0(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}',Y) + \mathcal{M}_1(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}',Y-1)]. \tag{2.84}$$

Now we define the generating function of $Y(t)$,

$$F_t(\mathcal{C},\gamma) = \sum_{Y=0}^{\infty} \mathrm{e}^{\gamma Y} P_t(\mathcal{C},Y). \tag{2.85}$$

Note that marginalising it over configurations gives the moment generating function of $Y(t)$,

$$\langle \mathrm{e}^{\gamma Y(t)} \rangle = \sum_{\mathcal{C}} F_t(\mathcal{C}, \gamma). \tag{2.86}$$

Multiplying the master equation of the joint probability distribution by $\mathrm{e}^{\gamma Y}$ and summing over Y, we obtain the evolution equation for the generating function,

$$\frac{\mathrm{d}F_t(\mathcal{C}, \gamma)}{\mathrm{d}t} = \sum_{\mathcal{C}'} [\mathcal{M}_0(\mathcal{C}, \mathcal{C}') + \mathrm{e}^{\gamma} \mathcal{M}_1(\mathcal{C}, \mathcal{C}')] F_t(\mathcal{C}', \gamma). \tag{2.87}$$

This has the same form as a master equation, but the Markov operator has been "deformed" by the parameter γ. This new time evolution operator,

$$\mathcal{M}_\gamma = \mathcal{M}_0 + \mathrm{e}^{\gamma} \mathcal{M}_1, \tag{2.88}$$

is called the deformed Markov operator. Note that in the limit $\gamma \to 0$, we recover the original (undeformed) Markov operator.

In the long time limit, $F_t(\mathcal{C}, \gamma)$ is dominated by the eigenvector corresponding to the largest eigenvalue of $\mathcal{M}_\gamma$, which we denote $\lambda(\gamma)$ (this is the only eigenvalue which will be important in the following, so this notation should not cause ambiguity). Denoting this eigenvector as $F_\infty(\mathcal{C}, \gamma)$, we have,

$$F_t(\mathcal{C}, \gamma) \sim \mathrm{e}^{\lambda(\gamma) t} F_\infty(\mathcal{C}, \gamma), \qquad \text{as} \quad t \to \infty. \tag{2.89}$$

Note that $\lambda(\gamma)$ is the eigenvalue that converges to the unique 0 eigenvalue of the undeformed Markov operator in the limit $\gamma \to 0$.

From (2.86), it follows that,

$$\lambda(\gamma) = \lim_{t \to \infty} \frac{\log \langle \mathrm{e}^{\gamma Y(t)} \rangle}{t}. \tag{2.90}$$

This means that the eigenvalue $\lambda(\gamma)$ can be identified with the scaled cumulant generating function of $Y(t)$ in the limit $t \to \infty$. Then, for instance, we can

express the mean steady-state current and diffusion constant as

$$J = \lim_{t\to\infty} \frac{\langle Y(t)\rangle}{t} = \left.\frac{\mathrm{d}\lambda(\gamma)}{\mathrm{d}\gamma}\right|_{\gamma=0} \tag{2.91a}$$

$$\Delta = \lim_{t\to\infty} \frac{\langle Y(t)^2\rangle - \langle Y(t)\rangle^2}{t} = \left.\frac{\mathrm{d}^2\lambda(\gamma)}{\mathrm{d}\gamma^2}\right|_{\gamma=0}. \tag{2.91b}$$

We can now use the Bethe ansatz to diagonalise the deformed Markov operator $\mathcal{M}_\gamma$, find its largest eigenvalue and thus obtain the exact long time particle displacement statistics.

The procedure to obtain the Bethe equations is identical with that in the previous section, with the only difference being that it is convenient to modify the ansatz in the following way,

$$\psi(x_1,\ldots,x_M) = \left(\prod_{i=1}^{M} \mathrm{e}^{\gamma x_i}\right) \sum_\sigma A_\sigma \prod_{i=1}^{M} z_{\sigma(i)}^{x_i}. \tag{2.92}$$

Then the equations (2.74a) and (2.74b) are identical and (2.74c) takes the form,

$$A_{i\ldots j} = \mathrm{e}^{L\gamma} A_{\ldots ji} z_i^L, \qquad i = 1,\ldots,M. \tag{2.93}$$

Then (2.81) becomes

$$\mathrm{e}^{ML\gamma} z_1^L \ldots z_M^L = 1. \tag{2.94}$$

As noted in the previous section, to obtain the "ground state" eigenvalue, we must choose the zero momentum root. This gives,

$$\mathrm{e}^{M\gamma} \prod_{i=1}^{M} z_i = 1. \tag{2.95}$$

Repeating the steps from the previous section to derive the Bethe equation, we obtain an equation of the same form as (2.79), with the only difference that the constant C is now given by,

$$C = (-)^M \mathrm{e}^{L\gamma} \prod_{j=1}^{M} (1 - z_j). \tag{2.96}$$

The Bethe equation is to be used together with (2.95) to solve for the Bethe roots

z_i. Once this is done, one needs to put them into (2.74a) to find $\lambda(\gamma)$.

2.4.4 Computation of cumulants

Although the task at hand is clear, it is still a nontrivial exercise to perform it, as one needs to solve a polynomial equation of a large degree with an unknown constant C. We also note that solving explicitly for all Bethe roots z_i would be overkill, as the expressions for the eigenvalue (2.74a) only involves certain symmetric combinations of them. A particularly elegant way to handle this problem is using a functional method. This approach also highlights the connection to the more difficult partially asymmetric case, which will be discussed in chapter 5.

Before we proceed with the functional approach, we observe that in the limit $\gamma \to 0$, the solution of the Bethe equation (2.79) and the "zero momentum" condition (2.95) is given by $z_i = 1$ for all i. As we slowly increase γ, the roots will evidently remain close to 1, with some distance that depends on γ. In particular, from (2.96) we see that this implies that $C \to 0$ as $\gamma \to 0$. Indeed, it turns out that $C = O(\gamma)$ and it will turn out to be convenient to build perturbative expansions using C, rather than γ as the perturbative parameter.

To reformulate the problem as a functional problem, we first introduce the slight change of variable

$$y_i = 1 - z_i. \tag{2.97}$$

Although in this case, this is mostly cosmetic, a generalisation of such a transformation will be crucial in the PASEP case (which will be discussed in chapter 5).

In terms of this new variable, the Bethe equation (2.79) becomes

$$(1 - y_i)^L C = y_i^M, \tag{2.98}$$

the "zero momentum" condition (2.95) becomes

$$\mathrm{e}^{M\gamma} \prod_{i=1}^{M} (1 - y_i) = 1, \tag{2.99}$$

and the equation for the rate function λ (2.74a) becomes,

$$\lambda = -M + \sum_{i=1}^{M} \frac{1}{1-y_i}. \tag{2.100}$$

A key role is played by the function

$$h(y) = \frac{(1-y)^L}{y^M}. \tag{2.101}$$

Indeed, from (2.98), we see that the Bethe roots $y_i = 1 - z_i$ are solutions of the equation

$$h(y)C = 1. \tag{2.102}$$

Now recall that the Bethe roots z_i are located close to $z = 1$. This means that after the change of variable, y_i are close to $y = 0$. Then we can exploit Cauchy's residue theorem to evaluate an arbitrary test function f at the Bethe roots. Specifically, if we take some function f that is analytic at $y = 0$, we have the identity,

$$\oint_\Gamma \frac{\mathrm{d}y}{2\pi\mathrm{i}} f(y) \frac{Ch'(y)}{Ch(y)-1} = -Mf(0) + \sum_{k=1}^{M} f(y_i), \tag{2.103}$$

where Γ is a small circle around the origin. The proof is a simple exercise in complex analysis. The intuition behind this identity is that by taking a contour around $y = 0$, we capture the Bethe roots, which are simple poles of $1/(Ch(y)-1)$ that lie close to $y = 0$.

Integrating (2.103) by parts, we obtain,

$$-Mf(0) + \sum_{k=1}^{M} f(y_i) = -\oint_\Gamma \frac{\mathrm{d}y}{2\pi\mathrm{i}} f'(y) \log[1 - Ch(y)]. \tag{2.104}$$

Then expanding the log in powers of C (recall that $C = O(\gamma)$, so this expansion is justified in the $\gamma \to 0$ limit), we finally obtain the following useful recipe,

$$-Mf(0) + \sum_{k=1}^{M} f(y_i) = \sum_{n=1}^{\infty} \frac{C^n}{n} \oint_\Gamma \frac{\mathrm{d}y}{2\pi\mathrm{i}} f'(y)[h(y)]^n. \tag{2.105}$$

We now apply this to (2.99) with $f(y) = \log(1-y)$ and to (2.100) with $f(y) = 1/(1-y)$ to obtain,

$$\gamma = \sum_{n=1}^{\infty} C^n \gamma^{(n)}, \qquad \lambda(\gamma) = \sum_{n=1}^{\infty} C^n \lambda^{(n)}, \tag{2.106}$$

where

$$\gamma^{(n)} = M^{-1} \frac{1}{n} \oint_\Gamma \frac{\mathrm{d}y}{2\pi \mathrm{i}} \frac{(1-y)^{Ln-1}}{y^{Mn}} \tag{2.107a}$$

$$\lambda^{(n)} = \frac{1}{n} \oint_\Gamma \frac{\mathrm{d}y}{2\pi \mathrm{i}} \frac{(1-y)^{Ln-2}}{y^{Mn}}. \tag{2.107b}$$

From this, one has to eliminate C between the two series order by order to obtain the cumulants. Inverting the γ expansion in (2.106) to second order, we have

$$C = \frac{\gamma}{\gamma^{(1)}} + \frac{\gamma^{(2)}}{\gamma^{(1)}} \left(\frac{\gamma}{\gamma^{(1)}} \right)^2 + \dots \tag{2.108}$$

We can use this to rewrite the expansion of λ in powers of C (2.106) as an expansion in powers of γ instead. Using (2.91a) and (2.91b), we obtain to second order,

$$J = \frac{\lambda^{(1)}}{\gamma^{(1)}}, \qquad \Delta = 2 \frac{\lambda^{(2)}\gamma^{(1)} - \lambda^{(1)}\gamma^{(2)}}{(\gamma^{(1)})^3}. \tag{2.109}$$

Finite-size expressions

One way to proceed is to identify the integrands in (2.107a) and (2.107b) as the generating functions of the binomial coefficients, giving

$$\gamma^{(n)} = (-)^{Mn-1} \frac{(Ln-1)!}{(Mn)!(Ln-Mn)!} \tag{2.110a}$$

$$\lambda^{(n)} = (-)^{Mn-1} \frac{M(Ln-2)!}{(Mn)!(Ln-Mn-1)!} \tag{2.110b}$$

This can be used to obtain exact finite-size expression.

For the current we obtain,

$$J = \frac{M(L-2)!(L-M)!}{(L-1)!(L-M-1)!} = \frac{M(L-M)}{L-1}. \tag{2.111}$$

Similarly, for the diffusion constant we obtain,

$$\begin{aligned}\Delta &= -2M\frac{\frac{(2L-2)!}{(2M)!(2L-2M-1)!}\frac{(L-1)!}{(M)!(L-M)!}-\frac{(L-2)!}{(M)!(L-M-1)!}\frac{(2L-1)!}{(2M)!(2L-2M)!}}{\left(\frac{(L-1)!}{(M)!(L-M)!}\right)^3}\\ &= \frac{M(M!)^2((L-M)!)^2(2L-3)!}{(2M)!((L-1)!)^2(2L-2M-1)!}\end{aligned} \tag{2.112}$$

Asymptotic expressions

We can also extract the asymptotic behaviour in the large system size limit $L \to \infty$, with the mean density $\rho = M/L$ fixed using the saddle point method (see 2.A). Taking, say, the integral in (2.107a), we can write it as,

$$\oint_\Gamma \frac{\mathrm{d}y}{2\pi\mathrm{i}}\frac{1}{1-y}\exp[Ln(\log(1-y)-\rho\log y)]. \tag{2.113}$$

This has a saddle point at $y_{\mathrm{sp}} = -\rho/(1-\rho)$. By inspection, it is evident that the integrals in (2.107b) will have the same saddle point, with the only difference being the prefactors. Using the standard expressions for saddle point integration (including the first correction for Δ), we obtain to leading order,

$$J \approx L\rho(1-\rho), \qquad \Delta \approx L^{3/2}\frac{\sqrt{\pi}}{2}[\rho(1-\rho)]^{3/2}. \tag{2.114}$$

2.5 Summary

In this chapter, we have introduced three methods for performing exact calculations in steady states of SEPs. We will use these methods in the remaining chapters to study various special cases of two-species ASEPs. Specifically, we will use the ZRP mapping in chapter 3, the matrix product ansatz in chapters 3 and 4, and the Bethe ansatz in chapters 5 and 6.

A natural question is which methods are appropriate for a given problem and what the respective advantages and disadvantages are.

Mapping to a ZRP has the advantage over the other techniques in that it is easy to write down without any heavy mathematical machinery. However, when other methods are available, it is not generally preferable, as the measure is local in the

ZRP picture, rather than the SEP picture. This can make calculations of local observables unintuitive and unnecessarily involved.

There are also two important situations in which the ZRP mapping fails. The first is when the particle number can vary, such as in open boundary problems. The mapping evidently does not work as it would require the number of sites in the ZRP to change. Although this excludes a large and important class of models, this drawback will not be of major concern in this thesis, as we will only discuss particle-conserving models with periodic boundary conditions.

The other situation is when overtaking is allowed. Here, the mapping fails as it would require the the interchange of sites in the ZRP. If in the SEP, particle i with a hopping rate p_1 overtakes particle $i+1$ with a hopping rate p_2, we would need to change the rates at which particles hop from site $i+1$ to i in the ZRP from p_2 to p_1. This precludes us from using the ZRP mapping for the model we will consider in chapters 4 and 6.

The matrix product approach involves, on the other hand, involves more formal manipulations, but can, at least in principle, be used for any two-species ASEP. However, in practice, one would like to make use of scalar matrix reduction relations, which somewhat limits the set of applicable models. The matrix product ansatz is more convenient than the ZRP mapping for calculating two-point functions (and higher order correlations) but less tractable than the Bethe ansatz for calculating current statistics beyond the mean.

The Bethe ansatz in principle constructs the steady-state measure but is primarily useful for calculating particle displacement statistics beyond the mean. The question of which two-species ASEPs are solvable using the Bethe ansatz (*i.e.* are integrable) is a complex one and will be addressed in chapter 5.

Another interesting question is the relation between solvability using the matrix product ansatz and the Bethe ansatz. In a few cases, it has been shown that the matrix product solution can be derived from the Bethe ansatz solution, if it exists [122–124]. However, whether such a construction is possible in general remains an open problem.

2.A Saddle point method

We will often need to extract the asymptotic behaviour of an integral of the form

$$I = \int_\gamma \mathrm{d}z\ A(z)\mathrm{e}^{L\phi(z)}, \tag{2.115}$$

where $A(z)$ and $\phi(z)$ are functions, γ is the contour of integration and L is some large parameter. This can often be done using the saddle point method (see, for instance [34]).

We wish to find the asymptotic behaviour of I as we take $L \to \infty$. To do this we exploit the property that the contours of integration for integrals in the complex plane can generally be deformed arbitrarily, provided the end-points remain fixed and one does not cross any singularities. For instance if the contour is closed (as will be the case for all the integrals we consider in this thesis), then I can be evaluated using the residue theorem as,

$$I = 2\pi\mathrm{i}\sum_j \mathrm{Res}(A(z)\mathrm{e}^{L\phi(z)}, z_j), \tag{2.116}$$

where z_j are the poles of the integrand inside the contour γ. Evidently, only the topology of γ relative to the location of the poles of the integrand is important.

Examining the form of the integrand, one sees that in the limit $L \to \infty$, it will generally be dominated by the values close to locations z_* at which

$$\phi'(z_*) = 0. \tag{2.117}$$

These points z_* are known as saddle points. The strategy to obtain the asymptotic behaviour of I is then to deform the original contour γ to pass through the saddle points z_* and Taylor expand around z_*. For simplicity, we consider the case of a unique saddle point and let $w = z - z_*$. Then we have,

$$I = \int_\gamma \mathrm{d}z\ [A(z_*) + wA'(z_*) + \dots]\exp\left(L\phi(z_*) + L\frac{w^2}{2}\phi''(z_*) + \dots\right). \tag{2.118}$$

Then as a first approximation, one may take,

$$I \;=\; A(z_*)\mathrm{e}^{L\phi(z_*)}\int_\gamma \mathrm{d}z\ \mathrm{e}^{L\frac{w^2}{2}\phi''(z_*)}. \tag{2.119}$$

As the parts of the integral away from the saddle point z_* are suppressed exponentially, one may approximate this as a complete Gaussian integral. In doing this, one subtlety is that one needs to choose the correct direction of integration, which ensures that $\phi''(z_*) < 0$ as one crosses the saddle point. This may introduce an overall phase $\mathrm{e}^{\mathrm{i}\varphi}$ to the integral for some $\varphi \in [0, 2\pi)$, though this will not be a cause of major concern in this thesis. Overall, the leading order approximation is

$$I \approx \mathrm{e}^{\mathrm{i}\varphi}\sqrt{\frac{2\pi}{L|\phi''(z_*)|}}A(z_*)\mathrm{e}^{L\phi(z_*)}. \tag{2.120}$$

We may extend this intuition to systematically add finite-size corrections in orders of L^{-1} by keeping more terms in the Taylor series expansion. Specifically, in doing this we Taylor expand the exponential as follows,

$$\begin{aligned} & \mathrm{e}^{L\phi(z_*)+L\frac{w^2}{2}\phi''(z_*)}\exp\left(L\frac{w^3}{6}\phi'''(z_*)+L\frac{w^4}{24}\phi''''(z_*)\dots\right) \\ \approx\ & \mathrm{e}^{L\phi(z_*)+L\frac{w^2}{2}\phi''(z_*)}\left[1+L\frac{w^3}{6}\phi'''(z_*)+L\frac{w^4}{24}\phi''''(z_*)+\dots\right. \\ & \left.+\frac{1}{2}\left(L\frac{w^3}{6}\phi'''(z_*)\right)^2+\dots\right] \end{aligned} \tag{2.121}$$

This needs to be multiplied by the Taylor series expansion of $A(z)$,

$$A(z) = A(z_*)\left[1+w\frac{A'(z_*)}{A(z_*)}+\frac{w^2}{2}\frac{A''(z_*)}{A(z_*)}+\dots\right]. \tag{2.122}$$

This generates many terms. To understand which terms are relevant at a given order in the expansion in L^{-1}, we need to consider the general Gaussian integral formula,

$$\int_{-\infty}^{\infty}\mathrm{d}z\ z^n\mathrm{e}^{-az^2} = \begin{cases} 0, & n \text{ odd} \\ \sqrt{\pi}\left(-\frac{\partial}{\partial a}\right)^{n/2}a^{-1/2}, & n \text{ even} \end{cases}. \tag{2.123}$$

Up to $n = 6$, we have,

$$\int_{-\infty}^{\infty} \mathrm{d}z\ \mathrm{e}^{-az^2} = \sqrt{\pi} a^{-1/2} \tag{2.124a}$$

$$\int_{-\infty}^{\infty} \mathrm{d}z\ z^2 \mathrm{e}^{-az^2} = \frac{\sqrt{\pi}}{2} a^{-3/2} \tag{2.124b}$$

$$\int_{-\infty}^{\infty} \mathrm{d}z\ z^4 \mathrm{e}^{-az^2} = \frac{3\sqrt{\pi}}{4} a^{-5/2} \tag{2.124c}$$

$$\int_{-\infty}^{\infty} \mathrm{d}z\ z^6 \mathrm{e}^{-az^2} = \frac{15\sqrt{\pi}}{8} a^{-7/2}. \tag{2.124d}$$

Thus for each term, one must carefully count the positive powers of L coming from the Taylor expansion (2.121) and the negative powers of L coming from the Gaussian integration. Overall, we find that there are four terms that contribute to the first correction term, which gives,

$$\begin{aligned} I \;\approx\; & \mathrm{e}^{\mathrm{i}\varphi} \sqrt{\frac{2\pi}{L|\phi''|}} A \mathrm{e}^{L\phi} \left[1 + \frac{1}{L} \left(\frac{A'\phi'''}{2A(\phi'')^2} - \frac{A''}{2A\phi''} + \frac{\phi''''}{8(\phi'')^2} \right.\right. \\ & \left.\left. - \frac{5(\phi''')^2}{24(\phi'')^3} \right) + O(L^{-2}) \right]_{z=z_*} . \end{aligned} \tag{2.125}$$

2.B Generating function approach to problems with a global constraint

When analysing problems with a global constraint, it is often convenient to introduce a formal parameter that effectively relaxes the constraint from a hard to a soft one. This relaxation makes it significantly easier to write down global statistics. Eventually, one has to fix the parameter that controls the softness of the constraint, which is achieved through complex contour integration.

In equilibrium statistical mechanics, this corresponds to solving a problem in a canonical ensemble by going through a fictitious grand canonical ensemble (see 9.1 in [14]). In that context, this approach is sometimes called the Darwin-Fowler method. However, we will keep the terminology in this section general to highlight that none of the features of this result are exclusive to equilibrium systems.

2.B.1 Problem statement

We consider a stochastic system with L components, each of which has a state described by $m_i \in \mathbb{Z}$ for $i \in \{1, \ldots, L\}$. Suppose that the states of each component are in some sense independent but we have the global constraint

$$\sum_{i=1}^{L} m_i = M. \tag{2.126}$$

As a simple example, one can consider an ideal gas of noninteracting, indistinguishable particles with L energy levels and M particles.

Then suppose we can write the probability of a state $\mathbf{m} = \{m_1, \ldots, m_L\}$ as

$$P(\mathbf{m}) = Z_{L,M} \prod_{i=1}^{L} f_i(m_i), \tag{2.127}$$

where f_i are some weight functions (which can be different for each component i) and $Z_{L,M}$ is a normalisation. For a noninteracting particle system, these are given by $f_i(m_i) = \mathrm{e}^{-\beta \epsilon_i m_i}$, where ϵ_i is the energy of one particle in energy level i.

The normalisation $Z_{L,M}$ generally allows us to get a handle on the measure $P(\mathbf{m})$. It is given by,

$$Z_{L,M} = \sum_{\mathbf{m}} \prod_{i=1}^{L} f_i(m_i) \delta_{\sum_{i=1}^{L} m_i, M}, \tag{2.128}$$

where the summation is carried out independently over all values of $m_1, \ldots, m_L$, with the constraint (2.126) being enforced by the Kronecker delta. The normalisation can be identified with the partition function in the equilibrium setting. In the nonequilibrium setting, the physical significance is less clear, although it can still be a useful object.

There are two sources of difficulty in working with the expression (2.128). Firstly, it contains an L-fold sum to account for all possible states. Secondly, the global constraint means that all L components are effectively interacting in some nontrivial manner. The generating function approach allows us to deal with these issues.

2.B.2 Solution

First, we define a counting variable, z, whose power we will use to keep track of the sum $\sum_{i=1}^{L} m_i$. In the equilibrium setting, it can be identified with the fugacity, $z = \mathrm{e}^{\beta\mu}$, where, as usual β is the inverse temperature $\beta = (k_{\mathrm{B}}T)^{-1}$ and μ is the chemical potential. Then we define the generating function of $Z_{L,M}$ (grand canonical partition function) as follows,

$$\mathcal{Z}_L(z) = \sum_{M=0}^{\infty} Z_{L,M} z^M. \tag{2.129}$$

We can view this object as the normalisation of a more general probability distribution, in which we allow the parameter M to take any value $0, 1, 2, \ldots$, with probability proportional to z^M. In this more general process, the global constraint has been softened, as any value of M is in principle allowed, just with a penalty, z^M. Now, another valid way to enumerate all configurations is to allow each component i to take any value m_i, provided we attach to it the weight z^{m_i}. In other words, we can write

$$\mathcal{Z}_L(z) = \prod_{i=1}^{L} \sum_{m_i=0}^{\infty} f_i(m_i) z^{m_i}. \tag{2.130}$$

This is generally a much simpler object to handle than $Z_{L,M}$, as it only involves one product and sum, with no global constraint.

Then to recover the original normalisation, we need to ensure that we pick only the terms in which the power of z is precisely M. This can be done using Cauchy's residue theorem. We multiply by $z^{-(M+1)}$, which means that the term with z^M in $\mathcal{Z}_L(z)$ will now have a simple pole at $z = 0$. This allows us to extract its coefficient by performing a complex contour integral along a small circle around the origin in the complex plane. Thus, we obtain the following expression for the original (canonical) normalisation,

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{\mathcal{Z}_L(z)}{z^{M+1}}. \tag{2.131}$$

One might now wish to evaluate the integral to obtain an exact explicit expression for $Z_{L,M}$, or extract asymptotics in the limit $L \to \infty$, $M \to \infty$ with $\rho = M/L$ fixed.

Asymptotic behaviour

The expression (2.131) is convenient for extracting large system size asymptotics. A typical limit is $L \to \infty$, $M \to \infty$, with $\rho = M/L$ fixed. Then we can write

$$\frac{\mathcal{Z}_L(z)}{z^M} = \exp[L\phi(z)], \tag{2.132}$$

where

$$\phi(z) = L^{-1} \sum_{i=1}^{L} \log \left(\sum_{m_i=0}^{\infty} f_i(m_i) z^{m_i} \right) - \rho \log z. \tag{2.133}$$

Then $\phi(z)$ typically becomes independent of L and the asymptotic behaviour of the integral can be derived using the saddle point method, with a saddle point at $\phi'(z_{\rm sp}) = 0$. This gives the following expression

$$Z_{L,M} \approx (2\pi L |\phi''(z_{\rm sp})|)^{-1/2} z_{\rm sp}^{-1} \exp[L\phi(z_{\rm sp})]. \tag{2.134}$$

Chapter 3

Driven tracer in a passive lattice gas

3.1 Introduction

In this chapter, we examine a simple exclusion process on a ring, in which a single totally asymmetric particle drives an otherwise symmetric bath out of equilibrium. Part of the work presented in this chapter has been published in [1].

3.1.1 Background

An interesting generalisation of the asymmetric simple exclusion process 1.5 is to consider a single asymmetric particle in a bath of symmetrically diffusing particles. This can be seen as a minimal model of a driven or active tracer particle in a narrow channel [50]. The symmetric bath represents the (passive) fluid in the channel. The tracer particle is being pushed in one direction with some external driving. See figure 3.1 for a visual aide.

The driving could be achieved, for example, with a uniform electric field if the tracer is charged but the bath is neutral. Alternatively, the tracer could be pushed with a high-precision device like optical tweezers. Yet another option is that it could generate the motion itself using an artificial or biological motor.

Such a setup is a natural probe to investigate the response of the fluid to a small perturbation in the form of the tracer. This method of investigation is known as microrheology [125]. By tracking the motion of the tracer, either

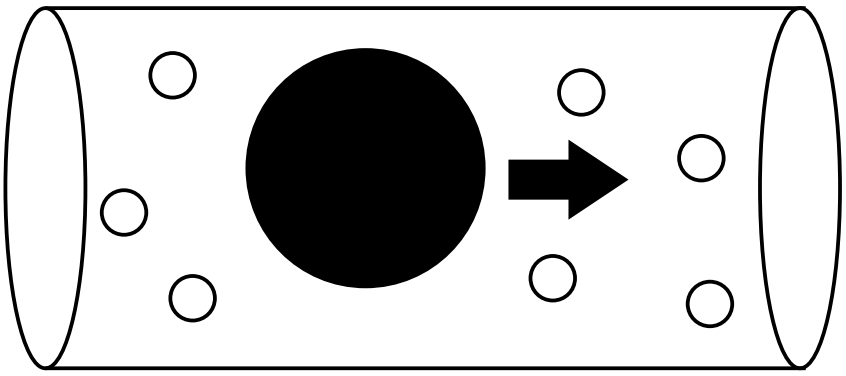

Figure 3.1 *Schematic illustration of driven tracer in a narrow channel.*

due to thermal fluctuations or some applied force, one aims to measure some characteristic properties of the fluid, such as viscosity or hydrodynamic resistance. This is particularly interesting for complex fluids, where response to an external force is generally not described by simple linear stress-strain relations.

We also remark that complex fluids and the motion of small particles through them are ubiquitous in biological systems [126]. Hence an understanding of these processes is crucial for biological applications. The setting of a narrow channel can be found in living matter, for instance in capillary blood vessels.

Lastly, the problem is interesting from the point of view of fundamental nonequilibrium physics. It is not immediately obvious whether this minimal source of activity is enough to drive the whole system out of equilibrium at a macroscopic level or whether the system will relax to a mostly equilibrium-like state with maybe some local perturbations near the driven tracer. We shall see, for the simple model we consider in this chapter, that the former is the case.

3.1.2 Overview of past literature

The first time a model of an asymmetric particle in a system of symmetric particles was examined was in [127]. There, it was included in a survey of various driven exclusion processes to investigate in which case an Einstein-type relation could be said to hold. The Einstein relation relates the mobility μ (ratio of driving force to terminal drift velocity $\mu = v/F$) to the diffusion constant Δ of a particle in a fluid,

$$\mu = \beta\Delta, \tag{3.1}$$

with the standard inverse temperature $\beta = (k_B T)^{-1}$. For an asymmetrically hopping particle, the asymmetry can be interpreted as resulting from a driving by a uniform external force F, which can be related to the ratio of forward and backward hopping probabilities,

$$P_{\text{fwd}}/P_{\text{bckwd}} = \mathrm{e}^{-2\beta F a}, \tag{3.2}$$

where a is the lattice spacing. Thus an Einstein relation holds if

$$\frac{v}{\Delta} = -(2a)^{-1} \log \frac{P_{\text{fwd}}}{P_{\text{bckwd}}}. \tag{3.3}$$

We remark that this line of inquiry relies on a specific interpretation of the model, namely that the driving is caused by a deterministic external force and the fluctuations are thermal. Yet, the model itself is more general and can equally describe an active (rather than driven) particle, and the origin of the fluctuations could be a noisy driving force.

Subsequently, this type of system was studied extensively by Oshanin *et al.* [76, 128–136]. In their initial work, they considered an infinite system with some mean density of symmetrically diffusing particles and an infinitely fast defect particle. They showed that the mean displacement of an infinitely fast tracer grows as $t^{1/2}$, as it builds up a growing blockage of environment particles in front of itself, which impedes its motion [128]. The calculation was done in a continuum limit by assuming independence of the tracer and bath particle positions, which essentially amounts to a mean-field theory. Using a similar approach, this result was extended to the case of a partially asymmetric tracer, again with the $t^{1/2}$ scaling [129]. The fact that this result is the same for an infinitely fast and partially asymmetric tracer is a consequence of the fact that without overtaking, the system will always be limited by the speed of the slowest component. Here, it is the diffusion of the high density blockage in front of the defect. What this means is that as long as the defect has some nonzero velocity, its drift speed at long times will be governed by the diffusion of the bath particles in the blockage and the specific parameters of the defect will not be important.

Further, in [76, 130, 131] calculations were done on a system in which bath particles are allowed to spontaneously be created and annihilated with some fixed rates. One difficulty in analysing the infinite system is that it does not settle into a steady state, since the blockage in front of the defect grows indefinitely. For a physical realisation, one might consider this as an effective model of a channel

in which the fluid particles can adsorb and desorb from the channel, which is assumed to be in contact with a large bath. This setup has the advantage that even in an infinite system, a steady state is reached, as the maximal size of the blockage is limited by the desorption rate. The calculations were extended beyond the simplest mean-field approximation using a third-order decoupling scheme. This means that instead of assuming that two-point functions can be written as products of one-point functions, one retains two-point functions, but assumes that three-point functions can be written in terms of two-point functions. Using this approach, a finite steady-state drift velocity was found for the defect. It was also confirmed that it creates a finite blockage in front of itself and a depletion zone behind itself, which have an exponential shape with finite characteristic lengths.

Another way to simplify the problem is to take a high density limit ($1-\rho \ll 1$) for the bath particles. This simplifies the calculations, as one can solve for the motion of holes (empty sites), which can be assumed to behave essentially independently. This was examined in [132], where detailed statistics of the defect displacement were calculated. In [136], this was considered with multiple tracers and it was shown that they exhibit bath-mediated entrainment and cooperation effects.

A first exact solution to the periodic case with a partially asymmetric tracer was presented in [50]. This was done using the mapping to the zero-range process (see section 2.2). The density profile was found to be exponential, with a characteristic length on the scale of the system size, and the steady state velocity was found to diminish as the inverse system size.

Lastly, we mention that some related problems have received attention in the literature as well. For instance, the random average process is a continuous space version of the simple exclusion process. Instead of hopping on a lattice, particles hop on a real line, with jump sizes determined by the distance to the nearest particle. It has been shown that the displacement of an asymmetric tracer in a symmetric bath grows as $t^{1/2}$, similarly to the lattice case [137].

3.1.3 Chapter summary

In this chapter we will focus on a system of symmetrically diffusing particles with a single totally asymmetric defect particle and no overtaking on a ring geometry. We will construct the exact nonequilibrium steady state in two different ways: first by mapping the process to the zero-range process in section 3.4, and then

by using a matrix product ansatz in section 3.5. We will use these results to calculate the mean steady-state current and the density profile in the reference frame of the defect.

We will then extend these results in two directions in section 3.6. First, we generalise to a system with multiple totally asymmetric defects. In this case, we will see that the system essentially factorises into subsystems with one defect, each of which behaves like an isolated one defect system. The densities in the subsystems are regulated by the condition that the mean drifts of all subsystems have to be equal.

The second generalisation we consider is to a system with a single partially asymmetric defect. This generalisation was first solved by Ayyer in [138]. We show that a similar, though slightly more complicated matrix product ansatz works in this case, allowing us to reproduce the results for the mean current and density profile.

We emphasise that for the case of a totally asymmetric tracer, an exact solution for the steady state measure, as well as expressions for the density profile were already presented in [50]. Also, a solution for the partially asymmetric case was obtained in [138]. These two works used a mapping to a ZRP and combinatorial analysis respectively. The novel results presented in this chapter are then: *(i)* the previously known results are re-derived using a matrix product approach (using a known matrix algebra) and *(ii)* the case of multiple totally asymmetric tracers is solved exactly.

3.2 Model definition

We consider a one-dimensional lattice with periodic boundary conditions (*i.e.* a ring) with $L+1$ sites. On it, we place M particles, which we will call normal particles. These particles perform normal diffusion. That means that after an exponential waiting time, with some fixed rate, they attempt to perform a hop, with equal probability of hopping left and right. We set the left and right hopping rates to unity without loss of generality. The particles interact via hard-core repulsion (also known as simple exclusion), which means that whenever a particle attempts a hop onto a site that is already occupied, the hop is forbidden.

A system of only symmetrically hopping particles is in equilibrium. This follows

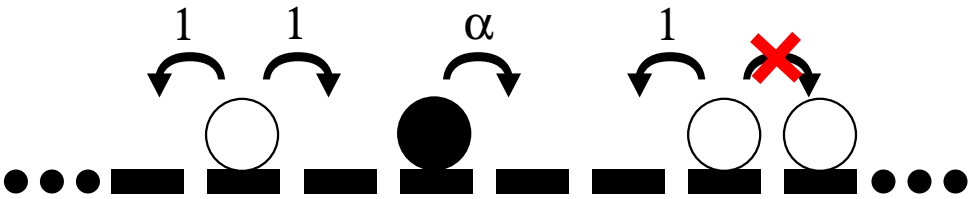

Figure 3.2 *Illustration of symmetric exclusion process with a totally asymmetric defect particle. The defect is shown in black and the normal particles in white.*

from the fact that the transition rates between any two states is equal to the transition rate of the inverse process, $W(X \to X') = W(X' \to X)$. Hence it is easy to see that Kolmogorov's criterion (see section 1.3.8) is satisfied.

To move the system away from equilibrium, we place on the lattice one additional particle, which we call the defect. The defect only hops to the right, with a rate α (in units of the hopping rates of the normal particles) and also obeys the simple exclusion rule with respect to the normal particles. We denote the number of empty sites by $N = L - M$ and the mean environment density by $\rho = M/L$.

The dynamics of the system can be summarised as follows, where we use 0, 1, 2 to denote empty sites, normal particles and the defect respectively,

$$10 \underset{1}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \overset{\alpha}{\to} 02.$$

See figure 3.2 for an illustration.

The defect is inherently far from equilibrium as it is driven. However, it is not entirely obvious *a priori* whether this is sufficient to drive the whole system out of equilibrium or whether the effect of the defect will only be microscopic.

To investigate this, one would like to analyse the steady-state density profile. Due to the periodicity of the system, the time-averaged density profile in the "laboratory frame" is always uniform. One can gain more insight by choosing the frame of reference in which the defect is always on site 0 and the normal particles diffuse on the sites $1, \ldots, L$. The hops of the defect (to the right) in this frame of reference become hops of all normal particles simultaneously to the left. Although this global jump might seem like a complication, it is easier to handle than a changing position of the defect.

Taking the local density view *à la* section 1.5.4 we can write an evolution equation

for the occupation of site i (relative to the defect), similar to (1.56),

$$\frac{\mathrm{d}\langle\eta_i\rangle}{\mathrm{d}t} = \langle\eta_{i-1}(1-\eta_i)\rangle + \langle\eta_{i+1}(1-\eta_i)\rangle - \langle\eta_i(1-\eta_{i-1})\rangle - \langle\eta_i(1-\eta_{i+1})\rangle \\ +\alpha\langle(1-\eta_1)(1-\eta_i)\eta_{i+1}\rangle - \alpha\langle(1-\eta_1)\eta_i(1-\eta_{i+1})\rangle. \quad (3.4)$$

The first four terms represent the diffusion of the normal particles. The last two terms represent the flow of probability into and out of site i due to hops by the defect. We can lose probability if the defect hops (for which site 1 must be empty), site i is occupied and site $i+1$ is not; and gain probability in the converse case. The evolution equation (3.4) simplifies to

$$\frac{\mathrm{d}\langle\eta_i\rangle}{\mathrm{d}t} = \langle\eta_{i-1}\rangle - 2\langle\eta_i\rangle + \langle\eta_{i+1}\rangle + \alpha\langle(1-\eta_1)(\eta_{i+1}-\eta_i)\rangle. \quad (3.5)$$

To complete the problem, we also have the boundary cases $i = 1$, L. After simplifying, these are

$$\frac{\mathrm{d}\langle\eta_1\rangle}{\mathrm{d}t} = \langle\eta_2\rangle - \langle\eta_1\rangle + \alpha\langle(1-\eta_1)\eta_2\rangle, \quad (3.6a)$$

$$\frac{\mathrm{d}\langle\eta_L\rangle}{\mathrm{d}t} = \langle\eta_{L-1}\rangle - \langle\eta_L\rangle - \alpha\langle(1-\eta_1)\eta_L\rangle. \quad (3.6b)$$

Thus the evolution equations for the one-point function $\langle\eta_i\rangle$ involve two-point functions $\langle\eta_i\eta_j\rangle$, which are, moreover, non-local. One approach at this point is to make a mean-field assumption. Later, we will turn to exact methods.

3.3 Mean-field theory

We take the mean-field approximation $\langle\eta_i\eta_j\rangle = \langle\eta_i\rangle\langle\eta_j\rangle$ and go to the continuum limit, $y = k/L$ (see section 1.5 for more detail). We apply this procedure to the equations for the density profile (3.5), (3.6a) and (3.6b). In doing this, we need to take the diffusive time scaling $t = L^2\tau$. Due to the boundary conditions in the present case, we need to be more careful in analysing these equations. In particular, it quickly becomes apparent that in order to balance terms appropriately, one needs to take

$$\eta(0,\tau) = 1 - \frac{\delta(\tau)}{L}, \quad (3.7)$$

where $\delta(\tau) = O(1)$ is some function. This is a sign of a separation of timescales. Physically we can understand it as follows. As the defect moves forward, it will eventually collide with a normal particle. As it cannot move backwards, it will be stuck in this configuration until the normal particle diffuses away. Thus the defect will almost always see a particle directly in front of itself. Substituting (3.7), the boundary conditions (3.6a) and (3.6b) become,

$$-\frac{\partial\delta(\tau)}{\partial\tau} = L(\alpha\delta(\tau) + \nabla\eta(0,\tau)) \tag{3.8a}$$

$$\frac{\partial\eta(1,\tau)}{\partial\tau} = -L(\alpha\delta(\tau)\eta(1,\tau) + \nabla\eta(1,\tau)) \tag{3.8b}$$

Note the appearance of L on the right hand side. This means that the two quantities $\delta(\tau)$, $\eta(1,\tau)$ evolve on a faster timescale than the bulk density profile $\eta(y,\tau)$ for $y \in (0,1)$. The density profile close to the defect ($y \to 0$ and $y \to 1$) equilibrates on a ballistic timescale, whereas the rest of the profile relaxes on a diffusive timescale. This means that for any finite τ (and $L \to \infty$), we may take the boundary conditions to be,

$$\alpha\delta(\tau) + \nabla\eta(0,\tau) = 0 \tag{3.9a}$$

$$\alpha\delta(\tau)\eta(1,\tau) + \nabla\eta(1,\tau) = 0. \tag{3.9b}$$

The bulk equation (3.5) becomes,

$$\frac{\partial\eta(y,\tau)}{\partial\tau} = \nabla^2\eta(y,\tau) + \alpha\delta(\tau)\nabla\eta(y,\tau). \tag{3.10}$$

We also need to enforce the global particle number constraint,

$$\int_0^1 \mathrm{d}y\ \eta(y,\tau) = \rho. \tag{3.11}$$

3.3.1 Steady-state solution

We now derive the steady-state solution of the mean-field equation (3.10), together with the boundary conditions (3.9a), (3.9b) and the constraint (3.11). Note that the unknowns in this problem are both the density field $\eta(y)$ and the constant δ. Setting the time derivative to 0 and integrating (3.10) once, we obtain

$$\nabla\eta(y) + \alpha\delta\eta = B, \tag{3.12}$$

where B is a constant of integration. This is now a first order ODE, which can be solved using the integrating factor method. Using the integrating factor $\mathrm{e}^{\alpha\delta y}$, we obtain the general solution,

$$\eta(y) = \frac{B}{\alpha\delta} + C\mathrm{e}^{-\alpha\delta y}, \tag{3.13}$$

where C is another constant of integration. Then using the conditions (3.7) and (3.9a), we find,

$$C = 1, \qquad B = -\frac{\alpha\delta^2}{L}. \tag{3.14}$$

This yields the solution,

$$\eta(y) = \mathrm{e}^{-\alpha\delta y} - \frac{\delta}{L}. \tag{3.15}$$

Now we plug this solution into (3.11) to obtain an equation for δ. Actually it turns out to be more convenient to express this in terms of $\zeta = \alpha\rho\delta$. Then (3.11) becomes,

$$\zeta = 1 - \mathrm{e}^{-\zeta/\rho}. \tag{3.16}$$

This can be solved (for instance numerically) for ζ. Note that it depends only on the mean density ρ. The density profile is then,

$$\eta(y) = \mathrm{e}^{-\zeta y/\rho} - \frac{\zeta}{\alpha\rho L}. \tag{3.17}$$

From this we may also deduce the mean drift velocity of the defect $\langle v\rangle$ (and therefore all particles, as there is no overtaking in this system). This is given simply by the hopping rate of the defect, α, multiplied by the probability that site 1 is empty, $(1 - \eta(0))$. This gives,

$$\langle v\rangle = \frac{\zeta}{\rho L}. \tag{3.18}$$

Interestingly, this is independent of α. We will return to the discussion of this once we have presented the exact solution.

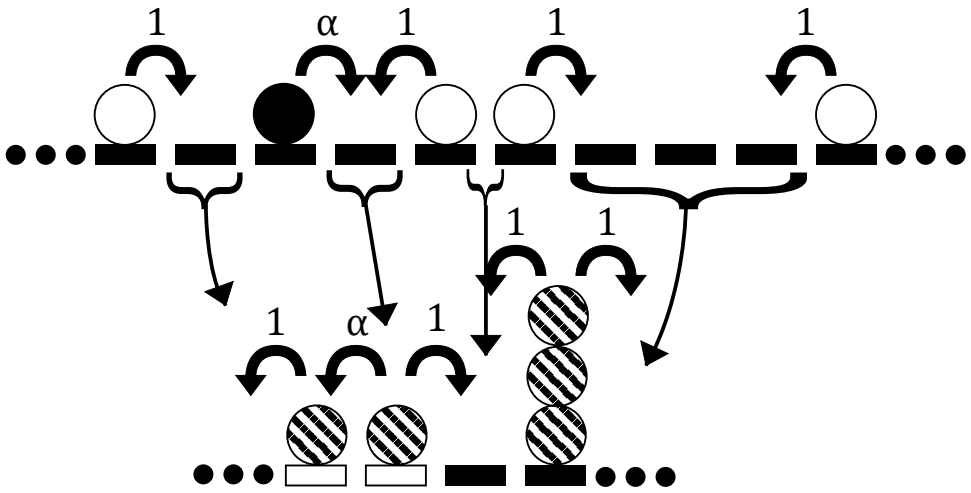

Figure 3.3 *Mapping between simple exclusion process (top) and zero-range process (bottom). The defect in the exclusion process is shown in black and the normal particles in white. The rates in the zero-range process are shown to emphasise that they do not depend on the occupations of other sites. Note that the zero-range process is inhomogeneous, as the rates on the sites corresponding to the empty spaces in front of and behind the defect (shown in white) are different from the rates of other sites.*

3.4 Solution by mapping to zero-range process

We now turn to exact methods. One way to obtain an exact solution for the steady state is by mapping it to a zero-range process (ZRP), as explained in section 2.2. To carry out the mapping, we first assign to each particle in the exclusion process a site in the ZRP, in order. The zeroth ZRP site corresponds to the defect, the first site corresponds to the first normal particle and so on. Then, we put as many particles on each ZRP site, as there are empty sites in front of the corresponding particle in the exclusion process. Note that the hopping rates in the ZRP will be constant (with the sites corresponding to the spaces next to the defect particle having different, but still constant rates). This justifies calling it a ZRP, as the escape rates from a given site do not depend on the occupation of any other site. See figure 3.3 for an illustration.

Thus we obtain a lattice with $M+1$ sites and N particles, where, as a reminder, M and N are respectively the numbers of environment particles and empty sites in the exclusion process. A configuration of a state in the ZRP is given by the occupation of each site $\{n_0, n_1, \ldots, n_M\}$, where $\sum_{i=0}^{M} n_i = N$.

3.4.1 Determination of weight functions

The advantage of reformulating the problem in terms of the ZRP is that it is known that the steady state of a ZRP has a factorised form. This means that the probability of a given configuration is

$$P(\{n_0, \ldots, n_M\}) = Z_{L,M}^{-1} \prod_{i=0}^{M} f_i(n_i), \tag{3.19}$$

where $f_i(n_i)$ are weight functions for $i = 0, 1, \ldots, M$, and $Z_{L,M}$ is the normalisation,

$$Z_{L,M} = \sum_{\{\mathbf{n}\}} \prod_{i=0}^{M} f_i(n_i) \delta_{\sum_{i=0}^{M} n_i, N}. \tag{3.20}$$

Taking (3.19) as an ansatz, we now use the steady-state form of the evolution equations to determine the weight functions. The local balance condition gives us the equations

$$\begin{aligned}
&f_{i-1}(n_{i-1}+1) f_i(n_i - 1) f_{i+1}(n_{i+1}) + f_{i-1}(n_{i-1}) f_i(n_i - 1) f_{i+1}(n_{i+1}+1) \\
&\qquad = 2 f_{i-1}(n_{i-1}) f_i(n_i) f_{i+1}(n_{i+1}), \qquad i = 1, \ldots, M-1, && \text{(3.21a)} \\
&f_M(n_M) f_0(n_0 - 1) f_1(n_1 + 1) = (1+\alpha) f_M(n_M) f_0(n_0) f_1(n_1), && \text{(3.21b)} \\
&\alpha f_{M-1}(n_{M-1}) f_M(n_M - 1) f_0(n_0 + 1) + f_{M-1}(n_{M-1}+1) f_M(n_M - 1) f_0(n_0) \\
&\qquad = f_{M-1}(n_{M-1}) f_M(n_M) f_0(n_0). && \text{(3.21c)}
\end{aligned}$$

As before, these equations come from requiring the positive and negative probability fluxes of a state with occupation n_i on site i to balance. The first equation describes this balance in the bulk, and the second and third for transitions involving the defect particle.

This system of equations a bit more complicated than the case considered in section 2.2. It is solved by a geometric ansatz [94],

$$f_i(n_i) = g_i^{n_i}, \tag{3.22}$$

for some constants g_i. Plugging this in, we get

$$g_i = \frac{g_{i-1} + g_{i+1}}{2}, \qquad i = 1, \ldots, M-1, \tag{3.23a}$$

$$g_0 = \frac{g_1}{1+\alpha}, \tag{3.23b}$$

$$g_M = \alpha g_0 + g_{M-1}. \tag{3.23c}$$

This system of equations is underdetermined with one degree of freedom. This allows us to set $g_0 = 1$. Then it is easily seen that the full solution is given by

$$g_i = 1 + \alpha i, \qquad i = 0, \ldots, M. \tag{3.24}$$

Putting this solution into (3.19), we see that the probabilities are given by,

$$P(\{n_0, \ldots, n_M\}) = Z_{L,M}^{-1} \prod_{i=0}^{M} (1 + i\alpha)^{n_i}. \tag{3.25}$$

3.4.2 Steady-state measure

We now analyse the normalisation $Z_{L,M}$ in more detail. We will see in section 3.5 that it plays a crucial role in the calculation of the steady-state density profile and velocity. From (3.25), we see that,

$$Z_{L,M} = \sum_{\{\mathbf{n}\}} \prod_{i=0}^{M} (1 + i\alpha)^{n_i} \delta_{\sum_{i=0}^{M} n_i, N}. \tag{3.26}$$

This normalisation resembles a partition function, as one might write down in an equilibrium problem. However, it is important to draw a distinction, as this object was not written down starting from a Hamiltonian. Without an explicit solution, we would not have known how to construct this *a priori*. However, it has some features similar to partition functions: it is the normalisation of a steady state measure and we will be able to derive correlation functions from it.

Although we now formally have an exact solution for the steady state, the normalisation is not very tractable, as it has the global constraint $\sum_{i=0}^{M} n_i = N$, where, again, N is the number of empty sites in the exclusion process. We use the generating function approach from section 2.B to deal with this. This allows

us to write

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(N+1)} \prod_{i=0}^{M} \frac{1}{1-(1+i\alpha)z}, \tag{3.27}$$

where the contour of integration is a small circle around the origin and we have used the geometric sum

$$\sum_{n_i=0}^{\infty} ((1+i\alpha)z)^{n_i} = \frac{1}{1-(1+i\alpha)z}. \tag{3.28}$$

We remark that even with an exact expression for the steady state probabilities, it is a nontrivial task to compute physical observables. For instance, to calculate the probability that site i is occupied, we would need to sum over all configurations of the system in which this is realised. This was done, for instance, in [50], but the resulting expression is quite complicated, with two contour integrals. We will show in section 3.5 that the matrix product formulation is better suited for addressing this question, as it makes it both more intuitive to write down formal expressions for expectations of local observables and gives expressions involving only a single contour integral.

3.4.3 Finite-size expression

Evaluating the residue of the pole at the origin is quite cumbersome. However, we can obtain an exact solution by using the result that a complex contour integral can also be evaluated by taking the residues of the poles outside the contour (including, potentially, a pole at infinity). In the case at hand, there is no pole at infinity, so we only need to consider the poles at $z = (1+i\alpha)^{-1}$ for $i = 0, \ldots, M$. Switching to the residues of the poles outside the contour also introduces a global minus sign. Evaluating the residues of (3.27) at those poles gives

$$\begin{aligned} Z_{L,M} &= -\sum_{i=0}^{M} \mathrm{Res}\left(z^{-(N+1)} \prod_{j=0}^{M} \frac{1}{1-(1+j\alpha)z}, (1-i\alpha)^{-1} \right) && (3.29) \\ &= \sum_{i=0}^{M} (1+i\alpha)^N \prod_{j=0, j\neq i}^{M} \frac{1}{1-(1+j\alpha)(1+i\alpha)^{-1}} && (3.30) \\ &= \sum_{i=0}^{M} \frac{(1+i\alpha)^L}{\alpha^M} \prod_{j=0, j\neq i}^{M} \frac{1}{i-j}. && (3.31) \end{aligned}$$

The product can be written as

$$\prod_{j=0,j\neq i}^{M} \frac{1}{i-j} = \frac{(-)^{M-i}}{M!}\binom{M}{i}. \tag{3.32}$$

Then expanding the binomial term $(1+i\alpha)^L$ and simplifying, we get

$$Z_{L,M} = \sum_{j=0}^{L} \frac{\alpha^{j-M}}{M!}\binom{L}{j} \sum_{i=0}^{M} (-)^{M-i}\binom{M}{i} i^j. \tag{3.33}$$

We can write this compactly by using the following identity for Stirling numbers of the second kind [1],

$$\left\{ {j \atop M} \right\} = \frac{1}{M!} \sum_{i=0}^{M} (-)^{M-i}\binom{M}{i} i^j. \tag{3.34}$$

This gives us the final result

$$Z_{L,M} = \sum_{j=0}^{L-M} \binom{L}{M+j} \left\{ {M+j \atop M} \right\} \alpha^j. \tag{3.35}$$

Thus we have obtained an expression for the normalisation as a finite sum of binomial coefficients, Stirling numbers and the defect parameter α. This can be readily evaluated numerically to obtain exact finite-size results. In section 3.5, we will show that physical observables like the density profile and the steady-state velocity can be expressed purely in terms of the normalisation.

3.4.4 Asymptotic behaviour

We now consider the limit $L \to \infty$, $M \to \infty$ with $\rho = M/L$ fixed. Then we can rewrite the integral in (3.27) as

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i} z} \exp\left\{ -L \left[L^{-1} \sum_{i=0}^{M} \log(1-(1+i\alpha)z) + (1-\rho)\log z \right] \right\}. \tag{3.36}$$

[1] These are defined as the number of ways to partition a set of j objects into M nonempty groups.

It is helpful to rescale the variable of integration by letting

$$z = \frac{\zeta}{M\alpha}. \tag{3.37}$$

Then, approximating the sum as an integral by setting $i = L\rho u$, we get to leading order in L,

$$\begin{aligned} L^{-1}\sum_{i=0}^{M}\log(1-(1+i\alpha)z) &\approx \rho\int_0^1 \log(1-u\zeta)\mathrm{d}u + o(L^0) &(3.38)\\ &\approx -\rho + \rho\left(1-\frac{1}{\zeta}\right)\log(1-\zeta) + o(L^0). &(3.39)\end{aligned}$$

Then the normalisation becomes,

$$Z_{L,M} = (M\alpha)^N \mathrm{e}^M \oint \frac{\mathrm{d}\zeta}{2\pi\mathrm{i}\zeta}\mathrm{e}^{-L\phi(\zeta)}, \tag{3.40}$$

where

$$\phi(\zeta) = \rho\left(1-\frac{1}{\zeta}\right)\log(1-\zeta) + (1-\rho)\log\zeta. \tag{3.41}$$

Note that ϕ does not depend on L, which justifies the definition of ζ (3.37).

The contour integral is dominated by a saddle point, which is defined by the condition $\phi'(\zeta_{\mathrm{sp}}) = 0$. Solving this, we get the following transcendental equation for the location of the saddle point

$$\zeta_{\mathrm{sp}} = 1 - \mathrm{e}^{-\zeta_{\mathrm{sp}}/\rho}. \tag{3.42}$$

The solution to this can be written using the Lambert W function,

$$\zeta_{\mathrm{sp}} = 1 + \rho W_0(-\mathrm{e}^{-1/\rho}/\rho), \tag{3.43}$$

where the subscript 0 indicates that we must take the principal branch. This is plotted as a function of ρ in Figure 3.4. Then to leading order in the saddle point expansion, we can write,

$$Z_{L,M} \approx A\exp\left\{-L\left[(1-\rho)\log\zeta_{\mathrm{sp}} - \zeta_{\mathrm{sp}}\right]\right\}\left[1+O(L^{-1})\right], \tag{3.44}$$

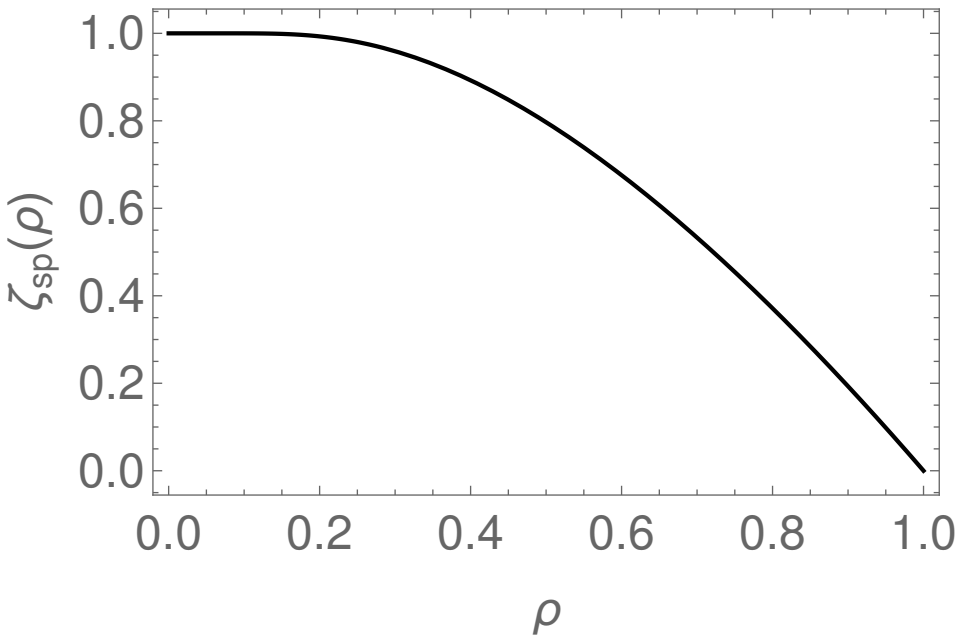

Figure 3.4 *Location of saddle point, ζ_{sp}, as defined in* (3.42)*, against mean density, ρ.*

where

$$A = \left[2\pi L \left(\frac{1}{1-\zeta_{\mathrm{sp}}} - \frac{1}{\rho}\right)\right]^{-1/2} \left(\frac{M\alpha}{\mathrm{e}}\right)^{N}. \tag{3.45}$$

However, it turns out that for the physical observables, like the mean steady-state velocity, only the location of the saddle point is important.

3.5 Solution by matrix product ansatz

We now approach the same problem using the matrix product formalism (see section 2.3). This involves more somewhat formal manipulations but lends itself quite naturally to calculating local physical observables, such as the occupation probability of a given site.

Specifically, we consider some configuration of the system $\{\tau_0, \tau_1, \tau_2, \dots\}$, where we now use the notation $\tau_i = 0,\ 1,\ 2$ for empty sites, normal particles and defects respectively. Then there exist matrices X_0, X_1, X_2 (representing empty sites, normal particles and the defect respectively) such that the weight of that

configuration is given by

$$P(\{\tau_0, \tau_1, \tau_2, \dots\}) = Z_{L,M}^{-1} \mathrm{tr}(X_{\tau_0} X_{\tau_1} X_{\tau_2} \dots), \tag{3.46}$$

where $Z_{L,M}$ is a normalisation. The use of the trace reflects the periodic boundary conditions.

For the system we are considering, the matrix product ansatz generates the correct steady state if the matrices satisfy the algebraic relations,

$$X_1 X_0 - X_0 X_1 = \alpha X_1, \tag{3.47a}$$

$$X_2 X_0 = X_2. \tag{3.47b}$$

A general proof of the validity of the matrix product ansatz is given in section 2.3.2. To obtain the relations (3.47a) and (3.47b) from the general case, we use equations (2.52a–2.52f) with the parameters $w_{01} = w_{10} = 1$, $w_{20} = \alpha$, $w_{02} = 0$ and $w_{12} = w_{21} = 0$. Then we see that all the relations are satisfied if we set $x_1 = x_2 = 0$, $x_0 = \alpha$.

We note that the algebra (3.47a), (3.47b) is one of the known two-species algebras with scalar reduction relations [106]. However, it has not been used before to perform calculations of physical observables.

The matrix product solution presented here is valid for any number of defects, but we leave that to section 3.6.1. For the remainder of this section, we restrict ourselves to the case of a single defect, which we place on site 0 without loss of generality.

As a consistency check, we can verify that the matrix product solution defined by (3.47a–3.47b) produces the same configuration probabilities as the mapping to the zero-range process (3.25). First, we note that from (3.47a), we have,

$$X_1 X_0 = (\alpha + X_0) X_1. \tag{3.48}$$

Using this repeatedly on $X_1^i X_0$ for some integer i, we obtain the helpful relation,

$$X_1^i X_0 = (i\alpha + X_0) X_1^i. \tag{3.49}$$

Then taking some general configuration with n_i empty sites in front of particle i,

and using (3.47b), we obtain,

$$\begin{aligned} X_2 X_0^{n_0} X_1 X_0^{n_1} X_1 X_0^{n_2} \dots &= X_2 X_1 X_0^{n_1} X_1 X_0^{n_2} \dots && (3.50) \\ &= (1+\alpha)^{n_1} X_2 X_1^2 X_0^{n_2} \dots && (3.51) \\ &\vdots \\ &= \left(\prod_{i=0}^{M} (1+i\alpha)^{n_i} \right) X_2 X_1^M. && (3.52) \end{aligned}$$

Taking the trace, this gives the same weight as (3.19), up to the constant factor $\mathrm{tr}(X_2 X_1^M)$, which can be set to unity without loss of generality.

To obtain the normalisation, we can use a technique similar to the generating function trick from section 2.B but with matrix, rather than scalar weights. We introduce the matrix

$$C = X_1 + zX_0, \tag{3.53}$$

where z is a dummy variable that counts empty sites, which can also be interpreted as a fugacity for empty sites. Then the normalisation can be written as

$$Z_{L,M} = \{z^N\} \mathrm{tr}(X_2 C^L), \tag{3.54}$$

where, again, the notation $\{z^N\}$ indicates that we take the coefficient of z^N in the expression that follows and N is the number of empty sites in the exclusion process.

3.5.1 Mean current

Before calculating the full density profile, it is helpful to calculate the occupation probability of site 1. By writing it as one minus the probability that site 1 is empty and using (3.47b), we get the expression,

$$\langle \eta_1 \rangle_{L,M} = 1 - \frac{\{z^N\} \mathrm{tr}(X_2 z X_0 C^{L-1})}{Z_{L,M}} = 1 - \frac{Z_{L-1,M}}{Z_{L,M}}. \tag{3.55}$$

This immediately gives us the mean velocity of the defect, which is simply the hopping rate of the defect, α, multiplied by the probability that site 1 is empty.

Using the exact finite-size expression for $Z_{L,M}$ (3.35), we obtain,

$$\langle v \rangle_{L,M} = \alpha \frac{Z_{L-1,M}}{Z_{L,M}} = \frac{\sum_{j=0}^{L-1-M} \binom{L-1}{M+j} \left\{ {M+j \atop M} \right\} \alpha^{j+1}}{\sum_{j=0}^{L-M} \binom{L}{M+j} \left\{ {M+j \atop M} \right\} \alpha^{j}}, \tag{3.56}$$

where, again, $\left\{ {a \atop b} \right\}$ denotes a Stirling number of the second kind. As there is no overtaking in the system, this will also equal the mean drift of the normal particles.

To obtain the asymptotic limit, recall that we have already analysed the asymptotic behaviour of the normalisation $Z_{L,M}$ in section 3.4. It is helpful to consider the form of the normalisation (3.27). The only difference between $Z_{L-1,M}$ and $Z_{L,M}$ is a factor of z, which will not change the location of the saddle point (3.42). Then at the leading order level in the saddle point, the ratio $Z_{L-1,M}/Z_{L,M}$ will simply be z evaluated at the saddle point, giving

$$\langle v \rangle_{L,M} \approx \frac{\zeta_{\text{sp}}}{M} + O(L^{-2}). \tag{3.57}$$

We remark that the mean velocity is diminished by L^{-1}, so it vanishes in the thermodynamic limit. It is also interesting to note that this expression does not depend on the defect hopping parameter α, as ζ_{sp} is defined purely in terms of ρ, (3.42). This means that for any finite defect velocity, the mean drift will determined solely by the diffusion rate and density of normal particles. Note that this expression is the same as the one obtained from mean-field theory, (3.18).

3.5.2 Density profile

The probability that site i is occupied can be written using the matrix product description as

$$\langle \eta_i \rangle_{L,M} = \frac{\{z^N\} \text{tr}(X_2 C^{i-1} X_1 C^{L-i})}{Z_{L,M}}. \tag{3.58}$$

We note that from (3.47a) and (3.53), we have,

$$CX_1 = X_1(C - \alpha z), \tag{3.59}$$

which gives

$$\langle \eta_i \rangle_{L,M} = \frac{\{z^N\}\mathrm{tr}(X_2 X_1 (C - \alpha z)^{i-1} C^{L-i})}{Z_{L,M}}. \tag{3.60}$$

Expanding the binomial term, this can be written as

$$\langle \eta_i \rangle_{L,M} = \sum_{j=0}^{i-1} \binom{i-1}{j} (-\alpha)^j \frac{\{z^{N-j}\}\mathrm{tr}(X_2 X_1 C^{L-1-j})}{Z_{L,M}}. \tag{3.61}$$

Now we note that the term in the numerator can be related to the mean density of site 1 in a system of size $L - j$,

$$\langle \eta_i \rangle_{L,M} = \sum_{j=0}^{i-1} \binom{i-1}{j} (-\alpha)^j \frac{Z_{L-j,M}}{Z_{L,M}} \langle \eta_1 \rangle_{L-j,M}. \tag{3.62}$$

Then using the result (3.55), we obtain,

$$\langle \eta_i \rangle_{L,M} = \sum_{j=0}^{i-1} \binom{i-1}{j} (-\alpha)^j \frac{Z_{L-j,M} - Z_{L-1-j,M}}{Z_{L,M}}. \tag{3.63}$$

Using the exact finite-size result (3.35), we can write this explicitly in the form,

$$\begin{aligned} \langle \eta_i \rangle_{L,M} &= Z_{L,M}^{-1} \sum_{j=0}^{i-1} \sum_{k=0}^{L-M-j} (-)^j \left(\binom{i-1}{j-1} + \alpha \binom{i-1}{j} \right) \\ &\quad \times \binom{L-j}{M+k} \left\{ {M+k \atop M} \right\} \alpha^{k+j-1}. \end{aligned} \tag{3.64}$$

This expression can be used to obtain the exact density profile for small system sizes.

To gain more physical intuition, we now proceed to extract the asymptotic behaviour. Rewriting the partition functions in (3.63) as integrals, we can perform the summation over j to obtain

$$\langle \eta_i \rangle_{L,M} = Z_{L,M}^{-1} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{(1 - \alpha z)^{i-1}}{z^{N+1}} (1 - z) \mathcal{Z}_M(z). \tag{3.65}$$

Then introducing the same scaling variable as in (3.37), $\zeta = M\alpha z$, we see that

$$(1 - \alpha z)^{i-1} = \exp\left[(i-1) \log\left(1 - \frac{\zeta}{L\rho} \right) \right] \approx \exp\left(-\frac{\zeta}{\rho} \frac{i-1}{L} \right). \tag{3.66}$$

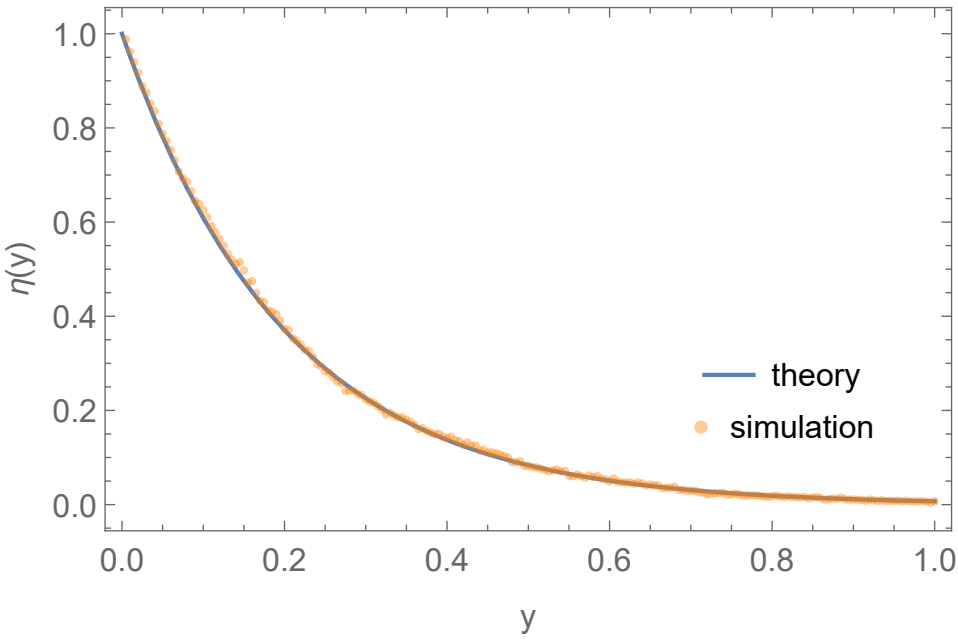

Figure 3.5 *Density profile as a function of $y = i/L$ for mean density $\rho = 0.2$, as given by* (3.67) *and as obtained from Monte Carlo simulations. Simulations were performed with $\alpha = 2$ and with a system size $L = 200$. Excellent agreement is seen between theory and simulations.*

Defining $y = i/L$, we see that this new term in the exponential is $-\xi y/\rho = O(1)$, which is subdominant compared to $L\phi(\zeta)$. Note that $i = \{1, \ldots, L\}$ means $y \in (0, 1]$. Hence the location of the saddle point (3.42) remains unchanged. Then letting $\eta(y) = \langle \eta_{Ly} \rangle$, we obtain to leading order in the saddle point,

$$\eta(y) \approx \exp\left(-\frac{\zeta_{\text{sp}}}{\rho} y\right) + O(L^{-1}). \tag{3.67}$$

Again, this expression is the same as the one obtained using mean-field theory, (3.17). It is an exponential decay with a characteristic length $\xi = \rho/\zeta_{\text{sp}}$. Note that it remains finite in the thermodynamic limit. The density profile is plotted as a function of y in figure 3.5.

The normal particles create a jam in front of the defect. The fact that the characteristic length is of the same scale as the system size means that the density profile is nonuniform at a macroscopic level. This is interesting, as the system has a macroscopic sign of nonequilibrium behaviour despite being driven by a single particle.

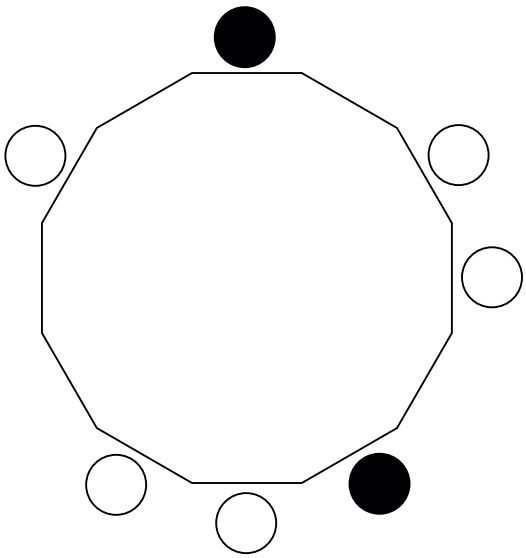

Figure 3.6 *Illustration of process with $k = 2$ defects (black), $L+k = 10+2$ sites, $M_1 = 2$ normal particles (white) in front of defect 1, and $M_2 = 3$ normal particles in front of defect 2.*

3.6 Extensions

3.6.1 Multiple defects

We can generalise the results from the previous section relatively straightforwardly to the case of multiple totally asymmetric defects. Suppose we have k defects, each hopping with rate α on a ring of $L + k$ sites with $M_1, \ldots, M_k$ normal particles between them, with $\sum_{i=1}^{k} M_i = M$. An illustration is given in figure 3.6

The number of normal particles between the defects is fixed, as there is no overtaking allowed. However, the number of empty sites between each defect can vary. Therefore the mean densities in the regions between the defects are not constant.

Let us denote the number of empty sites in front of particle j in the block in front of defect i by $n_{i,j}$. Moreover we denote $\sum_{j=0}^{M_i} n_{i,j} = N_i$ and $\sum_{i=1}^{k} N_i = N$. Using the matrix product ansatz with matrices defined by the algebra (3.47a)–(3.47b),

we see that the weight of a configuration is given by

$$\begin{aligned} &P(\{n_{1,0},\dots,n_{1,M_1};n_{2,0},\dots,n_{2,M_2};\dots;n_{k,0},\dots,n_{k,M_k}\}) \\ &\quad = Z^{-1}_{L,M_1,\dots,M_k}\mathrm{tr}(X_2X_0^{n_{1,0}}X_1X_0^{n_{1,1}}\dots X_2X_0^{n_{k,0}}X_1X_0^{n_{k,1}}\dots) \qquad (3.68) \\ &\quad = Z^{-1}_{L,M_1,\dots,M_k}\left(\prod_{i=1}^{k}\prod_{j=0}^{M_i}(1-j\alpha)^{n_{i,j}}\right)\mathrm{tr}(X_2X_1^{M_1}\dots X_2X_1^{M_k}). \qquad (3.69) \end{aligned}$$

As before, the constant term $\mathrm{tr}(X_2X_1^{M_1}\dots X_2X_1^{M_k})$ can be set to unity without loss of generality. The weights of configurations factorise around the defects. Thus, we can interpret the subsystems with one defect effectively as independent, except for the interaction through the constraint that they all share a common pool of empty sites. We can write the total partition function of the whole system as a sum of products of partition functions of subsystems with a single defect,

$$Z_{L,M_1,\dots,M_k} = \sum_{\{N_i\}}\prod_{i=1}^{k} Z_{M_i+N_i,M_i}\delta_{\sum_{i=1}^{k}N_i,N}. \qquad (3.70)$$

Evidently this is another case for the generating function trick from section 2.B. Similarly to (3.27), we can write

$$Z_{L,M_1,\dots,M_k} = \oint\frac{\mathrm{d}z}{2\pi\mathrm{i}}z^{-(N+1)}\prod_{i=0}^{k}\prod_{j=0}^{M_i}\frac{1}{1-(1+j\alpha)z}. \qquad (3.71)$$

Then repeating steps similar to those in section 3.4.4 (using the same scaling variable $\zeta = M\alpha z$ as in (3.37)), we can write this in the form

$$Z_{L,M_1,\dots,M_k} = (M\alpha)^N\mathrm{e}^M\oint\frac{\mathrm{d}\zeta}{2\pi\mathrm{i}\zeta}\mathrm{e}^{L\phi(\zeta)} \qquad (3.72a)$$

$$\phi(\zeta) = -\rho\sum_{i=1}^{k}m_i\left(1-\frac{1}{m_i\zeta}\right)\log(1-m_i\zeta) - (1-\rho)\log\zeta, \qquad (3.72b)$$

where we have introduced the fractions of normal particles in each one-defect subsystem,

$$m_i = M_i/M, \qquad (3.73)$$

and $\rho = M/L$ is the mean density of the whole system. Then the saddle point

equation (obtained from setting $\phi'(\zeta) = 0$ in (3.72b)) has the form,

$$\prod_{i=1}^{k}(1 - m_i\zeta_{\rm sp}) = \mathrm{e}^{-\zeta_{\rm sp}/\rho}, \tag{3.74}$$

which can be solved numerically for specific choices of m_i.

Then it is not hard to see that the mean speeds of the defects are given by,

$$\langle v_i \rangle_{L,M_1,\ldots,M_k} \approx \frac{\zeta_i}{M_i} = \frac{\zeta_{\rm sp}}{M}, \tag{3.75}$$

where $\zeta_i = m_i\zeta_{\rm sp}$. This can be interpreted as a statement of the necessary condition that in the steady state all defects have the same velocity. Conversely, we can see that this allows us to determine how many empty sites each one-defect subsystem will have on average. Namely, the densities of each subsystem must be adjusted such that all their velocities are equal.

The subsystems effectively behave like small versions of one-defect systems with fugacities $\zeta_i = m_i\zeta_{\rm sp}$. The steady-state measure is dominated by the configuration with the most likely distribution of empty sites between subsystems, $\{N_1^*, \ldots, N_k^*\}$. Then the normalisation (3.70) becomes,

$$Z_{L,M_1,\ldots,M_k} \approx \prod_{i=1}^{k} Z_{M_i+N_i^*,M_i}. \tag{3.76}$$

The values N_i^* are determined by first using (3.74) to solve for $\zeta_{\rm sp}$. This gives a value for $\zeta_i = m_i\zeta_{\rm sp}$. Then inverting the one-defect saddle point equation (3.42), we obtain the expression for the mean density in the steady-state of the i-th subsystem ρ_i,

$$\rho_i = -\frac{\zeta_i}{\log(1 - \zeta_i)}. \tag{3.77}$$

Then N_i^* is given simply by $N_i^* = (1 - \rho_i)M_i/\rho_i$.

From this, it follows, for instance, that, similarly to the one defect case, the density profile in each subsystem is an exponential decay, with $\langle \eta_{i,1} \rangle \approx 1 - O(L^{-1})$ for $i = 1, \ldots, k$; and a characteristic length $\xi_i = m_i\rho/\zeta_i$. An example for $k = 2$ is shown in figure 3.7.

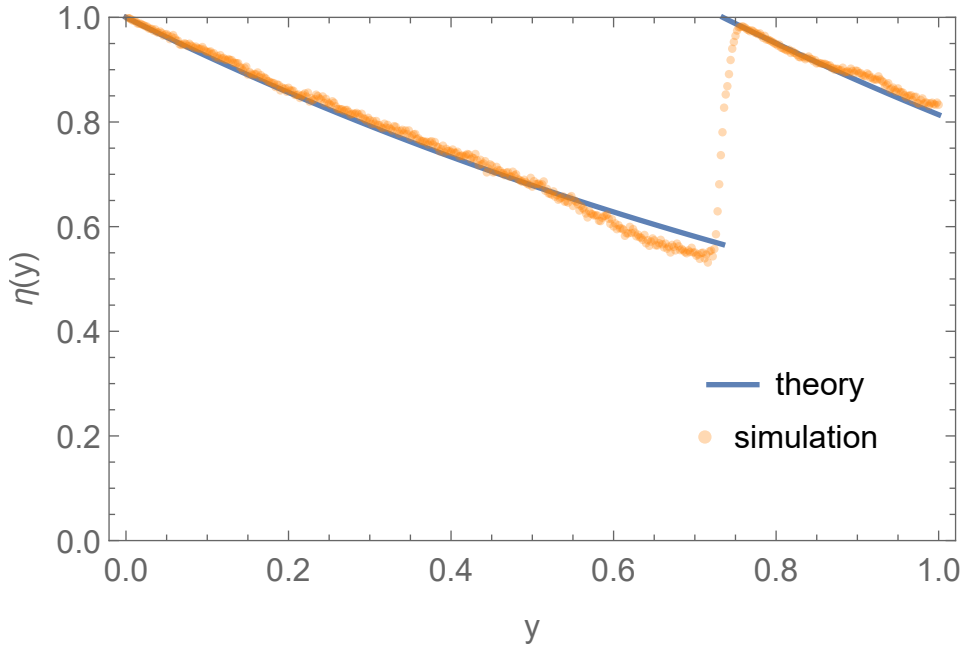

Figure 3.7 *Density profile for system with two defects as predicted by theory and as obtained from Monte Carlo simulations. Parameters used were* $\alpha = 2$, $\rho = 0.8$, $m_1 = 0.7$, $m_2 = 0.3$. *Monte Carlo simulations were performed for a system size of* $L = 500$. *Good agreement is seen between theory and simulations. The density profiles are exponential decays. However, with the chosen parameters, the decay lengths are relatively long compared to the system size, which makes it hard to distinguish from linear profiles by eye.*

Optimal partition of normal particles by defects

It is interesting to investigate the following question: given a fixed total number of normal particles, M, what partition by k defects, $\{M_1, \ldots, M_k\}$, will produce the optimal current? Essentially, we wish to optimise ζ_{sp}, defined implicitly by (3.74), subject to the constraint $\sum_{i=1}^{k} m_i = 1$. We can use the method of Lagrange multipliers. We have the Lagrangian function

$$\mathcal{L}(\mathbf{m}; \lambda) = \zeta_{\text{sp}}(\mathbf{m}) - \lambda \left(\sum_{i=1}^{k} m_i - 1 \right), \tag{3.78}$$

where $\mathbf{m} = \{m_1, \ldots, m_k\}$ and λ is the Lagrange multiplier. Then setting the partial derivatives to 0 (using implicit differentiation of (3.74) with respect to each m_i), we obtain

$$\frac{\zeta_{\text{sp}}\rho}{1 - m_i(\zeta_{\text{sp}} + \rho)} - \lambda = 0. \tag{3.79}$$

This can be solved to give the optimal value of m_i, which will not depend on i. Hence the optimal partition is the one in which $m_1 = \cdots = m_k = 1/k$.

Putting $m_i = 1/k$ for all i in (3.74), we can see that the equation takes the same form as the one-defect equation (3.42) but with the replacement $\zeta_{\text{sp}} \to \zeta_{\text{sp}}/k$. Then the mean velocity, which is proportional to ζ_{sp}, will be k times that of a one-defect system of the same total size and mean density.

The ratio of velocities for a two-defect system to that of a one-defect system of the same size and density is plotted in figure 3.8. We can see that the mean velocity equals that of a single defect system in the limits $m_1 \to 0$, $m_1 \to 1$ but is maximal (double the velocity of a one-defect system of the same size and total mean density) when $m_1 = 1/2$.

These results can be understood intuitively as follows. In the extreme case that there are no particles between two defects, the defect in front will eventually create a jam in front of itself and slow down. Then the defect behind it will always be right behind it, as it cannot overtake. Thus the two defects essentially behave in the same way as a single defect (that occupies two sites).

In the opposite limit, it makes sense intuitively that k defects should produce a current close to k times the current produced by one defect. The fact that the current cannot exceed this value is a manifestation of the fact that this system has

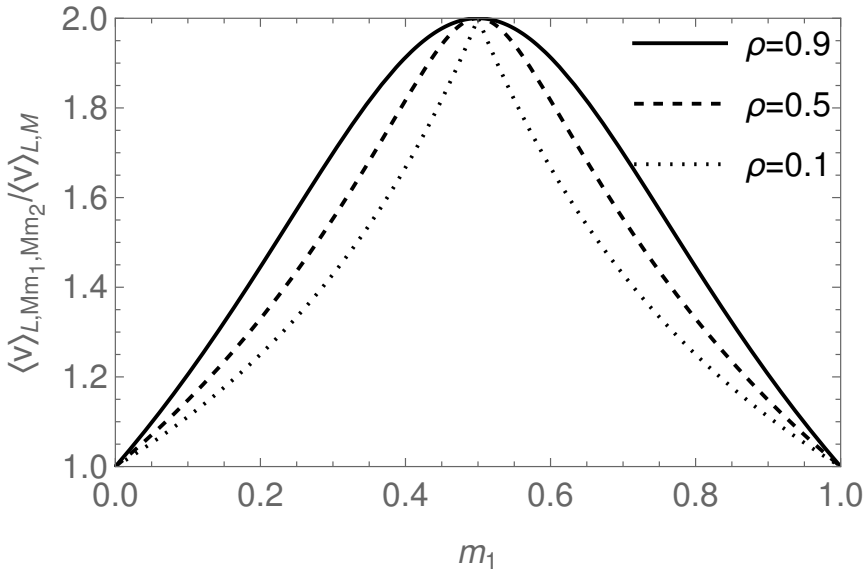

Figure 3.8 *Ratio of steady-state velocity of defect in a two-defect system (as given by* (3.75)*, with* ζ_{sp} *defined by* (3.74)*) to that in a one-defect system of the same size,* L*, and mean density,* ρ*, plotted against fraction of particles in front of defect 1,* m_1*, for various densities. Regardless of mean density, the ratio converges to 1 in the limits* $m_1 \to 0$ *and* $m_1 \to 1$*, and attains its maximal value of 2 for* $m_1 = 0.5$*.*

no pushing or overtaking. This means that the total current will always be limited by the speed of the slowest component. We can use this fact by treating each subsystem as independent and then choosing the velocity of the whole system to be the velocity of the slowest subsystem. As velocity is inversely related to system size (3.57), one can increase the speed of any one subsystem by putting fewer particles in it. However, this will come at the expense of putting more particles in the other subsystems, which will slow the system down overall. Hence the optimal distribution is to distribute the particles evenly. Again, recalling that velocity is inversely related to system size, this means that the velocity of this configuration will be k times that of a system with one tracer, as each subsystem is $1/k$ the size of the whole system.

3.6.2 Partially asymmetric tracer

Another way we can generalise the model is by again considering one defect but letting it be partially asymmetric. In other words, we let it hop to the right with rate α and to the left with some other rate, β, provided the target site is empty.

The dynamics can be summarised as follows,

$$10 \underset{1}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 02.$$

Defining

$$q = \beta/\alpha, \tag{3.80}$$

we can solve this case using the matrix product ansatz

$$X_1X_0 - X_0X_1 = \alpha X_1, \tag{3.81a}$$
$$X_2X_0 - qX_0X_2 = X_2. \tag{3.81b}$$

This can, again, be attained from the general case, as explained in section 2.3.3. We have $w_{10} = w_{01} = 1$, $w_{20} = \alpha$ and $w_{12} = w_{21} = 0$, as before, but now $w_{02} = \beta$. It is easy to check that the same choice of scalar auxiliaries, $x_1 = x_2 = 0$, $x_0 = \alpha$ is sufficient to satisfy the consistency relations (2.52a)–(2.52f).

Again, we emphasise that the algebra (3.81a), (3.81b) falls under one of the known cases of matrix algebras with scalar reduction relations [106]. However, it has not previously been used to compute physical observables.

We can obtain the weights using a procedure similar to that in section 3.5, though it is a bit heavier algebraically. From (3.81a), we have,

$$X_1X_0 = (X_0 + \alpha)X_1 \tag{3.82}$$
$$X_1^iX_0 = (X_0 + i\alpha)X_1^i. \tag{3.83}$$

And similarly from (3.81b),

$$X_2X_0 = (qX_0 + 1)X_2. \tag{3.84}$$

This allows us to write the weight of a general configuration as,

$$\mathrm{tr}(X_2X_0^{n_0}X_1X_0^{n_1}\dots) = \mathrm{tr}\left[\left(\prod_{i=0}^{M}(qX_0 + 1 + i\alpha)^{n_i}\right)X_2X_1^M\right]. \tag{3.85}$$

Now we would like to relate these weights to the weight of some reference state. To do this, we consider a state with exactly one empty site. By commuting the X_0 the whole way around the system, we see that the weight of this state must

satisfy

$$\mathrm{tr}(X_2 X_1^M X_0) = \mathrm{tr}(X_2 X_1^M (qX_0 + 1 + M\alpha)). \tag{3.86}$$

This allows us to establish a relation between the weight of this state and a state with no empty sites,

$$\mathrm{tr}(X_2 X_1^M X_0) = \frac{1 + M\alpha}{1 - q} \mathrm{tr}(X_2 X_1^M). \tag{3.87}$$

Thus we can finally write the weight of an arbitrary configuration as

$$\mathrm{tr}(X_2 X_0^{n_0} X_1 X_0^{n_1} \dots) = \left(\prod_{i=0}^{M} \left(q \frac{1 + M\alpha}{1 - q} + 1 + i\alpha \right)^{n_i} \right) \mathrm{tr}(X_2 X_1^M). \tag{3.88}$$

Then, as in section 3.5, we define the matrix

$$C = X_1 + zX_0, \tag{3.89}$$

though note that now X_0 and X_1 are defined by (3.81a) and (3.81b). Moreover, we define the normalisation as

$$Z_{L,M} = \{z^N\} \mathrm{tr}(X_2 C^L). \tag{3.90}$$

Then, similarly to (3.27), we can again use the generating function trick from section 2.B to write the normalisation in integral form,

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i} z} \exp \left\{ - \sum_{i=0}^{M} \log \left[1 - \left(q \frac{1 + M\alpha}{1 - q} + 1 + i\alpha \right) z \right] - N \log z \right\} \tag{3.91}$$

In this case, it is convenient to use the scaling variable

$$\zeta = \frac{M\alpha}{1 - q} z. \tag{3.92}$$

We again approximate the sum with an integral. After some algebra, we obtain,

$$Z_{L,M} = \left(\frac{M\alpha}{1 - q} \right)^N \mathrm{e}^M \oint \frac{\mathrm{d}\zeta}{2\pi \mathrm{i} \zeta} \mathrm{e}^{L\phi(\zeta)} \tag{3.93a}$$

$$\phi(\zeta) = \rho \left(\frac{1 - \zeta}{(1 - q)\zeta} \log(1 - \zeta) - \frac{1 - q\zeta}{(1 - q)\zeta} \log(1 - q\zeta) \right) - (1 - \rho) \log \zeta \tag{3.93b}$$

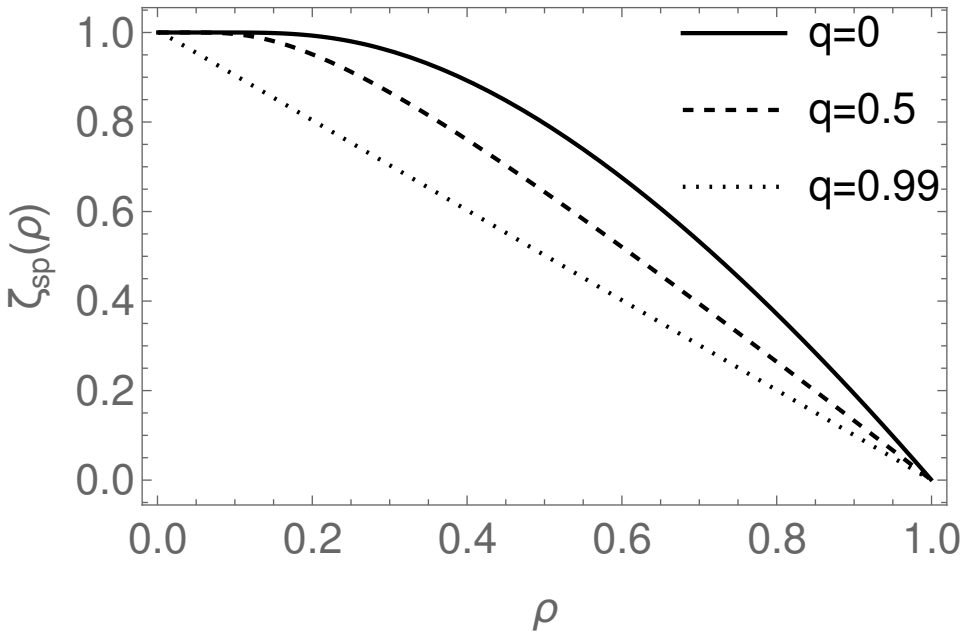

Figure 3.9 *Location of the saddle point, $\zeta_{\rm sp}$, (as defined by* (3.94)*) against mean density, ρ, for a partially asymmetric defect with various asymmetry ratios, $q = \beta/\alpha$. In the limit $q \to 0$, the curve for a totally asymmetric defect is reproduced. In the limit $q \to 1$, the curve approaches the line $1 - \rho$.*

Then the saddle point equation (obtained from setting $\phi'(\zeta) = 0$ in (3.88)) has the form,

$$\frac{1-\zeta_{\rm sp}}{1-q\zeta_{\rm sp}} = \mathrm{e}^{-(1-q)\zeta_{\rm sp}/\rho}. \tag{3.94}$$

The numerical solution for various values of q is plotted in figure 3.9.

The expression for the mean velocity takes a similar form to the totally asymmetric case, though note that we now also need to subtract the backwards hopping rate. Using (3.81b), we obtain,

$$\langle v\rangle_{L,M} = \frac{\{z^{N-1}\}[\alpha \mathrm{tr}(X_2X_0C^{L-1}) - \beta \mathrm{tr}(X_0X_2C^{L-1})]}{Z_{L,M}} = \alpha\frac{Z_{L-1,M}}{Z_{L,M}}. \tag{3.95}$$

Hence at first order in the saddle point, we again have an expression similar to (3.57),

$$\langle v\rangle_{L,M} \approx (1-q)\frac{\zeta_{\rm sp}}{M} + O(L^{-2}), \tag{3.96}$$

except now the location of the saddle point is defined by (3.94).

Deriving the density profile is somewhat harder in this case. We can repeat the calculations from the totally asymmetric defect case up to (3.62). At this point it is harder to proceed, as there is no easy way to obtain a nice expression for the density at site 1, $\langle \eta_1 \rangle_{L-j,M}$. To get an approximate result (which will be seen to be correct in the asymptotic limit), we make the assumption that $\langle \eta_1 \rangle_{L-j,M} \approx \langle \eta_1 \rangle_{L,M}$. This is certainly a bad approximation for large j, but the ratio of partition functions $Z_{L-j,M}/Z_{L,M}$ ensures that those terms are diminished by a factor of M^{-j}, so the error is of subleading order. Then, proceeding similarly to the totally asymmetric defect case, we obtain the expression,

$$\langle \eta_i \rangle_{L,M} \approx \langle \eta_1 \rangle_{L,M} \exp\left(-(1-q)\frac{\zeta_{\mathrm{sp}}}{\rho}\frac{i-1}{L}\right). \tag{3.97}$$

Similarly to the totally asymmetric defect case, this is an exponential decay with a characteristic length $\xi = L\rho/(\zeta_{\mathrm{sp}}(1-q))$. However, now the density in front of the defect is not close to 1 but instead takes some finite value. To determine it, we note that the relation (3.81b) implies the boundary condition

$$(1 - \langle \eta_1 \rangle_{L,M}) - q(1 - \langle \eta_L \rangle_{L,M}) = \alpha \frac{Z_{L-1,M}}{Z_{L,M}}. \tag{3.98}$$

This allows us to obtain the result

$$\langle \eta_1 \rangle_{L,M} \approx \frac{1-q}{1 - q\exp(-(1-q)\zeta_{\mathrm{sp}}/\rho)} = 1 - q\zeta_{\mathrm{sp}}, \tag{3.99}$$

which together with (3.97) gives the full density profile. This is shown in Figure 3.10.

We can verify that this reproduces the expected results in the limits $q \to 0$ and $q \to 1$. In the limit $q \to 0$, the defect becomes totally asymmetric. Then the saddle point equation (3.94) becomes identical to (3.42) and $\langle \eta_1 \rangle_{L,M} \approx 1$, which is consistent with the results from section 3.5.

In the limit $q \to 1$, the defect becomes symmetric and is therefore simply a tracked normal particle. We thus expect a flat density profile by symmetry. Indeed, the characteristic length of the profile diverges as $(1-q)^{-1}$, which is consistent with a flat profile. Expanding to leading order in $(1-q)$, the saddle point equation (3.94) becomes $\zeta_{\mathrm{sp}} = 1-\rho$. Then the density at site 1 becomes $\langle \eta_1 \rangle_{L,M} = 1-(1-\rho) = \rho$, which is also consistent with a flat profile.

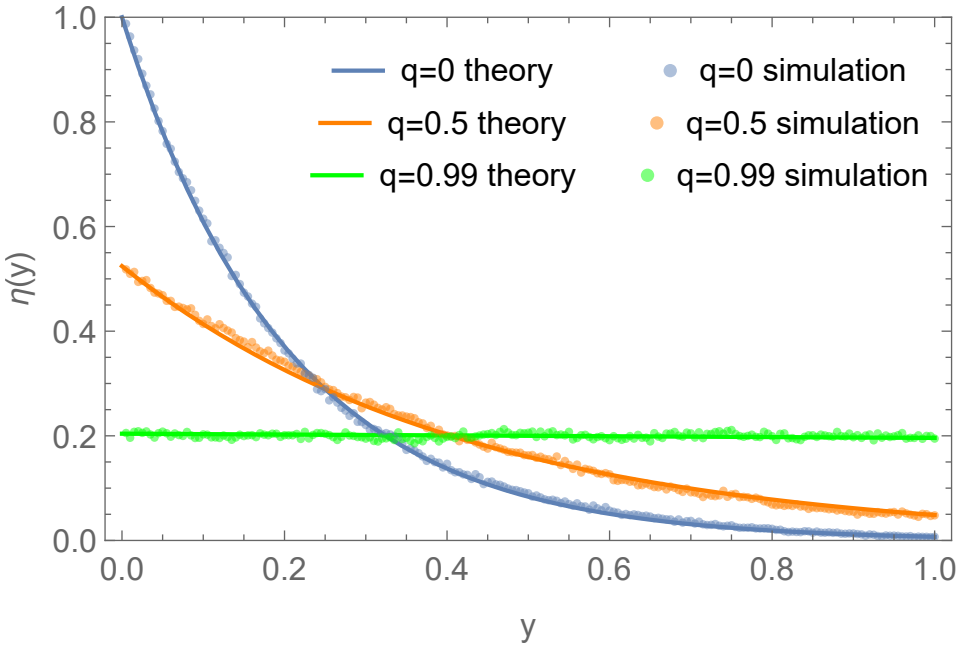

Figure 3.10 *Density profiles as given by* (3.97) *and as obtained from Monte Carlo simulations with mean density* $\rho = 0.2$ *for different defect asymmetry ratios,* q*. The system size used in simulations was* $L = 200$*. For* $q \to 0$*, the profile converges to the totally asymmetric result. For* $q \to 1$*, the profile becomes flat.*

3.7 Summary

In this chapter, we examined a symmetric simple exclusion process on a ring with a single driven tracer particle. We analysed it using mean-field theory and also constructed exact steady-state measures using a mapping to the zero-range process 3.4 and the matrix product formalism 3.5. This allowed us to calculate the steady-state tracer velocity 3.5.1 and density profile 3.5.2. We found that the tracer velocity (3.57) scales as L^{-1} in the large system size limit, which is typical of a driven system with diffusive scaling. At the same time, the density profile (3.67) is non-uniform on the scale of the system size, with a characteristic length $\xi \sim L$. This shows that even a single particle can be enough to drive a system out of equilibrium on a macroscopic scale.

For the system considered in this chapter, the mean-field expressions for the density profile (3.17) and the tracer velocity (3.18) agree with the exact expressions in the asymptotic limit $L \to \infty$, (3.67) and (3.57) respectively. However, the exact approaches using the ZRP mapping and the matrix product ansatz also give the full steady-state measure. In chapter 4, we will examine a

system in which mean-field theory yields incorrect expressions for lengthscales, even in the asymptotic limit.

We then generalised the system in two ways. The first way was to consider k totally asymmetric tracers 3.6.1. We showed that the steady-state measure for the k-defect system factorises into k measures for subsystems containing one defect, each of which is the same as the one-defect measure (3.70). This allowed us to use the measure from the one-defect case to analyse the k-defect case. We found that the steady-state velocity (3.75) depends on how the normal particles are distributed between the one-defect subsystems. The slowest configuration is when defects are all beside each other, in which case the system has the same velocity as a one-defect system. The fastest configuration is when the particles are distributed evenly between the defects, in which case one observes a cooperation effect and the velocity increases up to k times the velocity of a one-defect system of the same total system size and mean density.

The second generalisation we considered was to take a partially asymmetric defect 3.6.2. We were able to construct the steady-state measure (3.88) for this case, again using the matrix product formalism. Similarly to the totally asymmetric case, we found the steady-state velocity (3.96) and density profile (3.97). We showed how these results converge to the totally asymmetric tracer and symmetric tracer cases in the corresponding limits. We again found that the velocity scales with system size as L^{-1}. The density profile is also non-uniform with a characteristic length $\xi \sim L$. The main difference is that in the totally asymmetric defect case, the density on the first site in front of the defect is close to 1, whereas in the partially asymmetric defect case, it has a finite value (3.99).

It is interesting to note that in all three cases we examined, the steady-state velocity is given at leading order by an expression of the form

$$\langle v \rangle_{L,M} \approx \alpha z_{\mathrm{sp}}, \tag{3.100}$$

where z_{sp} is the location of the saddle point of a contour integral. The only difference between these cases is the location of the saddle point z_{sp}, which we summarise in table 3.1.

We remark that there is a substantial jump in complexity of the measure from the totally asymmetric tracer case (3.25) to the partially asymmetric tracer case (3.88). In particular, the weights in the latter case depend on the number of

	One TA tracer	Many TA tracers	One PA tracer
Scaling	$z = \zeta/(M\alpha)$	$z = \zeta/(M\alpha)$	$z = (1-q)\zeta/(M\alpha)$
Saddle point	$1 - \zeta_{\mathrm{sp}} = \mathrm{e}^{-\zeta_{\mathrm{sp}}/\rho}$	$\prod_{i=1}^{k}(1 - m_i\zeta_{\mathrm{sp}}) = \mathrm{e}^{-\zeta_{\mathrm{sp}}/\rho}$	$\frac{1-\zeta_{\mathrm{sp}}}{1-q\zeta_{\mathrm{sp}}} = \mathrm{e}^{-(1-q)\zeta_{\mathrm{sp}}/\rho}$

Table 3.1 *Locations of saddle points for one totally asymmetric tracer, many totally asymmetric tracers and one partially asymmetric tracer.*

normal particles M, whereas in the totally asymmetric case, even when we have multiple tracers, the weights only depend on their position relative to one defect. This is similar to what we saw in section 2.2.3. In totally asymmetric systems, components are in some sense independent, whereas in partially asymmetric systems, there emerges a sort of global interaction, with each local weight depending on all system parameters, as seen in (2.44). This increase in complexity has implications for how far we can extend the approach presented in this chapter. Although we could in principle use a similar matrix product ansatz to obtain a formal solution for the multiple partially asymmetric tracer case, it would be much more difficult to extract physically meaningful results.

In the remainder of this thesis, we will deal with partially asymmetric exclusion processes. This means that the driving is global, rather than coming from a single particle. Correspondingly, we will see ballistically, rather than diffusively scaling currents. The asymmetry of the normal particles adds an extra parameter to the model, which means that we can expect richer behaviour.

Chapter 4

Driven defect particle with priority in a driven lattice gas

4.1 Introduction

In the previous chapter, we studied a symmetrically diffusing gas with a driven defect particle. In this chapter, we turn to the more general case of a partially asymmetric gas with a defect particle that hops at a different rate and may overtake the gas particles. Part of the work presented in this chapter has been published in [2].

4.1.1 Background

As was discussed in section 1.5.3, the most general two species problem has 6 independent rates (one of which may be set to 1 by choosing appropriate time units),

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{q}{\overset{p}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} 21 \; . \tag{4.1}$$

In the literature, much attention has been given to certain special cases.

One such case is that of a TASEP with a defect particle that hops with a different rate and may be overtaken by normal particles (for instance [11, 12, 118]). In our notation, this corresponds to setting $x = q = \beta = 0$. This is essentially the most

general totally asymmetric case.

For partially asymmetric models, three classes of models have mainly been considered. These are:

(i) Models with no overtaking, $\alpha = \beta = 0$ (see section 2.2.3),

(ii) "Second-class" particle models, $p = \alpha = 1$, $q = \beta = x$ [68, 71, 103],

(iii) Models of the type studied by Evans *et al.* [139, 140] and later Arndt *et al.* [141, 142] and others [53, 143]. The most general version of this class requires $q = \beta = 0$ (see [53]).

We note that this list does not include models which satisfy all of the following conditions simultaneously: *(a)* both species are partially asymmetric, *(b)* the species have different hopping rates and *(c)* overtaking is allowed.

In this chapter, we examine a model that satisfies all these conditions. We make the simplifying assumption that the rate at which one of the species overtakes the other is the same as its usual hopping rates, $\alpha = p$, $\beta = q$. This simplifying assumption does not violate the conditions listed above. Also, with this assumption, one species may be seen as having priority in the dynamics over the other. Such hierarchical models (such as *(ii)*) tend to be amenable to exact methods.

As a further simplification, we take only one particle of one of the species. We make the somewhat unusual choice to take the particle with priority to be the defect. As a particle with lower priority does not affect the dynamics of the bath it is in, if we made the defect have lower priority, the bath would behave as a regular PASEP. Such "second-class" particle problems have been studied much in the literature. The idea with those systems is to use the second-class particle as a probe into the behaviour of a single-species PASEP. On the other hand, if we choose the defect to have higher priority, it will disrupt the local density of bath particles, thus creating the possibility of some novel collective phenomena. Thus this choice is more interesting from the point of view of one of the main goals of this thesis, namely to investigate the effects of defect particles on nonequilibrium steady states.

4.2 Model definition

We consider an $(L+1)$-site ring with $M+1$ particles. We define the mean density as $\rho = M/L$, as usual. M particles are "normal" partially asymmetric particles that hop to the right with rate 1 and to the left with rate x. We will take $x \leq 1$ without loss of generality. The last particle is a partially asymmetric defect that hops to the right with rate p and to the left with rate q. The normal particles obey the simple exclusion rule: when they attempt a hop onto a site that is already occupied by another normal particle or the defect, the hop is not allowed. The defect, on the other hand, is taken to have priority in the dynamics. When it attempts a hop onto a site that is occupied by a normal particle, the hop is allowed and the defect overtakes the normal particle.

Denoting empty sites, the defect and normal particles by 0, 1, 2 respectively, the dynamics can be summarised as follows,

$$10 \underset{q}{\overset{p}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{x}{\overset{1}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{q}{\overset{p}{\rightleftarrows}} 21. \tag{4.2}$$

Note that we have changed the notation from (4.1), with the defect now being denoted by 1 and the normal particles by 2. This is because the defect particle has priority in the dynamics, so it makes more sense to call it a "first-class" particle, with the normal particles being "second-class".

Note also that the defect overtakes normal particles with the same rates as it hops on an empty lattice. In other words, the defect does not distinguish between particles and empty sites. This means that the motion of the defect will be simply that of a free particle executing biased diffusion with hopping rates p, q. While the motion of the defect is trivial, it disrupts the local density of normal particles. The effect of this on the steady state of the normal particles is not obvious *a priori*, and indeed we will see in the remainder of this chapter that it generates surprisingly rich behaviour.

4.3 Mean-field theory

We first analyse the steady state of this system using mean-field theory. Although we will present an exact solution in section 4.4, the mean-field theory allows us

to build intuition. It also allows us to address a broader problem, as we will need to impose an extra condition on the defect parameters p, q to be able to write down an exact solution.

4.3.1 Hydrodynamics

We wish to obtain the steady-state density profile using mean-field theory. In a similar manner to chapter 3, this is most naturally done in the reference frame of the defect. Thus we take the defect always to be on site 0 and the remaining M particles on sites $1, \ldots, L$. We define η_i as a random variable that is equal to 1 if site i is occupied and 0 if it is empty.

The equations of motion for $\langle \eta_i \rangle$ are given as follows, for $i = 2, \ldots, L-1$:

$$\begin{aligned} \frac{\mathrm{d}\langle \eta_i \rangle}{\mathrm{d}t} &= -\langle \eta_i(1-\eta_{i+1})\rangle - x\langle (1-\eta_{i-1})\eta_i\rangle + \langle \eta_{i-1}(1-\eta_i)\rangle + x\langle (1-\eta_i)\eta_{i+1}\rangle \\ &\quad -p\langle \eta_i\rangle - q\langle \eta_i\rangle + p\langle \eta_{i+1}\rangle + q\langle \eta_{i-1}\rangle, \end{aligned} \tag{4.3a}$$

and for the boundary terms $i = 1$, L:

$$\begin{aligned} \frac{\mathrm{d}\langle \eta_1 \rangle}{\mathrm{d}t} &= -\langle \eta_1(1-\eta_2)\rangle + x\langle (1-\eta_1)\eta_2\rangle - p\langle \eta_1\rangle - q\langle \eta_1\rangle \\ &\quad +p\langle \eta_2\rangle + q\langle \eta_L\rangle \end{aligned} \tag{4.3b}$$

$$\begin{aligned} \frac{\mathrm{d}\langle \eta_L \rangle}{\mathrm{d}t} &= -x\langle (1-\eta_{L-1})\eta_L\rangle + \langle \eta_{L-1}(1-\eta_L)\rangle - p\langle \eta_L\rangle - q\langle \eta_L\rangle \\ &\quad +p\langle \eta_1\rangle + q\langle \eta_{L-1}\rangle. \end{aligned} \tag{4.3c}$$

We now take the continuum limit by setting $y = i/L$ and letting $L \to \infty$ (with mean density $\rho = M/L$ fixed). Then we need to expand each local density in gradients,

$$\langle \eta_i \rangle = \eta(y,t) \tag{4.4a}$$

$$\langle \eta_{i\pm1} \rangle = \eta(y,t) \pm \frac{1}{L}\nabla\eta(y,t) + \frac{1}{2L^2}\nabla^2\eta(y,t) + O(L^{-3}). \tag{4.4b}$$

As this is a case of finite asymmetry, we must take the ballistic scaling for time, $\tau = L\tau$ (see section 1.5). Putting this into (4.3a), we obtain,

$$\frac{\partial\eta(y,\tau)}{\partial\tau} = v'\nabla\eta(y,\tau) - v\nabla[\eta(y,\tau)(1-\eta(y,\tau))] + D\nabla^2\eta(y,\tau), \tag{4.5}$$

where we have introduced

$$v = 1 - x, \qquad v' = p - q, \qquad D = \frac{1 + x + p + q}{2L}. \tag{4.6}$$

Strictly speaking, the term with D is subdominant as it is of order L^{-1} compared to the other terms, unless we take the weak asymmetry limit, $v \sim O(L^{-1})$, $v' \sim O(L^{-1})$. We keep it here as it allows us to write a mean-field prediction for the microscopic structure of the shock, though we will see that this gives incorrect scaling for the characteristic shock width.

Setting the D term to 0, equation (4.5) becomes similar to the inviscid Burgers' equation 1.64, which one typically obtains for homogeneous PASEPs (see section 1.5). However, here we have an additional constant current term $v'\nabla\eta(y)$, due to the fact that we are essentially in a Galileo-boosted frame.

To complete the problem, we expand similarly the boundary conditions (4.3b), (4.3c), to obtain,

$$\begin{aligned}
\frac{\partial\eta(0,\tau)}{\partial\tau} &= L[-v\eta(0,\tau)(1-\eta(0,\tau)) - q\eta(0,\tau) + q\eta(1,\tau)] \\
&\quad + \left(p + \frac{1+x}{2}\right)\nabla\eta(0,\tau) - \frac{v}{2}\nabla[\eta(0,\tau)(1-\eta(0,\tau))] \qquad (4.7a) \\
\frac{\partial\eta(1,\tau)}{\partial\tau} &= L[v\eta(1,\tau)(1-\eta(1,\tau)) - p\eta(1,\tau) + p\eta(0,\tau)] \\
&\quad - \left(q + \frac{1+x}{2}\right)\nabla\eta(1,\tau) - \frac{v}{2}\nabla[\eta(1,\tau)(1-\eta(1,\tau))]. \qquad (4.7b)
\end{aligned}$$

The mismatch of powers of L is indicative of a separation of timescales. The density profile very close to the defect reaches its steady-state form on a super-ballistic timescale, such that the following relation holds for any finite $\tau > 0$,

$$\begin{aligned}
-v\eta(0,\tau)(1-\eta(0,\tau)) - q\eta(0,\tau) + q\eta(1,\tau) &= 0 \qquad (4.8a) \\
v\eta(1,\tau)(1-\eta(1,\tau)) - p\eta(1,\tau) + p\eta(0,\tau) &= 0. \qquad (4.8b)
\end{aligned}$$

4.3.2 Steady state

We wish to calculate the steady-state density profile $\eta(y)$. To do this, we set the time derivative in (4.5) to 0. Then integrating once with respect to y, we obtain

$$v\eta(y)(1-\eta(y)) - v'\eta(y) = J' = \text{const.} \tag{4.9}$$

We use J' here to denote the current in the reference frame of the defect. It is related to J, the current in the "lab frame", by,

$$J = J' + \rho v'. \tag{4.10}$$

The boundary conditions (4.8a), (4.8b) become,

$$J' \;=\; -(p\eta(0) - q\eta(1)). \tag{4.11}$$

Actually, we will later see that this definition for the current in the reference frame of the defect also holds at the exact level (see appendix 4.B). Indeed, as this expression only involves one-point functions and not two-point functions, the mean-field assumption does not enter here.

Then equation (4.9) suggests that we obtain uniform density profiles, as it does not involve any spatial derivatives. However, as it is quadratic in η, the same constant J' permits two different densities given by,

$$\rho_\pm = \frac{v - v' \pm \sqrt{(v-v')^2 - 4vJ'}}{2v}. \tag{4.12}$$

At this point, it is helpful to reintroduce the constant $D \sim L^{-1}$, as defined in (4.6). With this, the equation for the steady-state density profile becomes,

$$D\nabla\eta(y) = -v(\eta(y) - \rho_-)(\eta(y) - \rho_+). \tag{4.13}$$

From this, we see that $\eta(y) = \rho_\pm$ are fixed points in the sense that we necessarily have $\nabla\eta(y) = 0$ there.

We can solve this equation using an approach similar to the one used to solve the weakly asymmetric simple exclusion process in section 1.5. We obtain solutions

of the form

$$\eta(y) = \frac{v - v'}{2v}\left[1 + jf\left(\frac{v - v'}{2D} j(y - y_0)\right)\right], \tag{4.14}$$

where

$$f = \begin{cases} \tanh, & \rho_- < \rho < \rho_+ \\ \coth, & \text{otherwise} \end{cases}, \tag{4.15}$$

$$j = \sqrt{1 - 4\frac{vJ'}{(v - v')^2}}, \tag{4.16}$$

and y_0 is a constant of integration.

It is not entirely straightforward to obtain an explicit solution from this. The constants J', y_0 have to be fixed by the boundary condition (4.11) and the global density constraint, $\int_0^1 \mathrm{d}y\ \eta(y) = \rho$. Then one has to verify that the choice of f was consistent by putting J' into (4.12) and checking that ρ lies in the correct range. Overall, this approach is not very practical.

However, the form of this solution allows us to obtain an intuitive understanding of the types of density profiles and the phase diagram. In the coth case, y_0 must be fixed outside the range $[0, 1]$, as otherwise there would be a divergence in the density profile. Away from y_0, the profile is essentially an exponential decay to ρ_+ or ρ_- with a decay length $\xi = 2Dj/(v - v') \sim L^{-1}$. The profile is therefore almost uniform, with an exponentially small structure near $y = 0$ if $y_0 < 0$, or near $y = 1$ if $y_0 > 1$. We call these profiles right-localised ($\mathcal{L}_R$) and left-localised ($\mathcal{L}_L$) respectively. The mean-field prediction for the scaling of the decay length is correct, although its the precise value will be shown not to agree with the exact result.

In the tanh case, we have a profile that interpolates between the densities ρ_- and ρ_+, with a crossover at y_0 with a characteristic length $\xi = 2Dj/(v - v')$. Note that as $D \sim L^{-1}$, this crossover becomes a shock discontinuity in the continuum limit. From the exact solution, we will see that this scaling is incorrect and the width of the shocks in fact scales as $L^{-1/2}$.

The three types of profiles (left-localised, right-localised and shock) are shown schematically in figure 4.1.

Using this as a guide, we can analyse the density profiles more closely. Specifically,

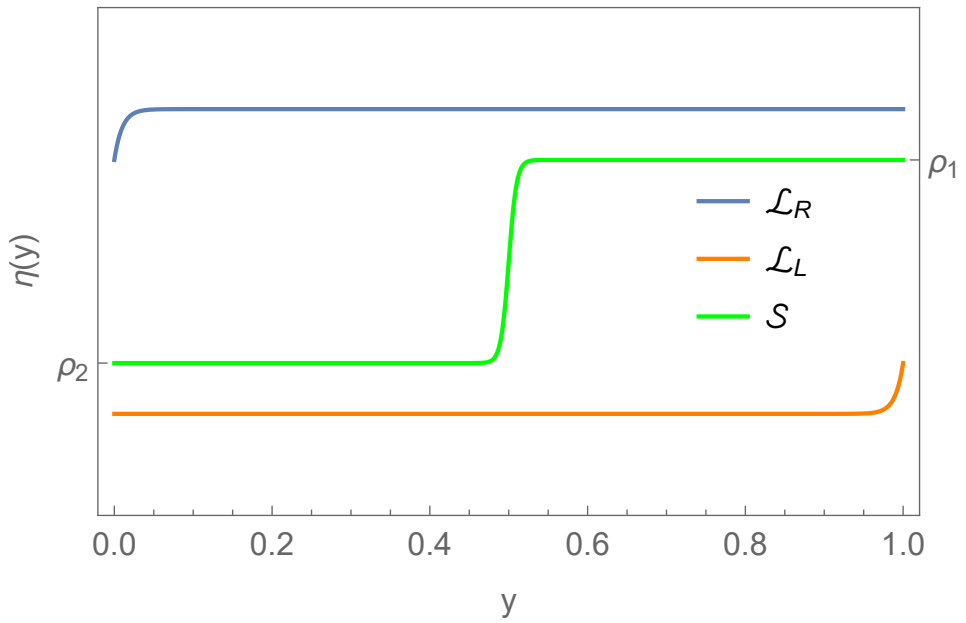

Figure 4.1 *Schematic representation of the three types of steady-state density profiles predicted by mean-field theory.*

we assume the shape of the profile (uniform for coth or shock for tanh) and then using the conditions (4.9), (4.11), (4.12), we can find the ranges of parameters in which these assumptions are valid, thereby determining the phase diagram.

Localised profiles

First, consider the profiles of the coth form. These profiles are mostly uniform,

$$\eta(y) = \rho. \tag{4.17}$$

The current (4.9) is given by,

$$J' = v\rho(1-\rho) + v'\rho. \tag{4.18}$$

Then we have the fixed points (4.12),

$$\rho_+ = \rho, \qquad \rho_- = 1 - \rho - \frac{v'}{v}. \tag{4.19}$$

Hence we must choose the solutions that decay to ρ_+.

Now if $y_0 > 1$, then there is a slight deviation from $\eta(y) = \rho$ close to $y = 1$, while $\eta(0) = \rho$. Conversely, if $y_0 < 1$, then $\eta(1) = \rho$ and there is a slight deviation close

to $y = 0$.

Right-localised phase. First, we consider the localised case when $\eta(1) = \rho$. This means that $\eta(0) \neq \rho$, so there is a small localised structure to the right of the defect (hence we call this the right-localised phase). From (4.11), we obtain,

$$\eta(0) = \rho - \frac{v}{p}\rho(1-\rho), \qquad \eta(1) = \rho. \tag{4.20}$$

For this solution to be self-consistent, we must have $\eta(0) > \rho_-$. This gives the consistency requirement,

$$\rho > \rho_1 = \frac{v - 2p + \sqrt{v^2 + 4pq}}{2v}. \tag{4.21}$$

Left-localised phase. The analysis of the other localised phase proceeds similarly. We take $\eta(0) = \rho$, which means that $\eta(1) \neq \rho$, so there is a small localised structure to the left of the defect (hence we call this the left-localised case). From (4.11), we obtain,

$$\eta(0) = \rho, \qquad \eta(1) = \rho + \frac{v}{q}\rho(1-\rho). \tag{4.22}$$

For this solution to be self-consistent, we must have $\eta(1) < \rho_-$. This gives the consistency requirement,

$$\rho < \rho_2 = \frac{v + 2q - \sqrt{v^2 + 4pq}}{2v}. \tag{4.23}$$

Shock profile

Now consider the profiles of the tanh form. These profiles have two bulk densities,

$$\eta(y) = \begin{cases} \rho_-, & y < y_0 \\ \rho_+, & y > y_0 \end{cases}, \tag{4.24}$$

where the crossover location y_0 is fixed by the global density restriction,

$$\rho_- y_0 + \rho_+(1 - y_0) = \rho. \tag{4.25}$$

This case is a little more tricky because the profile is not uniform, so we cannot immediately write down the current J'. We know that the two densities $\rho_\pm$ have the same current J', so from (4.9), we must have,

$$\rho_+ + \rho_- = 1 - \frac{v'}{v}, \qquad \rho_+\rho_- = \frac{J'}{v}. \tag{4.26}$$

Also from the boundary condition (4.11), we have

$$J' = -p\rho_- + q\rho_+. \tag{4.27}$$

We can solve this system of equations to obtain,

$$\rho_+ = \rho_1 = \frac{v - 2p + \sqrt{v^2 + 4pq}}{2v} \tag{4.28}$$

$$\rho_- = \rho_2 = \frac{v + 2q - \sqrt{v^2 + 4pq}}{2v} \tag{4.29}$$

$$J' = \frac{-vv' - 4pq + (p+q)\sqrt{v^2 + 4pq}}{2v}. \tag{4.30}$$

Interestingly, note that in this phase ρ_+, ρ_- and J' do not depend on the global mean density ρ. The mean density ρ only affects the location of the shock front, y_0.

To obtain the self-consistency condition, we note that for y_0 to have a solution in $y \in [0, 1]$, we must have,

$$\rho_2 < \rho < \rho_1. \tag{4.31}$$

This is consistent with the regimes of validity of the localised phases, (4.21), (4.23). From these three inequalities we thus obtain the phase diagram. This is shown in figure 4.2.

We can introduce a weaker condition that helps to understand the mechanism behind the shock formation. Note that a necessary requirement for the shock phase to make sense, is that $\rho_2 < \rho_1$, which can equivalently be expressed as,

$$|v'| < |v|. \tag{4.32}$$

This shows that the shock is caused by a the speed of the defect being in a "Goldilocks" zone: not too fast and not too slow. Specifically, v' must be small enough that the normal particles must try to overtake the defect, *i.e.* $v' < v$. At

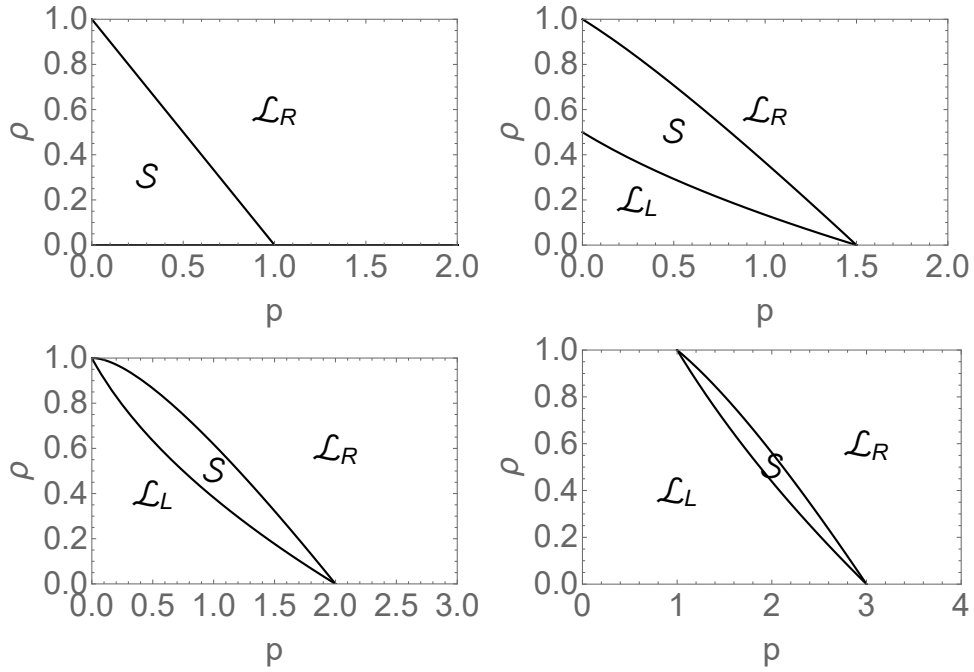

Figure 4.2 *Phase diagram as predicted by mean-field theory (*(4.21), (4.23), (4.31)*) with p and ρ as control parameters. The other parameters used were $x = 0$ (all); and $q = 0$ (top left), $q = 0.5$ (top right), $q = 1$ (bottom left), $q = 2$ (bottom right). The phases are labelled right-localised ($\mathcal{L}_R$), left-localised ($\mathcal{L}_L$) and shock ($\mathcal{S}$).*

the same time, if q is too large, the normal particles will overtake the defect too quickly and a shock will not form. This means that we must have $-v < -v'$.

The latter inequality can also be understood by thinking about the motion of holes. Note that holes in the system move with a mean drift velocity $-v$. A shock profile with densities ρ_1, ρ_2 for particles means an "anti-shock" profile with densities $(1 - \rho_1)$, $(1 - \rho_2)$ for holes. As a shock for particles can only occur if $v' < v$, equivalently a shock for holes can only occur if $-v' > -v$. As a shock for either species implies a shock for the other, both conditions must be satisfied for a shock to form.

4.4 Matrix product solution

We now present an exact solution for the steady state of a special case of the model discussed in the previous section. The matrix product solution presented

here is valid provided the following condition holds,

$$pq = x. \tag{4.33}$$

This means that the product of the hopping rates of the defect equals that of the normal particles. Whether or not the solution can be generalised to the case when this relation does not hold remains an open problem.

It is convenient to relabel the defect parameters using a single parameter, α. Specifically, we let

$$p = \alpha, \qquad q = x/\alpha. \tag{4.34}$$

The change of notation $p \to \alpha$ will help us differentiate results for the general case from the special case.

Initially, we may try to look for a matrix product solution with usual scalar matrix product relations. Recall the consistency relations (2.52a–2.52f). In that notation, we have $w_{10} = w_{12} = \alpha$, $w_{01} = w_{21} = x/\alpha$, $w_{20} = 1$, w_{02}. Then only relation (2.52b) is automatically satisfied. To satisfy all the conditions without any further restrictions on the parameters, we must therefore set at least $x_0 = x_2 = 0$. However, this implies the matrix reduction relation

$$X_2 X_0 - x X_0 X_2 = 0. \tag{4.35}$$

This is too restrictive and cannot be used to describe the actual steady state of this model.

To see this, note that given any configuration, we can use this relation to move all X_2 to the left and all X_0 to the left, only introducing some factors of x. We can check whether this is sufficient by calculating the steady-state measure for a small system exactly. For instance, taking $L = 4$, $M = 2$, we find using `Mathematica`,

$$\begin{aligned}
P(\{12200\}) &= Z_{4,2}^{-1}(\alpha^4 + x\alpha^3 + x^2\alpha^2 + x^2\alpha + x^2) && \text{(4.36a)}\\
P(\{12020\}) &= Z_{4,2}^{-1}(\alpha^4 + x\alpha^3 + x\alpha^2 + x^2\alpha + x^2) && \text{(4.36b)}\\
P(\{10220\}) &= Z_{4,2}^{-1}(\alpha^4 + \alpha^3 + x\alpha^2 + x^2\alpha + x^2) && \text{(4.36c)}\\
P(\{12002\}) &= Z_{4,2}^{-1}(\alpha^4 + x\alpha^3 + x\alpha^2 + x\alpha + x^2) && \text{(4.36d)}\\
P(\{10202\}) &= Z_{4,2}^{-1}(\alpha^4 + \alpha^3 + x\alpha^2 + x\alpha + x^2) && \text{(4.36e)}\\
P(\{10022\}) &= Z_{4,2}^{-1}(\alpha^4 + \alpha^3 + \alpha^2 + x\alpha + x^2), && \text{(4.36f)}
\end{aligned}$$

with the normalisation,

$$Z_{4,2} = 6\alpha^4 + 3x\alpha^3 + 3\alpha^3 + x\alpha^2 + 4x\alpha^2 + x^2\alpha^2 + 3x\alpha + 3x^2\alpha + 6x^2. \quad (4.36g)$$

Evidently the weights of different states differ by more than just factors of x.

Although we cannot construct matrices with the usual matrix reduction relations, the structure of the solution for small system sizes is simple enough to allow us to guess matrices that will generate the required weights. These are given by,

$$X_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} \alpha & 0 \\ \alpha & x \end{pmatrix}, \qquad X_0 = \begin{pmatrix} \alpha & 0 \\ \alpha & 1 \end{pmatrix}. \quad (4.37)$$

4.4.1 Proof of validity

The matrices (4.37) do not satisfy the typical scalar reduction relations. In fact, it is not even obvious whether there exist auxiliary matrices $\tilde{X}_0$, $\tilde{X}_1$, $\tilde{X}_2$, to complete the standard proof (see section 2.3). We therefore reconstruct the proof explicitly.

To satisfy the global balance condition (see section 1.3), we must have,

$$\sum_{i=0}^{L} [w_{\tau_i,\tau_{i+1}} \mathrm{tr}(\dots X_{\tau_i} X_{\tau_{i+1}} \dots) - w_{\tau_{i+1},\tau_i} \mathrm{tr}(\dots X_{\tau_{i+1}} X_{\tau_i} \dots)] = 0. \quad (4.38)$$

We examine separately the cases when one of τ_i, τ_{i+1} is equal to 1 (the defect) and when neither of them is 1.

Defect not involved. First, we consider $\tau_i = 2$, $\tau_{i+1} = 0$. Actually, for this case, a scalar-type reduction relation holds,

$$X_2 X_0 - x X_0 X_2 = x_0 X_2 - x_2 X_0, \quad (4.39)$$

with

$$x_0 = \alpha - x, \qquad x_2 = x(\alpha - 1). \quad (4.40)$$

Then the argument for the cancellation of these terms follows in the same way as the regular case (see section 2.3.2).

Defect involved. Now, we consider $\tau_i = 1$ and $\tau_{i+1} = 0$ or 2. Then we have

$$\alpha X_1 X_\tau - (x/\alpha) X_\tau X_1 = x_\tau X_1 - Y_\tau, \tag{4.41}$$

where

$$Y_0 = \begin{pmatrix} -\alpha^2 & 0 \\ -\alpha^2 & x/\alpha \end{pmatrix}, \qquad Y_2 = \begin{pmatrix} -\alpha^2 & 0 \\ -\alpha^2 & x^2/\alpha \end{pmatrix}. \tag{4.42}$$

The term $x_\tau X_1$ cancels with the adjacent terms telescopically, as usual.

To see how the terms with Y_τ cancel, we consider the terms preceding this one as well (*i.e.* all terms coming from the sequence $X_\tau X_1 X_{\tau'}$). We have,

$$-x_\tau X_1 X_{\tau'} + Y_\tau X_{\tau'} - X_\tau Y_{\tau'} + X_\tau (x_{\tau'} X_1). \tag{4.43}$$

The two outer terms will cancel with the preceding and following terms in the sequence (not included here) as usual, so we do not need to worry about them. For the two inner terms, it is readily checked by explicit calculation that

$$Y_\tau X_{\tau'} - X_\tau Y_{\tau'} = \begin{cases} \begin{pmatrix} 0 & 0 \\ x + \alpha^2 & 0 \end{pmatrix}, & \tau = 0 \\ \begin{pmatrix} 0 & 0 \\ x(x + \alpha^2) & 0 \end{pmatrix}, & \tau = 2 \end{cases}. \tag{4.44}$$

In particular, the only nonzero entry is on the bottom left. Furthermore, it is easily checked that as X_0, X_2 are lower triangular, they preserve this property on multiplication by them (*i.e.* the only nonzero entry will remain the bottom left one),

$$\begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix} \begin{pmatrix} b & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ ac & 0 \end{pmatrix}. \tag{4.45}$$

As there is only one defect, the only other matrices in the product will be X_0 and X_2. Hence the matrix product representing the whole state will be traceless.

This completes the proof that the matrices (4.37) are indeed a solution of the global balance equation (4.38).

4.4.2 Remarks on matrix product solution

Before using the matrix product solution presented in this section, we make some remarks on the solution, particularly in relation to other matrix product solutions.

The first remark we make is that it is very unusual for the matrices to have a finite-dimensional representation. Typically, deformed algebras require an infinite-dimensional representation (see [8]). Two-dimensional representations have been found for certain special cases, such as a PASEP on an infinite lattice with finely tuned bulk densities [144] and a TASEP with a parallel time-updating scheme [145]. Yet, two-dimensional representations are still uncommon. The fact that we have a simple explicit representation will allow us to perform calculations, despite the fact that the typical simplifications using matrix reduction relations cannot be performed.

Secondly, it can be checked, again by solving exactly some small examples, that the solution presented here is only valid for the case of exactly one defect. Indeed, in the proof of the validity of the solution, we made explicit use of the fact that there was only one matrix X_1 in the product. It is possible that some generalisation of the presented solution can solve the multiple defect case, though this remains an open problem.

Thirdly, one might wonder, having obtained this solution, whether it is possible to write a similar solution for the case with general p, q. Unfortunately, it seems that the approach presented here cannot be easily extended to that case. Solving some small systems exactly, one can see that the weights of the measure in the general case are much more complicated than in the special case (4.33). It is therefore not as straightforward to simply write down matrices which would generate those weights. This suggests that there may be no matrix product solution for the general case or, if it exists, it must have a significantly more complex structure.

Interestingly, in chapter 5, we will see that the case with independent p, q is also not solvable using the most general known form of the nested coordinate Bethe ansatz, whereas the case with $p = \alpha$, $q = x/\alpha$ is (for one defect). However, we hasten to reiterate that a link between models solvable by the matrix product approach and those solvable by the Bethe ansatz has not yet been rigorously established in the general case.

Rather than looking to solve the specific generalisations mentioned above (general

p, q and many defect particles), it would be interesting to explore generally whether the novel cancellation scheme presented in section 4.4.1 might uncover a new class of matrix product solutions. All previously known solutions have relied on the exact "telescoping" cancellation of terms, whereas the solution presented here shows that it is possible to have a non-vanishing, but nevertheless traceless, matrix product. This in principle leaves the door open for a wide range of new matrix product solutions and it would be interesting for future research to investigate this possibility in generality.

Finally, as the solution we have presented here takes the form of 2×2 matrices, it is of interest to investigate whether this solution can be interpreted as a transfer matrix of a classical Ising spin chain. We discuss this in appendix 4.A.

4.5 Normalisation

As in chapter 3, we first examine the "nonequilibrium partition function" or normalisation $Z_{L,M}$. This is defined as

$$Z_{L,M} = \{z^M\}\mathrm{tr}(X_1 C^L), \tag{4.46}$$

where

$$C = X_0 + zX_2 = \begin{pmatrix} \lambda_1 & 0 \\ \lambda_1 & \lambda_2 \end{pmatrix}, \tag{4.47}$$

and

$$\lambda_1 = \alpha(1+z), \qquad \lambda_2 = 1 + xz. \tag{4.48}$$

It is then straightforward to obtain a general expression for C^L. Doing the computation explicitly for $L = 1, 2, 3$, we postulate the following general form,

$$C^k = \begin{pmatrix} \lambda_1^k & 0 \\ \lambda_1 \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} & \lambda_2^k \end{pmatrix}, \tag{4.49}$$

Indeed, this is readily proved by induction,

$$C^{k+1} = \begin{pmatrix} \lambda_1^k & 0 \\ \lambda_1 \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} & \lambda_2^k \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ \lambda_1 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1^{k+1} & 0 \\ \lambda_1 \frac{\lambda_1^{k+1} - \lambda_2^{k+1}}{\lambda_1 - \lambda_2} & \lambda_2^{k+1} \end{pmatrix}. \tag{4.50}$$

Then using the generating function trick (see section 2.B), the normalisation (4.46) becomes,

$$Z_{L,M} \;=\; \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} z^{-(M+1)} \sum_{l=0}^{L} \lambda_1^{L-l} \lambda_2^l. \tag{4.51}$$

4.5.1 Finite-size expression

To obtain an explicit finite-size expression, we substitute the definitions of λ_1 and λ_2 into (4.51). Then expanding them using the binomial theorem gives,

$$Z_{L,M} = \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} \sum_{l=0}^{L} \sum_{k=0}^{L-l} \binom{L-l}{k} \alpha^{L-l} \sum_{m=0}^{l} \binom{l}{m} x^m z^{k+m-M-1}. \tag{4.52}$$

Evaluating the residues gives the following result after some simplification,

$$Z_{L,M} = \sum_{l=0}^{L} \sum_{m=0}^{l} \binom{L-l}{M-m} \binom{l}{m} \alpha^{L-l} x^m. \tag{4.53}$$

It is interesting to compare this to the result from the previous chapter, (3.35). On one hand, this expression is more complicated, as it involves a double sum and two parameters. On the other hand, it involves only simple combinatorial objects (binomial coefficients), whereas the previous result involved the somewhat more exotic Stirling numbers.

4.5.2 Asymptotic behaviour

To extract the asymptotic behaviour in the limit $L \to \infty$, with $\rho = M/L$ held fixed, we approximate the sum in (4.51) by an integral by letting $w = l/L$. Then the normalisation can be written as,

$$Z_{L,M} \approx \oint \frac{\mathrm{d}z}{2\pi \mathrm{i} z} L \int_0^1 \mathrm{d}w \; \mathrm{e}^{L\phi(z,w)}, \tag{4.54}$$

where

$$\phi(z,w) = (1-w)\log(\alpha(1+z)) + w\log(1+xz) - \rho\log z. \tag{4.55}$$

It is easiest to analyse this by first evaluating the z integral at the saddle point z_*. This is given implicitly as the solution of the equation

$$\left.\frac{\partial}{\partial z}\phi(z,w)\right|_{z=z_*} = \frac{1-w}{1+z_*} + \frac{xw}{1+xz_*} - \frac{\rho}{z_*} = 0. \tag{4.56}$$

Then at leading order in the saddle point approximation, we have,

$$Z_{L,M} \approx L\int_0^1 \mathrm{d}w\ A(w)\left[\frac{\mathrm{e}^{L\varphi(w)}}{z_*(w)} + O(L^{-1})\right], \tag{4.57}$$

where

$$\varphi(w) = \phi(z_*(w),w) \tag{4.58}$$

$$A(w) = [-2\pi L\partial_z^2\phi(z,w)|_{z=z_*(w)}]^{-1/2}. \tag{4.59}$$

Now the asymptotic behaviour of the w integral is determined by the function φ. The function φ has a maximum w_* at

$$\varphi'(w)|_{w=w_*} = 0. \tag{4.60}$$

Actually this equation is equivalent to

$$\left.\frac{\partial}{\partial w}\phi(z,w)\right|_{z=z_*(w_*),w=w_*} = \log(\alpha(1+z_*(w_*))) - \log(1+xz_*(w_*)) = 0, \tag{4.61}$$

as $\partial_z\phi(z,w)|_{z=z_*} = 0$ by assumption. This is easily solved for $z_*(w_*)$, which can be plugged back in to (4.56) to give,

$$z_*(w_*) = -\frac{\alpha-1}{\alpha-x}, \qquad w_* = \frac{\alpha((\alpha-1)+(1-x)\rho)}{(\alpha-1)(\alpha-x)}. \tag{4.62}$$

Hence φ has a unique turning point, w_*. After some tedious but straightforward algebra, we obtain,

$$\varphi''(w_*) = -\frac{(\alpha-1)^2(\alpha-x)^2}{\alpha((\alpha^2-x)(1-x)\rho + x(\alpha-1)^2)}. \tag{4.63}$$

It can be checked (for instance numerically) that $\varphi''(w_*) < 0$ for any choice of parameters α, x (which are assumed to lie in the ranges $\alpha > 0$, $0 < x < 1$). Therefore φ has a maximum at w_*.

So if $w_* < 0$, then φ is monotonically decreasing on the interval $[0, 1]$. Similarly, it is monotonically increasing if $w_* > 1$. Finally, if $w_* \in (0, 1)$, then the maximum of φ on $(0, 1)$ is at $w = w_*$. Depending on which of these cases we fall into, the w integral in (4.57) will be dominated by $w_{dom} = 0$, 1 or w_* respectively. Substituting the value for w_* from (4.62) into these inequalities, we get the following phase diagram in terms of ρ, x and α,

$$\begin{cases} \rho > \rho_1 & \Rightarrow w_{dom} = 1 \\ \rho_2 < \rho < \rho_1 & \Rightarrow w_{dom} = w_* \\ \rho < \rho_2 & \Rightarrow w_{dom} = 0 \end{cases}, \tag{4.64}$$

where the critical densities ρ_1 and ρ_2 are given by

$$\rho_1 = \frac{1-\alpha}{1-x}, \qquad \rho_2 = \frac{1-\alpha^{-1}}{1-x^{-1}}. \tag{4.65}$$

Note that these definitions of ρ_1, ρ_2 agree with those used in the mean-field analysis (4.21), (4.23) if one sets $p \to \alpha$, $q \to x/\alpha$ in the latter. We will show in the following section that these phases correspond to the left-localised, shock and right-localised density profiles respectively.

First, let us consider the cases when the normalisation is dominated by one of the boundary values, $w_{dom} = 0$ or $w_{dom} = 1$. From (4.56), we have

$$z_*(0) = \frac{\rho}{1-\rho}, \qquad z_*(1) = \frac{\rho}{x(1-\rho)}. \tag{4.66}$$

In the $w_{dom} = 0$ case, we can approximate

$$Z_{L,M} \approx LA(0)\frac{\mathrm{e}^{L\varphi(0)}}{z_*(0)} \int_0^1 \mathrm{d}w \, \left[\mathrm{e}^{L\varphi'(0)w} + O(L^{-1})\right] \tag{4.67}$$

$$\approx -A(0)(\varphi'(0))^{-1}\frac{\mathrm{e}^{L\varphi(0)}}{z_*(0)} + O(L^{-1}). \tag{4.68}$$

Similarly, in the $w_{dom} = 1$ case, we have,

$$Z_{L,M} \approx LA(1)\frac{\mathrm{e}^{L\varphi(1)}}{z_*(1)}\int_0^1 \mathrm{d}w \left[\mathrm{e}^{L\varphi'(1)(w-1)} + O(L^{-1})\right] \tag{4.69}$$

$$\approx A(1)(\varphi'(1))^{-1}\frac{\mathrm{e}^{L\varphi(1)}}{z_*(1)} + O(L^{-1}). \tag{4.70}$$

Finally, in the $w_{dom} = w_*$ case, we have,

$$Z_{L,M} \approx LA(w_*)\frac{\mathrm{e}^{L\varphi(w_*)}}{z_*(w_*)}\int_0^1 \mathrm{d}w \left[\mathrm{e}^{\frac{1}{2}L\varphi''(w_*)(w-w_*)^2} + O(L^{-1})\right] \tag{4.71}$$

$$\approx LA(w_*)\sqrt{\frac{2\pi}{-L\varphi''(w_*)}}\frac{\mathrm{e}^{L\varphi(w_*)}}{z_*(w_*)} + O(L^{-1}). \tag{4.72}$$

4.6 Density profiles and currents

Using the calculations in the previous section as a guide, we can now calculate the steady-state density profiles and currents.

4.6.1 Density profiles

We define the density at site i to the right of the defect as

$$\langle \eta_i \rangle = \frac{\{z^M\}\mathrm{tr}(X_1 C^{i-1}(zX_2)C^{L-i})}{Z_{L,M}}. \tag{4.73}$$

Using the general expression for C^k, (4.49), we can express this as,

$$\langle \eta_i \rangle = Z_{L,M}^{-1}\oint \frac{\mathrm{d}z}{2\pi\mathrm{i}} z^{-M}\left[\alpha\sum_{l=0}^{i-1}\lambda_1^{L-1-l}\lambda_2^l + x\sum_{l=i}^{L}\lambda_1^{L-l}\lambda_2^{l-1}\right]. \tag{4.74}$$

To obtain an exact finite-size expression, we follow a similar approach to the normalisation. Substituting the expressions for λ_1 and λ_2, expanding and

evaluating the residues yields,

$$\langle \eta_i \rangle \;=\; Z_{L,M}^{-1} \left[\left(\sum_{l=0}^{i-1} + \frac{x}{\alpha} \sum_{l=i-1}^{L-1} \right) \sum_{m=0}^{l} \binom{L-1-l}{M-1-m} \binom{l}{m} \alpha^{L-l} x^m \right]. \quad (4.75)$$

To get the asymptotic behaviour, we approximate the sums in (4.74) as integrals with $w = l/L$ and $y = i/L$. Letting $\eta(y) = \langle \eta_{Ly} \rangle$, as usual, we obtain,

$$\eta(y) \approx Z_{L,M}^{-1} \oint \frac{\mathrm{d}z}{2\pi \mathrm{i}} \left(\frac{\alpha}{\lambda_1} L \int_0^y \mathrm{d}w + \frac{x}{\lambda_2} L \int_y^1 \mathrm{d}w \right) \exp[L\phi(z,w)], \quad (4.76)$$

where $\phi(z,w)$ is as defined in (4.55). Again, evaluating the z integrals first, we obtain

$$\eta(y) \approx Z_{L,M}^{-1} \left(L \int_0^y \mathrm{d}w \frac{1}{1+z_*(w)} + L \int_y^1 \mathrm{d}w \frac{x}{1+xz_*(w)} \right) A(w) \times [\mathrm{e}^{L\varphi(w)} + O(L^{-1})]. \quad (4.77)$$

We can evaluate this in each phase.

Right-localised phase

For $w_{dom} = 0$, we have,

$$\begin{aligned} \eta(y) &\approx \frac{z_*(0)}{1+z_*(0)}(1 - \mathrm{e}^{L\varphi'(0)y}) + \frac{xz_*(0)}{1+xz_*(0)}(\mathrm{e}^{L\varphi'(0)y} - 0) + O(L^{-1}) \\ &\approx \rho - \frac{\rho(1-\rho)(1-x)}{1-\rho+x\rho} \mathrm{e}^{-y/\xi_R}, \end{aligned} \quad (4.78)$$

where

$$\xi_R = -[L\varphi'(0)]^{-1} = -[L \log(\alpha(1-\rho+x\rho)^{-1})]^{-1}. \quad (4.79)$$

This is the right-localised phase, as the profile is mostly uniform with a small localised structure immediately to the right of the defect with a decay length ξ_R. It has the scaling $\xi_R \sim L^{-1}$. This profile is shown in figure 4.3.

Left-localised phase

For $w_{dom} = 1$, we have

$$\begin{aligned}\eta(y) &\approx \frac{z_*(1)}{1+z_*(1)}(\mathrm{e}^{-L\varphi'(1)(1-y)} - 0) + \frac{xz_*(1)}{1+xz_*(1)}(1 - \mathrm{e}^{-L\varphi'(1)(1-y)}) + O(L^{-1}) \\ &\approx \rho + \frac{\rho(1-\rho)(1-x)}{x(1-\rho+x^{-1}\rho)}\mathrm{e}^{-(1-y)/\xi_L}, \end{aligned} \tag{4.80}$$

where

$$\xi_L = [L\varphi'(1)]^{-1} = [L\log(\alpha(1-\rho+x^{-1}\rho))]^{-1}. \tag{4.81}$$

This is the left-localised phase, as the profile is mostly uniform with a small localised structure immediately to the left of the defect with a decay length ξ_L. It has the scaling $\xi_L \sim L^{-1}$. This profile is shown in figure 4.3.

Shock phase

For $w_{dom} = w_*$, we have

$$\begin{aligned}\eta(y) &\approx \left(\frac{z_*(w_*)}{1+z_*(w_*)}\int_0^y \mathrm{d}w + \frac{xz_*(w_*)}{1+xz_*(w_*)}\int_y^1 \mathrm{d}w\right)[\mathrm{e}^{\frac{1}{2}L\varphi''(w_*)(w-w_*)^2} + O(L^{-1})] \\ &\approx \frac{\rho_1+\rho_2}{2} + \frac{\rho_1-\rho_2}{2}\mathrm{erf}[(y-w_*)/\xi_S] + O(L^{-1}), \end{aligned} \tag{4.82}$$

where

$$\xi_S = [-L\varphi''(w_*)/2]^{-1/2}, \tag{4.83}$$

and $\varphi''(w_*)$ is given in (4.63). This is the shock phase, as we have a profile that interpolates between two densities, ρ_1 and ρ_2, with a front width given by ξ_S. It has the scaling $\xi_S \sim L^{-1/2}$. This profile is shown in figure 4.3.

Note that the two bulk densities ρ_1, ρ_2 do not depend on the global mean density, ρ. Instead, the global density constraint is fixed by the location of the shock front, $y = w_*$. Indeed, from (4.62) and (4.65), it is easily verified that w_*, ρ_1 and ρ_2 satisfy the consistency relation

$$(1-w_*)\rho_1 + w_*\rho_2 = \rho. \tag{4.84}$$

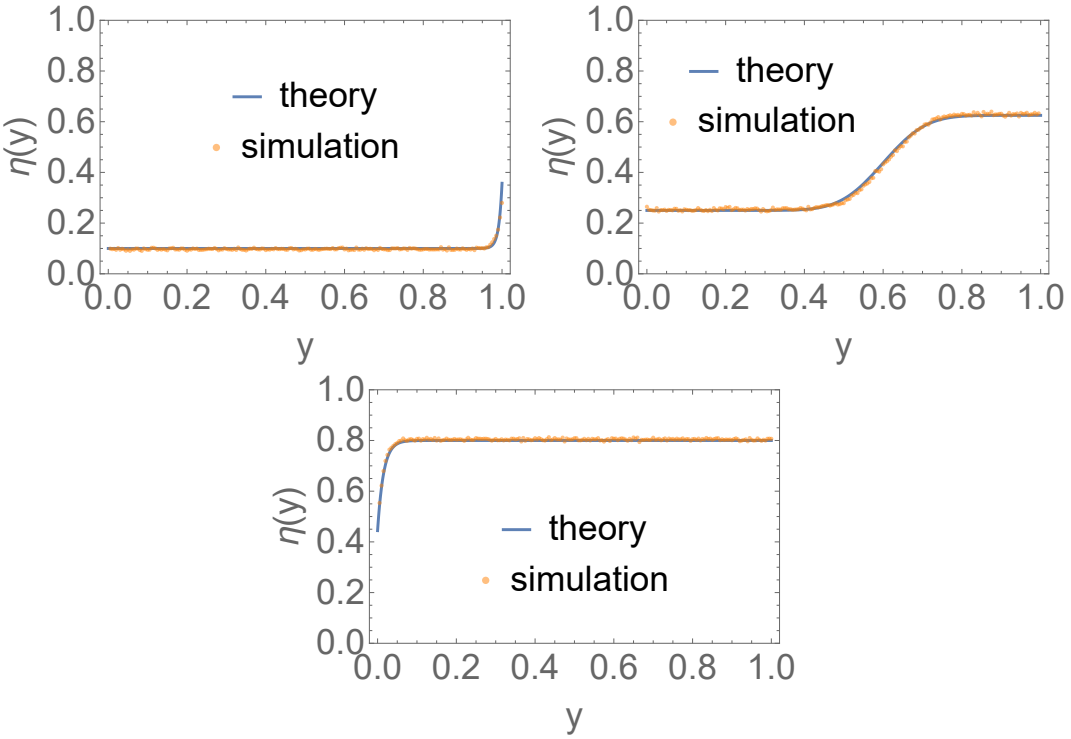

Figure 4.3 *Density profiles as given by* (4.78), (4.80), (4.82), *and as obtained from Monte Carlo simulations. Parameters used were* $L = 200$, $x = 0.2$, $\alpha = 0.5$. *The densities used were* $\rho = 0.1$ *(top left, left-localised phase),* $\rho = 0.4$ *(top right, shock phase), and* $\rho = 0.8$ *(bottom, right-localised phase). Excellent agreement is seen between theory and simulations.*

We can compare the exact results for the density profiles to the mean-field prediction. The mean-field theory correctly predicts the phase diagram and qualitatively captures the density profiles (localised and shock shapes). Also, in the localised phases, it predicts the correct scaling for the decay lengths, $\xi \sim L^{-1}$.

However, the specific expressions for the decay lengths in terms of the system parameters are incorrect. In particular, the exact expressions for the decay lengths in the localised phases diverge as one approaches the critical densities, whereas the lengths predicted by the mean-field theory remain finite. Moreover, in the shock phase, the mean-field theory gives an incorrect functional form (tanh, rather than erf) and gives the wrong front width, even at the level of the scaling ($\xi_S \sim L^{-1}$, rather than $\xi_S \sim L^{-1/2}$).

4.6.2 Currents

The fastest way to obtain the current is as follows. As the defect does not distinguish between particles and holes, it simply moves with a mean velocity v'. Then the rate at which it overtakes particles must equal the negative current of normal particles in the reference frame of the defect,

$$J' = -(\alpha\langle\eta_1\rangle - (x/\alpha)\langle\eta_L\rangle). \tag{4.85}$$

Indeed, this is precisely the boundary condition (4.11) which we derived in the mean-field context. As this expression only involves one-point functions, it makes sense that mean-field theory is exact here.

We remark that one might alternatively define the current as the net average hopping rates by normal particles across all bonds. This is given by the following expression,

$$J = \frac{1}{L}\sum_{i=1}^{L-1}[\langle\eta_i(1-\eta_{i+1})\rangle - x\langle(1-\eta_i)\eta_{i+1}\rangle]. \tag{4.86}$$

Actually, the two definitions (4.85) and (4.86) are related exactly by the Galilean boost (4.10), though this is not immediately obvious. We give a proof in appendix 4.B.

To obtain a the exact finite-size expression for the current, we simply substitute

(4.75) into (4.85), which gives,

$$J' = -Z_{L,M}^{-1}\left[\binom{L-1}{M-1}(\alpha^{L+1} - (x/\alpha)x^M) + (\alpha-1)xZ_{L-1,M-1}\right]. \quad (4.87)$$

Note that this is not simply a ratio of normalisations, as is typical in ASEPs. We can trace the cause of this back to the lack of scalar matrix reduction relations. Writing (4.85) using matrices gives

$$J' = -Z_{L,M}[\alpha\{z^M\}\mathrm{tr}(X_1(zX_2)C^{L-1}) - (x/\alpha)\{z^M\}\mathrm{tr}(zX_2X_1C^{L-1})]. \quad (4.88)$$

Now as the difference $\alpha X_1X_2 - (x/\alpha)X_2X_1$ is not given by a typical scalar reduction relation (see (4.41)), we are left with a term that does not have the same form as a normalisation $Z_{L,M}$.

To obtain asymptotic expressions for the current, we substitute the asymptotic expressions for the densities $\eta(0)$, $\eta(1)$ into (4.85). The expressions are particularly simple if we transform to the "lab" frame using the Galilean boost (4.10). Then we have,

$$J \approx \begin{cases} v\rho(1-\rho), & \text{localised} \\ \rho v' + \frac{x(1-\alpha)^2}{\alpha(1-x)}, & \text{shock} \end{cases}. \quad (4.89)$$

These expressions actually agree with the mean-field predictions, (4.18) and (4.30) if one sets $p \to \alpha$, $q \to x/\alpha$ in the latter. Current is plotted against density in figure 4.4.

The expression for the current in the localised phases is the same as for a pure (defectless) PASEP. Indeed, this is not surprising, as the effect of the defect on the density profile is localised, so one would not expect a global effect on the current.

In the shock phase, the defect throttles the current, restricting the current to below a certain value. This is clearly seen in 4.4. We remark that apart from the boost contribution $\rho v'$, the current has no density dependence in the shock phase.

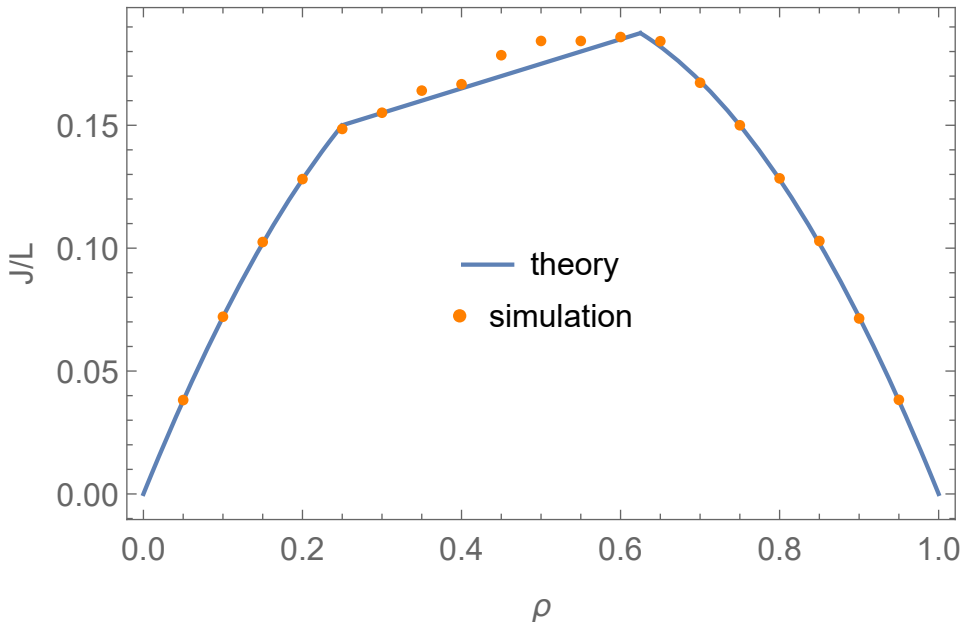

Figure 4.4 *Current plotted against density, as given by* (4.89) *and as obtained from Monte Carlo simulations. Parameters used were* $x = 0.2$, $\alpha = 0.5$. *Simulations were carried out with a system size* $L = 1000$. *Good agreement is seen between theory and simulations.*

4.7 Special cases

Having obtained the exact result, we can now see how we recover familiar special cases by taking specific limits.

4.7.1 Symmetric environment

The first limit we consider is to make the environment symmetric by taking $x = 1$. In this case, the normalisation (4.53) can be simplified. The m summation can be performed using the following binomial identity,

$$\sum_{m=0}^{l} \binom{L-l}{M-m}\binom{l}{m} = \binom{L}{M}. \tag{4.90}$$

Then the normalisation becomes,

$$Z_{L,M} = \binom{L}{M}\frac{\alpha^{L+1}-1}{\alpha-1}. \tag{4.91}$$

We can similarly simplify the density profile (4.75) and we obtain,

$$\langle \eta_i \rangle = M/L, \tag{4.92}$$

so we always have a uniform density profile. Hence the current vanishes and there are no phase transitions.

It is perhaps somewhat surprising that the density profile is uniform at an exact level. This is in stark contrast to the model we examined in chapter 3, where there were macroscopically non-uniform density profiles. There has been some recent interest in phase transitions in systems of symmetrically diffusing particles with a single driven particle and overtaking [146], so this result has important implications for this problem.

4.7.2 Totally asymmetric process

Another limit we can consider is the totally asymmetric case $x = 0$. In this case, both the normal particles and the defect can only hop to the right. The normalisation (4.53) simplifies to

$$Z_{L,M} = \sum_{l=0}^{L} \binom{L-l}{M} \alpha^{L-l}. \tag{4.93}$$

Similarly the density profile becomes,

$$\langle \eta_i \rangle = Z_{L,M}^{-1} \sum_{l=0}^{i-1} \binom{L-1-l}{M-1} \alpha^{L-l}. \tag{4.94}$$

From the asymptotics, we see that the right-localised phase vanishes. The phase transition between the left-localised phase and the shock phase occur on the line $\rho = 1 - \alpha$. The currents in the two phases are $J_{\mathcal{L}} \approx \rho(1-\rho)$ and $J_{\mathcal{S}} \approx \alpha\rho$.

This is a bit simpler to understand than the partially asymmetric case. The normal particles execute a TASEP, which has a current $\rho(1-\rho)$. If the defect moves slower than the mean drift velocity of a normal particle, $(1-\rho)$, then it creates a jam and all particles move at its speed, α. Note that $\rho_2 = 0$, which indicates that there is a large completely empty region in front of the defect in the shock phase.

This limit is very similar to the example we examined in section 2.3, though there are some slight differences in the exact expressions due to the fact that there, the defect was not allowed to overtake the normal particles.

4.7.3 "First-class" limit

The final limit we can consider is to make the defect a typical "first-class" particle. This is done by setting $\alpha = 1$. Then the defect moves at the same velocity as the normal particles. This is the inverse of the much-studied "second-class" particle problem [68].

The normalisation can again be simplified using a binomial identity, namely,

$$\sum_{l=m}^{L} \binom{L-l}{M-m}\binom{l}{m} = \binom{L+1}{M+1}. \tag{4.95}$$

The normalisation then becomes,

$$Z_{L,M} = \binom{L+1}{M+1}\frac{1-x^{L+1}}{1-x}. \tag{4.96}$$

Unfortunately, the density profile cannot be simplified with the binomial identity and we have

$$\langle \eta_i \rangle = Z_{L,M}^{-1}\left[\left(\sum_{l=0}^{i-1} + x\sum_{l=i-1}^{L-1}\right)\sum_{m=0}^{l}\binom{L-1-l}{M-1-m}\binom{l}{m}x^m\right]. \tag{4.97}$$

Similarly to the symmetric environment limit, the phase transitions disappear. In this case, the system is always in the right-localised phase.

It is interesting that we have produced here a finite-dimensional matrix product solution for this problem that does not satisfy matrix reduction relations, when there is also known to exist an infinite dimensional representation, which satisfies matrix reduction relations (see [68]).

4.8 Summary

In this chapter, we examined a partially asymmetric driven lattice gas with a defect particle that hops with different rates to the normal particles and has priority in the dynamics. We presented a general mean-field analysis, as well as an exact solution using a matrix product ansatz, which is valid provided condition (4.33) is satisfied.

We showed that the steady state density profile may take one of three forms depending on the system parameters: right-localised, left-localised and shock. We derived the phase diagram and computed the exact and asymptotic forms of the density profiles and currents in all phases.

It is interesting to compare this nonequilibrium phase transition to those encountered in equilibrium systems. One may view the right-localised phase as a high density phase and the left-localised phase as a low density phase, with the shock phase representing a region of phase coexistence between the two phases. This suggests an analogy with an equilibrium liquid-gas phase transition. The role of the free energy is here played by the current in the reference frame of the defect, J', which must be matched between the two phases. The current is continuous at the phase transition, with its first derivative being discontinuous, which is consistent with a first order transition. Interestingly, the lengthscales ξ_R and ξ_L as given by (4.79) and (4.81) respectively also diverge as the phase transition is approached.

Recall that in section 2.3.6, we showed that for systems solvable by a matrix product ansatz, fluxes may typically be expressed as ratios of normalisations. This is another point of similarity between currents in nonequilibrium systems and free energies in equilibrium systems. However, in the case we studied in this chapter, the matrices did not satisfy algebraic relations, so the current is not expressed as a simple ration of normalisations. It is interesting that it still seems to play a role similar to a free energy despite this.

4.A Interpretation as transfer matrices in an Ising spin chain

The matrix product solution (4.37) involves products of 2×2 matrices. This is reminiscent of the familiar transfer matrix solution for the one-dimensional Ising spin chain (see for instance [147]). As a reminder, this model is defined as a set of spins σ_i, each of which can be in one of two states, $\sigma_i = \pm 1$, with nearest neighbour interactions. For the purpose of the comparison here, we consider a finite chain. In the absence of an external field, the Hamiltonian is given by,

$$H(\{\sigma_i\}) = \sum_{i=1}^{L-1} U(\sigma_i, \sigma_{i+1}) + a(\sigma_1) + b(\sigma_L), \tag{4.98}$$

where $U(\sigma, \sigma')$ is the pairwise interaction and a, b represent the boundary conditions. Typically, one chooses $U(\sigma, \sigma') = -u\sigma\sigma'$ for some constant u but we will leave it general. The canonical partition function (see section 1.2) is given by,

$$Z = \sum_{\{\sigma_i\}} \exp\left(-\beta \sum_{i=1}^{L-1} U(\sigma_i, \sigma_{i+1}) - \beta a(\sigma_1) - \beta b(\sigma_L) \right). \tag{4.99}$$

The transfer matrix solution is obtained as follows. We can write all possible values of the weight $\mathrm{e}^{-\beta U(\sigma,\sigma')}$ in a matrix T (called the transfer matrix) in which columns correspond to values of σ and rows correspond to values of σ'. As σ, σ' can each take two values (± 1), we get the 2×2 matrix,

$$T = \begin{pmatrix} \mathrm{e}^{-\beta U(-1,-1)} & \mathrm{e}^{-\beta U(-1,+1)} \\ \mathrm{e}^{-\beta U(+1,-1)} & \mathrm{e}^{-\beta U(+1,+1)} \end{pmatrix}. \tag{4.100}$$

Then the partition function can be obtained simply by multiplying the transfer matrices with $\sigma = \sigma_i$, $\sigma' = \sigma_{i+1}$ for $i = 1, 2, \ldots, L-1$ and imposing the appropriate boundary conditions by using the vectors

$$\langle V| = \begin{pmatrix} \mathrm{e}^{-\beta a(-1)}, & \mathrm{e}^{-\beta a(+1)} \end{pmatrix}, \qquad |W\rangle = \begin{pmatrix} \mathrm{e}^{-\beta b(-1)} \\ \mathrm{e}^{-\beta b(+1)} \end{pmatrix}. \tag{4.101}$$

Then the partition function is given by,

$$Z = \langle V | T^L | W \rangle. \tag{4.102}$$

Now we note that in the matrix product solution (4.37), the matrix X_1 can be written as,

$$X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0, & 1 \end{pmatrix}. \tag{4.103}$$

For simplicity, we consider the grand-canonical version of the normalisation (*i.e.* leaving the parameter z unfixed). This is given by,

$$\mathcal{Z}_{L,M} = \begin{pmatrix} 0, & 1 \end{pmatrix} C^L \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{4.104}$$

where C is given by (4.47). Comparing directly with (4.100) and (4.101), we see that this matrix product solution is equivalent to a transfer matrix solution of an Ising spin chain with

$$U(-1,-1) = U(+1,-1) = \beta^{-1} \log \lambda_1 \tag{4.105a}$$
$$U(+1,+1) = \beta^{-1} \log \lambda_2 \tag{4.105b}$$
$$U(-1,+1) = \infty \tag{4.105c}$$
$$a(-1) = \infty \tag{4.105d}$$
$$a(+1) = b(-1) = b(+1) = 0. \tag{4.105e}$$

This is a somewhat unusual Hamiltonian as the pairwise interactions are anisotropic (*i.e.* $U(\sigma, \sigma') \neq U(\sigma', \sigma)$). The boundary conditions essentially mean that the left boundary forces $\sigma_1 = +1$, while the right boundary allows $\sigma_L = \pm 1$ with equal probability. In the bulk of the chain, starting from the left end, we expect a string of $+1$s. Then, at some point a flip occurs and some spin takes the value -1. Once this has occurred, the infinite $U(-1,+1)$ component means that the remaining spins will all be -1. Thus the whole configuration can essentially be described by a single variable, namely the location i at which the flip occurs. Then the energy of a configuration with the flip occurring at i is given by

$$H(i) = \beta^{-1}[i \log \lambda_2 + (L-i) \log \lambda_1]. \tag{4.106}$$

Thus the model parameters x, α and the "fugacity" z control the weights λ_1, λ_2 and determine where the flip is likely to occur.

The picture of the string of $+1$s followed by an irreversible flip to -1s is somewhat reminiscent of the shock density profile in the lattice gas interpretation, where

we have a region with bulk density ρ_2, followed by a region with bulk density ρ_1. It is interesting that similar phenomenology can be reproduced in both pictures, despite the fact that one is an equilibrium spin chain and the other a nonequilibrium lattice gas.

4.B Proof of equivalence of definitions of current

In this appendix, we prove that the definitions of the current (4.85) and (4.86) are equivalent up to the Galilean boost (4.10). We start with the following relation, which is reminiscent of a discrete continuity equations,

$$j_k + (x/\alpha)d_k = j_{k+1} + \alpha d_{k+1}, \tag{4.107}$$

where

$$\begin{aligned} j_k &= Z_{L,M}^{-1}[\{z^M\}\mathrm{tr}(X_1 C^{k-1} X_2 X_0 C^{L-k-1}) \\ &\quad -x\{z^M\}\mathrm{tr}(X_1 C^{k-1} X_0 X_2 C^{L-k-1})] \qquad (4.108) \\ d_k &= \langle \eta_k \rangle - \langle \eta_{k+1} \rangle. \qquad (4.109) \end{aligned}$$

The relations (4.107) hold for $k = 1, \ldots, L-2$. Close to the defect, we have,

$$j_{L-1} + (x/\alpha)d_{L-1} = \alpha(\langle \eta_L \rangle - \langle \eta_1 \rangle) \tag{4.110}$$

$$(x/\alpha)(\langle \eta_L \rangle - \langle \eta_1 \rangle) = j_1 + \alpha d_1. \tag{4.111}$$

To prove (4.107), it is sufficient to check that

$$\begin{aligned} X_2 X_0 C - x X_0 X_2 C + (x/\alpha) X_2 C C - \alpha C X_2 C = C X_2 X_0 - x C X_0 X_2 \\ +(x/\alpha) C X_2 C - \alpha C C X_2, \end{aligned} \tag{4.112}$$

which can be done explicitly by substituting X_0, X_2, C.

Now we multiply (4.107) by k and (4.110) by $L-1$ and add them together to obtain,

$$\sum_{k=1}^{L-1} k[j_k + (x/\alpha)d_k] = \sum_{k=2}^{L-1}(k-1)[j_k + \alpha d_k] + (L-1)\alpha(\langle \eta_L \rangle - \langle \eta_1 \rangle). \tag{4.113}$$

We can rearrange this to give,

$$\sum_{k=1}^{L-1} j_k = \sum_{k=1}^{L-1}[(k-1)\alpha - k(x/\alpha)]d_k + (L-1)\alpha(\langle\eta_L\rangle - \langle\eta_1\rangle). \tag{4.114}$$

The sum on the left hand side equals (4.86) up to a factor of L, so we just need to simplify the right hand side. In the sum on the right hand side, after substituting the definition of d_k (4.109), there will be some telescopic cancellation, after which we are left with,

$$\sum_{k=1}^{L-1} j_k = -(x/\alpha)\langle\eta_1\rangle + (\alpha - (x/\alpha))\sum_{k=2}^{L-1}\langle\eta_k\rangle - [(L-2)\alpha - (L-1)(x/\alpha)]\langle\eta_L\rangle \\ +(L-1)\alpha(\langle\eta_L\rangle - \langle\eta_1\rangle). \tag{4.115}$$

After some manipulation, we can rewrite it as,

$$\sum_{k=1}^{L-1} j_k = (\alpha - (x/\alpha))\sum_{k=1}^{L}\langle\eta_k\rangle + L((x/\alpha)\langle\eta_L\rangle - \alpha\langle\eta_1\rangle). \tag{4.116}$$

Now note that the sum on the right hand side is simply the total number of normal particles, M. Then dividing by L, this becomes the Galilean boost,

$$J = \rho v' + J', \tag{4.117}$$

with J given by (4.86) and J' given by (4.85), as required.

Chapter 5

Integrability of two-species driven lattice gases

5.1 Introduction

So far in this thesis, we have used the mapping to the zero-range process and the matrix product ansatz (see section 2.2 and 2.3 respectively) to obtain exact steady states for exclusion processes. In this chapter, we take a step back to consider more generally what makes certain problems solvable using a given method.

This question is particularly important for exact methods, such as the matrix product formalism and the Bethe ansatz. In contrast, with approximate methods, such as mean field theory, it is understood that the solution will contain some error. They are therefore often used without a rigorous proof of their correctness. This has the advantage that they can be used in almost any context, with the understanding that the solution will be an approximation. However, if a method is to produce an exact solution, one naturally requires strict validity. Thus, although exact methods are very powerful, they also tend to have a more limited range of applicability.

For the matrix product formalism, we have already seen in section 2.3.2 that requiring simple matrix reduction relations to exist may impose some restrictions on the model parameters. Also, in chapter 4, we saw that the presented matrix product solution was valid only provided the parameters satisfied a certain condition.

In this chapter, we turn to similar considerations for the Bethe ansatz (see section 2.4). The literature on this subject is extensive. Indeed, a whole field has developed dedicated to analysing which systems are solvable using the Bethe ansatz (are integrable) [148]. One of the most powerful tools in assessing whether a given model is integrable is the Yang-Baxter equation [148]. This is a set of consistency relations that must be satisfied in order for the Bethe ansatz to form a valid description of the dynamics of a given system. It has been shown that certain specific two-species asymmetric exclusion processes satisfy the Yang-Baxter equations [71, 118, 149] and a set of sufficient conditions on the model parameters has been proposed [150]. Yet, a clear definitive answer to which two-species models on a ring are integrable is lacking. In this chapter, we add to this effort by using the Yang-Baxter equations to systematically obtain a list of integrable models, which we conjecture to be exhaustive.

We comment that the Yang-Baxter equations are typically derived by using an algebraic Bethe ansatz. Here, we instead use a coordinate Bethe ansatz, as this is conceptually simpler and it is common in the literature concerning exclusion processes. Specifically, to study two-species exclusion processes, we will need to make use of a nested coordinate Bethe ansatz. The nested Bethe ansatz is a generalisation of the simple Bethe ansatz introduced in section 2.4. Its origins go back to the works of Baxter [115, 151] and Yang [152], though our treatment will follow mostly [71, 149].

We will only consider exclusion processes on a ring. Open boundary conditions introduce more parameters and therefore the question of integrability becomes more involved. We will not concern ourselves with this, as there is sufficient complexity to be found in processes on a ring for our purposes.

Actually an investigation of the range of three-state Hamiltonians solvable using a nested coordinate Bethe ansatz was previously carried out in [153, 154]. In particular, two-species exclusion processes can be mapped to a subclass of models considered in [154], so the results presented in this chapter have some overlap with that work. However, importantly, the ansatz used in those works was less general than the one used here, so some solvable models were not recovered. We will comment on this further once the ansatz is introduced.

Finally, we mention that recently an algebraic, differential approach was used to derive the Yang-Baxter equations and therefore classify all integrable $(n+1) \times (2n+1)$ vertex models for $n = 1, 2, \ldots$ [155]. In particular, 15-vertex models (*i.e.*

$n = 2$) can be mapped to two-species exclusion processes, so the results of [155] may have relevance in this context.

5.2 Bethe ansatz solution of two-species PASEP

5.2.1 Model definition

We consider the general two-species PASEP problem on a ring. This is defined by the following rates, where 0, 1, 2 denote holes, particles of type 1 and particles of type 2 respectively, as usual,

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{q}{\overset{p}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{\beta_{21}}{\overset{\beta_{12}}{\rightleftarrows}} 21 \, . \tag{5.1}$$

Note that we have set the rate of the process $10 \to 01$ to 1, which may always be done by a redefinition of units of time. The rates x, p, q, β_{12} and β_{21} are at this point left unconstrained. We may define this to be on an L-site ring with M_1 particles of species 1 and M_2 particles of species 2, with $M_1 + M_2 = M$, though this will not be important in this chapter. The main point of considering systems on a ring is that we will not have to deal with the integrability of the boundary conditions.

5.2.2 Eigenvalue problem

The time evolution of this problem is given by a master equation (see section 1.3.4),

$$\frac{\mathrm{d}P_t(\mathcal{C})}{\mathrm{d}t} = \sum_{\mathcal{C}'} \mathcal{M}(\mathcal{C}, \mathcal{C}') P_t(\mathcal{C}'), \tag{5.2}$$

where $\mathcal{C}$ denotes the configuration of the system, $P_t(\mathcal{C})$ denotes the probability of configuration $\mathcal{C}$ at time t and $\mathcal{M}$ is the transition rate matrix.

We denote configurations by $\mathcal{C} = \{Q_1, \ldots, Q_M; x_1, \ldots, x_M\}$, where Q_i is the species (1 or 2) of particle i and x_i is its position. We can write the eigenvalue-eigenvector problem in a form similar to the TASEP (see section 2.4.2), though we will have more terms to account for the partial asymmetry and two species of

particles. The bulk equation is given by,

$$\lambda\psi^{Q_1\dots Q_M}(x_1,\dots,x_M) = -[M_1(1+x)+M_2(p+q)]\psi^{Q_1\dots Q_M}(x_1,\dots,x_M) + \sum_{i=1}^{M}[p_{Q_i}\psi^{Q_1\dots Q_M}(\dots,x_i-1,\dots)+q_{Q_i}\psi^{Q_1\dots Q_M}(\dots,x_i+1,\dots)], \quad (5.3)$$

where

$$p_Q = \begin{cases} 1, & Q=1 \\ p, & Q=2 \end{cases}, \qquad q_Q = \begin{cases} x, & Q=1 \\ q, & Q=2 \end{cases}. \quad (5.4)$$

The collision equations read,

$$-(p_Q+q_Q)\psi^{\dots QQ\dots}(\dots,x_i,x_i+1,\dots)+p_Q\psi^{\dots QQ\dots}(\dots,x_i,x_i,\dots) + q_Q\psi^{\dots QQ\dots}(\dots,x_i+1,x_i+1,\dots)=0, \quad (5.5a)$$

for collisions involving particles of the same species and

$$-(p_Q+q_{Q'})\psi^{\dots QQ'\dots}(\dots,x_i,x_i+1,\dots)+p_{Q'}\psi^{\dots QQ'\dots}(\dots,x_i,x_i,\dots) + q_Q\psi^{\dots QQ'\dots}(\dots,x_i+1,x_i+1,\dots)=-\beta_{QQ'}\psi^{\dots QQ'\dots}(\dots,x_i,x_i+1,\dots) + \beta_{Q'Q}\psi^{\dots Q'Q\dots}(\dots,x_i,x_i+1,\dots), \quad (5.5b)$$

for collisions involving particles of different species.

Note that unlike the simple TASEP example we examined in section 2.4, these equations have a nested structure. The state space consists of blocks with a fixed species order $\{Q_1,\dots,Q_M\}$. The collision equations (5.5b) introduce mixing between blocks. The idea of the nested Bethe ansatz is then to write a (simple) Bethe ansatz solution for each sector with a fixed species order. One then constructs the total trial eigenstate by a linear combination of all permutations of species order, similarly to how the standard Bethe ansatz is constructed as a linear combination of permutations of plane waves. It is in this sense that the ansatz is "nested".

Finally, to complete the problem, we implement the periodic boundary conditions through the requirement,

$$\psi^{Q_1Q_2\dots}(1,x_2,\dots)=\psi^{Q_2\dots Q_1}(x_2,\dots,L+1), \quad (5.6)$$

although we will not use it in this chapter.

5.2.3 Deformed eigenvalue problem

For the aims of this thesis, the main purpose of the Bethe ansatz is to compute the large deviations of particle current. To this end, we reformulate the eigenvalue-eigenvector problem from the previous section for the generating function of the particle displacements, as was demonstrated in section 2.4.3. Although this is not obvious *a priori*, the integrability of the models we consider in this chapter is not affected by this deformation, so it makes sense to examine only the more general deformed case.

As we have two species of particles, we need to consider three random variables, $Y_1(t)$, $Y_2(t)$ and $Y_{12}(t)$. The $Y_i(t)$ for $i = 1, 2$ are defined as the number of processes $i0 \to 0i$ minus the number of processes $0i \to i0$ that have occurred up to time t. The $Y_{12}(t)$ is correspondingly defined as the number of processes $12 \to 21$ minus the number of processes $21 \to 12$ that have occurred up to time t. Then we define the joint distribution of the three variables with some configuration $\mathcal{C}$ as $P_t(\mathcal{C}; Y_1, Y_2, Y_{12})$. From (5.2), its evolution equation is given by,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}P_t(\mathcal{C}; Y_1, Y_2, Y_{12}) &= \sum_{\mathcal{C}'}\mathcal{M}_0(\mathcal{C},\mathcal{C}')P_t(\mathcal{C}'; Y_1, Y_2, Y_{12}) \\ &\quad + \sum_{\mathcal{C}'}\sum_{i=1,2,12}[\mathcal{M}_{i,+}P_t(\mathcal{C}'; \dots Y_i - 1 \dots) \\ &\qquad + \mathcal{M}_{i,-}P_t(\mathcal{C}'; \dots Y_i + 1 \dots)], \end{aligned} \tag{5.7}$$

where we have split the transition rate matrix $\mathcal{M}$ such that $\mathcal{M}_0$ contains the diagonal terms and $\mathcal{M}_{i,+(-)}$ contain the off-diagonal terms corresponding to the rates of processes which increase (decrease) $Y_i(t)$ by 1.

We now define the joint moment generating function for the three variables,

$$F_t(\mathcal{C}; \gamma_1, \gamma_2, \gamma_{12}) = \sum_{Y_1=-\infty}^{\infty}\sum_{Y_2=-\infty}^{\infty}\sum_{Y_{12}=-\infty}^{\infty} \mathrm{e}^{\gamma_1 Y_1 + \gamma_2 Y_2 + \gamma_{12} Y_{12}} P_t(\mathcal{C}; Y_1, Y_2, Y_{12}). \tag{5.8}$$

From (5.7), its evolution is given by,

$$\frac{\mathrm{d}}{\mathrm{d}t}F_t(\mathcal{C}; \gamma_1, \gamma_2, \gamma_{12}) = \sum_{\mathcal{C}'}\mathcal{M}_\gamma(\mathcal{C},\mathcal{C}')F_t(\mathcal{C}'; \gamma_1, \gamma_2, \gamma_{12}), \tag{5.9}$$

where the deformed transition rate matrix is given by

$$\mathcal{M}_\gamma = \mathcal{M}_0 + \sum_{i=1,2,12} [\mathrm{e}^{\gamma_i}\mathcal{M}_{i,+} + \mathrm{e}^{-\gamma_i}\mathcal{M}_{i,-}]. \tag{5.10}$$

Whence we obtain the eigenvalue-eigenvector problem for the deformed operator $\mathcal{M}_\gamma$,

$$\begin{aligned}\lambda\psi^{Q_1\dots Q_M}(x_1,\dots,x_M) &= -[M_1(1+x)+M_2(p+q)]\psi^{Q_1\dots Q_M}(x_1,\dots,x_M)\\ &+\sum_{i=1}^{M}[\mathrm{e}^{\gamma_{Q_i}}p_{Q_i}\psi^{Q_1\dots Q_M}(\dots,x_i-1,\dots)+\mathrm{e}^{-\gamma_{Q_i}}q_{Q_i}\psi^{Q_1\dots Q_M}(\dots,x_i+1,\dots)],\end{aligned} \tag{5.11}$$

Analogously, the collision equations read,

$$\begin{aligned}-(p_Q+q_Q)\psi^{\dots QQ\dots}(\dots,x_i,x_i+1,\dots)+\mathrm{e}^{\gamma_Q}p_Q\psi^{\dots QQ\dots}(\dots,x_i,x_i,\dots)\\ +\mathrm{e}^{-\gamma_Q}q_Q\psi^{\dots QQ\dots}(\dots,x_i+1,x_i+1,\dots)=0,\end{aligned} \tag{5.12a}$$

for collisions involving particles of the same species and

$$\begin{aligned}-(p_Q+q_{Q'})\psi^{\dots QQ'\dots}(\dots,x_i,x_i+1,\dots)+\mathrm{e}^{\gamma_{Q'}}p_{Q'}\psi^{\dots QQ'\dots}(\dots,x_i,x_i,\dots)\\ +\mathrm{e}^{-\gamma_Q}q_Q\psi^{\dots QQ'\dots}(\dots,x_i+1,x_i+1,\dots)=-\beta_{QQ'}\psi^{\dots QQ'\dots}(\dots,x_i,x_i+1,\dots)\\ +\mathrm{e}^{\gamma_{Q'Q}}\beta_{Q'Q}\psi^{\dots Q'Q\dots}(\dots,x_i,x_i+1,\dots),\end{aligned} \tag{5.12b}$$

for collisions involving particles of different species, where γ_{21} is to be interpreted as $-\gamma_{12}$.

The boundary condition (5.6) remains unchanged.

5.2.4 Bethe ansatz solution

We now present a (nested) coordinate Bethe ansatz solution to the deformed eigenvalue-eigenvector problem defined by the bulk equation (5.11) together with the collision equations (5.12a)–(5.12b) and the boundary condition (5.6). A general form for the coordinate Bethe ansatz is known provided the parameters satisfy the condition

$$pq = x. \tag{5.13}$$

This is the same condition as the one we required for the existence of a matrix product solution in chapter 4. It is interesting that it appears seemingly independently in this context. Although the precise nature of the link between the existence of a matrix product solution and a Bethe ansatz solution is not known, it has been shown in certain cases that the Bethe ansatz may be formulated as a matrix product state [156, 157]. It would thus be interesting to investigate whether a similar construction could be carried out in the present case, which could explain the simultaneous appearance of condition (5.13) in both contexts.

The condition (5.13) means that we are not able to analyse models of the type studied by Arndt *et al* (see *e.g.* [141, 142]). However, these are thought to be non-integrable [158], so we do not consider this a major issue.

As in chapter 4, we introduce the parameter α such that

$$p = \alpha, \qquad q = x/\alpha. \tag{5.14}$$

Then the problem is solved by the following ansatz,

$$\psi^{Q_1 \dots Q_M}(x_1, \dots, x_M) = \left(\prod_{i=1}^{M} (\mathrm{e}^{\gamma_{Q_i}} p_{Q_i})^{x_i} \right) \sum_{\sigma} A_{\sigma}^{Q_1 \dots Q_M} \prod_{i=1}^{M} z_{\sigma(i)}^{x_i}, \tag{5.15}$$

where σ indicates the permutations of the indices $\{1, \dots, M\}$, A are (undetermined) amplitudes and z are (undetermined) Bethe roots. If condition (5.13) is not satisfied, one may still write down a trial function of the form (5.15). However, it can then be verified that this function does not generally diagonalise the Markov operator.

Plugging this ansatz into the bulk equation (5.11), we can verify that the problem is diagonalised and we obtain the following expression for the eigenvalue,

$$\lambda(\gamma_1, \gamma_2, \gamma_3) = -[M_1(1 + x) + M_2(\alpha + x/\alpha)] + \sum_{i=1}^{M} (z_i^{-1} + x z_i). \tag{5.16}$$

Plugging the ansatz (5.15) into the collision equations (5.12a)–(5.12b), we obtain

the following relations between amplitudes,

$$[w_{ij} - (1+x)z_j]A^{\cdots 11\cdots}_{\cdots ij\cdots} + [w_{ij} - (1+x)z_i]A^{\cdots 11\cdots}_{\cdots ji\cdots} = 0 \tag{5.17a}$$

$$[w_{ij} - (\alpha + x/\alpha)z_j]\, A^{\cdots 22\cdots}_{\cdots ij\cdots} + [w_{ij} - (\alpha + x/\alpha)z_i]A^{\cdots 22\cdots}_{\cdots ij\cdots} = 0 \tag{5.17b}$$

$$\begin{aligned}&\big[(w_{ij} - k_1 z_j)A^{\cdots 12\cdots}_{\cdots ij\cdots} - \mathrm{e}^{\bar\gamma}\alpha^{-1}\beta_{21}z_j A^{\cdots 21\cdots}_{\cdots ij\cdots}\big] \\ &\quad + [(w_{ij} - k_1 z_i)A^{\cdots 12\cdots}_{\cdots ji\cdots} - \mathrm{e}^{\bar\gamma}\alpha^{-1}\beta_{21}z_i A^{\cdots 21\cdots}_{\cdots ji\cdots}] = 0\end{aligned} \tag{5.17c}$$

$$\begin{aligned}&\big[(w_{ij} - k_2 z_j)A^{\cdots 21\cdots}_{\cdots ij\cdots} - \mathrm{e}^{-\bar\gamma}\alpha\beta_{12}z_j A^{\cdots 12\cdots}_{\cdots ij\cdots}\big] \\ &\quad + [(w_{ij} - k_2 z_i)A^{\cdots 21\cdots}_{\cdots ji\cdots} - \mathrm{e}^{-\bar\gamma}\alpha\beta_{12}z_i A^{\cdots 12\cdots}_{\cdots ji\cdots}] = 0,\end{aligned} \tag{5.17d}$$

where we have introduced the notation,

$$w_{ij} = 1 + x z_i z_j \tag{5.18}$$

$$k_1 = 1 + x/\alpha - \beta_{12}, \qquad k_2 = \alpha + x - \beta_{21} \tag{5.19}$$

$$\bar\gamma = \gamma_1 - \gamma_2 - \gamma_{12}. \tag{5.20}$$

The relations (5.17a)–(5.17d) can be rewritten compactly by introducing the amplitude vector notation,

$$\vec{A}_{\cdots ij\cdots} = \left(A^{\cdots 11\cdots}_{\cdots ij\cdots},\ \ A^{\cdots 12\cdots}_{\cdots ij\cdots},\ \ A^{\cdots 21\cdots}_{\cdots ij\cdots},\ \ A^{\cdots 22\cdots}_{\cdots ij\cdots}\right)^{\mathrm{T}}. \tag{5.21}$$

In this notation, we can write,

$$[w_{ij}1_{4\times 4} - z_i H]\vec{A}_{\cdots ij\cdots} + [w_{ij}1_{4\times 4} - z_j H]\vec{A}_{\cdots ji\cdots} = 0, \tag{5.22}$$

where

$$H = \begin{pmatrix} 1+x & 0 & 0 & 0 \\ 0 & k_1 & \mathrm{e}^{\bar\gamma}\alpha^{-1}\beta_{21} & 0 \\ 0 & \mathrm{e}^{-\bar\gamma}\alpha\beta_{12} & k_2 & 0 \\ 0 & 0 & 0 & \alpha + x/\alpha \end{pmatrix}. \tag{5.23}$$

By inverting the matrix in front of one of the amplitude vectors, we can write (5.22) as,

$$\vec{A}_{\cdots ji\cdots} = -[w_{ij}1_{4\times 4} - z_j H]^{-1}[w_{ij}1_{4\times 4} - z_i H]\vec{A}_{\cdots ij\cdots}, \tag{5.24}$$

where $[\bullet]^{-1}$ denotes the matrix inverse. Note that (5.24) has a similar form to the simple TASEP case we examined in section 2.4. The difference here is that the amplitudes have been lifted from scalars to vectors.

From (5.22), we can obtain explicit relations between the components $A^{\cdots Q_1 Q_2 \cdots}_{\cdots ij \cdots}$ and $A^{\cdots Q_1 Q_2 \cdots}_{\cdots ji \cdots}$. We define the 4×4 matrix

$$\mathbf{S}(z_i, z_j) = -[w_{ij} 1_{4\times 4} - z_j H]^{-1} [w_{ij} 1_{4\times 4} - z_i H]. \tag{5.25}$$

Writing out (5.24) in component form, we obtain relations of the form

$$A^{\cdots Q_1 Q_2 \cdots}_{\cdots ji \cdots} = \sum_{Q_1', Q_2' = 1,2} S^{Q_1 Q_2}_{Q_1' Q_2'}(z_i, z_j) A^{\cdots Q_1' Q_2' \cdots}_{\cdots ij \cdots}, \tag{5.26}$$

where the entries $S^{Q_1 Q_2}_{Q_1' Q_2'}(z_i, z_j)$ of $\mathbf{S}(z_i, z_j)$ are called form factors. Although there are in principle 16 form factors (corresponding to $Q_1, Q_2, Q_1', Q_2' = 1, 2$), many of them vanish due to species conservation. The explicit expressions for the non-zero form factors are

$$S^{11}_{11}(z_i, z_j) = -\frac{w_{ij} - (1+x) z_j}{w_{ij} - (1+x) z_i} \tag{5.27a}$$

$$S^{22}_{22}(z_i, z_j) = -\frac{w_{ij} - (\alpha + x/\alpha) z_j}{w_{ij} - (\alpha + x/\alpha) z_i} \tag{5.27b}$$

$$S^{12}_{12}(z_i, z_j) = -\mathcal{D}_{ij}^{-1} [(w_{ij} - k_2 z_i)(w_{ij} - k_1 z_j) - \beta_{12}\beta_{21} z_i z_j] \tag{5.27c}$$

$$S^{21}_{21}(z_i, z_j) = -\mathcal{D}_{ij}^{-1} [(w_{ij} - k_1 z_i)(w_{ij} - k_2 z_j) - \beta_{12}\beta_{21} z_i z_j] \tag{5.27d}$$

$$S^{12}_{21}(z_i, z_j) = -\mathcal{D}_{ij}^{-1} w_{ij} \mathrm{e}^{\bar{\gamma}} \alpha^{-1} \beta_{21} (z_i - z_j) \tag{5.27e}$$

$$S^{21}_{12}(z_i, z_j) = -\mathcal{D}_{ij}^{-1} w_{ij} \mathrm{e}^{-\bar{\gamma}} \alpha \beta_{12} (z_i - z_j), \tag{5.27f}$$

where

$$\mathcal{D}_{ij} = (w_{ij} - k_1 z_i)(w_{ij} - k_2 z_i) - \beta_{12}\beta_{21} z_i^2. \tag{5.28}$$

Note that $\mathcal{D}_{ij} \neq \mathcal{D}_{ji}$. For clarity, we write out the matrix $\mathbf{S}(z_i, z_j)$ explicitly,

$$\mathbf{S}(z_i, z_j) = \begin{pmatrix} S^{11}_{11}(z_i, z_j) & 0 & 0 & 0 \\ 0 & S^{12}_{12}(z_i, z_j) & S^{12}_{21}(z_i, z_j) & 0 \\ 0 & S^{21}_{12}(z_i, z_j) & S^{21}_{21}(z_i, z_j) & 0 \\ 0 & 0 & 0 & S^{22}_{22}(z_i, z_j) \end{pmatrix}. \tag{5.29}$$

5.3 Yang-Baxter equations

5.3.1 Derivation of Yang-Baxter equations

It is generally not easy to know whether a given model is integrable or not. One approach to answering this question that has become popular follows the formalism of Yang and Baxter [148]. The main idea is that if a model is integrable (*i.e.* the Bethe ansatz description is valid), then the form factors must satisfy a certain set of nontrivial consistency relations called the Yang-Baxter equations.

We remark that this is a necessary but not sufficient condition. It might be possible, for instance, that the Bethe ansatz is valid but does not generate all eigenvectors of the linear operator of interest. It has been proven rigorously only in certain cases that vectors of the Bethe ansatz type are general enough to generate all eigenvectors [159–161] or how they may be extended to generate all eigenvectors [162, 163]. We will not concern ourselves with this difficulty, as our motivations for using the Bethe ansatz are purely practical and we only need one eigenvector to achieve our goal. If our calculations match numerical simulation, we will assume that the solution is valid.

The Yang-Baxter equations are derived as follows. The relations (5.22) relate the amplitude vectors $\vec{A}_{\ldots ij\ldots}$ and $\vec{A}_{\ldots ji\ldots}$, which has the effect of permuting any two adjacent indices. In order for this operation to be consistent, any two ways of permuting a sequence of indices that produces the same final sequence must give the same relation between the corresponding amplitudes. The simplest nontrivial example of this statement is seen by considering a string of three indices, $\{ijk\}$. To represent permutations of two indices with the third index not involved, we introduce the following notation,

$$\vec{A}_{\ldots ikj\ldots} = (1_{2\times 2} \otimes \mathbf{S}(z_j, z_k))\vec{A}_{\ldots ijk\ldots}. \tag{5.30}$$

Now consider the following case. We can transform $\{ijk\}$ into $\{kji\}$ through two different sequences of nearest-neighbour index swaps: *(i)* $\{ijk\} \to \{jik\} \to \{jki\} \to \{kji\}$, and *(ii)* $\{ijk\} \to \{ikj\} \to \{kij\} \to \{kji\}$. Equating the results obtained from these two sequences gives the following equation,

$$\begin{aligned}(\mathbf{S}(z_j, z_k) \otimes 1_{2\times 2})(1_{2\times 2} \otimes \mathbf{S}(z_i, z_k))(\mathbf{S}(z_i, z_j) \otimes 1_{2\times 2}) = \\ (1_{2\times 2} \otimes \mathbf{S}(z_i, z_j))(\mathbf{S}(z_i, z_k) \otimes 1_{2\times 2})(1_{2\times 2} \otimes \mathbf{S}(z_j, z_k)).\end{aligned} \tag{5.31}$$

This is the famed Yang-Baxter equation. For the models we are considering here, it is an 8×8 matrix equation, though, importantly, it consists of 4 disjoint blocks due to species conservation. The blocks correspond to situations with different numbers of particles of each species among the three particles involved in the collisions. We therefore have two 1×1 blocks, corresponding to 111 and 222, and two 3×3 blocks corresponding to 112 and 122.

The latter decomposition into blocks is particularly important for stochastic particle models with species conservation. We will see that we can find a combination of parameters for which the Yang-Baxter equations hold in the 111, 222 and 112 blocks but not the 122 block. This means that the Bethe ansatz can be used to solve a system with exactly one particle of species 2 (and an arbitrary number of particles of species 1) but not systems with an arbitrary number of particles of species 2. These kinds of partial "solutions" of the Yang-Baxter equations have not received much attention in the literature, as one typical seeks solutions that satisfy all blocks. In the context of spin chains, such solutions are arguably less physically significant, as the blocks correspond to different total magnetisation.

5.3.2 Analysis of Yang-Baxter equations

By plugging in the explicit forms of the form factors (5.27a)–(5.27f) into (5.31), we can investigate which restrictions on the model parameters x, α, β_{12}, β_{21} ensure that the Yang-Baxter equations are satisfied. The 111 and 222 blocks are actually satisfied automatically, so for the remainder of this section we concern ourselves only with the 112 and 122 blocks. These correspond to the matrix entries (r, s) with $r, s \in \{2, 3, 5\}$ and $\{4, 6, 7\}$ respectively.

Manipulating the expressions explicitly becomes algebraically very cumbersome. It is best performed with the aid of a symbolic programming language such as `Mathematica`. Even so, the calculations can be hard to manage. We propose the following procedure. First, we bring both matrices in (5.31) to one side of the equation and combine the terms in each entry into single fractions (this can be done using the `Together` function in `Mathematica`). Now note that the form factors are rational functions of z_i, z_j, z_k. Then to ensure identical equality, the coefficients of all terms of the form $z_i^{\nu_i} z_j^{\nu_j} z_k^{\nu_k}$ for ν_i, ν_j, $\nu_k \in \mathbb{Z}$ in the numerators of all entries must vanish. Examining each coefficient of each entry is evidently very laborious. Fortunately, it turns out that via judicious choices of coefficients,

we can perform only a few checks and obtain strict enough restrictions to satisfy the Yang-Baxter equations fully (at least in certain blocks). This can be done in many ways. We outline one approach now.

We begin with the observation that the numerators of all entries are multiples of $(z_i - z_j)(z_j - z_k)(z_k - z_i)$. Therefore, we always divide by this first for simplicity. This will be assumed to have been done in all following expressions. Setting $z_i = z_j = z_k = 0$ (*i.e.* looking at the constant term) in the $(3, 2)$ and $(6, 7)$ entries gives respectively,

$$\alpha^4\beta_{12}(\beta_{12}\beta_{21} - x), \qquad \mathrm{e}^{\bar{\gamma}}\alpha^4\beta_{21}(\beta_{12}\beta_{21} - x). \tag{5.32}$$

We will be ignoring options that require $\alpha = 0$ as this makes our ansatz (5.15) singular. Therefore we have two options: *(a)* set $\beta_{12}\beta_{21} = x$, or *(b)* set $\beta_{12} = \beta_{21} = 0$.

Case *(a)*: $\beta_{12}\beta_{21} = x$

We may define a parameter β such that,

$$\beta_{12} = \beta, \qquad \beta_{21} = x/\beta. \tag{5.33}$$

Then setting $z_i = z_j = z_k = 0$ in the $(2, 3)$ entry we have

$$\mathrm{e}^{\bar{\gamma}}\alpha x\beta^{-2}(\alpha - 1)(x + \beta)(x - \alpha\beta). \tag{5.34}$$

We then have three options: *(i)* set $x = 0$, *(ii)* set $\alpha = 1$, *(iii)* set $\beta = x/\alpha$.

(i) Setting $x = 0$ (note that this also implies β_{21}=0) is sufficient to satisfy all Yang-Baxter equations. We call this solution 1.

(ii) We set $\alpha = 1$. Then we may consider the coefficient of $z_i^0 z_j z_k$ in the $(2, 2)$ entry. This is given by,

$$x(1 + x)(1 - \beta)(x - \beta). \tag{5.35}$$

Setting $x = 0$ reverts to case *(i)*, which gives solution 1. The other options we have is to set $\beta = 1$ or $\beta = x$. Actually these two are equivalent up to relabelling of species. Doing this is sufficient to satisfy all Yang-Baxter equations. We call this solution 2.

(iii) We set $\beta = x/\alpha$. This is sufficient to satisfy the Yang-Baxter equations in the 112 block. We call this solution 3.

We may inquire what further restrictions are needed to satisfy the Yang-Baxter equations in the 122 block as well. To this end, we may set $z_j = z_k = 0$, $z_i = 1$ in the numerator of the $(6, 7)$ entry, which becomes,

$$\mathrm{e}^{\bar{\gamma}}\alpha^3 x(\alpha^2 + x)(\alpha - 1)(\alpha - x). \tag{5.36}$$

Choosing $x = 0$ gives solution 1, while $\alpha = 1$ and $\alpha = x$ give solution 2. So setting $\beta = x/\alpha$ does not give any additional solutions unless we restrict ourselves to block 112.

Case *(b)*: $\beta_{12} = \beta_{21} = 0$

Setting $z_i = z_j = z_k = 0$ in the $(2, 2)$ entry, we obtain

$$\alpha^2 x^2(\alpha - 1). \tag{5.37}$$

Setting $\alpha = 1$ gives a one-species PASEP. Setting $x = 0$ gives a special case of solution 1 with $\beta = 0$. Therefore the case $\beta_{12} = \beta_{21} = 0$ does not produce any additional solutions.

We summarise the procedure for obtaining solutions which we have outlined in figure 5.1.

5.3.3 Solutions of Yang-Baxter equations

In the previous section, we obtained three distinct classes of solutions. For the sake of conciseness, we do not count separately one-species solutions and solutions that are equivalent to the presented ones up to a relabelling of species.

We remark that all solutions can be derived from case *(a)*, which involves the

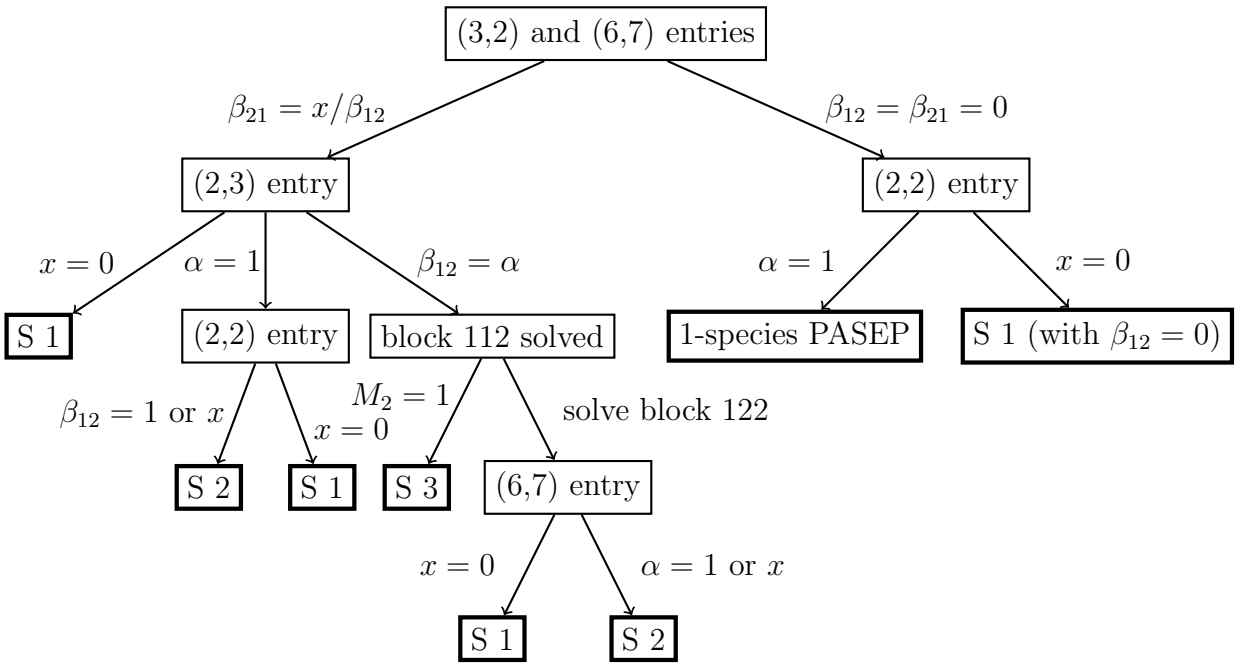

Figure 5.1 *Representation of the procedure to obtain solutions to the Yang-Baxter equations. Solutions are outlined with thick boundaries. S stands for solution. See section 5.3.2 for coefficients to examine in the listed entries.*

restriction $\beta_{12}\beta_{21} = x$. This is equivalent to the condition (5.13) under a relabelling of species. It is interesting that the former condition was derived as a requirement from the Yang-Baxter equations, rather than from the form of the coordinate Bethe ansatz. This supports the claim that these conditions are necessary for integrability.

Solution 1: Two-species TASEP

The first solution involves the restrictions

$$x = \beta_{21} = 0. \tag{5.38}$$

This corresponds to the process

$$10 \xrightarrow{1} 01 \quad ; \quad 20 \xrightarrow{\alpha} 02 \quad ; \quad 12 \xrightarrow{\beta_{12}} 21 \, . \tag{5.39}$$

This is a TASEP with two species of particles that move at different rates 1, α, with one species overtaking the other with a rate β_{12}. This model has been studied much in the literature. In particular, its Yang-Baxter integrability was

proved using an algebraic Bethe ansatz approach [118, 150].

Solution 2: PASEP with first- and second-class particles

The second solution involves the restrictions

$$\alpha = \beta_{12} = 1, \qquad \beta_{21} = x. \tag{5.40}$$

This corresponds to the process

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{x}{\overset{1}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{x}{\overset{1}{\rightleftarrows}} 21 \, . \tag{5.41}$$

This is a PASEP with two species of particles that move with the same rates. Species 1 particles may also overtake species 2 particles with their usual hopping rates, essentially treating them as if they were holes. This class of models has also been studied much in the literature, where the species 2 particles are usually referred to as "second-class", due to their reduced priority in the dynamics compared to species 1 particles. In fact, in [71] the integrability of a somewhat broader class of models was shown, namely the extension of this model to an arbitrary number of species and with particles of arbitrary size.

Solution 3: PASEP with a single first-class defect particle

The third solution involves the restrictions

$$\beta_{12} = x/\alpha, \qquad \beta_{21} = \alpha. \tag{5.42}$$

This corresponds to the process

$$10 \underset{x}{\overset{1}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{x/\alpha}{\overset{\alpha}{\rightleftarrows}} 02 \quad ; \quad 21 \underset{x/\alpha}{\overset{\alpha}{\rightleftarrows}} 12 \, . \tag{5.43}$$

Furthermore, the Yang-Baxter equations are only solved in the 112 block, which we may express as the restriction

$$M_2 = 1. \tag{5.44}$$

This is the model that was examined in chapter 4. It is a PASEP with a single defect particle that hops with different rates to the normal particles. The defect

Model	Parameter restrictions	Free parameters
TASEP with defect species	$x = q = \beta_{12} = 0$	$\beta_{12}\,, M_1,\ M_2$
Second-class PASEP	$p = \beta_{12} = 1,\ q = \beta_{21} = x$	$x,\ M_1,\ M_2$
PASEP with one defect	$p = \beta_{21} = \alpha,\ q = \beta_{12} = x/\alpha,\ M_1 = 1$	$x,\ \alpha,\ M_1$

Table 5.1 *All Bethe ansatz integrable sub-classes of the two-species partially asymmetric exclusion process with overtaking on a ring. The particle numbers M_1, M_2 are included as free parameters to highlight that M_2 is fixed in the third case.*

particle may overtake the normal particles with its usual hopping rates, meaning that it has priority in the dynamics. It is therefore arguably more natural to label it as a 1 and the normal particles as 2. This was the notation used in chapter 4 and we will revert to it in chapter 6. To our knowledge, this model has not received much attention in the literature. It is the only known integrable partially asymmetric model with particles hopping at different rates. The restriction on particle number (5.44) is also peculiar, probably due to the fact that it is uncommon to look at such "partial" solutions to the Yang-Baxter equations (*i.e.* models that satisfy certain blocks of the Yang-Baxter equations but not others).

5.4 Summary

We used a nested coordinate Bethe ansatz to obtain solutions for the eigenvalue problems corresponding to the conditioned time-evolution of a broad class of two-species exclusion processes on a ring. By verifying for which cases the Yang-Baxter equations hold, we obtained a list of integrable two-species exclusion processes on a ring. This is summarised in table 5.1.

Out of the three classes of models found, the third has not yet been studied in the context of the Bethe ansatz. This is actually the model whose steady state was already solved in chapter 4 using the matrix product approach. In the next chapter, we continue our analysis of this model, and in particular its current fluctuations in the long time limit, using the Bethe ansatz solution presented in this chapter.

Chapter 6

Current fluctuations of a driven lattice gas with a driven defect particle with priority

6.1 Introduction

In the previous chapter, we proposed a classification of integrable two-species partially asymmetric exclusion processes. Of these, two have been solved using the Bethe ansatz, whereas the third one has not. In fact, this is the same model whose steady state we solved using a matrix product approach in chapter 4. In this chapter, we derive the Bethe equations for this model and use them to calculate the current statistics up to second order in a perturbative expansion. This allows us to reproduce the mean current, which was already calculated in chapter 4, and calculate the diffusion constant. The work presented in this chapter has been published in [3] and [4].

The Bethe equations we will derive and then analyse in this chapter have a more complicated structure than those for the simple TASEP case, which we examined in section 2.4.2. This will require us to use a more refined mathematical approach. Namely, we reformulate the usual (algebraic) Bethe equations for the Bethe roots as a functional equation for a single function that contains the same information as the Bethe roots. Physical observables may then be expressed in closed form in terms of this function, which is to be found by solving the

(functional) Bethe equations perturbatively. A similar functional reformulation of the Bethe equations has been used for the pure PASEP [164].

Functional methods in the context of the Bethe ansatz actually go back to Baxter's original solution of the eight vertex model [151]. This approach, known as the functional Bethe ansatz, was later developed further and applied to a range of integrable problems, including spin chains [165–167] and also ASEPs [168]. However, we do not use the full technical machinery of the functional Bethe ansatz, as our problem is readily diagonalised using a coordinate Bethe ansatz. Rather, we simply convert the Bethe equations into a functional form to make them easier to solve.

In the literature, calculations of current fluctuations have been performed for the pure TASEP [10], the TASEP with defect particles [12, 118] and the pure PASEP [119, 164, 169]. The Bethe equations for the PASEP with a second-class species are known [170, 171], although we remark that the current fluctuations have not yet been calculated for that case. Thus, in accordance with the list of integrable systems we presented in chapter 5, the calculations performed in this chapter represent the last integrable case among two-species processes on a ring.

6.1.1 Schematic overview of calculations

As the calculations presented in this chapter are quite involved and somewhat formal, we provide a schematic overview for reference. This should allow the reader to keep track of the main goals at each stage in the calculations.

1. *Coordinate Bethe ansatz.* In section 6.3, we diagonalise the deformed Markov operator using the coordinate Bethe ansatz introduced in chapter 5. The Bethe equations are then derived using a procedure similar to the one used for the TASEP in section 2.4. However, in this case, the Bethe equations involve as many unknown constants as the number of lattice sites, so a more sophisticated approach than the one used for the TASEP is required.

2. *Polynomial formulation.* The first key idea is that instead of dealing with each constant individually, one combines them all by making them the coefficients of a polynomial. This is done in section 6.4.1. The polynomial is defined in such a way that the Bethe roots are its zeros. This polynomial

thus contains the same information as the Bethe roots. The Bethe equations then take the form of a functional equation for this polynomial and another, unknown polynomial.

3. *Going "beyond the equator".* Although the polynomial formulation of the Bethe equation is a significant simplification, for the problem at hand it is still somewhat cumbersome to deal with because one must solve for two unknown polynomials. However, importantly, these two polynomials are not independent. Their relation can be exploited to produce another equivalent form of the Bethe equation, in terms of another polynomial. This is done in section 6.4.2. In the literature, this form of the Bethe equation is known as the Bethe equation "beyond the equator", as the two different Bethe equations encode the relations of the Bethe ansatz solutions above and below half-filling of the lattice.

4. *One function formulation.* In section 6.4.3, we use the results from the previous two steps, to combine the polynomials of the two Bethe equations into a single function. The Bethe equation then becomes a functional equation for this single unknown function. This is the form of the Bethe equation that we use for practical calculations.

5. *Perturbation theory.* In the remainder of the chapter, we solve the one-function formulation of the Bethe equation in perturbation theory to calculate particle displacement statistics.

6.2 Problem statement

6.2.1 Model definition

Before we derive the Bethe equations, we recall the precise definition of the model we will be considering in this chapter. The process takes place on a ring, whose total number of sites we take to be $L + 1$. On the ring are M normal particles, which hop to the right and left with rates 1 and x, and one defect particle, which hops to the right and left with rates α and x/α. The defect is also allowed to overtake normal particles with its usual hopping rates. Thus, it does not distinguish between normal particles and holes. We can summarise the dynamics

of this process as follows,

$$10 \underset{x/\alpha}{\overset{\alpha}{\rightleftarrows}} 01 \quad ; \quad 20 \underset{x}{\overset{1}{\rightleftarrows}} 02 \quad ; \quad 12 \underset{x/\alpha}{\overset{\alpha}{\rightleftarrows}} 21 \,, \tag{6.1}$$

where we denote empty sites by 0, the defect by 1 and normal particles by 2. Note that we have reverted to the notation of chapter 4, which is different from the notation in chapter 5, where the labels 1 and 2 were swapped. We do this because in this model, the defect has priority in the dynamics over the normal particles. Therefore, it makes more sense to call it a "first-class" particle, with the normal particles being "second-class".

6.2.2 Large deviation function of particle displacement

We now recall the formulation of the long time current statistics problem, which was already derived in chapter 5. The central objects we are interested in are random variables that count the number of hops of a given type. We define $Y_1(t)$ as the variable that counts the processes $10 \to 01$ minus the reverse processes, $Y_2(t)$ as the variable that counts the processes $20 \to 02$ minus the reverse processes, and finally $Y_{12}(t)$ as the variable that counts the processes $12 \to 21$ minus the reverse processes, all up to time t. Strictly speaking, these variables should be conditioned on some initial condition. However, as the process we are considering is ergodic and we will be taking the long time limit, we may neglect this.

We define the joint moment generating function of $Y_1(t)$, $Y_2(t)$, $Y_{12}(t)$,

$$F_t(\mathcal{C};\gamma_1,\gamma_2,\gamma_{12}) = \sum_{Y_1=-\infty}^{\infty}\sum_{Y_2=-\infty}^{\infty}\sum_{Y_{12}=-\infty}^{\infty} \mathrm{e}^{\gamma_1 Y_1+\gamma_2 Y_2+\gamma_{12}Y_{12}} P_t(\mathcal{C};Y_1,Y_2,Y_{12}), \tag{6.2}$$

where $\mathcal{C}$ denotes the configuration of the system (*i.e.* the positions of all particles) and $P_t(\mathcal{C};Y_1,Y_2,Y_{12})$ denotes the joint probability of observing configuration $\mathcal{C}$, together with $Y_i(t) = Y_i$ for $i = 1, 2, 12$. The evolution of F_t is given by a deformed master equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}F_t(\mathcal{C};\gamma_1,\gamma_2,\gamma_{12}) \;=\; \sum_{\mathcal{C}'}\mathcal{M}_\gamma(\mathcal{C},\mathcal{C}')F_t(\mathcal{C}';\gamma_1,\gamma_2,\gamma_{12}). \tag{6.3}$$

For the purpose of analysing particle current statistics, we actually do not need

to consider the general joint moment generating function. Instead, we will focus on the two key cases: *(i)* the motion of the defect ($\gamma_1 = \gamma_{12}$, $\gamma_2 = 0$) and *(ii)* the motion of the normal particles ($\gamma_1 = \gamma_{12} = 0$). It makes sense to introduce notation with a single conjugate parameter,

$$\gamma_i = a_i \gamma, \tag{6.4}$$

where a_i are constants. We may set $a_1 = a_{12} = 1$ and $a_2 = 0$ to track the motion of the defect, or $a_2 = 1$ and $a_1 = a_{12} = 0$ to track the motion of the normal particles.

We note that as there is only one particle of species 1, we do not need to keep track of the order of species when specifying a configuration. Instead, it is easier to always choose the particle labelling in such a way that the defect is the first particle. We therefore denote configurations by $\{x_0, x_1, \dots, x_M\}$, where x_0 is the position of the defect and x_i for $i = 1, \dots, M$ are the positions of the normal particles. The eigenvalue-eigenvector problem for the deformed transition rate matrix $\mathcal{M}_\gamma$ is then given by the equation

$$\begin{aligned}\lambda(\gamma)\psi(x_0, \dots, x_M) = &-[M(1+x) + (\alpha + x/\alpha)]\psi(x_0, \dots, x_M) \\ &+\mathrm{e}^{\gamma_1}\alpha\psi(x_0 - 1, \dots) + \mathrm{e}^{-\gamma_1}(x/\alpha)\psi(x_0 + 1, \dots) \\ &+\sum_{i=1}^{M}[\mathrm{e}^{\gamma_2}\psi(\dots, x_i - 1, \dots) + \mathrm{e}^{-\gamma_2}x\psi(\dots, x_i + 1, \dots)],\end{aligned} \tag{6.5}$$

together with the collision equations

$$\begin{aligned}-(1+x)\psi(\dots, x_i, x_i + 1, \dots) + \mathrm{e}^{\gamma_2}\psi(\dots, x_i, x_i, \dots) \\ +\mathrm{e}^{-\gamma_2}x\psi(\dots, x_i + 1, x_i + 1, \dots) = 0,\end{aligned} \tag{6.6a}$$

for $i = 1, \dots, M-1$, and for $i = 0, M$,

$$\begin{aligned}&-x\psi(x_0, x_0 + 1, \dots) + \mathrm{e}^{\gamma_2}\psi(x_0, x_0, \dots) + \mathrm{e}^{-\gamma_1}(x/\alpha)\psi(x_0 + 1, x_0 + 1, \dots) \\ &\quad = \mathrm{e}^{-\gamma_{12}}(x/\alpha)\psi(x_0 + 1, \dots, x_0 + L + 1)\end{aligned} \tag{6.6b}$$

$$\begin{aligned}&-\psi(x_0, \dots, x_0 + L) + x\mathrm{e}^{-\gamma_2}\psi(x_0, \dots, x_0 + L + 1) + \mathrm{e}^{\gamma_1}\alpha\psi(x_0 - 1, \dots, x_0 + L) \\ &\quad = \mathrm{e}^{\gamma_{12}}\alpha\psi(x_0 - 1, x_0, \dots),\end{aligned} \tag{6.6c}$$

where we have used the periodic boundary condition (5.6) to ensure that the position of the defect is always the first argument.

Now the key observation is that the particle displacements $Y_i(t)$ satisfy a large deviation principle in the long time limit (see section 1.4 and section 2.4),

$$\langle \exp[\gamma_1 Y_1(t) + \gamma_2 Y_2(t) + \gamma_{12} Y_{12}(t)] \rangle \sim \mathrm{e}^{\mu(\gamma_1,\gamma_2,\gamma_{12})t}, \qquad \text{as } t \to \infty, \tag{6.7}$$

where $\mu(\gamma_1, \gamma_2, \gamma_{12})$ is the joint scaled cumulant generating function of $Y_i(t)$ for $i = 1, 2, 12$. Now as the left hand side of this expression is simply F_t marginalised over configurations $\mathcal{C}$, the evolution equation (6.3) implies that $\mu(\gamma_1, \gamma_2, \gamma_{12})$ can be identified with the eigenvalue $\lambda(\gamma)$ of the deformed transition rate matrix $\mathcal{M}_\gamma$ with the largest real part. Henceforth, whenever we write $\lambda(\gamma)$, we shall refer exclusively to this eigenvalue.

Note that the problem defined by (6.5)–(6.6c) can be used to obtain multiple eigenvalues. To ensure that the eigenvalue we find is the correct one, we must check that it satisfies

$$\lambda(0) = 0. \tag{6.8}$$

This is because it must converge to the unique 0 eigenvalue of the undeformed transition rate matrix. In fact, this can be ensured by requiring that the corresponding eigenvector satisfies the condition

$$\psi(x_0 + 1, \ldots, x_M + 1) = 1. \tag{6.9}$$

See section 2.4 for an explanation of why this is sufficient. Note that this essentially means that this eigenvector is translationally invariant. We shall refer to this as the periodicity condition.

Once $\lambda(\gamma)$ is found, rewriting (6.7), we have

$$\lambda(\gamma) = \lim_{t\to\infty} \frac{\log\langle \exp[\gamma(a_1 Y_1(t) + a_2 Y_2(t) + a_{12} Y_{12}(t))] \rangle}{t}. \tag{6.10}$$

Thus it allows us to calculate the cumulants of $Y_i(t)$ in the long time limit. For instance setting $a_2 = 1$ and $a_1 = a_{12} = 0$, we have

$$\left.\frac{\mathrm{d}\lambda(\gamma)}{\mathrm{d}\gamma}\right|_{\gamma=0} = J = \lim_{t\to\infty} \frac{\langle Y_2(t) \rangle}{t} \tag{6.11a}$$

$$\left.\frac{\mathrm{d}^2\lambda(\gamma)}{\mathrm{d}\gamma^2}\right|_{\gamma=0} = \Delta = \lim_{t\to\infty} \frac{\mathrm{Var}(Y_2(t))}{t}. \tag{6.11b}$$

We will use these expressions as the definitions of the mean current J and diffusion constant Δ of the normal particles.

Note that the definition of the current J is slightly different from the one used in chapter 4. There, current was defined as the net flow of particles across a given bond, whereas here it is defined as the net total displacement of all particles. Therefore the current from the matrix product definition J_{MP} is related to the current from the Bethe ansatz definition J_{BA} as

$$J_{BA} = LJ_{MP}. \tag{6.12}$$

6.3 Bethe ansatz solution

6.3.1 Bethe ansatz

As in chapter 5, we solve the eigenvalue problem using a simple form of the nested coordinate Bethe ansatz,

$$\psi(x_0, \ldots, x_M) = (\alpha \mathrm{e}^{\gamma_1})^{x_0} \left(\prod_{i=1}^{M} \mathrm{e}^{\gamma_2 x_i} \right) \sum_{\sigma} A_\sigma \prod_{i=0}^{M} z_{\sigma(i)}^{x_i}, \tag{6.13}$$

where σ indicates the permutations of $\{0, \ldots, M\}$. Note that we have simplified the notation from chapter 5 by dropping the species order dependence, $A_\sigma^{Q_0 \cdots Q_M} \to A_\sigma$. This can be done without ambiguity as the defect is taken to be particle 0, so the order is fixed to be $\{Q_0, \ldots, Q_M\} = \{1, 2, 2, \ldots, 2\}$.

Plugging the ansatz (6.13) into (6.5), we obtain,

$$\lambda(\gamma) = -[M(1+x) + \alpha + x/\alpha] + \sum_{i=0}^{M} [z_i^{-1} + x z_i]. \tag{6.14}$$

The periodicity condition (6.9) is then expressed as

$$\alpha \mathrm{e}^{\gamma_1 + M\gamma_2} \prod_{i=0}^{M} z_i = 1. \tag{6.15}$$

From this, we see that in the undeformed limit ($\gamma \to 0$), all Bethe roots z_i converge to 1, except for one root, which converges to α^{-1}. Without loss of generality, we

may say

$$\gamma \to 0 \Rightarrow \begin{cases} z_0 \to \alpha^{-1} \\ z_i \to 1, \qquad i \geq 1 \end{cases}. \tag{6.16}$$

Given this, it is easy to see from (6.14) that $\lambda(0,0,0) = 0$, as required.

6.3.2 Derivation of Bethe equations

Plugging the ansatz (6.13) into (6.6a)–(6.6c), we obtain,

$$A_{\dots ji\dots} = S_{ij} A_{\dots ij\dots} \tag{6.17a}$$

$$\begin{aligned} &\left[(w_{ij} - xz_j)A_{ij\dots} - \mathrm{e}^{\bar{\gamma}}(\mathrm{e}^{\gamma_2} z_j)^{L+1} x z_i A_{i\dots j}\right] \\ &\quad + \left[(w_{ij} - xz_i)A_{ji\dots} - \mathrm{e}^{\bar{\gamma}}(\mathrm{e}^{\gamma_2} z_i)^{L+1} x z_j A_{j\dots i}\right] = 0 \end{aligned} \tag{6.17b}$$

$$\begin{aligned} &\left[(w_{ij} - z_i)(\mathrm{e}^{\gamma_2} z_j)^{L+1} A_{i\dots j} - \mathrm{e}^{-\bar{\gamma}} z_j A_{ij\dots}\right] \\ &\quad + \left[(w_{ij} - z_j)(\mathrm{e}^{\gamma_2} z_i)^{L+1} A_{j\dots i} - \mathrm{e}^{-\bar{\gamma}} z_i A_{ji\dots}\right] = 0, \end{aligned} \tag{6.17c}$$

where, we have introduced the shorthand notation,

$$S_{ij} = -\frac{w_{ij} - (1+x)z_j}{w_{ij} - (1+x)z_i} \tag{6.18}$$

$$w_{ij} = 1 + x z_i z_j \tag{6.19}$$

$$\bar{\gamma} = \gamma_1 - \gamma_2 - \gamma_{12}. \tag{6.20}$$

Now, similarly to the procedure in section 2.4, the task is to eliminate the amplitudes A to obtain a closed equation for just the Bethe roots z. We apply (6.17a) $M-1$ times to (6.17b) and (6.17c) to transform $A_{i\dots j} \to A_{ij\dots}$ and $A_{j\dots i} \to A_{ji\dots}$. This gives

$$\left[(w_{ij} - xz_j) + x z_i S_{ij} E_j\right] A_{ij\dots} + \left[(w_{ij} - xz_i) + x z_j S_{ji} E_i\right] A_{ji\dots} = 0 \tag{6.21a}$$

$$\left[-(w_{ij} - z_i) S_{ij} E_j - z_j\right] A_{ij\dots} + \left[-(w_{ij} - z_j) S_{ji} E_i - z_i\right] A_{ji\dots} = 0, \tag{6.21b}$$

where

$$E_i = \mathrm{e}^{\bar{\gamma}}(\mathrm{e}^{\gamma_2} z_i)^{L+1} \prod_{k=0}^{M} S_{ik}. \tag{6.22}$$

Now as we cannot have vanishing amplitudes, the simultaneous equations (6.21a) and (6.21b) must be degenerate. This means that

$$\begin{aligned}&[(w_{ij} - xz_j) + xz_i S_{ij} E_j]\,[-(w_{ij} - z_j) S_{ji} E_i - z_i] \\ &- [(w_{ij} - xz_i) + xz_j S_{ji} E_i]\,[-(w_{ij} - z_i) S_{ij} E_j - z_j] = 0. \end{aligned} \tag{6.23}$$

Multiplying this out and simplifying a few expressions gives,

$$[w_{ij} - (1+x)z_i]E_i - [w_{ij} - (1+x)z_i]E_j + (z_j - z_i)(1 + xE_iE_j) = 0. \tag{6.24}$$

At this point it makes it easier to introduce the change of variable,

$$y_i = \frac{1 - z_i}{1 - xz_i}, \tag{6.25}$$

which has the inverse transformation

$$z_i = \frac{1 - y_i}{1 - xy_i}. \tag{6.26}$$

Note that in terms of y_i, the term E_i becomes

$$E_i = \mathrm{e}^{\bar{\gamma}} (\mathrm{e}^{\gamma_2} z_i)^{L+1} \prod_{k=0}^{M} \frac{xy_i - y_k}{y_i - xy_k}. \tag{6.27}$$

Also, in terms of y_i, equation (6.24) becomes,

$$(xy_j - y_i)E_i - (xy_i - y_j)E_j + (y_j - y_i)(1 + xE_iE_j) = 0. \tag{6.28}$$

This can now be factorised as follows,

$$y_j(1 + E_j)(1 + xE_i) - y_i(1 + E_i)(1 + xE_j) = 0. \tag{6.29}$$

This implies the Bethe equations

$$y_i \frac{1 + E_i}{1 + xE_i} = B, \tag{6.30}$$

where B is a constant. Rewriting this to isolate E_i and using the explicit form (6.27), we finally get the following form for the Bethe equations,

$$\mathrm{e}^{\bar{\gamma}} \left(\mathrm{e}^{\gamma_2} \frac{1 - y_i}{1 - xy_i} \right)^{L+1} \prod_{k=0}^{M} \frac{xy_i - y_k}{y_i - xy_k} = -\frac{B - y_i}{Bx - y_i}. \tag{6.31}$$

6.3.3 Auxiliary equations

To complete the formulation of the problem, we now also explicitly write the auxiliary equations that are needed to form a closed system of equations. First of all, we note that the periodicity condition (6.9), together with the form of the ansatz (6.13) implies

$$\alpha \mathrm{e}^{\gamma_1 + M\gamma_2} \prod_{i=0}^{M} \frac{1 - y_i}{1 - x y_i} = 1. \tag{6.32}$$

Secondly, we need to fix the constant B. We can do this by multiplying (6.31) for all i, which gives

$$\mathrm{e}^{(M+1)\bar{\gamma}} \left(\prod_{i=0}^{M} \mathrm{e}^{\gamma_2} \frac{1 - y_i}{1 - x y_i} \right)^{L+1} \prod_{k=0}^{M} \prod_{i=0}^{M} \frac{x y_i - y_k}{y_i - x y_k} = (-)^{M+1} \prod_{i=0}^{M} \frac{B - y_i}{B x - y_i}. \tag{6.33}$$

On the left hand side, the term with the power $L + 1$ can be simplified using the periodicity condition (6.15). In the double product, for all pairs $i \neq k$, there will be a corresponding factor with the indices swapped. Taking both factors together, we have

$$\frac{x y_i - y_k}{y_i - x y_k} \frac{x y_k - y_i}{y_k - x y_i} = 1. \tag{6.34}$$

The remaining factors are ones with $k = i$, which simplify to -1. As there are $M + 1$ such factors in total, this will cancel with the $(-)^{M+1}$ on the right hand side. Then we are left with,

$$\prod_{i=0}^{M} \frac{B - y_i}{B x - y_i} = \alpha^{-(L+1)} \mathrm{e}^{-(M+1)\gamma_{12} - (L-M)(\gamma_1 - \gamma_2)}. \tag{6.35}$$

Finally, with a little algebra we can rewrite the eigenvalue (6.14) in terms of y_i as follows,

$$\lambda(\gamma) = 1 + x - \alpha - x/\alpha + (1 - x) \sum_{i=0}^{M} \left[\frac{1}{1 - y_i} - \frac{1}{1 - x y_i} \right]. \tag{6.36}$$

The equations (6.31), (6.15) and (6.35) now form a complete set of equations for the Bethe roots y_i. Once these are solved, one may calculate the eigenvalue $\lambda(\gamma)$ through (6.36). However, this system of equations has quite a complex structure.

It is a set of polynomial equations of degree $L + M + 3$, in which the coefficients themselves are functions of the roots, and it contains the undetermined constant B, which is also a function of the roots. This is significantly more complex than the Bethe equations for the TASEP case, which we examined in section 2.4. To deal with this problem, we reformulate it as a functional equation. We shall proceed to this in section 6.4.

6.3.4 Undeformed limit

For completeness, we also state explicitly the undeformed limit, $\gamma_1,\ \gamma_2,\ \gamma_{12} \to 0$. From (6.16), we have that the Bethe roots y_i take the values,

$$\begin{cases} y_0 \to y_0^{(0)} = \frac{\alpha-1}{\alpha-x} \\ y_i \to 0, \qquad\qquad i \geq 1 \end{cases}. \tag{6.37}$$

Using this, we then also obtain from (6.35),

$$B \to B^{(0)} = \frac{\alpha - 1}{\alpha - x} \frac{\alpha^{L+1} - x^M}{\alpha^{L+1} - x^{M+1}}. \tag{6.38}$$

6.4 Functional reformulation of Bethe ansatz

We wish to reformulate the Bethe equations (6.31) in such a way that we will not need to solve for each root y_i individually. This approach is an extension of the functional formulation that was introduced in section 2.4 but it is more involved as we need to incorporate a more complex structure. These calculations are guided by the approach developed for the pure PASEP case [164], which was inspired by the functional Bethe ansatz.

6.4.1 Polynomial formulation

First, we define the single-variable polynomial $Q(T)$ as

$$Q(T) = \prod_{k=0}^{M} (T - y_k). \tag{6.39}$$

By construction, all Bethe roots y_i are roots of $Q(T)$. Then the Bethe equation (6.31) can be expressed as

$$\Gamma h(T)Q(xT) + x^M h(xT)Q(T/x) = 0, \tag{6.40}$$

where

$$\Gamma = \mathrm{e}^{\gamma_1 + L\gamma_2 - \gamma_{12}} \tag{6.41}$$

$$h(T) = (1-T)^{L+1}(Bx - T). \tag{6.42}$$

However, note that this is polynomial equation of degree $L + M + 3$, whereas we expect only $M + 1$ Bethe roots. This means that the left hand side is divisible by $Q(T)$ as a polynomial. In other words, there exists some polynomial $R(T)$ of degree $L + 2$ such that

$$R(T)Q(T) = \Gamma h(T)Q(xT) + x^M h(xT)Q(T/x). \tag{6.43}$$

This is now a functional equation for $Q(T)$. Note that $Q(T)$ actually contains all the same information as the Bethe roots y_i. Therefore solving (6.43) for $Q(T)$ is equivalent to solving (6.31) for y_i.

Auxiliary equations

We can complete the reformulation of the problem in terms of $Q(T)$ as follows. The periodicity condition (6.32) becomes,

$$\alpha \mathrm{e}^{\gamma_1 + M\gamma_2} \frac{Q(1)}{x^{M+1}Q(x^{-1})} = 1. \tag{6.44}$$

The B-fixing equation (6.35) becomes,

$$\frac{Q(B)}{Q(Bx)} = \alpha^{-(L+1)} \mathrm{e}^{-(M+1)\gamma_{12} - (L-M)(\gamma_1 - \gamma_2)}. \tag{6.45}$$

We can also express $\lambda(\gamma)$ in terms of $Q(T)$ as,

$$\lambda(\gamma) = 1 + x - \alpha - x/\alpha + (1-x)\left(\frac{Q'(1)}{Q(1)} - \frac{Q'(x^{-1})}{xQ(x^{-1})}\right). \tag{6.46}$$

6.4.2 Going "beyond the equator"

The polynomial formulation (6.43) is an improvement over (6.31). However, it is still somewhat cumbersome as it involves an additional polynomial $R(T)$, which we do not require for $\lambda(\gamma)$ but which it is still necessary to calculate to find $Q(T)$. We proceed to modify the polynomial formulation of the Bethe equation in a procedure that is sometimes called going "beyond the equator" [172]. This will eventually lead to us being able to express the equations in terms of a single function, for which we can then solve directly.

To obtain the "beyond the equator" formulation, we first divide (6.43) by $Q(T)Q(xT)Q(T/x)$,

$$\frac{R(T)}{Q(xT)Q(T/x)} = \Gamma \frac{h(T)}{Q(T)Q(T/x)} + x^M \frac{h(xT)}{Q(T)Q(xT)}. \tag{6.47}$$

By the rules of polynomial division, we can write

$$\frac{h(T)}{Q(T)Q(T/x)} = \frac{U(T)}{Q(T/x)} + \frac{V(T)}{Q(T)} + W(T), \tag{6.48}$$

where $U(T)$ and $V(T)$ are polynomials of degree at most M and $W(T)$ is a polynomial of degree at most $L-2M$. Using this decomposition of $h(T)$, we can write (6.47) as

$$\begin{aligned} \frac{R(T)}{Q(xT)Q(T/x)} &= \Gamma \left(\frac{U(T)}{Q(T/x)} + \frac{V(T)}{Q(T)} + W(T) \right) \\ &\quad + x^M \left(\frac{U(xT)}{Q(T)} + \frac{V(xT)}{Q(xT)} + W(xT) \right). \end{aligned} \tag{6.49}$$

Now crucially, we observe that, as $Q(xT)$ and $Q(T/x)$ generically do not share any roots with $Q(T)$, there are only two terms in this equation that have poles at the roots of $Q(T)$. Both of them have polynomials of degree at most M in the numerator, whereas $Q(T)$ is of degree $M+1$. The only way this can be satisfied identically is if

$$\Gamma V(T) + x^M U(xT) = 0. \tag{6.50}$$

We thus have found a relation between $U(T)$ and $V(T)$. Using this to eliminate

$V(T)$ in (6.48), we obtain

$$\frac{h(T)}{Q(T)Q(T/x)} = \frac{U(T)}{Q(T/x)} - \Gamma^{-1}x^M\frac{U(xT)}{Q(T)} + W(T). \tag{6.51}$$

Now as $W(T)$ is a polynomial, we can always write it as

$$W(T) = \Gamma^{-1}x^M\tilde{W}(T) - \tilde{W}(T/x), \tag{6.52}$$

where $\tilde{W}(T)$ is simply the polynomial with the coefficients w_k of $W(T)$ rescaled by $[\Gamma x^{-M} - x^{-k}]^{-1}$. Then we can write

$$\begin{aligned} h(T) &= [U(T) - \tilde{W}(T/x)Q(T/x)]Q(T) \\ &\quad -\Gamma^{-1}x^M[U(xT) - \tilde{W}(T)Q(T)]Q(T/x). \end{aligned} \tag{6.53}$$

We now define the polynomial

$$P(T) = 2C[U(xT) - \tilde{W}(T)Q(T)], \tag{6.54}$$

where C is a constant. We will later see that it is convenient to choose

$$C = \frac{1 - \Gamma^{-1}x^M}{2Bx}y_0Q(0). \tag{6.55}$$

Note that this implies

$$P(0) = -y_0. \tag{6.56}$$

Then we obtain the Bethe equation beyond the equator

$$2Ch(T) = P(T/x)Q(T) - \Gamma^{-1}x^MP(T)Q(T/x), \tag{6.57}$$

where $h(T)$ is defined by (6.42).

Undeformed limit

For future calculations, at this point it is helpful to consider the undeformed limit of (6.57). From the behaviour of the Bethe roots (6.37), we see that $Q(T)$ will behave as

$$Q(T) \to Q^{(0)}(T) = T^M(T - y_0^{(0)}). \tag{6.58}$$

Then in (6.57), we see that the right hand side is proportional to T^M but the left hand side is proportional to $(1-T)^{L+1}$. Given that $P(T)$ is a polynomial, this is incompatible unless

$$C \to C^{(0)} = 0. \tag{6.59}$$

Then we see by inspection that

$$P(T) \to P^{(0)}(T) = T - y_0^{(0)}. \tag{6.60}$$

It might seem strange that $P^{(0)}(T)$ is a polynomial of degree 1, when we expect $P(T)$ to be a polynomial of degree $L - M + 1$. In fact, this indicates that the coefficients of the higher powers in $P(T)$ are of higher order in γ.

6.4.3 One function reformulation

Currently, the Bethe equation (6.57) is expressed in terms of two unknown functions, $P(T)$ and $Q(T)$. We can re-express it as an equation for a single function using the following procedure.

Firstly, we note that from (6.58) and remembering the form of $Q(T)$, (6.39), we can write $Q(T)$ as follows,

$$Q(T) = T^M(T - y_0)(1 + q(T)), \tag{6.61}$$

where $q(T)$ is a polynomial in negative powers of T that is of order γ.

Similarly, $P(T)$ must have one root $\bar{y}_0$ close to $y_0^{(0)}$, although note that $\bar{y}_0 \neq y_0$, as otherwise the right hand side of (6.57) would have y_0 as a root, whereas the left hand side would not. The remaining roots of $P(T)$ must diverge as $\gamma \to 0$, which corresponds to the vanishing of the coefficients. It is helpful to introduce the rational function

$$p(T) = \frac{P(T)}{T - y_0} - 1. \tag{6.62}$$

This function is analytic except for a simple pole at $T = y_0$. Also, by the definition of C (6.55), it has the property $p(0) = 0$. Moreover, from (6.60), we see that we

must have $p(T) = O(\gamma)$. Rearranging this definition, we have

$$P(T) = (T - y_0)(1 + p(T)). \tag{6.63}$$

With this in mind, we now divide (6.57) by $T^M(T - y_0)(T/x - y_0)$, which yields,

$$2C\tilde{h}(T) = (1 + p(T/x))(1 + q(T)) - \Gamma^{-1}(1 + p(T))(1 + q(T/x)), \tag{6.64}$$

where

$$\tilde{h}(T) = \frac{(1 - T)^{L+1}(Bx - T)}{T^M(T - y_0)(T/x - y_0)}. \tag{6.65}$$

We now introduce the two functions

$$\begin{aligned} w(T) = \frac{1}{2}[&\log(1 + q(T)) - \log(1 + q(T/x)) - \log(1 + p(T)) \\ &+ \log(1 + p(T/x)) + \log\Gamma] \end{aligned} \tag{6.66}$$

$$\begin{aligned} \tilde{w}(T) = \frac{1}{2}[&\log(1 + q(T)) + \log(1 + q(T/x)) + \log(1 + p(T)) \\ &+ \log(1 + p(T/x)) - \log\Gamma]. \end{aligned} \tag{6.67}$$

Using these, we see that (6.64) becomes

$$2C\tilde{h}(T) = \mathrm{e}^{\tilde{w}(T)+w(T)} - \mathrm{e}^{\tilde{w}(T)-w(T)}. \tag{6.68}$$

This can be rewritten as

$$w(T) = \operatorname{arcsinh}\left(C\tilde{h}(T)\mathrm{e}^{-\tilde{w}(T)}\right). \tag{6.69}$$

Now to make this an equation for just one function, we need to establish a relation between $w(T)$ and $\tilde{w}(T)$. We can do this by introducing a linear operator X that acts on power series in T as follows. Given a power series $u(T) = \sum_k u_k T^k$, we define

$$X[u](T) = u(T) + 2\sum_{k=-\infty}^{\infty} \frac{x^{|k|}}{1 - x^{|k|}} u_k T^k, \tag{6.70}$$

with the convention

$$\frac{x^{|0|}}{1 - x^{|0|}} = -1. \tag{6.71}$$

The key property of this operator is that due to the absolute value in the exponent, it acts differently on series with only negative and only positive powers. Specifically, if $u^-(T)$ is a series with only negative powers, we have

$$\begin{aligned} X[u^-](T) - X[u^-](T/x) &= u^-(T) - u^-(T/x) \\ &\quad +2\sum_{k=-\infty}^{-1}\left(\frac{x^{-k}}{1-x^{-k}} - \frac{x^{-2k}}{1-x^{-k}}\right) u_k^- T^k \\ &= u^-(T) - u^-(T/x) + 2\sum_{k=-\infty}^{-1} u_k^- x^{-k} T^k \\ &= u^-(T) + u^-(T/x). \end{aligned} \tag{6.72}$$

On the other hand, if $u^+(T)$ is a series with only positive powers, we have,

$$\begin{aligned} X[u^+](T) - X[u^+](T/x) &= u^+(T) - u^+(T/x) \\ &\quad +2\sum_{k=1}^{\infty}\left(\frac{x^{k}}{1-x^{k}} - \frac{1}{1-x^{k}}\right) u_k^+ T^k \\ &= u^+(T) - u^+(T/x) - 2\sum_{k=1}^{\infty} u_k^+ T^k \\ &= -u^+(T) - u^+(T/x). \end{aligned} \tag{6.73}$$

Now consider the definitions (6.66) and (6.67). As $p(T) = O(\gamma)$ and $q(T) = O(\gamma)$, we can expand the logs in powers of p and q. Now $q(T)$ is a polynomial in negative powers of T. On the other hand, $p(T)$ is analytic at $T = 0$ with $p(0) = 0$, meaning that its power series expansion at $T = 0$ will have only positive powers of T. Hence, if we apply X to $w(T)$, we get,

$$\begin{aligned} X[w(T)] &= \frac{1}{2}\sum_{n=1}^{\infty}\frac{(-)^n}{n}\left[X[q^n](T) - X[q^n](T/x) - X[p^n](T) + X[p^n](T/x)\right] \\ &\quad +\frac{1}{2}X[\log\Gamma] && (6.74) \\ &= \frac{1}{2}\left[\sum_{n=1}^{\infty}\frac{(-)^n}{n}\left[q^n(T) + q^n(T/x) + p^n(T) + p^n(T/x)\right] - \log\Gamma\right] && (6.75) \\ &= \tilde{w}(T). && (6.76) \end{aligned}$$

Then we can finally write a Bethe equation in terms of a single function, $w(T)$ as follows,

$$w(T) = \operatorname{arcsinh}\left(C\tilde{h}(T)\mathrm{e}^{-X[w](T)}\right). \tag{6.77}$$

Auxiliary equations

To complete the one function formulation, we would like to write the auxiliary equations (6.44) and (6.45), as well as the eigenvalue λ in terms of $w(T)$. Note that those equations are given in terms of $Q(T)$. To isolate the correct component of $w(T)$, it is useful to consider a linear operator $\mathcal{P}$ defined as follows,

$$\mathcal{P}[w](T) = 2 \oint' \frac{w(T')}{T - T'}, \tag{6.78}$$

where the contour of integration is a small circle around the origin and we have suppressed a factor of $\mathrm{d}T'/(2\pi\mathrm{i})$. Similarly to the argument relating $w(T)$ to $\tilde{w}(T)$ through the X operator, we need to expand the logarithms in $w(T)$ (see the definition (6.66)) in powers of p and q. Performing this expansion in (6.78), we see that, as p is analytic at $T = 0$, the terms containing p will not contribute. Similarly, the constant term $(1/2)\log\Gamma$ does not contribute. Then we have

$$\mathcal{P}[w](T) = \oint' \frac{\log(1 + q(T')) - \log(1 + q(T'/x))}{T - T'}. \tag{6.79}$$

This integral can be evaluated using the residue theorem. Specifically, a complex contour integral can be evaluated as $-2\pi\mathrm{i}$ times the residues of the poles of the integrand outside the contour. As q is a (finite) polynomial in negative powers of T, the only pole outside the contour is the simple pole at $T' = T$. Thus we obtain the important identity

$$\mathcal{P}[w](T) = \log(1 + q(T)) - \log(1 + q(T/x)). \tag{6.80}$$

Strictly speaking, we have only shown this to hold provided we expand the logarithms up to a finite number of powers of q. However, this is sufficient for our purposes.

Periodicity condition. Now consider the periodicity condition (6.44). Note that

$$Q(1) = (1 - y_0)(1 + q(1)), \qquad x^M Q(x^{-1}) = (1 - xy_0)(1 + q(x^{-1})). \tag{6.81}$$

Then we obtain,

$$\alpha \mathrm{e}^{\gamma_1 + M\gamma_2} \frac{1 - y_0}{1 - xy_0} \exp[\log(1 + q(1)) - \log(1 + q(x^{-1}))] = 1. \tag{6.82}$$

We now use (6.80) and rearrange the equation to solve for y_0. This yields,

$$y_0 = \frac{\alpha \mathrm{e}^{\delta} - 1}{\alpha \mathrm{e}^{\delta} - x}, \tag{6.83}$$

where

$$\delta = \gamma_1 + M\gamma_2 + \mathcal{P}[w](1). \tag{6.84}$$

B-fixing condition. Similarly, consider the B-fixing equation (6.45). This can be written as,

$$\begin{aligned} \frac{(B - y_0)}{x^M (Bx - y_0)} \exp[\log(1 + q(B)) - \log(1 + q(Bx))] \\ = \alpha^{-(L+1)} \mathrm{e}^{-(M+1)\gamma_{12} - (L-M)(\gamma_1 - \gamma_2)}. \end{aligned} \tag{6.85}$$

Using (6.80) and solving for B yields,

$$B = y_0 \frac{\alpha^{L+1} \mathrm{e}^{\delta_B} - x^M}{\alpha^{L+1} \mathrm{e}^{\delta_B} - x^{M+1}}, \tag{6.86}$$

where

$$\delta_B = (M+1)\gamma_{12} - (L-M)(\gamma_1 - \gamma_2) - \mathcal{P}[w](Bx). \tag{6.87}$$

Note that this is self-referential, as B is expressed in terms of δ_B, which is in turn defined in terms of B. However we will see that this equation allows us to compute B order by order in perturbation theory.

Normalisation. Actually, in this formulation, we need an additional equation. This essentially comes from the fact that we need to fix the constant C. As C multiplied the whole Bethe equation (6.57), this can be seen as a normalisation condition. We can obtain this by examining the form of $w(T)$, (6.66). As p has an expansion in only positive powers of T and q has an expansion in only negative powers of T at $T = 0$, the only contribution to the constant term (the coefficient of T^0) will be $(1/2)\log\Gamma$. Hence by the residue theorem, we have the equality

$$\oint \frac{w(T)}{T} = \frac{1}{2}\log\Gamma, \tag{6.88}$$

where Γ is given by (6.41).

Eigenvalue. Finally, we express the eigenvalue $\lambda(\gamma)$ in terms of $w(T)$. From (6.46), we have,

$$\begin{aligned}\lambda(\gamma) &= 1 + x - \alpha - x/\alpha + (1-x)\left.\frac{\mathrm{d}}{\mathrm{d}T}[\log Q(T) - \log Q(T/x)]\right|_{T=1} \quad (6.89)\\ &= 1 + x - \alpha - x/\alpha + (1-x)\left.\frac{\mathrm{d}}{\mathrm{d}T}\log\frac{T-y_0}{T-xy_0}\right|_{T=1}\\ &\quad + (1-x)\left.\frac{\mathrm{d}}{\mathrm{d}T}[\log(1+q(T)) - \log(1+q(T/x))]\right|_{T=1}. \quad (6.90)\end{aligned}$$

Then using (6.80) and (6.83), we can write this as,

$$\lambda(\gamma) = (1-\mathrm{e}^{-\delta})(\alpha\mathrm{e}^{\delta} - x/\alpha) + (1-x)\left.\frac{\mathrm{d}}{\mathrm{d}T}\mathcal{P}[w](T)\right|_{T=1}, \quad (6.91)$$

where δ is as defined in (6.84).

We now have a complete formulation of the large deviation problem of the particle current. Specifically, one needs to solve the functional equation (6.77) for $w(T)$ (together with the auxiliary equations (6.83), (6.86) and (6.88)) and then use (6.91) to compute the scaled cumulant generating function $\lambda(\gamma)$. In the next two sections, we perform this for the two cases we mentioned in section 6.2.2, namely the motion of the defect and the motion of the normal particles. In the former case, we will see that the calculation becomes trivial and we recover the expected result. In the latter case, we will have to perform the calculations in perturbation theory, which we will carry out up to second order.

6.5 Defect particle current statistics

We now turn to the computation of the current statistics of the defect. Actually, from the definition of the dynamical rules, it is clear that the defect simply performs biased diffusion with rates α and x/α, as it does not distinguish between particles and holes. Therefore, even without doing any calculations, it should be obvious that its displacement will follow a Skellam distribution (*i.e.* the difference of two Poissonian random variables) with parameters α and x/α. Thus this calculation is merely a proof of validity of the Bethe ansatz approach.

Setting $a_1 = a_{12} = 1$ and $a_2 = 1$ and recalling the definition of Γ (6.41), we see

that (6.88) becomes

$$\oint \frac{w(T)}{T} = 0. \tag{6.92}$$

From the Bethe equation (6.77), we see that this can only hold if we have identically $w(T) = 0$.

Then (6.84) becomes,

$$\delta = \gamma, \tag{6.93}$$

and consequently,

$$\lambda(\gamma) = (1 - \mathrm{e}^{-\gamma})(\alpha \mathrm{e}^{\gamma} - x/\alpha). \tag{6.94}$$

This is precisely the cumulant generating function of a Skellam distribution (see section 1.4), as expected. The cumulants are given by

$$\left.\frac{\mathrm{d}^n \lambda(\gamma)}{\mathrm{d}\gamma^n}\right|_{\gamma=0} = \begin{cases} \alpha - x/\alpha, & n \text{ odd} \\ \alpha + x/\alpha, & n \text{ even} \end{cases}. \tag{6.95}$$

6.6 Normal particle current statistics

To calculate the current of the normal particles, we set $a_1 = a_{12} = 0$ and $a_2 = 0$. In this case, we must proceed perturbatively. Although nominally, the small parameter is γ, it turns out to be more convenient to expand in terms of the constant C. Note that from (6.59), we have $C = O(\gamma)$, so this choice of perturbative parameter is justified.

For an arbitrary object Ξ, we denote by $\Xi^{(k)}$ the coefficient of C^k in its expansion in powers of C. One important consequence of this change of perturbative parameter is how the relation between λ and γ is expressed. Expanding both, we have,

$$\lambda = C\lambda^{(1)} + C^2\lambda^{(2)} + \ldots \tag{6.96}$$

$$\gamma = C\gamma^{(1)} + C^2\gamma^{(2)} + \ldots. \tag{6.97}$$

Inverting the expansion of γ order by order, we obtain

$$C = \gamma(\gamma^{(1)})^{-1} - \gamma^2 \frac{\gamma^{(2)}}{(\gamma^{(1)})^3} + \dots . \tag{6.98}$$

Then we can express λ as follows,

$$\lambda = \gamma \frac{\lambda^{(1)}}{\gamma^{(1)}} + \gamma^2 \frac{\lambda^{(2)}\gamma^{(1)} - \lambda^{(1)}\gamma^{(2)}}{(\gamma^{(1)})^3} + \dots . \tag{6.99}$$

Then recalling the definitions of the mean current and diffusion constant (6.11a), (6.11b), we have

$$J = \frac{\lambda^{(1)}}{\gamma^{(1)}}, \qquad \Delta = 2\frac{\lambda^{(2)}\gamma^{(1)} - \lambda^{(1)}\gamma^{(2)}}{(\gamma^{(1)})^3}. \tag{6.100}$$

6.6.1 Perturbation theory

We now need to expand $w(T)$, y_0 and B in powers of C. We show these expansions only up to the order that we need to calculate the diffusion constant.

The Bethe equation (6.77) becomes

$$\begin{aligned} w^{(0)}(T) + Cw^{(1)}(T) + C^2 w^{(2)}(T) + \dots = C\tilde{h}^{(0)}(T)\mathrm{e}^{-X[w^{(0)}](T)} \\ +C^2[\tilde{h}^{(1)}(T) - \tilde{h}^{(0)}(T)X[w^{(1)}](T)]\mathrm{e}^{-X[w^{(0)}](T)} + \dots . \end{aligned} \tag{6.101}$$

From this, we read off,

$$w^{(0)}(T) = 0 \tag{6.102a}$$

$$w^{(1)}(T) = \tilde{h}^{(0)}(T) \tag{6.102b}$$

$$w^{(2)}(T) = \tilde{h}^{(1)}(T) - \tilde{h}^{(0)}(T)X[\tilde{h}^{(0)}](T). \tag{6.102c}$$

Note that from (6.42), $\tilde{h}^{(0)}$ is simply

$$\tilde{h}^{(0)}(T) = \frac{(1-T)^{L+1}(B^{(0)}x - T)}{T^M(T - y_0^{(0)})(T/x - y_0^{(0)})}, \tag{6.103}$$

where the values for $y_0^{(0)}$ and $B_0^{(0)}$ are easily found and will be given shortly. On the other hand, $w^{(2)}$ depends on $\tilde{h}^{(1)}$. To fully determine $\tilde{h}^{(1)}$, we need to know $y_0^{(1)}$ and $B^{(1)}$, which must be computed from a perturbative expansion of (6.83)

and (6.86). Evidently this procedure becomes algebraically quite involved, so it is wise to proceed systematically.

Expanding (6.88), we obtain

$$\gamma^{(0)} = 0 \tag{6.104a}$$

$$\gamma^{(1)} = \frac{2}{L}\oint \frac{w^{(1)}(T)}{T} \tag{6.104b}$$

$$\gamma^{(2)} = \frac{2}{L}\oint \frac{w^{(2)}(T)}{T}. \tag{6.104c}$$

Expanding (6.83) and (6.84), we obtain

$$y_0^{(0)} = \frac{\alpha - 1}{\alpha - x} \tag{6.105a}$$

$$y_0^{(1)} = \frac{(1-x)\alpha}{(\alpha - x)^2}\delta^{(1)}, \tag{6.105b}$$

and

$$\delta^{(0)} = 0 \tag{6.106a}$$

$$\delta^{(1)} = 2\oint g(T)\frac{w^{(1)}(T)}{T} \tag{6.106b}$$

$$\delta^{(2)} = 2\oint g(T)\frac{w^{(2)}(T)}{T}, \tag{6.106c}$$

where

$$g(T) = \rho + \frac{T}{1-T}. \tag{6.107}$$

Expanding (6.86) and (6.87), we obtain,

$$B^{(0)} = \frac{\alpha - 1}{\alpha - x}\frac{\alpha^{L+1} - x^M}{\alpha^{L+1} - x^{M+1}} \tag{6.108a}$$

$$B^{(1)} = \frac{B^{(0)}}{y_0^{(0)}}y_0^{(1)} + y_0^{(0)}\frac{(1-x)x^M\alpha^{L+1}}{(\alpha^{L+1} - x^{M+1})^2}\delta_B^{(1)}, \tag{6.108b}$$

and

$$\delta_B^{(0)} = 0 \tag{6.109a}$$

$$\delta_B^{(1)} = -2\oint g_B(T)\frac{w^{(1)}(T)}{T}, \tag{6.109b}$$

where

$$g_B(T) = 1 - \rho + \frac{T}{B^{(0)}x - T}. \tag{6.110}$$

Finally, expanding (6.91), we obtain,

$$\lambda^{(0)} = 0 \tag{6.111a}$$

$$\lambda^{(1)} = (\alpha - x/\alpha)\delta^{(1)} - 2(1-x)\oint \frac{w^{(1)}(T)}{(1-T)^2} \tag{6.111b}$$

$$\lambda^{(2)} = (\alpha - x/\alpha)\delta^{(2)} + \frac{\alpha + x/\alpha}{2}(\delta^{(1)})^2 - 2(1-x)\oint \frac{w^{(2)}(T)}{(1-T)^2}. \tag{6.111c}$$

6.6.2 Current

To calculate the current from (6.100), we need to determine $\lambda^{(1)}$ and $\gamma^{(1)}$. From (6.104b), (6.102b) and (6.65), we have

$$\gamma^{(1)} = \frac{2}{L} Z_{L,M}, \tag{6.112}$$

where $Z_{L,M}$ is the integral

$$Z_{L,M} = \oint \frac{(1-T)^{L+1}(B^{(0)}x - T)}{T^{M+1}(T - y_0^{(0)})(T/x - y_0^{(0)})}. \tag{6.113}$$

Explicit computation of integrals

As $Z_{L,M}$ and other similar integrals will appear consistently throughout our calculations, it is instructive to evaluate it explicitly. This can be done as follows. We first expand the binomial in the numerator, giving,

$$\begin{aligned} Z_{L,M} = & \oint \frac{1}{(T - y_0^{(0)})(T/x - y_0^{(0)})} \\ & \times \sum_{n=0}^{L+1} (-)^n \binom{L+1}{n} (B^{(0)} x T^{n-M-1} - T^{n-M}). \end{aligned} \tag{6.114}$$

Note that only terms with negative powers of T contribute, as the other terms do not have a residue at $T = 0$. Hence we can truncate the sum, giving,

$$\begin{aligned} Z_{L,M} &= \oint \frac{1}{(T - y_0^{(0)})(T/x - y_0^{(0)})} \\ &\quad \times \left[\sum_{n=0}^{M} (-)^n \binom{L+1}{n} B^{(0)} x T^{n-M-1} - \sum_{n=0}^{M-1} (-)^n \binom{L+1}{n} T^{n-M} \right]. \end{aligned} \tag{6.115}$$

It is easiest to evaluate these integrals using the residues of the poles outside the contour, which are the two simple poles at $T = y_0^{(0)}$ and $T = x y_0^{(0)}$ (note that if we had not truncated the sums, we would have had further poles at infinity). This gives,

$$\begin{aligned} Z_{L,M} &= -\frac{x}{(1-x) y_0^{(0)}} \left[\sum_{n=0}^{M} (-)^n \binom{L+1}{n} B^{(0)} x (1 - x^{n-M-1}) (y_0^{(0)})^{n-M-1} \right. \\ &\quad \left. - \sum_{n=0}^{M-1} (-)^n \binom{L+1}{n} (1 - x^{n-M}) (y_0^{(0)})^{n-M} \right]. \end{aligned} \tag{6.116}$$

We can extend the upper limit of the second sum to M, as this spurious term will vanish anyway as $1 - x^{M-M} = 0$. Then we can combine the two sums, giving,

$$\begin{aligned} Z_{L,M} &= -\frac{x}{(1-x) y_0^{(0)}} \left[\sum_{n=0}^{M} (-)^n \binom{L+1}{n} \right. \\ &\quad \left. \times \left(\frac{B^{(0)}}{y_0^{(0)}} (x - x^{n-M}) - (1 - x^{n-M}) \right) (y_0^{(0)})^{n-M} \right]. \end{aligned} \tag{6.117}$$

Using (6.108a), we can simplify this to give the final expression,

$$Z_{L,M} = \frac{x}{y_0^{(0)}} \sum_{n=0}^{M} (-)^n \binom{L+1}{n} \frac{\alpha^{L+1} - x^n}{\alpha^{L+1} - x^{M+1}} (y_0^{(0)})^{n-M}. \tag{6.118}$$

Finite-size expression for current

From (6.111b) and (6.106b), we can express $\lambda^{(1)}$ as,

$$\lambda^{(1)} = 2 \oint \left[(\alpha - x/\alpha) \left(\rho + \frac{T}{1-T} \right) - \frac{(1-x)T}{(1-T)^2} \right] \frac{w^{(1)}(T)}{T}. \tag{6.119}$$

Then from (6.100), the current is given by

$$\frac{J}{L} = \rho(\alpha - x/\alpha) + Z_{L,M}^{-1} \oint f(T) \frac{w^{(1)}(T)}{T}, \tag{6.120}$$

where

$$f(T) = \frac{T}{1-T}\left(\alpha - x/\alpha - \frac{1-x}{1-T}\right). \tag{6.121}$$

Evaluating this integral similarly to the calculation of $Z_{L,M}$, we obtain the following exact expression,

$$\begin{aligned} \frac{J}{L} &= \rho(\alpha - x/\alpha) \\ &\quad + \frac{\sum_{n=0}^{M-1}(-)^n \left[(\alpha - x/\alpha)\binom{L}{n} - (1-x)\binom{L-1}{n}\right](\alpha^{L+1} - x^{n+1})(y_0^{(0)})^{n+1}}{\sum_{n=0}^{M}(-)^n \binom{L+1}{n}(\alpha^{L+1} - x^n)(y_0^{(0)})^n}. \end{aligned} \tag{6.122}$$

This is alternative expression for the current to the one obtained in chapter 4 using the matrix product formalism. There, the exact expression of the current, as given in (4.87), is

$$\begin{aligned} J_{MP} &= \rho(\alpha - x/\alpha) \\ &\quad - [Z_{L,M}^{(MP)}]^{-1} \left[\binom{L-1}{M-1}(\alpha^{L+1} - (x/\alpha)x^M) + (\alpha - 1)xZ_{L-1,M-1}^{(MP)}\right] \end{aligned} \tag{6.123a}$$

$$Z_{L,M}^{(MP)} = \sum_{l=0}^{L}\sum_{m=0}^{l} \binom{L-l}{M-m}\binom{l}{m}\alpha^{L-l}x^m. \tag{6.123b}$$

Importantly, recall (6.12). The two expressions (6.122) and (6.123a) can be checked to produce the same result numerically. In particular, choosing rational numbers for all parameters allows one to perform the comparison exactly, without any machine precision error, in a symbolic programming language such as `Mathematica`. In this regard, one can convince oneself that the two expressions agree. An explicit analytic proof of identical equality remains an open problem. We remark that the expression obtained through the Bethe ansatz is simpler than the one obtained through the matrix product formalism, as it only involves single sums, whereas the latter contains double sums.

Asymptotic behaviour of current

To gain a physical understanding of the result (6.122), it is helpful to extract its asymptotic behaviour in the limit $L \to \infty$, with ρ held fixed. Consider the expression (6.120). Evidently the key to understanding the asymptotic behaviour is the object $Z_{L,M}$. We can rewrite (6.113) as

$$Z_{L,M} = \oint A(T) \mathrm{e}^{L\phi(T)}, \tag{6.124}$$

where

$$A(T) = \frac{(1-T)(B^{(0)}x - T)}{T(T - y_0^{(0)})(T/x - y_0^{(0)})} \tag{6.125}$$

$$\phi(T) = \log(1-T) - \rho \log(T). \tag{6.126}$$

In this form, it is clear that the integral $Z_{L,M}$ has a saddle point at T_*, defined by $\phi'(T_*) = 0$. Solving this equation, we find,

$$T_* = -\frac{\rho}{1-\rho}. \tag{6.127}$$

However, $Z_{L,M}$ is not necessarily dominated by the saddle point, as $A(T)$ also has the two poles at $T = y_0^{(0)}$ and $T = xy_0^{(0)}$. We can proceed as follows.

The original contour is a small circle around the origin. Regardless of where the poles are located in relation to the saddle point, we deform the contour to pass through the saddle point. Then if the poles are located between the origin and the saddle point, we subtract their residues. This contribution will actually generally dominate over the saddle point. If the poles are not located between the origin and the saddle point, they do not come into play and the integral is simply dominated by the saddle point.

From (6.105a) we can determine the ranges of values in which the poles are located inside or outside the contour. We find that there are two critical densities,

$$\rho_1 = \frac{1-\alpha}{1-x}, \qquad \rho_2 = \frac{1-\alpha^{-1}}{1-x^{-1}}. \tag{6.128}$$

The location of the poles is as follows,

$$\begin{cases} \rho > \rho_1, & \text{both poles outside} \\ \rho_2 < \rho < \rho_1, & xy_0^{(0)} \text{ inside, } y_0^{(0)} \text{ outside} \\ \rho < \rho_2, & \text{both poles inside} \end{cases}. \tag{6.129}$$

It is important to observe that the sum of the residues of the two poles, which is given by

$$\frac{x(1-y_0^{(0)})^{L+1}(B^{(0)}x - y_0^{(0)})}{(1-x)y_0^{(0)}(y_0^{(0)})^{M+1}} - \frac{x^2(1-xy_0^{(0)})^{L+1}(B^{(0)} - y_0^{(0)})}{(1-x)y_0^{(0)}(xy_0^{(0)})^{M+1}}, \tag{6.130}$$

vanishes exactly (this can be verified using (6.105a) and (6.108a)). Thus when both poles are inside the contour, the integral is actually also dominated by the saddle point. It can be verified that this precise cancellation of the residues also occurs in the integral $\oint(\mathrm{d}T/(2\pi\mathrm{i}))f(T)\tilde{h}^{(0)}(T)/T$, which occurs in (6.120).

Therefore overall, we have that when both poles are inside or outside the contour, the integral is dominated by the saddle point. When the pole $xy_0^{(0)}$ is inside the contour and $y_0^{(0)}$ is outside, then the integral is dominated by the pole at $xy_0^{(0)}$. Note that this matches exactly with the phase diagram we derived in chapter 4. The localised phases correspond to the integral being dominated by the saddle point and the shock phase corresponds to the integral being dominated by the pole at $xy_0^{(0)}$.

Using this, we can deduce the asymptotic behaviour of J from (6.120), which is given by,

$$\frac{J}{L} \approx \begin{cases} \rho(\alpha - x/\alpha) + f(T_*), & \rho_2 < \rho < \rho_1 \\ \rho(\alpha - x/\alpha) + f(xy_0^{(0)}), & \text{else} \end{cases}, \tag{6.131}$$

with corrections of order $O(L^{-1})$. Evaluating these expressions, we obtain

$$\frac{J}{L} \approx \begin{cases} (1-x)\rho(1-\rho), & \rho_2 < \rho < \rho_1 \\ \rho(\alpha - x/\alpha) + \frac{x(\alpha-1)^2}{\alpha(\alpha-x)^2}, & \text{else} \end{cases}, \tag{6.132}$$

in agreement with the result from chapter 4.

6.6.3 Diffusion constant

Finite-size expression for diffusion constant

We now proceed to the calculation of the diffusion constant. Using (6.111c) and (6.106c), $\lambda^{(2)}$ becomes,

$$\lambda^{(2)} = \frac{\alpha + x/\alpha}{2}(\delta^{(1)})^2 + 2\oint[(\alpha - x/\alpha)\rho + f(T)]\frac{w^{(2)}(T)}{T}. \tag{6.133}$$

Then also using (6.104c), we have

$$\begin{aligned}\lambda^{(2)}\gamma^{(1)} - \lambda^{(1)}\gamma^{(2)} &= L^{-1}(\alpha + x/\alpha)(\delta^{(1)})^2 \oint \frac{w^{(1)}(T)}{T} \\ &\quad + \frac{4}{L}\oint[(\alpha - x/\alpha)\rho + f(T)]\frac{w^{(2)}(T)}{T}\oint \frac{w^{(1)}(T)}{T} \\ &\quad - \frac{4}{L}\oint[(\alpha - x/\alpha)\rho + f(T)]\frac{w^{(1)}(T)}{T}\oint \frac{w^{(2)}(T)}{T}\end{aligned} \tag{6.134}$$

The terms proportional to $(\alpha - x/\alpha)\rho$ will cancel. Then we obtain for the diffusion constant from (6.100),

$$\begin{aligned}\frac{\Delta}{L^2} &= Z_{L,M}^{-3}\left[\oint f(T)\frac{w^{(2)}(T)}{T}\oint \frac{w^{(1)}(T)}{T} - \oint f(T)\frac{w^{(1)}(T)}{T}\oint \frac{w^{(2)}(T)}{T}\right] \\ &\quad + Z_{L,M}^{-2}(\alpha + x\alpha)\left(\oint g(T)\frac{w^{(1)}(T)}{T}\right)^2.\end{aligned} \tag{6.135}$$

Now using (6.102c), we can write this out $w^{(2)}$ a bit more explicitly as,

$$\begin{aligned}w^{(2)}(T) &= \tilde{h}^{(0)}(T)\left[\frac{B^{(1)}x}{B^{(0)}x - T} + y_0^{(1)}\left(\frac{1}{T - y_0^{(0)}} + \frac{1}{T/x - y_0^{(0)}}\right)\right] \\ &\quad - [\tilde{h}^{(0)}(T)]^2 + 2\tilde{h}^{(0)}(T)K[\tilde{h}^{(0)}](T)\end{aligned} \tag{6.136}$$

where

$$K[u](T) = \oint' \frac{u(T')}{T'}\sum_{k=-\infty}^{\infty}\frac{x^{|k|}}{1 - x^{|k|}}\left(\frac{T}{T'}\right)^k. \tag{6.137}$$

Then using (6.108b), and (6.105b), we can write this as

$$w^{(2)}(T) = \tilde{h}^{(0)}(T)\left[\frac{y_0^{(0)}}{B^{(0)}x-T}\frac{(1-x)x^{M+1}\alpha^{L+1}}{(\alpha^{L+1}-x^{M+1})^2}\delta_B^{(1)} + f_2(T)\frac{(1-x)\alpha}{(\alpha-x)^2}\delta^{(1)}\right] - [\tilde{h}^{(0)}(T)]^2 + 2\tilde{h}^{(0)}(T)K[\tilde{h}^{(0)}](T), \tag{6.138}$$

where

$$f_2(T) = \frac{1}{T-y_0^{(0)}} + \frac{1}{T/x-y_0^{(0)}} + \frac{B^{(0)}x/y_0^{(0)}}{B^{(0)}x-T}. \tag{6.139}$$

Plugging this into (6.135) and using (6.106b) and (6.109b), the resulting expression for Δ is very big, so we split it into parts

$$\frac{\Delta}{L^2} = I_1 + I_2 + I_3 + I_4 + I_5, \tag{6.140}$$

where

$$I_1 = -Z_{L,M}^{-3}\frac{2(1-x)x^{M+1}\alpha^{L+1}y_0^{(0)}}{(\alpha^{L+1}-x^{M+1})^2}\oint g_B\frac{\tilde{h}^{(0)}}{T}\left[\oint \frac{f}{B^{(0)}x-T}\frac{\tilde{h}^{(0)}}{T}\oint\frac{\tilde{h}^{(0)}}{T} - \oint f\frac{\tilde{h}^{(0)}}{T}\oint\frac{1}{B^{(0)}x-T}\frac{\tilde{h}^{(0)}}{T}\right] \tag{6.141a}$$

$$I_2 = Z_{L,M}^{-3}\frac{2(1-x)\alpha}{(\alpha-x)^2}\oint g\frac{\tilde{h}^{(0)}}{T}\left[\oint f f_2\frac{\tilde{h}^{(0)}}{T}\oint\frac{\tilde{h}^{(0)}}{T} - \oint f\frac{\tilde{h}^{(0)}}{T}\oint f_2\frac{\tilde{h}^{(0)}}{T}\right] \tag{6.141b}$$

$$I_3 = -Z_{L,M}^{-3}\left[\oint f\frac{[\tilde{h}^{(0)}]^2}{T}\oint\frac{\tilde{h}^{(0)}}{T} - \oint f\frac{\tilde{h}^{(0)}}{T}\oint\frac{[\tilde{h}^{(0)}]^2}{T}\right] \tag{6.141c}$$

$$I_4 = -Z_{L,M}^{-3}2\left[\oint f\frac{\tilde{h}^{(0)}K[\tilde{h}^{(0)}]}{T}\oint\frac{\tilde{h}^{(0)}}{T} - \oint f\frac{\tilde{h}^{(0)}}{T}\oint\frac{\tilde{h}^{(0)}K[\tilde{h}^{(0)}]}{T}\right] \tag{6.141d}$$

$$I_5 = Z_{L,M}^{-2}(\alpha+x/\alpha)\left(\oint g\frac{\tilde{h}^{(0)}}{T}\right)^2, \tag{6.141e}$$

where f is defined in (6.121), g is defined in (6.107), g_B is defined in (6.110), f_2 is defined in (6.139), $\tilde{h}^{(0)}$ is given in (6.103), K is defined in (6.137), $y_0^{(0)}$ is given in (6.105a), $B^{(0)}$ is given in (6.108a), $Z_{L,M}$ is given in (6.118), and we have suppressed the arguments (T) of all functions for compactness. This expression for the diffusion constant is plotted numerically in figure 6.1.

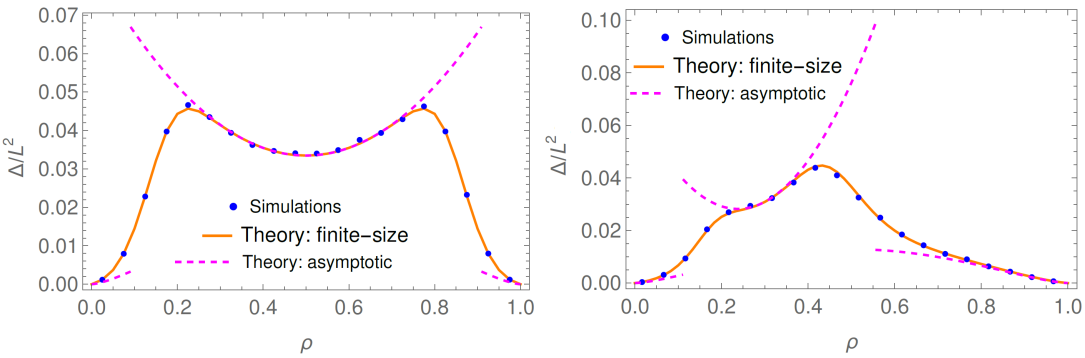

Figure 6.1 *Diffusion constant against density from theory (*(6.140) *for finite-size values, and* (6.147) *and* (6.151) *for asymptotic values) and as obtained from Monte Carlo simulations. The parameters used were* $x = 0.01$, $\alpha = 0.1$, $L = 40$ *(left), and* $x = 0.1$, $\alpha = 0.5$, $L = 60$ *(right). The asymptotic analysis predicts phase transitions at* $\rho \approx 0.1$, $\rho \approx 0.9$ *(left) and* $\rho \approx 0.1$, $\rho \approx 0.5$ *(right). Excellent agreement is seen between simulations and finite-size theoretical predictions. There is good agreement between finite-size and asymptotic values deep in each phase, but there are noticeable deviations near the phase transitions for the given system sizes.*

Although this expression for the diffusion constant is quite cumbersome, it has two positive aspects. Firstly, it is completely explicit. One just needs to implement all the integrals numerically on a computer and one can evaluate this exactly. Secondly, we will see that only few of the parts I_i contribute to the asymptotic behaviour.

Asymptotic analysis

We now turn to the asymptotic behaviour of the diffusion constant Δ, as given by (6.140). We will rely on similar analysis of integrals as we considered in section 6.6.2. The integrals still have a unique saddle point at $T = -\rho/(1-\rho)$ and poles at $T = y_0^{(0)}$ and $T = xy_0^{(0)}$. However, the justification for picking the relevant terms in each phase is more complicated than in the asymptotic analysis of $Z_{L,M}$. We deduce the terms that contribute at leading order in each phase guided by extensive numerical analysis, Monte Carlo simulations, as well as comparison to known results in similar systems (for instance the TASEP with a defect particle). All these lines of reasoning suggest the scaling $\Delta \sim L^{3/2}$ in the localised phases and $\Delta \sim L^2$ in the shock phase.

The analysis is complicated by the fact that in some integrals that occur in the full expression for Δ, such as $\oint g\tilde{h}^{(0)}/T$, the residues of the two poles do not

cancel exactly. This seems to suggest that these terms have super-dominant contributions in the phase when both poles are inside the contour. Actually, these spurious terms cancel exactly with contributions from other terms in the full expression for Δ. However, to see the full cancellation, all terms must be taken into account. Although we do not have a rigorous analytic proof of the mechanism of this cancellation, our argument is supported by extensive numerical verification as well as agreement of the final results with Monte Carlo simulations.

The conclusion of this analysis is that the terms I_1 and I_4 do not contribute in any phase, the terms I_2 and I_5 contribute only in the shock phase, and the term I_3 contributes in both phases.

Asymptotic behaviour in localised phases

In the localised phase, only term I_3 contributes to the asymptotic behaviour and it is to be evaluated using the saddle point approximation. This means that we can write,

$$\frac{\Delta_{\mathcal{S}}}{L^2} \approx -Z_{L,M}^{-3}\left[\oint f\frac{[\tilde{h}^{(0)}]^2}{T}\oint\frac{\tilde{h}^{(0)}}{T}-\oint f\frac{\tilde{h}^{(0)}}{T}\oint\frac{[\tilde{h}^{(0)}]^2}{T}\right]. \tag{6.142}$$

Then using (6.124), we can write this as,

$$\frac{\Delta_{\mathcal{S}}}{L^2} \approx -\left(\oint A\mathrm{e}^{L\phi}\right)^{-3}\left[\oint fTA^2\mathrm{e}^{2L\phi}\oint A\mathrm{e}^{L\phi}-\oint fA\mathrm{e}^{L\phi}\oint TA^2\mathrm{e}^{2L\phi}\right] \tag{6.143}$$

The first order saddle point approximations of the integrals in the numerator cancel, so we need to consider the first correction (see 2.A). We give the general formula using $Z_{L,M}$ for illustration (where it is important to recall the suppressed factor of $(2\pi)^{-1}$),

$$\begin{aligned}\oint A\mathrm{e}^{L\phi} \approx{}& (-2\pi L\phi'')^{-1/2}A\mathrm{e}^{L\phi}\left[1+\frac{1}{L}\left(-\frac{A''}{2A\phi''}+\frac{\phi''''}{8(\phi'')^2}\right.\right.\\ &\left.\left.+\frac{A'\phi'''}{2A(\phi'')^2}-\frac{5(\phi''')^2}{24(\phi'')^3}\right)+O(L^{-2})\right],\end{aligned} \tag{6.144}$$

where all expressions are to be evaluated at the saddle point. Now examining the form of the integrals in (6.143), one eventually sees that the terms coming from the correction that only involve derivatives of ϕ will cancel exactly. Hence

we only have remaining,

$$\frac{\Delta_{\mathcal{S}}}{L^2} \approx -\left(-\frac{\pi\phi''}{L}\right)^{1/2}\frac{fT}{4\phi''}\left[-\frac{(fTA^2)''}{fTA^2}+\frac{(fTA^2)'\phi'''}{ftA^2\phi''}-2\frac{A''}{A}+2\frac{A'\phi'''}{A\phi''}\right.$$
$$\left.+2\frac{(fA)''}{fA}-2\frac{(fA)'\phi'''}{2fA\phi''}+\frac{(TA^2)''}{TA^2}-\frac{(TA^2)'\phi'''}{TA^2\phi''}\right]. \tag{6.145}$$

Simplifying this, we obtain,

$$\frac{\Delta_{\mathcal{S}}}{L^2} \approx \left(-\frac{\pi\phi''}{L}\right)^{1/2}\frac{T}{4\phi''}\left(f''-2\frac{f'}{T}-f'\frac{\phi'''}{\phi''}\right). \tag{6.146}$$

Now using (6.121), (6.126) and (6.127) yields the expression,

$$\Delta_{\mathcal{S}} \approx L^{3/2}\frac{\sqrt{\pi}}{2}(1-x)[\rho(1-\rho)]^{3/2}. \tag{6.147}$$

This is shown in figure 6.1.

Asymptotic behaviour in shock phase

In the shock phase, the terms I_2, I_3 and I_5 all contribute to the leading order asymptotic expression and they are to be evaluated using the residue at the pole at $T = xy_0^{(0)}$. In these calculations, it is helpful to define

$$\bar{h}(T) = (T - xy_0^{(0)})\tilde{h}(T), \qquad \bar{f}_2(T) = (T - xy_0^{(0)})f_2(T). \tag{6.148}$$

It is also important to remember that the poles are initially always outside the contour, so the residues should be multiplied by -1. This did not matter in the first order calculations, as there the minus signs always cancelled between the numerator and denominator. Here, however, this introduces an extra minus sign for I_3 as there are three integrals in the denominator but only two in the

numerator. Then we have,

$$Z_{L,M} \approx \frac{\bar{h}(xy_0^{(0)})}{xy_0^{(0)}} \tag{6.149a}$$

$$I_2 \approx Z_{L,M}^{-3} \frac{2(1-x)\alpha}{(\alpha-x)^2} \frac{g\bar{h}}{T} \left[\left(\frac{f\bar{f}_2\bar{h}}{T} \right)' \frac{\bar{h}}{T} - f\frac{\bar{h}}{T} \left(\frac{\bar{f}_2\bar{h}}{T} \right)' \right]_{T=xy_0^{(0)}} \tag{6.149b}$$

$$I_3 \approx Z_{L,M}^{-3} \left[\left(\frac{f\bar{h}^2}{T} \right)' \frac{\bar{h}}{T} - \frac{f\bar{h}}{T} \left(\frac{\bar{h}^2}{T} \right)' \right]_{T=xy_0^{(0)}} \tag{6.149c}$$

$$I_5 \approx Z_{L,M}^{-2}(\alpha + x/\alpha) \left(\frac{g\bar{h}}{T} \right)^2_{T=xy_0^{(0)}}. \tag{6.149d}$$

Simplifying this, we obtain

$$\frac{\Delta_{\mathcal{S}}}{L^2} \approx \left[\frac{2(1-x)\alpha}{(\alpha-x)^2} g\bar{f}_2 f' + Tf' + (\alpha + x/\alpha)g^2 \right]_{T=xy_0^{(0)}}. \tag{6.150}$$

Now using (6.121), (6.107) and (6.139), we can write this as

$$\Delta_{\mathcal{S}} \approx L^2(\rho_1 - \rho_2)^2[\alpha(1-w_*)^2 + (x/\alpha)w_*^2], \tag{6.151}$$

where w_* is the location of the shock front, defined by (4.62). This is shown in figure 6.1.

Convergence to asymptotic results

We comment briefly on the convergence of the finite-size expression for Δ, (6.140), to the asymptotic expressions, (6.147) and (6.151). As shown in figure 6.1, the agreement between finite-size and asymptotic expressions is generally good deep in each phase, but there are some deviations near the phase transitions. This is expected, as finite-size effects are typically strongest near phase transitions. Moreover, in the present case this is exacerbated by the fact that there is a discontinuity at the phase transition due to the different scaling regimes ($\Delta \sim L^{3/2}$ in the localised phases but $\Delta \sim L^2$ in the shock phase). However, on increasing the system size, it can be seen that the finite-size values converge to the asymptotic ones. This is shown in figure 6.2.

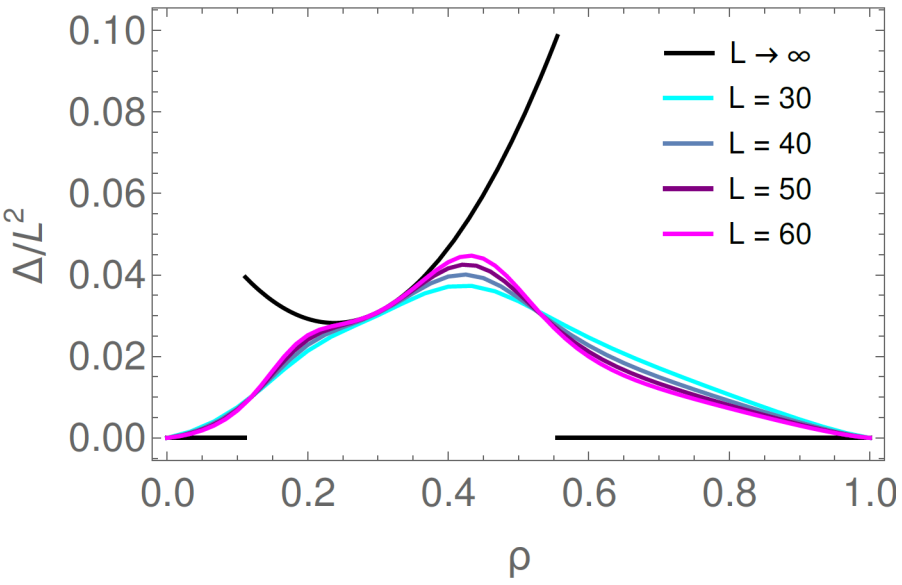

Figure 6.2 *Diffusion constant against density for various system sizes. Parameters used were $x = 0.1$, $\alpha = 0.5$. Asymptotic analysis predicts phase transitions at $\rho \approx 0.1$ and $\rho \approx 0.5$. Convergence to asymptotic results can be seen for increasing L.*

6.6.4 Physical interpretation of results

Having derived the asymptotic behaviour of the current fluctuations, (6.147) and (6.151), we now aim to understand these expressions using simple physical intuition.

The expression in the localised phases (6.147) is actually the same as one would obtain for a regular PASEP (without a defect). This continues the trend that we saw with the current: in the localised phases, the defect has only a microscopic effect on the system, so it does not affect the large scale behaviour. Thus higher order cumulants in the localised phases are expected to follow regular PASEP statistics.

The expression in the shock phase (6.151) looks a bit more involved, but it can be understood intuitively together with the expression for the current, (6.132). In particular, we note that we may write the current in the shock phase in the form

$$J_{\mathcal{S}} \approx L(\rho_1 - \rho_2)[\alpha(1 - w_*) + (x/\alpha)w_*]. \tag{6.152}$$

These expressions suggest the following interpretation. Recall that the mean density profile in the shock phase comprises a low density region with density

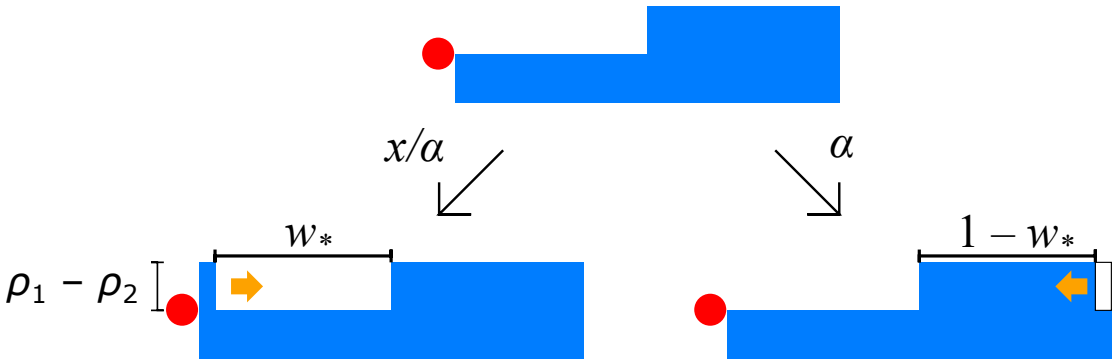

Figure 6.3 *Schematic illustration of how fluctuations are generated in the shock phase. The defect is shown in red and the density field of normal particles is shown in blue.*

ρ_2 over the region of length w_* in front of the defect, and a high density region with density ρ_1 over a region of length $(1 - w_*)$ behind the defect. Now at a mean-field level, we may treat the density field as continuous. Then when the defect hops forward, it creates a "hole" of density ρ_2 behind itself [1]. This "hole" must travel backwards a distance $L(1 - w_*)$ to restore the stationary profile. Similarly, when the defect hops backwards, it creates a "particle" of density ρ_1 in front of itself, which must travel forward a distance Lw_* to restore the stationary profile. Thus we may see the current of normal particles as a combination of two independent processes, which create currents of size $L(\rho_1 - \rho_2)(1 - w_*)$ with rate α and $L(\rho_1 - \rho_2)w_*$ with rate x/α. This is illustrated in figure 6.3.

Then using the additive property for cumulants of independent random variables, this suggests the following form for the cumulants in the shock phase,

$$\lim_{t\to\infty} \frac{\langle Y_2^n \rangle_c}{t} \approx L^n(\rho_1 - \rho_2)^n[\alpha(1 - w_*)^n + (x/\alpha)w_*^n]. \tag{6.153}$$

As we have shown, the expressions for the first two cumulants, (6.152) and (6.151), agree with this. It would be of interest to investigate whether this holds for higher order statistics, though this remains an open problem.

6.7 Summary

In this chapter, we used a coordinate Bethe ansatz to study the long time limit of current fluctuations in a partially asymmetric exclusion process on a ring

[1] We call this a hole as it is a region of lower density ρ_2 in an environment of higher density ρ_1.

with a defect particle. Using a functional reformulation of the Bethe equations, we calculated the mean and variance of the current exactly both for finite-size systems and in the asymptotic limit $L \to \infty$ with mean density ρ held fixed. This complements the calculation of the mean current and steady-state density profile in chapter 4, which was done using a matrix product ansatz.

It is interesting to observe that the diffusion constant exhibited different dependence on system size in the localised and shock phases, scaling as $\Delta \sim L^{3/2}$ in the localised phases and $\Delta \sim L^2$ in the shock phase. This is in contrast to the current, which was continuous across the phase transition (with a discontinuity in the first derivative). The two different scaling regimes can be understood through the physical picture explained in section 6.6.4. Curiously, if the conjecture for the asymptotic behaviour presented in section 6.6.4 is true, the third cumulant would again scale in the same way ($\sim L^3$) in both phases, although it would be negative in the localised phases and positive in the shock phase.

It would be a good result to extend the calculations presented in this chapter to arbitrary order in perturbation theory, especially as this would allow one to calculate the full large deviation function of particle displacement. Calculations to arbitrary orders have been performed for the one-species TASEP [10], the TASEP with a defect particle [12] and the one-species PASEP [164]. However, for the model considered in this chapter, this endeavour would be quite challenging. Even the calculations for the one-species PASEP to arbitrary order involved some quite complicated combinatorial objects. Moreover, the second-order calculations done in this chapter were already significantly more complex than those for the one-species PASEP. This makes the extension of the calculations to arbitrary order a very formidable task, unless a simpler alternative approach is found.

We emphasise that this is the first time that the Bethe ansatz was used to calculate the particle displacement statistics for a partially asymmetric system in which particles move with different rates. As such, it required us to build on the techniques that have been developed for the two-species TASEP [12, 118] and the one-species PASEP [119, 164].

Chapter 7

Conclusion

7.1 Summary of results

In this thesis, we presented novel results on four research projects involving simple exclusion processes on a ring. We give a brief summary of the results in each.

7.1.1 Driven tracer in a passive lattice gas

In chapter 3, we studied a symmetric process with a single totally asymmetric defect particle. We obtained the exact steady-state measure using the mapping to the zero-range process and the matrix product ansatz. This allowed us to compute the density profile and mean current. In the large system size limit, it was found that these quantities were predicted correctly by mean-field theory. However, the exact approaches also allowed us to produce finite-size expressions, which the simple continuum mean-field approach did not.

Furthermore, we extended the calculations of steady-state measures to the case of multiple totally asymmetric defect particles and a single partially asymmetric defect particle. We found that with multiple totally asymmetric defects, the system essentially behaves as a sequence of independent subsystems with one defect. The density of each subsystem is set such that the mean drift velocities of all subsystems are equal. With a single partially asymmetric defect, the current and density profile were modified by an additional parameter given by the ratio of the defect's forward and backward hopping rates.

The phenomenology of the system studied in this chapter was quite simple, which makes it a good candidate to understand the minimal ingredients of a nonequilibrium steady state. One can see that the steady state is of the nonequilibrium type by the presence of a nonzero current, which breaks detailed balance. At the microscopic level, the particles possess a net drift velocity with the scaling $\langle v \rangle \sim L^{-1}$, which is a characteristic of the diffusive time scale. One also observes that the steady-state density profile is non-uniform on the scale of the system size, which is another consequence of the driving by the defect. This is significant as it shows that a nonequilbrium steady state can be generated even if the system is driven by a single particle.

7.1.2 Driven defect particle with priority in a driven lattice gas

In chapter 4, we studied a partially asymmetric process with a defect particle that could overtake normal particles with its usual hopping rates. For a special choice of the defect hopping rates, the steady state could be solved exactly using an unusual matrix product ansatz. This revealed a phase diagram with three distinct types of density profiles: left-localised, right-localised and shock. In the localised profiles, the defect only disrupted the normal particle density locally. In the shock phase, on the other hand, the defect throttled the normal particle current and created two regions with different bulk densities.

The exact solution showed that the mean current was continuous, but its first derivative was discontinuous at the points of phase transition. This supports the view that currents in nonequilibrium systems may play a similar role to free energies in equilibrium systems. Indeed, the phase transitions in this system may be compared to the first-order liquid-gas transition, with the left- and right-localised phases representing low and high density phases and the shock phase being a region of phase coexistence. At the same time, in the exclusion process, we also saw diverging length scales at the phase transitions.

Unlike the system studied in chapter 3, in this chapter we saw mean-field theory giving quantitatively incorrect predictions. In particular, in the shock phase, mean-field theory predicted an incorrect scaling for the width of the shock profile.

The matrix product solution presented in this chapter had some very unusual features. The matrices had dimension 2×2, whereas typically matrix product

solutions involve infinite-dimensional matrices. Despite the fact that the matrices were only 2×2, they were able to capture a phase transition in the system. Moreover, they did not satisfy typical tilted algebraic relations with scalar auxiliaries. This meant that this matrix product solution fell outside the standard classification of two-species models on a ring solvable by a matrix product ansatz.

7.1.3 Integrability of two-species driven lattice gases

In chapter 5, we studied a broad class of two-species exclusion processes with overtaking. Starting from a generalised, nested coordinate Bethe ansatz, we used the Yang-Baxter formalism to determine which models in this class could be solved with this ansatz. Three classes of solvable models were thus found: *(i)* a two-species TASEP, *(ii)* a PASEP with a "second-class" species and *(iii)* a PASEP with a single generalised "first-class" particle.

Of these three classes, *(i)* and *(ii)* had received much attention in the literature, but *(iii)* had received comparatively little. In fact, this was the same model that we studied in chapter 4. The fact that it had not received much attention in literature related to matrix product states can be explained by the fact that the matrix product solution has an unusual form and is not derived from typical scalar reduction relations. From the point of view of Bethe ansatz integrability, this model is also unusual as it satisfies the Yang-Baxter equations only if exactly one particle of species 1 is present.

7.1.4 Current fluctuations of a driven lattice gas with a driven defect particle with priority

In chapter 6, we studied the long-time current fluctuations of the model defined in chapter 4. Using a coordinate Bethe ansatz, the large deviation function of particle displacement was expressed implicitly in terms of the solution of the Bethe equations. These were solved up to second order in perturbation theory using a functional approach, yielding the mean current and diffusion constant.

The expression for the mean current was found to agree numerically with the one obtained using the matrix product approach. The diffusion constant was shown to obey different scaling in the localised and shock phases. The different behaviours was explained by an intuitive physical picture. In the localised

phases, the system behaved essentially as a pure PASEP. In the shock phase, the current was controlled by hops of the defect, which creates density wavelets. The corresponding fluctuations therefore scale with the system size.

7.2 Discussion

We began this thesis with a general consideration of exactly solvable models of nonequilibrium systems. It is now worth examining the implications that the results we have presented have in this regard.

One recurring theme in this thesis has been the limits of applicability of exact methods. It is universally understood that integrability is a fragile property and generally requires fine-tuning of parameters. Indeed, we saw in chapter 5 that among two-species partially asymmetric processes, there is a very limited set of models that are solvable via the Bethe ansatz. In fact, two out of the three cases had already been solved in the literature, even before any classification had been carried out. Similarly, we saw in chapter 4 that the exact solution through the matrix ansatz was only valid provided the model parameters satisfied a specific condition and there was only one defect particle.

However, we also saw evidence against an overly cautious perspective on the limits of applicability of exact methods. The matrix product solution presented in 4 did not fit into the usual classification scheme. Also, the Bethe ansatz solution presented in 6 only satisfied the Yang-Baxter equations in a specific block, meaning that a very strict classification might categorise it as non-integrable.

These observations point to the symptom that general surveys of exactly solvable systems are always limited by the specific solutions known at the time. For instance, the classification of models solvable by a matrix product ansatz assumes a deformed matrix algebra with scalar auxiliaries. Also, the analysis of integrable systems in chapter 5 relied on a particular form of the coordinate Bethe ansatz. In an algebraic Bethe ansatz approach, one might instead postulate a form for the R-matrices and then determine which models they describe. This still requires one to know what a "reasonable" form for an R-matrix is based on known examples, which may not give an exhaustive list of solvable models.

This feature of exactly solvable systems leaves room for progress if a new way to simplify the problem is found. For instance, the matrix product solution in

chapter 4 was simple enough to perform explicit calculations even in the absence of the traditional scalar matrix reduction mechanism. This suggests an alternative approach to matrix product states and may open an avenue for a new class of matrix product measures.

It should also be noted that exactly solvable models are important beyond the specific fine-tuned parameter regimes in which they are applicable. For instance, the one-species TASEP and PASEP are both exactly solvable and the results obtained in studying them help to build intuition for more complex cases, which may themselves not be exactly solvable. In this thesis we saw, for instance, in chapter 4 that the model was exactly solvable for a special choice of parameters. However, in the mean-field analysis, that choice of parameters did not correspond to a singular limit. It is therefore reasonable to believe that the results obtained for that case are also relevant for the general case.

Another major theme of this thesis has been the effect of defects on lattice gases. In chapter 3, nonequilibrium behaviour was achieved by adding a single driven particle to a bath of passive particles. In chapter 4, a defect in a partially asymmetric system created a rich phase diagram.

It is interesting to consider this from the point of view of system control using defects. In the model in chapter 3, although the totally asymmetric defect had an effect on the steady state on a macroscopic level, the density profile and mean current did not depend on the hopping rate of the defect. On the other hand, with a partially asymmetric defect, the mean current and density profile could be controlled through the ratio of the defect's forward and backward hopping rates.

In the system considered in chapter 4, the defect did not have macroscopic effects in the localised phases, but affected the system on a global level in the shock phase. For this to be achieved, the parameters had to be tuned to the shock phase. Interestingly, in the shock phase, the bulk densities of the density profile depended only on microscopic parameters, namely the hopping rates of the defect and the normal particles, whereas the global mean density only entered via the location of the shock front.

We remark that the fact that a defect can have a macroscopic effect on a system is a characteristic feature of nonequilibrium systems. For equilibrium systems, one would expect a local perturbation to have only localised effects, unless the system is critical.

7.3 Outlook

To build on the main results of this thesis, it would be of interest to explore further the integrability of multi-species exclusion processes. One avenue to explore is to consider processes with three or more species. Certain special cases are known to be integrable, such as the partially asymmetric multi-species process in which all particles have equal hopping rates [71]. It remains to be seen whether general criteria for integrability may be formulated. Also, it would be interesting to investigate the integrability of processes with two or more species with open boundary conditions. This would require the algebraic Bethe ansatz approach, as the coordinate ansatz only works for processes with a fixed particle number.

The other direction in which the results of this thesis could be generalised is to consider different choices of parameters for defects in symmetric and partially asymmetric processes. For instance, in symmetric processes, defects with overtaking have recently been considered [146], but such models have yet only been investigated with mean field theory. Exact solutions are still lacking.

We also mention that with the recent progress in exact solutions for macroscopic fluctuation theory [79–81], it would be interesting to see whether it could be used to gain a deeper understanding of the models considered in this thesis.

Bibliography

[1] Lobaskin I and Evans M R 2020 *Journal of Statistical Mechanics: Theory and Experiment* **2020** 053202

[2] Lobaskin I, Evans M R and Mallick K 2022 *Journal of Physics A: Mathematical and Theoretical* **55** 205002

[3] Lobaskin I, Evans M R and Mallick K 2023 *Journal of Physics A: Mathematical and Theoretical* **56** 165003

[4] Lobaskin I, Evans M R and Mallick K 2023 Current fluctuations in a partially asymmetric simple exclusion process with a defect particle (*Preprint* 2307.06770)

[5] Ben-Naim A 2019 *Entropy* **21** 1170

[6] Zhang X J, Qian H and Qian M 2012 *Physics Reports* **510** 1–86

[7] Chou T, Mallick K and Zia R K 2011 *Reports on progress in physics* **74** 116601

[8] Blythe R A and Evans M R 2007 *Journal of Physics A: Mathematical and Theoretical* **40** R333

[9] Bethe H 1931 *Zeitschrift für Physik* **71** 205–226

[10] Derrida B and Lebowitz J L 1998 *Physical review letters* **80** 209

[11] Mallick K 1996 *Journal of Physics A: Mathematical and General* **29** 5375

[12] Derrida B and Evans M R 1999 *Journal of Physics A: Mathematical and General* **32** 4833

[13] Kardar M 2007 *Statistical physics of particles* (Cambridge University Press)

[14] Huang K 2008 *Statistical mechanics* (John Wiley & Sons)

[15] Landau L D and Lifshitz E M 2013 *Statistical Physics: Volume 5* vol 5 (Elsevier)

[16] Rigol M, Dunjko V, Yurovsky V and Olshanii M 2007 *Physical review letters* **98** 050405

[17] Vidmar L and Rigol M 2016 *Journal of Statistical Mechanics: Theory and Experiment* **2016** 064007

[18] Gillespie D T 1991 *Markov processes: an introduction for physical scientists* (Elsevier)

[19] Van Kampen N G 1992 *Stochastic processes in physics and chemistry* vol 1 (Elsevier)

[20] Ibe O 2013 *Markov processes for stochastic modeling* (Newnes)

[21] Crick F H 1958 On protein synthesis *Symp Soc Exp Biol* 138-63 p 8

[22] Jülicher F and Bruinsma R 1998 *Biophysical journal* **74** 1169–1185

[23] Leibler S and Huse D A 1993 *The Journal of cell biology* **121** 1357–1368

[24] Prost J, Chauwin J F, Peliti L and Ajdari A 1994 *Physical Review Letters* **72** 2652

[25] Peskin C, Ermentrout G and Oster G 1994 The correlation ratchet: a novel mechanism for generating directed motion by atp hydrolysis *Cell Mechanics and Cellular Engineering* (Springer) pp 479–489

[26] Derényi I and Vicsek T 1996 *Proceedings of the National Academy of Sciences* **93** 6775–6779

[27] Jülicher F, Ajdari A and Prost J 1997 *Reviews of Modern Physics* **69** 1269

[28] Chakraborti S, Dhar A, Goldstein S, Kundu A and Lebowitz J L 2022 *Journal of Physics A: Mathematical and Theoretical* **55** 394002

[29] Buchete N V and Hummer G 2008 *The Journal of Physical Chemistry B* **112** 6057–6069

[30] Täuber U C 2014 *Critical dynamics: a field theory approach to equilibrium and non-equilibrium scaling behavior* (Cambridge University Press)

[31] Touchette H 2009 *Physics Reports* **478** 1–69

[32] Gärtner J 1977 *Theory of Probability & Its Applications* **22** 24–39

[33] Ellis R S 1984 *The Annals of Probability* **12** 1–12

[34] Bender C M and Orszag S A 1999 *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory* vol 1 (Springer Science & Business Media)

[35] Skellam J G 1946 *Journal of the Royal Statistical Society Series A: Statistics in Society* **109** 296–296

[36] Ellis R S 2006 *Entropy, large deviations, and statistical mechanics* vol 1431 (Taylor & Francis)

[37] Gallavotti G and Cohen E G D 1995 *Physical review letters* **74** 2694

[38] Taniguchi T and Cohen E 2007 *Journal of Statistical Physics* **126** 1–41

[39] Van Zon R and Cohen E 2003 *Physical Review E* **67** 046102

[40] Harris R J and Schütz G M 2007 *Journal of Statistical Mechanics: Theory and Experiment* **2007** P07020

[41] Graham R and Tél T 1985 *Physical Review A* **31** 1109

[42] Freidlin M I, Wentzell A D, Freidlin M and Wentzell A 1998 *Random perturbations* (Springer)

[43] Kipnis C, Olla S and Varadhan S S 1989 *Communications on Pure and Applied Mathematics* **42** 115–137

[44] Derrida B, Lebowitz J L and Speer E 2003 *Journal of statistical physics* **110** 775–810

[45] Enaud C and Derrida B 2004 *Journal of statistical physics* **114** 537–562

[46] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2015 *Reviews of Modern Physics* **87** 593

[47] MacDonald C T, Gibbs J H and Pipkin A C 1968 *Biopolymers: Original Research on Biomolecules* **6** 1–25

[48] MacDonald C T and Gibbs J H 1969 *Biopolymers: Original Research on Biomolecules* **7** 707–725

[49] Zia R, Dong J and Schmittmann B 2011 *Journal of Statistical Physics* **144** 405–428

[50] Cividini J, Mukamel D and Posch H 2017 *Physical Review E* **95** 012110

[51] Miron A, Mukamel D and Posch H A 2021 *Physical Review E* **104** 024123

[52] Wolf D E, Schreckenberg M and Bachem A 1996 *Traffic and granular flow* (World Scientific)

[53] Schadschneider A 2000 *Physica A: Statistical Mechanics and its Applications* **285** 101–120

[54] Chowdhury D, Santen L and Schadschneider A 2000 *Physics Reports* **329** 199–329

[55] Schadschneider A 2002 *Physica A: Statistical Mechanics and its Applications* **313** 153–187

[56] Schütz G M 1997 *Journal of statistical physics* **88** 427–445

[57] Tracy C A and Widom H 2008 *Communications in Mathematical Physics* **279** 815–844

[58] Tracy C A and Widom H 2008 *Journal of Statistical Physics* **132** 291–300

[59] Ferrari P L 2010 *Journal of Statistical Mechanics: Theory and Experiment* **2010** P10016

[60] Derrida B, Domany E and Mukamel D 1992 *Journal of statistical physics* **69** 667–687

[61] Derrida B, Evans M R, Hakim V and Pasquier V 1993 *Journal of Physics A: Mathematical and General* **26** 1493

[62] Schütz G and Domany E 1993 *Journal of statistical physics* **72** 277–296

[63] Essler F H L and Rittenberg V 1996 *Journal of Physics A: Mathematical and General* **29** 3375

[64] Sasamoto T 1999 *Journal of Physics A: Mathematical and General* **32** 7109

[65] Andjel E, Bramson M and Liggett T 1988 *Probability Theory and Related Fields* **78** 231–247

[66] Boldrighini C, Cosimi G, Frigio S and Grasso Nunes M 1989 *Journal of Statistical Physics* **55** 611–623

[67] Ferrari P A 1992 *Probability theory and related fields* **91** 81–101

[68] Derrida B, Janowsky S A, Lebowitz J L and Speer E R 1993 *Journal of statistical physics* **73** 813–842

[69] Ferrari P A and Fontes L R 1994 *Probability theory and related fields* **99** 305–319

[70] Mallick K, Mallick S and Rajewsky N 1999 *Journal of Physics A: Mathematical and General* **32** 8399

[71] Alcaraz F and Bariev R 2000 *Brazilian Journal of Physics* **30** 655–666

[72] Evans M R 1996 *Europhysics Letters* **36** 13

[73] Benjamini I, Ferrari P A and Landim C 1996 *Stochastic processes and their applications* **61** 181–204

[74] Krug J and Ferrari P A 1996 *Journal of Physics A: Mathematical and General* **29** L465

[75] Harris T E 1965 *Journal of Applied Probability* **2** 323–338

[76] Bénichou O, Cazabat A, Lemarchand A, Moreau M and Oshanin G 1999 *Journal of statistical physics* **97** 351–371

[77] Bénichou O, Moreau M and Oshanin G 2000 *Physical Review E* **61** 3388

[78] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2002 *Journal of Statistical Physics* **107** 635–675

[79] Krajenbrink A and Le Doussal P 2021 *Physical Review Letters* **127** 064101

[80] Bettelheim E, Smith N R and Meerson B 2022 *Physical Review Letters* **128** 130602

[81] Mallick K, Moriya H and Sasamoto T 2022 *Physical Review Letters* **129** 040601

[82] Doyon B, Perfetto G, Sasamoto T and Yoshimura T 2022 *arXiv preprint arXiv:2206.14167*

[83] Kardar M, Parisi G and Zhang Y C 1986 *Physical Review Letters* **56** 889

[84] Halpin-Healy T and Zhang Y C 1995 *Physics reports* **254** 215–414

[85] Huse D A and Henley C L 1985 *Physical review letters* **54** 2708

[86] Kardar M and Zhang Y C 1987 *Physical review letters* **58** 2087

[87] Derrida B 1990 *Physica A: Statistical Mechanics and its Applications* **163** 71–84

[88] Alcaraz F C, Droz M, Henkel M and Rittenberg V 1994 *Annals of Physics* **230** 250–302

[89] Schütz G and Sandow S 1994 *Physical Review E* **49** 2726

[90] Alexander S and Holstein T 1978 *Physical Review B* **18**(1) 301–302

[91] Gwa L H and Spohn H 1992 *Physical Review A* **46** 844

[92] Kim D 1995 *Physical Review E* **52** 3512

[93] Ciandrini L, Stansfield I and Romano M 2010 *Physical Review E* **81** 051904

[94] Evans M R 2000 *Brazilian Journal of Physics* **30** 42–57

[95] Evans M R and Hanney T 2005 *Journal of Physics A: Mathematical and General* **38** R195

[96] Großkinsky S, Schütz G M and Spohn H 2003 *Journal of statistical physics* **113** 389–410

[97] Godrèche C 2003 *Journal of Physics A: Mathematical and General* **36** 6313

[98] Weisstein E W Negative binomial series URL `https://mathworld.wolfram.com/NegativeBinomialSeries.html`

[99] Juhász R, Santen L and Iglói F 2005 *Physical review letters* **94** 010601

[100] Affleck I, Kennedy T, Lieb E H and Tasaki H 1987 *Physical Review Letters* **59** 799

[101] Affleck I, Kennedy T, Lieb E H and Tasaki H 1988 *Communications in Mathematical Physics* **115** 477–528

[102] Derrida B, Evans M R and Mukamel D 1993 *Journal of Physics A: Mathematical and General* **26** 4911

[103] Prolhac S, Evans M R and Mallick K 2009 *Journal of Physics A: Mathematical and Theoretical* **42** 165004

[104] Derrida B 1998 *Physics Reports* **301** 65–83

[105] Krebs K and Sandow S 1997 *Journal of Physics A: Mathematical and General* **30** 3165

[106] Isaev A P, Pyatov P N and Rittenberg V 2001 *Journal of Physics A: Mathematical and General* **34** 5815

[107] Sandow S 1994 *Physical review E* **50** 2660

[108] Mallick K and Sandow S 1997 *Journal of Physics A: Mathematical and General* **30** 4513

[109] de Gier J and Essler F H L 2005 *Physical review letters* **95** 240601

[110] de Gier J and Essler F H L 2006 *Journal of Statistical Mechanics: Theory and Experiment* **2006** P12011

[111] de Gier J and Essler F H L 2008 *Journal of Physics A: Mathematical and Theoretical* **41** 485002

[112] Lieb E H and Liniger W 1963 *Physical Review* **130** 1605

[113] Lieb E H and Wu F Y 1968 *Physical Review Letters* **20** 1445

[114] Lieb E H 1967 *Physical Review* **162** 162

[115] Baxter R J 1971 *Physical Review Letters* **26** 832

[116] Dhar D 1987 *Phase Transitions* **1** 51

[117] Essler F H L, Korepin V E and Schoutens K 1992 *Nuclear Physics B* **372** 559–596

[118] Cantini L 2008 *Journal of Physics A: Mathematical and Theoretical* **41** 095001

[119] Prolhac S and Mallick K 2008 *Journal of Physics A: Mathematical and Theoretical* **41** 175002

[120] Prolhac S and Mallick K 2009 *Journal of Physics A: Mathematical and Theoretical* **42** 175001

[121] de Gier J and Essler F H L 2011 *Physical review letters* **107** 010602

[122] Sasamoto T and Wadati M 1997 *Journal of the Physical Society of Japan* **66** 279–282

[123] Golinelli O and Mallick K 2006 *Journal of Physics A: Mathematical and General* **39** 10647

[124] Katsura H and Maruyama I 2010 *Journal of Physics A: Mathematical and Theoretical* **43** 175003

[125] Furst E M and Squires T M 2017 Introduction *Microrheology* (Oxford University Press) ISBN 9780199655205

[126] Vasquez P A and Forest M G 2015 Complex fluids and soft structures in the human body *Complex Fluids in Biological Systems: Experiment, Theory, and Computation* ed Spagnolie S E (New York, NY: Springer New York) pp 53–110 ISBN 978-1-4939-2065-5

[127] Ferrari P A, Goldstein S and Lebowitz J L 1985 *Diffusion, mobility and the Einstein relation* (Springer)

[128] Burlatsky S, Oshanin G, Mogutov A and Moreau M 1992 *Physics Letters A* **166** 230–234

[129] Burlatsky S, Oshanin G, Moreau M and Reinhardt W 1996 *Physical Review E* **54** 3165

[130] Bénichou O, Cazabat A, De Coninck J, Moreau M and Oshanin G 2000 *Physical Review Letters* **84** 511

[131] Bénichou O, Cazabat A, De Coninck J, Moreau M and Oshanin G 2001 *Physical Review B* **63** 235413

[132] Bénichou O, Illien P, Oshanin G and Voituriez R 2013 *Physical Review E* **87** 032164

[133] Illien P, Bénichou O, Oshanin G and Voituriez R 2014 *Physical review letters* **113** 030603

[134] Illien P, Bénichou O, Oshanin G and Voituriez R 2015 *Journal of Statistical Mechanics: Theory and Experiment* **2015** P11016

[135] Bénichou O, Illien P, Oshanin G, Sarracino A and Voituriez R 2016 *Physical Review E* **93** 032128

[136] Poncet A, Benichou O, Démery V and Oshanin G 2019 *Physical Review Research* **1** 033089

[137] Cividini J, Kundu A, Majumdar S N and Mukamel D 2016 *Journal of Statistical Mechanics: Theory and Experiment* **2016** 053212

[138] Ayyer A 2022 *Ann. Inst. Henri Poincaré Comb. Phys. Interact.*

[139] Evans M R, Foster D P, Godrèche C and Mukamel D 1995 *Physical review letters* **74** 208

[140] Evans M R, Foster D P, Godrèche C and Mukamel D 1995 *Journal of statistical physics* **80** 69–102

[141] Arndt P F, Heinzel T and Rittenberg V 1998 *Journal of statistical physics* **90** 783–815

[142] Arndt P F, Heinzel T and Rittenberg V 1998 *Journal of Physics A: Mathematical and General* **31** L45

[143] Jafarpour F H 2000 *Journal of Physics A: Mathematical and General* **33** 8673

[144] Derrida B, Lebowitz J and Speer E 1997 *Journal of statistical physics* **89** 135–167

[145] Jafarpour F, Ghafari F and Masharian S 2005 *Journal of Physics A: Mathematical and General* **38** 4579

[146] Miron A, Mukamel D and Posch H A 2020 *Journal of Statistical Mechanics: Theory and Experiment* **2020** 063216

[147] Baxter R J 2016 *Exactly solved models in statistical mechanics* (Elsevier)

[148] Kulish P P and Sklyanin E K 1982 *Journal of Soviet Mathematics* **19** 1596–1620

[149] Alcaraz F C and Bariev R Z 1999 *Physical Review E* **60** 79

[150] Popkov V, Fouladvand M E and Schütz G M 2002 *Journal of Physics A: Mathematical and General* **35** 7187

[151] Baxter R J 1972 *Annals of Physics* **70** 193–228

[152] Yang C N 1967 *Physical Review Letters* **19** 1312

[153] Crampé N, Frappat L and Ragoucy E 2013 *Journal of Physics A: Mathematical and Theoretical* **46** 405001

[154] Crampé N, Frappat L, Ragoucy E and Vanicat M 2016 *Journal of Mathematical Physics* **57**

[155] Vieira R S and Lima-Santos A 2021 *Journal of Statistical Mechanics: Theory and Experiment* **2021** 053103

[156] Alcaraz F C and Lazo M J 2003 *Journal of Physics A: Mathematical and General* **37** L1

[157] Alcaraz F C and Lazo M J 2004 *Journal of Physics A: Mathematical and General* **37** 4149

[158] Sasamoto T and Cantini L Personal communications

[159] Kirillov A 1985 *Journal of Soviet Mathematics* **30** 2298–2310

[160] Kirillov A N and Liskova N A 1997 *Journal of Physics A: Mathematical and General* **30** 1209

[161] Baxter R 2002 *Journal of statistical physics* **108** 1–48

[162] Essler F H L, Korepin V E and Schoutens K 1991 *Physical review letters* **67** 3848

[163] Essler F H L, Korepin V E and Schoutens K 1994 *Exactly Solvable Models Of Strongly Correlated Electrons* **18** 133

[164] Prolhac S 2010 *Journal of Physics A: Mathematical and Theoretical* **43** 105002

[165] Yung C and Batchelor M T 1995 *Nuclear Physics B* **446** 461–484

[166] Nepomechie R I 2003 *Journal of statistical physics* **111** 1363–1376

[167] Bazhanov V V, Łukowski T, Meneghelli C and Staudacher M 2010 *Journal of Statistical Mechanics: Theory and Experiment* **2010** P11002

[168] Lazarescu A and Pasquier V 2014 *Journal of Physics A: Mathematical and Theoretical* **47** 295202

[169] Prolhac S 2008 *Journal of Physics A: Mathematical and Theoretical* **41** 365003

[170] Alcaraz F C and Bariev R Z 2000 *Brazilian Journal of Physics* **30** 13–26

[171] Alcaraz F C and Bariev R Z 2000 *Brazilian Journal of Physics* **30** 655–666

[172] Pronko G P and Stroganov Y G 1999 *Journal of Physics A: Mathematical and General* **32** 2333