Nonequilibrium steady states of matrix-product form: a solver's guide

2007 J. Phys. A: Math. Theor. 40 R333

(http://iopscience.iop.org/1751-8121/40/46/R01)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) R333–R441 doi:10.1088/1751-8113/40/46/R01

TOPICAL REVIEW

Nonequilibrium steady states of matrix-product form: a solver's guide

R A Blythe and M R Evans

SUPA, School of Physics, University of Edinburgh, Mayfield Road, Edinburgh, EH9 3JZ, UK

Received 12 June 2007, in final form 26 September 2007
Published 31 October 2007
Online at stacks.iop.org/JPhysA/40/R333

Abstract

We consider the general problem of determining the steady state of stochastic nonequilibrium systems such as those that have been used to model (among other things) biological transport and traffic flow. We begin with a broad overview of this class of driven-diffusive systems—which includes exclusion processes—focusing on interesting physical properties, such as shocks and phase transitions. We then turn our attention specifically to those models for which the exact distribution of microstates in the steady state can be expressed in a matrix-product form. In addition to a gentle introduction to this matrix-product approach, how it works and how it relates to similar constructions that arise in other physical contexts, we present a unified, pedagogical account of the various means by which the statistical mechanical calculations of macroscopic physical quantities are actually performed. We also review a number of more advanced topics, including nonequilibrium free-energy functionals, the classification of exclusion processes involving multiple particle species, existence proofs of a matrix-product state for a given model and more complicated variants of the matrix-product state that allow various types of parallel dynamics to be handled. We conclude with a brief discussion of open problems for future research.

PACS numbers: 05.40.−a, 05.70.Fh, 02.50.Ey, 64.60.−i

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction

1.1. What is a nonequilibrium steady state?

Consider a finite volume that contains interacting particles and is coupled at opposite boundaries to particle reservoirs held at *different* chemical potentials, as shown in figure 1. In this situation one anticipates a net flux of particles from the reservoir with the greater chemical potential through the system and into the opposite reservoir. After some initial transients we expect a state to arise in which there is a nonzero mean flux that is constant over space and time. This flux reveals the system to be in a *nonequilibrium steady state*. More generally, we have in mind any physical system all of whose macroscopic observables do not change with time, but nevertheless exhibits an irreversible exchange of heat, particles, volume or some other physical quantity with its environment.

When heat, particle or volume exchanges are reversible, the system is at equilibrium with its environment and, as is well established, its microstates have a Gibbs–Boltzmann distribution. This knowledge allows one then to predict the macroscopic physics of a many-body system given a microscopic model couched in terms of energetic interactions. Furthermore, these Gibbs–Boltzmann statistics emerge from the application of a very simple principle, namely that of equal *a priori* probabilities for the system combined with its environment.

By contrast, when a system is out of equilibrium, very little is known about the statistics of the microstates, not even in the steady state. This is despite a wide range of approaches aimed at addressing this deficiency in our understanding of nonequilibrium physics. These approaches fall roughly into two broad categories.

First, one has macroscopic theories, such as Onsager and Machlup's pioneering work on near-equilibrium fluctuations [1, 2]. In this regime, it is assumed that the thermodynamic forces that restore an equilibrium are linear, and it is possible to determine the probability of witnessing certain fluctuations in the equilibrium steady state, the most probable trajectory given by a principle of minimum energy dissipation. In this work, we are interested in steady states that are far from equilibrium, i.e., where the system is driven beyond the linear response regime. Recently there has been some success in extending the Onsager–Machlup

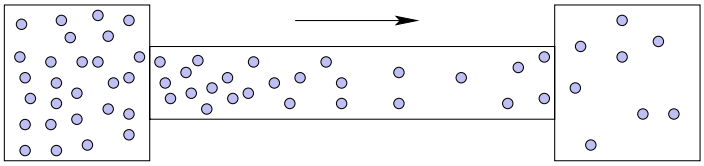


Figure 1. A channel held out of equilibrium through the application of a chemical potential gradient. It is anticipated that a nonequilibrium steady state will be reached in which particles flow from one reservoir to another at a constant rate.

theory of equilibrium fluctuations to such nonequilibrium steady states, as long as one has been able to derive macroscopic (hydrodynamic) equations of motion for the system [3, 4]. Furthermore, over the past few years fluctuation theorems have been derived under a wide range of conditions. Typically, these relate the probabilities of entropy changes of equal magnitude but opposite sign occurring in a system driven arbitrarily far from equilibrium (see [5, 6] for a brief introductory overview). Although some connections between these various topics have been established (see, e.g., [7]), it is fair to say that a complete coherent picture of the macroscopic properties of nonequilibrium steady states is still lacking.

Somewhere between this macroscopic approach and a truly microscopic approach (i.e., one that would follow directly from some microscopic equations of motion) lie a range of mesoscopic models that are intended to capture the essential features of nonequilibrium dynamics—irreversibility, currents, dissipation of heat and so on—but are nevertheless simple enough that analytical treatment is possible. In particular, it is possible to explore the macroscopic consequences of the underlying dynamics, in the process identifying both similarities with equilibrium states of matter and the novel features peculiar to nonequilibrium systems.

It is this approach that is the focus of this review article. Specifically, we will discuss a set of models that have a steady-state distribution of microstates that can be expressed mathematically in the form of a matrix product. We will explain in detail how these expressions can be used to calculate collective steady-state properties exactly. These will include particle currents, density profiles, correlation functions, the distribution of macroscopic fluctuations and so on. These calculations will reveal phenomena that occur purely because of the far-from-equilibrium conditions: for example, boundary-induced phase transitions and shock fronts. Meanwhile, we will also find conceptual connections with equilibrium statistical physics: in some of the models we discuss, for example, it is possible to construct partition functions and free-energy-like quantities that are meaningful both for equilibrium and nonequilibrium systems. We begin by explaining this modelling approach in more detail.

1.2. Modelling nonequilibrium dynamics

In figure 2 we have sketched some molecular motors (kinesins) that are attached to a microtubule. This is essentially a track along which the motors can walk by using packets of energy carried by ATP molecules present in the surroundings. Clearly, at the most microscopic level, there are a number chemical processes at play that combine to give rise to the motor's progress along the track. At a more coarse-grained, mesoscopic level, we can simply observe that the motor takes steps of a well-defined size (approximately 8 nm for a kinesin [8]) at a time, and thus model the system as a set of particles that hop *stochastically* between

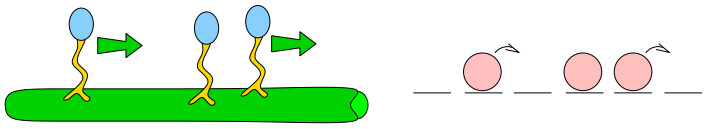


Figure 2. Kinesins attached to a microtubule that move in a preferred direction by extracting chemical energy from the environment. At a mesoscopic level, this can be modelled as a stochastic process in which particles hop along a one-dimensional lattice as shown.

sites of a lattice. This stochastic prescription is in part intended to reflect the fact that the internal degrees of freedom that govern when a particle hops are not explicitly included in the model.

The simplest such stochastic model for the particle hops is a Poisson process that occurs at some prescribed rate. Let $\mathcal{C}$ and $\mathcal{C}'$ be two configurations of the lattice that differ by a single particle hop. We then define $W(\mathcal{C} \to \mathcal{C}')$ as the rate at which this hop occurs, such that in an infinitesimal time interval $\mathrm{d}t$ the probability that a hop takes place is $W(\mathcal{C} \to \mathcal{C}')\,\mathrm{d}t$. The rate of change of the probability $P(\mathcal{C}, t)$ for the system to be in configuration $\mathcal{C}$ at time t is then the solution of the master equation

$$\frac{\partial}{\partial t} P(\mathcal{C}, t) = \sum_{\mathcal{C}' \neq \mathcal{C}} P(\mathcal{C}', t) W(\mathcal{C}' \to \mathcal{C}) - \sum_{\mathcal{C}' \neq \mathcal{C}} P(\mathcal{C}, t) W(\mathcal{C} \to \mathcal{C}') \tag{1.1}$$

subject to some initial condition $P(\mathcal{C}, 0)$. The first term on the right-hand side of this equation gives contributions from all possible hops (transitions) into the configuration $\mathcal{C}$ from other configurations $\mathcal{C}'$; the second term gives contributions from hops out of $\mathcal{C}$ into other configuration $\mathcal{C}'$. We are concerned here with steady states, where these gain and loss terms exactly balance causing the time derivative to vanish. Such solutions of the master equation we denote by $P(\mathcal{C})$. In this work, nearly all the models we discuss have the property that a single steady state is reached from any initial condition, i.e., they are ergodic [9, 10]. Given then that the distribution $P(\mathcal{C})$ is unique for these models, we concentrate almost exclusively on how it is obtained analytically and how expectation values of observables are calculated. The remaining models that are not ergodic over the full space of configurations turn out to have a unique steady-state distribution over some subspace of configurations. We will see that such a distribution of this type can also be handled using the methods described in this work.

1.3. A catalogue of steady states

Solutions of the master equation (1.1) can assume a number of structures. Some general classes of steady state—not intended to be mutually exclusive—can be summarized as follows:

Equilibrium (Gibbs–Boltzmann) steady state. If the internal energy of a microstate $\mathcal{C}$ is $E(\mathcal{C})$ and the system as a whole is in equilibrium with a heat bath with inverse temperature β, the microstates have a Gibbs–Boltzmann distribution $P(\mathcal{C}) \propto \mathrm{e}^{-\beta E(\mathcal{C})}$. The dynamics is reversible in that the probability of witnessing any particular trajectory through phase space is equal to that of its time reversal. This then implies that stochastic dynamics expressed in terms of transition rates $W(\mathcal{C} \to \mathcal{C}')$ must satisfy the detailed balance condition [11, 12]

$$P(\mathcal{C}) W(\mathcal{C} \to \mathcal{C}') = P(\mathcal{C}') W(\mathcal{C}' \to \mathcal{C}) \tag{1.2}$$

where $P(\mathcal{C})$ is the Gibbs–Boltzmann distribution. A simple consequence of (1.2) is that there are no fluxes in the steady state. Hence any system that does exhibit currents in the steady state,

i.e., nonequilibrium systems, generally does not have a stationary distribution that satisfies detailed balance.

Factorized steady state. Sometimes, whether in or out of equilibrium, one has a factorized steady state. Typically one has in mind a lattice (or graph) with N sites with configurations specified by occupancy variables n_i (i.e., site i contains n_i particles). A factorized steady state then takes the form $P(\mathcal{C}) \propto \prod_{i=1}^{N} f_i(n_i)$. For such a structure to arise, certain constraints on the transition rates $W(\mathcal{C} \to \mathcal{C}')$ must be satisfied. For example, in an equilibrium system, one would need the energy of the system to be the sum of single-site energies. Out of equilibrium, one finds factorized steady states in zero-range processes, in which particles hop at a rate that depends only on the occupation number at the departure site. Moreover, necessary and sufficient conditions for factorization in a broad class of models which includes the ZRP have been established [13]. Zero-range processes and various generalizations have been reviewed recently in a companion paper [14] and so we do not discuss them further here.

Matrix-product steady state. A matrix-product steady state is an extension to a factorized steady state that is of particular utility for one-dimensional models. The rough idea is to replace the scalar factors $f_i(n_i)$ with matrices X_{n_i}, the steady-state probability then being given by an element of the resulting matrix product $\prod_{i=1}^{N} X_{n_i}$. Since the matrices X_n for different occupancy numbers n need not commute, one opens up the possibility for correlations between the occupancy of different sites (above those that emerge from global constraints, such as a fixed particle number, that one sees in factorized states). It turns out that quite a number of nonequilibrium models have a matrix-product steady state and we will encounter them all in the course of this review.

Most general steady state. In principle, the (assumed unique) stationary solution of the master equation (1.1) can always be found if the number of configurations is finite, since (1.1) is a system of linear equations in the probabilities $P(\mathcal{C})$. One way to express this solution is in terms of *statistical weights* $f(\mathcal{C})$, each of which is given by the determinant of the matrix of transition rates obtained by removing the row and column corresponding to the configuration $\mathcal{C}$ (see, e.g., [15, 16]). To arrive at the probability distribution $P(\mathcal{C})$ one requires the *normalization* $Z = \sum_{\mathcal{C}} f(\mathcal{C})$, so that then $P(\mathcal{C}) = f(\mathcal{C})/Z$. This normalization has some interesting properties: it is uniquely defined for any ergodic Markov process with a finite number of configurations; it can be shown to be equal to the product of all nonzero eigenvalues of the matrix of transition rates; and it is a partition function of a set of trees on the graph of transitions between different microscopic configurations. In this latter case, the densities of particular edges in the ensemble of trees are controlled by tuning the transition rates corresponding to those edges, so in this interpretation these transition rates are equivalent to equilibrium fugacities. Although this interpretation of the normalization is rather abstract, concrete connections between transition rates in certain nonequilibrium models that have a matrix-product steady state and fugacities in an equilibrium ensemble have been made, and these we shall discuss later in this review. We also remark that the equilibrium theory of partition function zeros to characterize phase transitions also carries over to the normalization as just defined, at least for those models that have been tested [17–19].

1.4. Purpose of this review

We have three principal aims in this work. First, we wish to illustrate the insights into nonequilibrium statistical mechanics that have been gained from exactly solvable models, particularly those that have a steady state of the matrix-product form. Secondly, we seek

to provide a self-contained pedagogical account of the various analytical methods and calculational tools that can be used to go from the matrix-product expressions to predictions for the macroscopic physics. Finally, we wish to present a thorough review of the progress that has been made using the matrix-product approach over the last few years. As far as we are aware, the matrix-product approach has not been the focus of a review for nearly a decade [20, 21], and we feel that it is high time that the significant developments that have occurred in the meantime should be collected together in one place. In order to prevent this review from becoming unmanageably long, we focus purely on static physical properties exhibited in the steady states of models solvable by the matrix-product method. This means we have unfortunately had to omit discussion of some very interesting topics, for example how some dynamical properties, such as fluctuation phenomena, have been elucidated through the use of the Bethe ansatz [22–24], determinental solutions and their connection to random matrix theory [25–27] and dynamical matrix products [28, 29]. Some of these topics, however, have recently been reviewed elsewhere [30, 31]. We also direct the reader to established reviews, such as [21, 32–37], for any other background that we have been forced to leave out here.

To realize the aims stated above, we provide in the next section a general account of the physics one expects to see in nonequilibrium dynamical systems and outline the essential ideas underlying the matrix-product approach. Thereafter, we go into the details of how the simplest models are solved and show a number of contrasting (but equivalent) approaches that have found application to more complex problems. These cases then form the material of the remainder of the review, which we round off by posing some open problems for future research.

2. An overview of driven-diffusive systems

2.1. A paradigmatic example: the asymmetric simple exclusion process (ASEP)

The asymmetric simple exclusion process (ASEP) is a very simple model of a driven system in one dimension that has biased diffusion of hard-core particles. We shall discuss two versions: the periodic system and the open system. The latter is coupled to particle reservoirs at either end so that, as described in the previous section, there is a steady state with a constant particle flux. The open system has a long history, having first appeared in the literature to our knowledge as a model of biopolymerization [38] and transport across membranes [39]. In the mathematical literature meanwhile diffusion with collisions between particles was first studied by Harris [40] and the terminology simple exclusion was first defined by Spitzer [41]. Over the years, applications to other transport processes have appeared, e.g., as a general model for traffic flow [42] and various other theoretical and experimental studies of biophysical transport [43–48]. Whilst interesting and important, these many applications are not our primary concern here. Also we shall not do justice to the many rigorous mathematical results which have been summarized in the books by Liggett [49, 50]. Rather, our interest in the ASEP lies in its having acquired the status of a fundamental model of nonequilibrium statistical physics in its own right in much the same way that the Ising model has become a paradigm for equilibrium critical phenomena. In particular, the ASEP and related models show—despite their superficial simplicity—a range of nontrivial macroscopic phenomena, such as phase transitions, spontaneous symmetry breaking, shock fronts, condensation and jamming.

In common with the vast majority of the models we discuss in this review, the ASEP is defined as a stochastic process taking place in continuous time on a discrete one-dimensional lattice of N sites. Hopping is totally asymmetric: a particle sitting to the left of an unoccupied

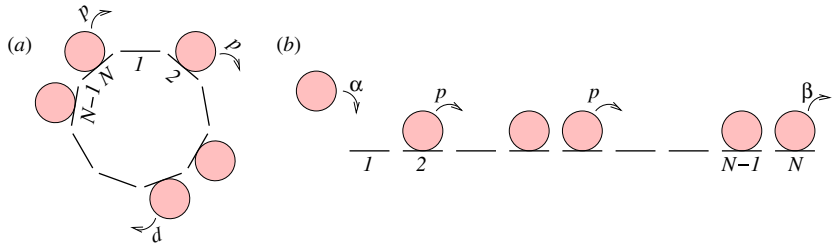


Figure 3. Asymmetric exclusion process with (*a*) periodic and (*b*) open-boundary conditions. The labels indicate the rates at which the various particle moves can occur.

site hops forwards one site (as in figure 2), each hop being a Poisson process with rate p. That is, in an infinitesimal time interval $\mathrm{d}t$, there is a probability $p\mathrm{d}t$ that one of the particles that can hop, does hop, the identity of that particle being chosen at random. In numerical simulations a *random-sequential* updating scheme is often used. That is, bonds between particles are chosen at random and then, if there is a particle at the left-hand end of the bond and a vacancy at the right, the particle is moved forwards. In this scheme, each update corresponds (on average) to $1/(Np)$ units of time; in the following we take $p = 1$ without loss of generality. The definition of the model is completed by specifying the boundary conditions.

2.1.1. Periodic boundary conditions. The simplest version of the model has a ring of N sites as depicted in figure 3(*a*): a particle hopping to the right on site N lands on site 1 if the receiving site is empty. Under these dynamics, the total particle number M is conserved. Furthermore, as we will show in section 3.1, the steady state is very simple: all configurations (with the allowed number of particles) are equally likely. As there are $\binom{N}{M}$ allowed configurations the probability of any one is $1\big/\binom{N}{M}$. The steady-state current of particles, J, through a bond $i, i+1$ is given by p (here equal to 1) multiplied by the probability that there is a particle at site i and site $i+1$ is vacant. One finds

$$J = \frac{\binom{N-2}{M-1}}{\binom{N}{M}} = \frac{M}{N}\frac{(N-M)}{(N-1)} \to \bar{\rho}(1-\bar{\rho}) \tag{2.1}$$

in the thermodynamic limit where $N, M \to \infty$ at a constant ratio $M/N = \bar{\rho}$.

One may also ask, starting from a known configuration, how long the system takes to relax to the steady state. It turns out that this timescale T depends on the system size N via the relation $T \sim N^z$ which defines a *dynamic exponent* z. One can show using the Bethe ansatz (not covered here, but see, e.g., [22, 51, 52]) that for the ASEP on a ring, $z = 3/2$. Also in the symmetric case where particles hop both to the left and to the right with equal rates, but still with exclusion, one obtains $z = 2$ which is the value one would find for a purely diffusive process.

2.1.2. Open-boundary conditions. To model the interaction of an open system with reservoirs at different densities, Poisson processes acting at the boundary sites are added. Specifically, a particle may be inserted onto the leftmost site with rate α (if it is vacant) and may leave the system from the rightmost site at rate β: see figure 3(*b*). Here we will consider only $\alpha, \beta < 1$ although there is no problem in considering rates without this range. Then we may think of reservoirs at site 0 with density α and a reservoir at site $N+1$ with density $1-\beta$.

To include these moves in the simulation scheme described above, one would additionally allow the bonds between the system and the reservoirs to be chosen, and perform the particle updates with probability α (entry) or β (exit) according to the bond chosen (assuming α and β both less than 1). These moves admit the possibility of a nonzero current in the steady state.

The open system is more interesting in that phase transitions may occur in the steady state. In the following we shall explore the origins of these phase transitions, both through approximate treatments and an exact solution. In this context we recognize—by direct analogy with equilibrium phase transitions—a phase transition as a sudden change in form of global ensemble-averaged quantities such as the particle current through the system or the particle density in the bulk of the system. It turns out that the particle current plays a central role, analogous to the free energy of an equilibrium system, in determining the nature of the phase transition, i.e., the order of the transition is determined by which the derivative of the current with respect to some external parameter, such as α or β, exhibits a discontinuity. We will expand on this idea in the next section, where we derive the phase diagram for the ASEP.

We remark that other types of updating scheme can also be considered. For example, in applications to traffic flow or pedestrian dynamics [35, 53], it is more natural to consider parallel dynamics where many particles can hop in concert. The ASEP remains solvable under certain classes of parallel updating schemes which will be discussed towards the end of this review (section 10).

2.2. *Mean-field and hydrodynamic approaches*

2.2.1. Lighthill–Whitham theory of kinematic waves. We begin by reviewing a classical phenomenological theory, first applied to traffic flow [54], which serves as a first recourse in our understanding of the phase diagrams of driven-diffusive systems. The idea is to postulate a continuity equation for the local density ρ

$$\frac{\partial \rho}{\partial t} + \frac{\partial J}{\partial x} = 0 \tag{2.2}$$

where J is the current of particles. (Note we use here a different nomenclature and notation to [54] which discusses concentration and flow rather density and current.)

The analysis rests on the key assumption that there is a unique relation, $J(\rho)$, between the current and local density. Also, it is assumed that there is a maximum current J_m (the capacity of the road) at density ρ_m. The first assumption implies that (2.2) becomes

$$\frac{\partial \rho}{\partial t} + v_g \frac{\partial \rho}{\partial x} = 0 \tag{2.3}$$

where

$$v_g(\rho) = \frac{\mathrm{d}J(\rho)}{\mathrm{d}\rho}. \tag{2.4}$$

Now, an implicit solution of (2.3) is of travelling-wave form

$$\rho(x,t) = f(x - v_g(\rho)t) \tag{2.5}$$

as can readily be checked by substitution into (2.3). The arbitrary function f is determined by the initial density profile $\rho(x,0) = f(x)$. The interpretation of the solution is that a patch with local density ρ propagates with velocity $v_g(\rho)$. The propagation of such a patch is a collective phenomenon known as a *kinematic wave* and v_g is a group velocity. The velocity c of a single particle in an environment of density ρ, on the other hand, is defined through

$$J(\rho) = c(\rho)\rho. \tag{2.6}$$

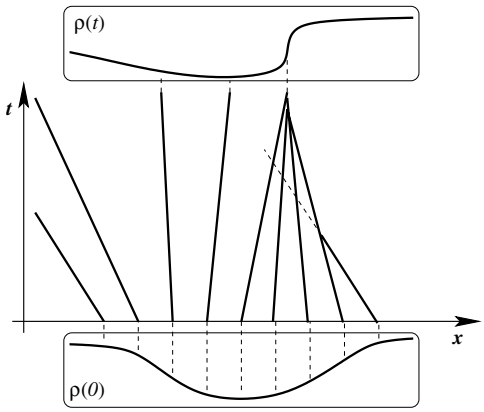


Figure 4. Kinematic waves and sharpening of a shock: the figure sketches the evolution of a density profile $\rho(t)$ from an initial profile $\rho(0)$ according to (2.2) and illustrates how sharp discontinuities in the profile may develop (see the text for discussion).

One sees that

$$v_g = \frac{\mathrm{d}}{\mathrm{d}\rho}(\rho c) = c + \rho\frac{\mathrm{d}c}{\mathrm{d}\rho}. \tag{2.7}$$

Thus, if the single particle velocity is a decreasing function of density, which is what we expect, we find $v_g < c$ and kinematic waves travel backwards in the frame of a moving particle.

If v_g is a decreasing function of ρ we have the phenomenon of shock formation (see figure 4). That is, since the patches of an initial density profile travel with different speeds, the low-density regions catch up with higher density regions and discontinuities in the density profile, known as shocks, develop. Strictly, after the formation of a shock the description of the density profile by the first-order equation (2.2) breaks downs and one has to supplement (2.2) with second-order spatial derivatives to describe the shock profile.

To deduce the velocity of a shock consider two regions of density ρ_1 to the left and ρ_2 to the right separated by a shock. Mass conservation implies that the velocity (positive to the right) of the shock is given by

$$v_s = \frac{J(\rho_2) - J(\rho_1)}{\rho_2 - \rho_1}. \tag{2.8}$$

Note that if the current-density relation has a maximum then one can have a stationary shock, $v_s = 0$.

A particular form for the current-density relationship studied in [54] and which is relevant to the ASEP is

$$J(\rho) = \rho(1 - \rho). \tag{2.9}$$

In this case $J_m = 1/4$ at $\rho_m = 1/2$ and J is symmetric about the maximum. Expression (2.9) may be easily understood for the ASEP by noting that in order for a particle to hop across a bond, the site to the left must be occupied and that to the right empty. For a bond at position x, the probability of the former is $\rho(x)$ and that of the latter is $1 - \rho(x)$. Assuming these events are uncorrelated (which they are not—but we shall come to this later), we have $J(\rho) = \rho(1 - \rho)$.

In the case (2.9) one can compute the kinematic wave and shock velocities from (2.4), (2.8):

$$v_g = 1 - 2\rho \qquad v_s = 1 - \rho_1 - \rho_2 \tag{2.10}$$

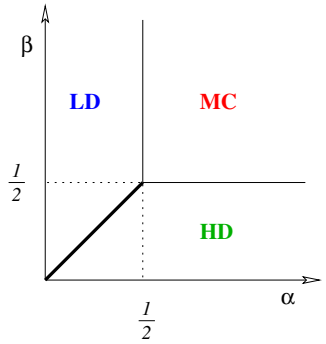


Figure 5. Mean-field phase diagram for the ASEP showing the range of α and β in which the low-density (LD), high-density (HD) and maximal-current (MC) phases are seen.

where, as above, ρ_1 and ρ_2 are the densities on either side of the shock. Thus, the kinematic wave velocity is negative when $\rho > 1/2$ and the shock velocity is negative when $\rho_1 > 1 - \rho_2$.

The kinematic wave theory can be used to predict the phase diagram of the open-boundary ASEP, in which distinct steady-state behaviours are demarcated (figure 5). The left-hand boundary $x = 0$ is considered as a reservoir of particle of density $\rho_l = \alpha$ and the right-hand boundary ($x = N + 1$) as a reservoir of density $\rho_r = 1 - \beta$. Associated with these boundary densities are kinematic waves with velocities

$$v_l = 1 - 2\alpha \qquad v_r = 2\beta - 1. \tag{2.11}$$

In the case where $\alpha < 1/2$ and $\beta < 1/2$ both kinematic waves propagate into the system. So, for example, from an initially empty system the kinematic waves of densities α and $1 - \beta$ will enter from the left and right of the system and meet somewhere in the middle forming a shock which then moves with velocity

$$v_s = \beta - \alpha. \tag{2.12}$$

If $\beta > \alpha$ the shock moves to the right-hand boundary and density associated with the left-hand boundary, $\rho_l = \alpha$, is adopted throughout the bulk of the system. On the other hand if $\alpha < \beta$ the shock moves to the left-hand boundary and density associated with the right-hand boundary, $\rho_r = 1 - \beta$, is adopted throughout the bulk of the system.

In the case $\alpha = \beta < 1/2$, $v_s = 0$ and the shock is stationary. In the stochastic system the shock, although on average stationary, diffuses around the system and effectively reflects off the boundaries. The result is that the shock is equally likely to be anywhere in the system.

In the case where one of α or $\beta > 1/2$, the kinematic wave associated with that boundary does not propagate into the system and the kinematic wave which does propagate from the other boundary controls the bulk density. Thus, the boundary with α or $\beta < 1/2$ controls the bulk density

Finally if both $\alpha, \beta > 1/2$, the kinematic waves from both boundaries do not penetrate. To describe this phase one needs to add a diffusive contribution to the current (2.9), i.e., to consider second-order spatial derivative of ρ which we shall do in the next section. The result is that the steady state of the system has density $1/2$; the system adopts the maximal current density $\rho_m = 1/2$ which is the density associated with kinematic wave velocity zero.

The resulting bulk densities and currents are then as shown in table 1 which corresponds to the phase diagram given in figure 5. Here we have adopted what has become the standard nomenclature for the three phases—the terminology should hopefully be self-explanatory.

Table 1. Properties of the ASEP obtained under a mean-field approximation for the density-current relationship.

Region	Phase	Current J	Bulk density $\bar{\rho}$
$\alpha < \beta, \alpha < \frac{1}{2}$	Low-density (LD)	$\alpha(1-\alpha)$	α
$\beta < \alpha, \beta < \frac{1}{2}$	High-density (HD)	$\beta(1-\beta)$	$1-\beta$
$\alpha > \frac{1}{2}, \beta > \frac{1}{2}$	Maximal-current (MC)	$\frac{1}{4}$	$\frac{1}{2}$

We see that both the current and the bulk density exhibit discontinuities across the phase boundaries. To be explicit, across the transition lines LD–MC and HD–MC, the second derivative of the current is discontinuous whereas across the LD–HD transition line, the first derivative is discontinuous. Thus we may identify these transitions as being of second and first orders, respectively.

This simple kinematic wave theory using the first-order equation (2.2) correctly predicts the phase diagram, as we shall see from the exact solution. The theory also implies two sub-phases of the high- and low-density phases according to whether the boundary which does not control the bulk density has α or β greater than or equal to $1/2$.

A more refined theory is the domain wall theory [55–58] which is able to make more general predictions (such as for multispecies problems) and is reviewed in [37]. We remark that kinematic waves are in fact the characteristic curves of the differential equation (2.3). For more general first-order differential equations modelling driven-diffusive systems one can use the method of characteristics to predict the phase diagram as we briefly discuss in appendix A.

2.2.2. Comparison of mean-field and hydrodynamic approaches. In the previous subsection, we showed how a simple phenomenological description by a first-order differential equation can predict the correct phase diagram of the ASEP. Generally, the assumption that the current is solely a function of the local density is an assumption which is *mean field* in nature. Mean-field theories are usually considered as a general class of approximations where correlations within the configuration of the system are ignored at some level. Thus, a very simple mean-field theory is to only keep information about the density profile. A more sophisticated mean-field theory might keep information about the two-point correlation functions but will still ignore correlations at some level. In some cases or limits a mean-field description may become an exact description. See, for example, [59] for further discussion of mean-field theories.

At this point, it is useful to compare the idea of a mean-field description with the hydrodynamic limit usually considered in the mathematical literature. In the latter case one defines, for example, a local density by averaging or coarse graining over some scale. Then one seeks to define a limit where this scale is allowed to become large but the ratio of this scale to the system size is small. The hydrodynamic limit is then proven by showing that the variables coarse-grained on this scale satisfy a deterministic equation where the noise has been scaled out. The procedure involves determining the local stationarity of the system on the coarse-grained scale. As in mean-field theory, the simplest case results in describing the system by the density profile alone without correlations.

Therefore, at a practical level, the results of a mean-field description often coincide with those obtained in a hydrodynamic limit—both approaches result in the same differential equations involving the density or some small number of variables. But, it is important to stress that the approaches differ significantly in spirit: the hydrodynamic limit specifies a limit

or scale where an exact equation is proven. On the other hand, mean-field theory is often used as a rough and ready approximation regardless of whether the hydrodynamic limit exists. In addition, the hydrodynamic limit introduces coarse-grained densities whereas the equivalent mean-field equations are often obtained without such an explicit procedure as we show below. Finally, mean-field or hydrodynamic equations can often be written phenomenologically as with the Lighthill–Witham theory.

2.2.3. Mean-field treatment of the density profile. We start our considerations of mean-field approximations by keeping space discrete and implementing a simple procedure which is to assume a factorized form for the stationary distribution. That is, in terms of the indicator variables τ_i which take values $\tau_i = 1$ if site i is occupied and $\tau_i = 0$ otherwise, we suppose

$$P(\tau_1, \tau_2, \ldots, \tau_N) = \prod_{i=1}^{N} \mu_i(\tau_i) \qquad \text{where} \qquad \mu_i(\tau) = \begin{cases} \rho_i & \text{if} \quad \tau = 1 \\ 1 - \rho_i & \text{if} \quad \tau = 0 \end{cases} \tag{2.13}$$

and $\rho_i = \langle \tau_i \rangle$. Note that all higher order correlations vanish, i.e., $\langle \tau_i \tau_j \rangle$ is approximated in this mean-field theory by $\rho_i \rho_j$.

Using the master equation (1.1), one can derive an exact equation for the evolution of the density at site i:

$$\frac{\partial \langle \tau_i \rangle}{\partial t} = \langle \tau_{i-1}(1 - \tau_i) \rangle - \langle \tau_i (1 - \tau_{i+1}) \rangle. \tag{2.14}$$

A full derivation is given, for example, in [20]. To understand this equation note that the right-hand side contains the particle current from site $i - 1$ to i minus the particle current from i to $i + 1$, which by particle conservation gives the rate of change of density at site i. In the steady state, where the density is independent of time, the right-hand side of (2.14) must vanish which implies the exact steady-state result

$$J = \langle \tau_i (1 - \tau_{i+1}) \rangle \tag{2.15}$$

where the current J between two neighbouring sites is independent of position.

In the mean-field approximation equation (2.14) becomes

$$\frac{\partial \rho_i}{\partial t} = \rho_{i-1}(1 - \rho_i) - \rho_i (1 - \rho_{i+1}) \tag{2.16}$$

and the steady-state mean-field current becomes

$$J = \rho_i (1 - \rho_{i+1}). \tag{2.17}$$

The density profile and phase diagram can be obtained in the mean-field approximation by considering the mapping

$$\rho_{i+1} = 1 - \frac{J}{\rho_i}. \tag{2.18}$$

The mapping has fixed points at

$$\rho_\pm = \tfrac{1}{2}[1 \pm \sqrt{1 - 4J}]. \tag{2.19}$$

The structure of possible solutions is illustrated in figures 6 and 7 and correspond to the three phases. In the high-density phase the map iterates to a value of ρ_i arbitrarily close to the stable ρ_+ fixed point in some finite number of steps. In the low-density phase the map starts close to the unstable ρ_- fixed point. In the maximal-current phase $J = 1/4 + O(N^{-2})$ so that there are no real fixed points and the density profile takes on its characteristic maximal-current shape: the gradient of the density profile is always negative and far away from both boundaries the density approaches the value $1/2$ (see figure 7).

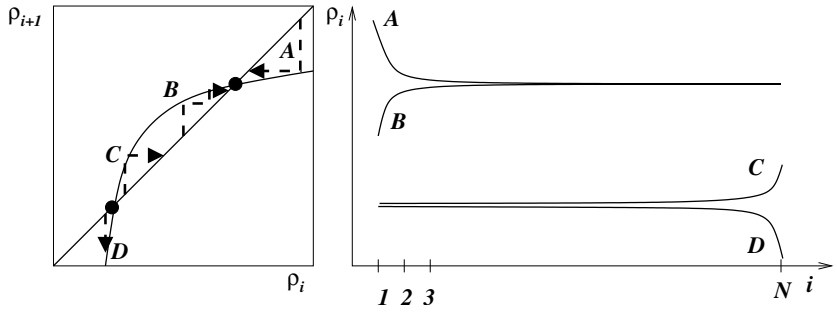


Figure 6. Iteration of the map (2.18) in the high- and low-density phases. Profiles A and B iterate towards the high-density fixed point; profiles C and D begin infinitesimally close to the low-density fixed point and iterate away from it as the right boundary is approached.

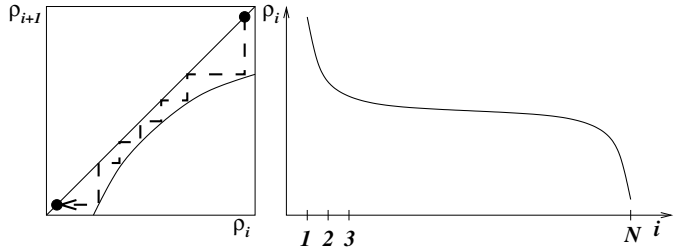


Figure 7. Iteration of the map (2.18) to give the density profile in the maximal-current phase.

The phase diagram can be fully analysed within this discrete mean-field approximation [60]. Here, we choose to make a further continuum approximation as this approach can be easily generalized to other systems as we shall see in the next subsection. One can take a continuum limit of (2.16) in a simple way by replacing $\rho_i(t)$ by $\rho(x, t)$ where $x = ai$ and a is introduced as the lattice spacing. Expanding to second order in a

$$\rho_{i\pm1} = \rho \pm a\frac{\partial\rho}{\partial x} + \frac{a^2}{2}\frac{\partial^2\rho}{\partial x^2} \tag{2.20}$$

yields

$$\frac{\partial\rho}{\partial t} = -a\frac{\partial}{\partial x}\left[J_0(\rho) - D\frac{\partial\rho}{\partial x}\right] \tag{2.21}$$

where

$$J_0(\rho) = \rho(1-\rho) \qquad \text{and} \qquad D = \frac{a}{2}. \tag{2.22}$$

$J_0(\rho)$ represents a contribution to the current due to the asymmetry of the dynamics, i.e., the external drive and $-D\frac{\partial}{\partial x}\rho(x)$ is the diffusive contribution to the current (down the concentration gradient). An equation of this form was postulated by Krug [61]. Ignoring the diffusive part recovers the Lighthill–Whitham phenomenological equation (2.2) or more formally rescaling time to $t' = t/a$ and letting $a \to 0$ recovers the hydrodynamic limit on the Euler scale. Retaining the diffusive term in (2.21) allows one to calculate the (approximate) density profiles explicitly. To see this one considers the steady state of (2.21)

$$J = J_0(\rho) - D\frac{\partial\rho}{\partial x} \tag{2.23}$$

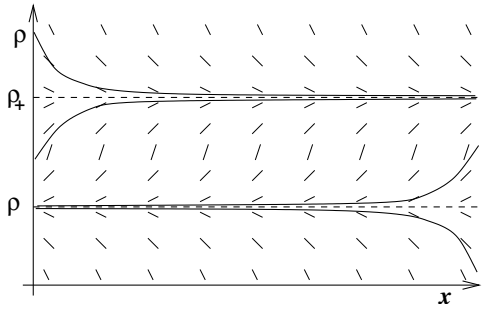


Figure 8. Sketches of the possible solutions of equation (2.24) with $\rho_\pm$ real, which corresponds to J less than the maximum of $J_0(\rho)$.

where J is a constant (the current) to be determined. Rearranging (2.23), (2.9) yields

$$\frac{\partial \rho}{\partial x} = -\frac{(\rho - \rho_+)(\rho - \rho_-)}{D} \tag{2.24}$$

where $\rho_\pm$ are given by (2.19). One can separate variables of this first-order differential equation and integrate, for example, from the left boundary

$$-\int_{\rho(0)}^{\rho(x)} \mathrm{d}\rho \frac{D}{(\rho - \rho_+)(\rho - \rho_-)} = x, \tag{2.25}$$

which yields

$$\ln\left[\frac{(\rho - \rho_+)(\rho(0) - \rho_-)}{(\rho - \rho_-)(\rho(0) - \rho_+)}\right] = -\frac{(\rho_+ - \rho_-)x}{D}. \tag{2.26}$$

The boundary conditions to be imposed on (2.23) are implied by the reservoir densities:

$$\rho(0) = \alpha \qquad \rho(N+1) = 1 - \beta \tag{2.27}$$

the second of these will fix J from (2.26):

$$\left[\frac{(1 - \beta - \rho_+)(\alpha - \rho_-)}{(1 - \beta - \rho_-)(\alpha - \rho_+)}\right] = \exp\left(-\frac{(\rho_+ - \rho_-)(N+1)}{D}\right). \tag{2.28}$$

One can rearrange (2.26) to give an explicit expression for the profile, but to deduce the phase diagram it is simplest to return to (2.24). One can easily sketch the possible solution curves to this equation as is done in figures 8 and 9. The phase diagram in the limit of a large system can be deduced as follows:

High-density phase. The bulk density is $\rho_+ = \rho(N+1) = 1 - \beta$ which yields $J = \beta(1 - \beta)$ and $\rho_- = \beta$. Then, the condition $\rho_+ > \rho_-$ implies $\beta < 1/2$ and the condition that $\rho(0) > \rho_-$ yields $\alpha < \beta$.

Low-density phase. The bulk density is $\rho_- = \rho(0) = \alpha$ which yields $J = \alpha(1 - \alpha)$ and $\rho_+ = \alpha(1 - \alpha)$. Then, the condition $\rho_+ > \rho_-$ implies $\alpha < 1/2$ and the condition that $\rho(N+1) < \rho_+$ yields $\alpha > \beta$.

Coexistence line. When $\alpha = \beta < 1/2$ the profile begins infinitesimally close to $\rho_- = \alpha$ and finishes infinitesimally close to $\rho_+ = 1 - \beta$. In the bulk of the system there is a shock front in

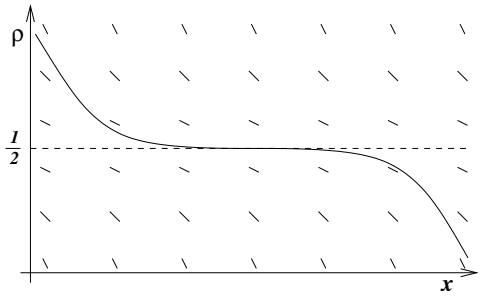


Figure 9. Sketches of the solutions of equation (2.24) with $\rho_\pm$ complex, which corresponds to J greater than the maximum of $J_0(\rho)$.

which the density switches over. Thus, on this line the high-density and low-density phases coexist.

Maximal-current phase: In this case $J > 1/4$ and the profile has the shape shown in figure 9. One can show that to satisfy (2.28) one must take

$$J \approx \frac{1}{4} + \frac{(D\pi)^2}{N^2} \tag{2.29}$$

Then setting $\rho(x) = 1/2 + \eta(x)$, (2.24) becomes

$$\frac{\partial \eta}{\partial x} = -\frac{\eta^2}{D} - O(1/N^2) \tag{2.30}$$

which implies that in the large $N \to \infty$ limit the decay of the density away from the left boundary follows, for large x,

$$\rho(x) - \frac{1}{2} \sim \frac{D}{x}. \tag{2.31}$$

2.2.4. Extremal current principle. Let us consider more general equations of the form (2.23) where the current of particles J across each bond is the sum of a systematic current $J_0(x)$ and a diffusive current J_D. The former is taken to be a known function $J_0(\rho)$ of the local density $\rho(x)$ arising from the nature of the particle interactions whilst the latter is taken to cause motion down a density gradient. Thus the diffusive current should depend only on the local gradient, take opposite sign to its argument and vanish in regions of constant density: e.g. $J_D = -D(\rho)\rho'$ where $D(\rho)$ is some (positive) diffusion constant.

The steady-state profile is given by (2.24) where J is a constant and D may now depend on ρ. Generally, if J_0 has a single maximum $J_0^{\max}$ as a function of ρ we obtain the density profiles with the same qualitative structure as just deduced for the ASEP. To see this we can again sketch the solution curves as in figures 8 and 9. If $J < J_0^{\max}$ there are two fixed points $\rho_\pm$ where $\frac{\partial \rho(x)}{\partial x} = 0$ and the high- and low-density profiles iterate to and away from these fixed points. If $J > J_0^{\max}$ there are no real fixed points and $\frac{\partial \rho(x)}{\partial x} = 0$ yielding a maximal-current profile. As the boundary reservoir densities α and $1 - \beta$ are varied, the system undergoes transitions, of the same type as those exhibited by the ASEP, between regimes characterized by different functional forms of the particle current and density.

These phase transitions can be understood qualitatively using a macroscopic extremal current principle proposed by Krug [61] and later developed into a more general description

of driven systems [56, 57]. The argument runs as follows. From figures 8 and 9 one sees that the density profile has the following properties: it is monotonic in x; it is bounded from above and below due to the finite values of the boundary reservoir densities, and in the limit of an infinite system there is a bulk region in which the density is roughly constant $\rho(x) = \bar{\rho}$. In the bulk region, the diffusive current J_D vanishes, and so here $J = J_0(\bar{\rho})$. If the boundary conditions are such that the density at the left boundary ρ_L is greater than that at the right ρ_R, we have regions near the boundaries where the density is a decreasing function of x (as a consequence of monotonicity). In these regions, the diffusive current $J_D(\rho)$ is positive, and hence the systematic current is reduced compared to that in the bulk. This implies that

$$J = J_0(\bar{\rho}) = \max_{\rho\in[\rho_R,\rho_L]} J_0(\rho) \qquad \text{when} \quad \rho_R < \rho_L \tag{2.32}$$

since every possible density ρ in the range $[\rho_R, \rho_L]$ is represented somewhere in the system, and as we just argued, the systematic current in the bulk must be greater than that at any other point. A similar argument goes through for the case $\rho_L < \rho_R$, except here the diffusive current is negative in the boundary regions and hence

$$J = J_0(\bar{\rho}) = \min_{\rho\in[\rho_L,\rho_R]} J_0(\rho) \qquad \text{when} \quad \rho_L < \rho_R. \tag{2.33}$$

The pair of equations (2.32) and (2.33) encapsulates the extremal current principle.

2.3. The exact solution of the open-boundary ASEP

So far we have discussed mean-field approaches to deducing the phase diagram of the ASEP. These approaches yield the correct phase diagram but do not correctly predict correlation functions. For example, the density profiles discussed in the previous subsections are only qualitatively correct. The time has now come to summarize the main ideas behind the exact solution of the ASEP that uses the matrix-product formulation briefly introduced in section 1 [62]. (We note that an exact solution using recursion relations for the stationary weights was found independently [63], building on earlier work [60, 64] for the case $\alpha = \beta = 1$.)

2.3.1. Nonequilibrium matrix-product steady states. To recap, a matrix-product state is constructed as a product of matrices, one for each site, chosen according to the state of the site. Then, a scalar probability is obtained by combining in some way the elements of this product. The precise prescription for the ASEP with open boundaries was determined in [62]. It is summarized as follows.

First, it is convenient to work with an unnormalized weight $f(\tau_1, \ldots, \tau_N)$ so that

$$P(\tau_1, \tau_2, \ldots, \tau_N) = \frac{f(\tau_1, \tau_2, \ldots, \tau_N)}{Z_N} \tag{2.34}$$

in which Z_N is the normalization obtained by summing the weights f over all 2^N possible configurations of the N-site lattice. It is these weights that take the matrix-product form

$$f_N(\tau_1, \tau_2, \ldots, \tau_N) = \langle W|X_{\tau_1}X_{\tau_2}\cdots X_{\tau_N}|V\rangle \tag{2.35}$$

in which X_{τ_i} is a matrix D if site i is occupied by a particle ($\tau_i = 1$) or a matrix E otherwise. Observe that this is a very visual way of representing a particle configuration—see figure 10. Note further that the matrix appearing at a particular point in the product does not depend explicitly on the site label i, only the state of that site, and that the vectors $\langle W|$ and $|V\rangle$ perform the necessary conversion of the matrix product to a scalar.

With the notation established, expressions for macroscopic quantities quickly follow. In all of these, one requires the normalization which can be written as

$$Z_N = \langle W|(D + E)^N|V\rangle \equiv \langle W|C^N|V\rangle \tag{2.36}$$

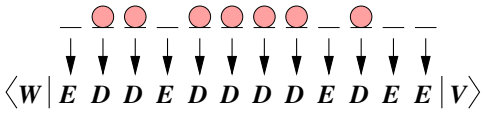


Figure 10. The procedure by which a particle configuration is transliterated to a string of matrices.

in which the matrix $C = D + E$ has been introduced for convenience. The quantity Z_N is simply the sum of the weights of all possible configurations. The density at site i is then

$$\rho_i = \langle \tau_i \rangle = \frac{\langle W|C^{i-1}DC^{N-i}|V\rangle}{Z_N} \tag{2.37}$$

and the current between bonds i and $i+1$ is

$$J_{i,i+1} = \langle \tau_i(1-\tau_{i+1})\rangle = \frac{\langle W|C^{i-1}DEC^{N-i-1}|V\rangle}{Z_N}. \tag{2.38}$$

In order actually to evaluate these expressions, one needs to know a set of algebraic relations between D, E, $|V\rangle$ and $\langle W|$. In [62] it was shown that sufficient conditions for the stationary solution of the master equation of the ASEP via (2.35) are

$$DE = D + E \tag{2.39}$$

$$D|V\rangle = \frac{1}{\beta}|V\rangle \tag{2.40}$$

$$\langle W|E = \frac{1}{\alpha}\langle W|. \tag{2.41}$$

In section 3.1, we will prove that these relations indeed give the desired solution of the master equation. In the meantime, we can at least check the necessary condition that the steady-state particle current $J_{i,i+1}$ across the ith bond is independent of i, and further equals the currents of particles into and out of the system at its boundaries. To see this note first of all that the combination DE appearing in (2.38) can be replaced with $D + E = C$ using (2.39). Then,

$$J_{i,i+1} \equiv J = \frac{\langle W|C^{N-1}|V\rangle}{Z_N} = \frac{Z_{N-1}}{Z_N}. \tag{2.42}$$

At the left boundary we have

$$J_L = \alpha(1-\rho_1) = \alpha\frac{\langle W|EC^{N-1}|V\rangle}{Z_N} = \frac{\langle W|C^{N-1}|V\rangle}{Z_N} = J \tag{2.43}$$

where we have used (2.41). Similarly, at the right boundary one has, using (2.40)

$$J_R = \beta\rho_N = \beta\frac{\langle W|C^{N-1}D|V\rangle}{Z_N} = \frac{\langle W|C^{N-1}|V\rangle}{Z_N} = J. \tag{2.44}$$

Relations (2.39)–(2.41) form what is referred to as a *quadratic algebra* which amounts to a set of *reduction relations* for matrix products. The power of the quadratic algebra is that it allows one to consider any arbitrary product U of D and E matrices and determine the element $\langle W|U|V\rangle$ without recourse to an explicit representation. This is because if one has a product in which a D matrix appears before an E matrix, it will be possible to use (2.39) to reduce to the sum of two shorter products, one of which is missing the D and the other the E. If either of these terms themselves has a D before an E, (2.39) can again be used to generate further terms

in the sum. Once all terms contain only D matrices appearing after E matrices or completely lack either Ds or Es no further reduction is possible and one has found an expression for U in the form

$$U = \sum_{n,m} a_{n,m} E^n D^m \tag{2.45}$$

where $a_{n,m}$ is a positive coefficient arising from the reduction process. Once this has been achieved, evaluating the element $\langle W|U|V\rangle$ using (2.40) and (2.41) is simplicity itself:

$$\langle W|U|V\rangle = \left(\sum_{n,m} a_{n,m} \frac{1}{\alpha^n} \frac{1}{\beta^m}\right) \langle W|V\rangle. \tag{2.46}$$

Conventionally, the combination $\langle W|V\rangle$ is taken as unity, although its actual value never enters into any physical quantities, since it appears as a prefactor in both the numerator and denominator of (2.34).

One might ask whether the order with which (2.39) is applied to different DE pairs appearing in U could affect the final outcome (i.e., whether (2.39)–(2.41) describes a non-associative algebra). That this is not the case can be seen from a straightforward argument. Let U comprise at least two DE pairs: $U = U_1(DE)U_2(DE)U_3$ where $U_{1,2,3}$ are products of D and E matrices that may, or may not, contain DE pairs. Clearly, the same matrix-product expression is obtained by applying (2.39) to the two DE pairs in either order. Therefore, any non-uniqueness would have to come at the next stage where there are four strings to reduce. Now, each of these strings contains either no DE pairs (which is unique), one DE pair (which can be reduced only in one way), or is in the form given for U above. In the latter case, one can iterate the argument just given to demonstrate that the reduction is unique. Hence, the quadratic algebra (2.39)–(2.41) involving only two matrices D and E is associative. However, as we shall see in section 7, with more than two matrices the requirement for uniqueness under reduction is important and imposes restrictions on the class of models that can be solved using the matrix-product method.

The unconvinced reader may take refuge in the fact that explicit representations of the D and E matrices can be constructed—for example, that presented in section 3.2.2. In general, these matrices do not commute and have infinite dimension, as was shown in [62]. An exception to this can be determined by assuming that D and E do commute: then, using (2.39)–(2.41) one finds that

$$\langle W|(DE - ED)|V\rangle = \left(\frac{1}{\alpha} + \frac{1}{\beta} - \frac{1}{\alpha\beta}\right) \langle W|V\rangle = 0 \tag{2.47}$$

which can be true only if $\alpha+\beta = 1$. Along this line, the relations (2.39)–(2.41) can be satisfied by the scalar choices $D = 1/\beta$, $E = 1/\alpha$, $\langle W| = |V\rangle = 1$. This is equivalent to putting $\rho_i = \alpha/(\alpha+\beta)$ in (2.13), and thus along the line $\alpha+\beta = 1$ (and only this line) the mean-field theory is exact.

2.3.2. Nonequilibrium partition functions and phase transitions. In section 3.3, we will show how the reduction process just described can be applied systematically to evaluate the matrix expressions (2.37) and (2.38) for the density and current in the ASEP. Here, we quote the result for the normalization Z_N

$$Z_N = \sum_{p=1}^{N} \frac{p(2N-p-1)!}{N!(N-p)!} \frac{\left(\frac{1}{\alpha}\right)^{p+1} - \left(\frac{1}{\beta}\right)^{p+1}}{\frac{1}{\alpha} - \frac{1}{\beta}} \tag{2.48}$$

from which the current follows via (2.42). We now discuss aspects of this expression which have generated recent interest.

The relationship (2.42) between the current and the normalization is intriguing. We may interpret Z_N as the partition function for the nonequilibrium steady state. Then assuming that Z_N grows exponentially with N we find that the analogue of the thermodynamic free energy is

$$-\lim_{N\to\infty}\frac{\ln Z_N}{N}=\lim_{N\to\infty}\ln J \tag{2.49}$$

in the sense that in an equilibrium system the (Helmholtz) free-energy per site is given by

$$f=-kT\lim_{N\to\infty}\frac{\ln Z_N}{N}. \tag{2.50}$$

Thus the current plays the role of the free energy.

The phase transitions in the α–β plane, manifested by different forms for the current $J(\alpha,\beta)$, arise from different large-N asymptotic forms for Z_N and we shall determine these explicitly in section 3. Moreover, recognizing the normalization Z_N as a nonequilibrium partition function, one can analyse the Lee–Yang zeros of the partition function and show that they predict precisely the location and order of the phase transitions in the driven system [19, 65]. The interesting point is that the zeros of the partition function are in the complex plane of transition rates α, β, as compared to equilibrium problems where partition functions zeros are in the plane of some intensive thermodynamic variable such as fugacity or temperature. Also, physical behaviour characteristic of a first-order phase transition (i.e., phase coexistence) is seen at the locations where $\ln J$ has a discontinuity in its first derivative, and likewise lengthscales diverge where the second derivative is discontinuous.

As we will discuss in section 3.6, one can map the normalization (2.48) of the ASEP to the partition function of a surface in thermal equilibrium with a heat bath. In this mapping, the transition rates α and β become equilibrium fugacities, and the phase transitions in the driven system correspond to adsorption–desorption transitions in the equilibrium interpretation.

The expression for the normalization (2.48) involves combinatorial coefficients known as ballot numbers [66]. There has been considerable recent interest in developing a combinatorial understanding of the origin of these numbers and other aspects of the steady state of the ASEP [67–70].

2.3.3. Comparison between matrix-product states of equilibrium and nonequilibrium systems. In the latter sections of this review we shall discuss in detail other nonequilibrium systems whose steady states are of matrix-product form. For systems with open boundaries the ansatz used is of the form (2.35) where the number of matrices equals the number of possible states for each site. For systems with periodic boundary conditions a trace operation is used to obtain a scalar from the matrix product rather than using boundary vectors, see section 4. At this point, it is useful to establish what the differences are between these nonequilibrium steady states and the equilibrium steady states which may also be written in matrix-product form. To illustrate this we rewrite the familiar transfer matrix solution of the Ising model (an equilibrium system) as a matrix-product state.

Recall the one-dimensional Ising ferromagnet with N spins, $s_i=\pm1$ and Hamiltonian

$$\mathcal{H}=-J\sum_{\langle i,j\rangle}s_is_{i+1}. \tag{2.51}$$

To this chain let us add two fixed spins, s_L to the left of site 1 and s_R to the right of site N. The Boltzmann weight of a configuration within this system when at equilibrium with a heat bath

at inverse temperature β can be written as a matrix product analogous to that used above for the ASEP (2.35). It takes the form

$$f(s_1, s_2, \ldots, s_N; s_L, s_R) = \mathrm{e}^{-\beta\mathcal{H}} = \langle s_L | X_{s_1} X_{s_2} \cdots X_{s_N} | s_R \rangle \tag{2.52}$$

where

$$X_{+1} = \begin{pmatrix} \mathrm{e}^{\beta J} & 0 \\ \mathrm{e}^{-\beta J} & 0 \end{pmatrix}, \qquad X_{-1} = \begin{pmatrix} 0 & \mathrm{e}^{-\beta J} \\ 0 & \mathrm{e}^{\beta J} \end{pmatrix} \tag{2.53}$$

and

$$\langle s_L = +1| = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad \langle s_L = -1| = \begin{pmatrix} 0 & 1 \end{pmatrix}, \tag{2.54}$$

$$|s_R = +1\rangle = \begin{pmatrix} \mathrm{e}^{\beta J} \\ \mathrm{e}^{-\beta J} \end{pmatrix}, \qquad |s_R = -1\rangle = \begin{pmatrix} \mathrm{e}^{-\beta J} \\ \mathrm{e}^{\beta J} \end{pmatrix}. \tag{2.55}$$

Note that the algebraic relations obeyed by the matrices, such as

$$\begin{aligned} &W(S_{i-1}S_iS_{i+1} \to S_{i-1}S_i'S_{i+1})X_{S_{i-1}}X_{S_i}X_{S_{i+1}} \\ &\quad = W(S_{i-1}S_i'S_{i+1} \to S_{i-1}S_iS_{i+1})X_{S_{i-1}}X_{S_i'}X_{S_{i+1}}, \end{aligned} \tag{2.56}$$

where $W(S_{i-1}S_iS_{i+1} \to S_{i-1}S_i'S_{i+1})$ is a transition rate for flipping spin S_i, do not imply any reduction relations, rather they simply state the condition of detailed balance in the steady state.

The partition function is given by

$$Z = \langle s_L | C^N | s_R \rangle \tag{2.57}$$

where $C = X_{+1} + X_{-1}$. Calculating Z, for example by diagonalizing C, one finds the free-energy per spin

$$f = -\frac{1}{\beta} \lim_{N\to\infty} \frac{\ln Z}{N} = -\frac{1}{\beta} \ln(2 \cosh \mathrm{e}^{\beta J}). \tag{2.58}$$

Here, we note that the boundary conditions contribute only sub-extensively to the free energy and that the free energy is an analytic function of temperature—that is, there is no ordering phase transition in the 1D equilibrium Ising model. This contrasts strongly with the driven nonequilibrium steady state of the ASEP, where varying the boundary conditions induces phase transitions in bulk quantities such as the density.

Mathematically this comes as a consequence of the D and E matrices for the ASEP being infinite dimensional whilst the 1D Ising transfer matrices are finite dimensional and, because their elements are Boltzmann weights, non-negative. For matrices of this latter class one knows from the Perron–Frobenius theorem that the largest eigenvalue is nondegenerate [71–73]. This means that is not possible, as temperature is varied, for the largest and second-largest eigenvalues to cross and the functional form of the dominant contribution to the partition function to change, which would give rise to a nonanalyticity in the free energy. This state of affairs applies, in fact, to all 1D spin systems with short-range interactions [74]. On the other hand, if the interactions become long ranged, or one goes into two or more dimensions, the transfer matrices become infinite dimensional and phase transitions become a possibility.

To reiterate, the fact that one is with the matrix-product construction solving a master equation, rather than simply packaging Boltzmann factors in an attractive way, means that one-dimensional nonequilibrium steady states can show much richer and more interesting behaviour than their equilibrium counterparts. As we shall see in the course of this review, the matrices that appear in various nonequilibrium models often violate the conditions for the Perron–Frobenious theorem to hold.

2.3.4. Comparison between matrix-product states of quantum spin chains and stochastic systems. Historically, matrix-product states first appeared in the context of quantum spin chains [75, 76]. There, the wavefunction $|\Psi\rangle$ for the quantum spin chain is written as a product of matrices. Matrix-product states also serve as a starting point for approximations and variational approaches such as density matrix renormalization group (DMRG) [77–79]. Recently, matrix-product states have received a lot of attention in the context of quantum information as they are able to describe highly entangled states [80].

The connection between the formalism for stochastic systems and that for quantum systems has been clearly reviewed in [81]. Here we summarize some important features. We first note that the master equation (1.1) can be written in a suggestive way as

$$\frac{\partial}{\partial t}|P(t)\rangle = -H|P(t)\rangle \tag{2.59}$$

where $|P(t)\rangle$ is the vector of probabilities and H is a matrix formed from the transition rates. Clearly one can think of H as a Hamiltonian, $|P(t)\rangle$ as a wavefunction and (2.59) as a Schrödinger equation—this is referred to as the 'quantum Hamiltonian formalism' for stochastic systems (see, e.g., [34, 36]). The steady state of the stochastic system becomes the ground state of the quantum system and has eigenvalue 0. However, it is important to note some distinguishing features of (2.59) which are the consequences of its true stochastic nature. First, the elements of $|P(t)\rangle$ are probabilities therefore all must be positive (this contrasts with the true quantum case where the amplitude squared of the elements are probabilities). Also, as the master equation conserves probability the sum of each column of the matrix H adds to zero whereas in the true quantum case there is no such restriction.

For both a quantum spin chain and a nonequilibrium system the Hamiltonian can be written as a sum of local terms, for example representing the exchanges between neighbouring sites,

$$H = \sum_i h_{i,i+1}. \tag{2.60}$$

Now in the matrix-product states constructed for quantum spin chains (see, e.g., [82]) one finds

$$h_{i,i+1}|\Psi\rangle = 0 \tag{2.61}$$

for each local operator $h_{i,i+1}$. For a stochastic system, the simple mechanism (2.61) to obtain a zero eigenvalue for H and hence the steady state is only possible in the case of detailed balance, due to conservation of probability for each $h_{i,i+1}$. So to obtain a true nonequilibrium steady state a more complicated mechanism is required. It is this more general mechanism which is at the heart of the matrix-product formalism for nonequilibrium systems and gives rise to the algebraic relations, such as (2.39)–(2.41), which are a main focus of this review.

2.4. Shocks and second-class particles

So far we have considered the ASEP with only type of particle. In this section we discuss systems with two types of particles: first- and second-class particles. The solution of the steady state of a system containing second-class particles was the second major success using the matrix-product approach [83]. We shall discuss in detail the matrix-product solution in section 4, here we motivate the study of second-class particles by discussing some physical properties of interest.

A system containing first-class particles (denoted 1) and second-class particles (denoted 2) has dynamics

$$1\,0 \to 0\,1 \qquad 2\,0 \to 0\,2 \qquad 1\,2 \to 2\,1, \tag{2.62}$$

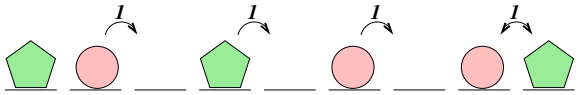


Figure 11. Dynamics of a system of first-class particles (circles) and second-class particles (pentagons).

all processes occurring with rate 1 (see figure 11). Thus, a first-class particle treats the second-class particle as vacancy but from the point of view of a vacancy the second-class particle behaves like an ordinary (first-class) particle.

Let us consider the velocity of a second-class particle in a local environment of density ρ of first-class particles. The velocity of the second-class particle, denoted as v, is given by the probability of the second-class particle hopping forward minus the probability of it hopping backward. Ignoring correlations between the second-class particle and local density the velocity is given by the density of a vacancy (in front), $1-\rho$, minus the density of first-class particles (behind), ρ,

$$v = 1 - 2\rho. \tag{2.63}$$

Beware that this approximation is not at all correct!—there is structure in the local density profile around the second-class particle, as we shall see in section 4. However, the expression (2.63) turns out to be exact and was proved by Ferrari [84]. We shall understand this result below. For the moment, note that (2.63) is precisely the velocity of a kinematic wave of density ρ discussed in section 2.2. Thus the second-class particle travels along with the kinematic wave.

As described in subsection 2.2, when two kinematic waves of different densities meet a shock may form. The velocity is $1-\rho_1-\rho_2$ where ρ_1 is the density to the left and ρ_2 is the density to the right of the shock and $\rho_2 > \rho_1$. A second-class particle will be attracted to the position of the shock since when it is to the left of the shock $1-2\rho_1 > 1-\rho_1-\rho_2$ and it catches up with the shock and when it is to the right of the shock $1-\rho_1-\rho_2 > 1-2\rho_1$ so that the shock catches up to it. Therefore we expect the second-class particle to track the shock. Moreover, one can use the second-class particle to actually define the position of the shock at the microscopic level. Then, the density profile as seen from the moving frame of the second-class particle can describe the microscopic structure of the shock [85–87].

On an infinite system exhibiting a shock, it has been shown that the density as viewed from the second-class particle approaches a steady-state distribution with density ρ_+ at $+\infty$ and density ρ_- at $-\infty$ [84, 86]. The exact structure of this steady distribution has been calculated using a matrix-product approach on an infinite system which we will review in sections 4 and 7.2.3. It demonstrates that the density profiles decay to their limits exponentially quickly in the distance from the second-class particle. An interesting case is when $\rho_+ = \rho_-$ so that there is a uniform density of particles and no apparent shock. However, in the moving frame of a second-class particle there is structure in the density profile: in front of the second-class particle the density is higher and decays to the asymptotic density according to a power law in the distance from the second-class particle; similarly behind the second-class particle the density is reduced and increases to the asymptotic density according to a power law.

Another interesting aspect of the second-class particle is how its dynamics are related to the spreading of excess mass. The central idea, termed coupling [50], is well known in the mathematical community but less so within physics. Consider two systems containing only first-class particles, identical except that one system has M particles and the other $M-1$ particles (one can consider either a finite system or an infinite system). The two systems start

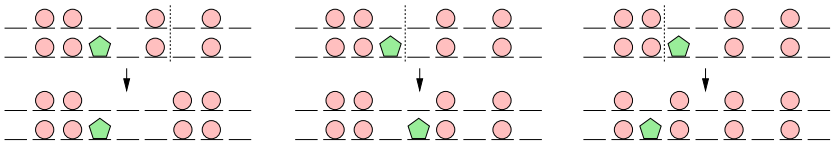


Figure 12. Illustration of coupling: the evolution of two systems one with an excess particle. The vertical lines indicate the bond which is updated in the two coupled systems. The dynamics of the excess particle (pentagon) is equivalent to that of a second-class particle—see the text for explanation.

from initial conditions differing only by the position of the extra particle in the system with M particles. In order to implement the dynamics one can consider at each timestep randomly choosing a pair of sites $i, i+1$ to update; then if there is particle at site i and a hole at site $i+1$ the particle is moved forward. In the dynamics let us choose the same pairs of sites in the two systems at each update (one can think of using the same random numbers in a Monte Carlo program).

It is easy to convince oneself that after any length of time the configurations of the two systems will differ only by the position of the extra particle (note that if we label the particles, the label of the extra particle will change under the dynamics). To see this consider the situation where the extra particle is at site i (see also figure 12). Then, if any bond except $i-1, i$ or $i, i+1$ is updated, the result will be the same for the two systems and the extra particle will remain at site i. If site $i+1$ is empty and bond $i, i+1$ is updated then the extra particle at i will hop forward. If site $i-1$ is occupied and bond $i-1, i$ is updated then the particle at $i-1$ will hop forward in the system with $M-1$ particle but not in the system with M particles. This results in the extra particle effectively hopping backwards. The latter two events result in the dynamics of the extra particle and the dynamics are precisely those of a second-class particle.

Thus, the dynamics of a system comprising $M-1$ first-class particles and one second-class particle describes the motion of an extra particle added to a system of $M-1$ particles and the dynamics of second-class particles tells us about the dynamics of mass fluctuations.

This idea can be used to recover the velocity (2.63) [83, 86]. Consider adding an extra particle to a system of density ρ, increasing the density by $\Delta\rho$. The overall current increases from $\rho(1-\rho)$ to $\rho(1-\rho)+\Delta\rho(1-2\rho)+O(\Delta\rho^2)$. Now, since the system can be thought of as a system of first-class particles and an excess particle which behaves as a second-class particle we deduce that on a large system where $\Delta\rho \to 0$ the speed of the second-class particle must be given by (2.63).

An interesting result related to the spreading of mass fluctuations is that the second-class particle exhibits *superdiffusive* behaviour. That is, with y_t defined as the distance travelled by the second-class particle, on an infinite system the variance in y_t grows as [88]

$$\langle y_t^2 \rangle - \langle y_t \rangle^2 \sim t^{4/3}. \tag{2.64}$$

On a finite periodic system of size N, with a density ρ of first-class particles, an exact calculation has shown [89]

$$\langle y_t^2 \rangle - \langle y_t \rangle^2 \simeq t N^{1/2} (\pi\rho(1-\rho))^{1/2}. \tag{2.65}$$

Thus on a finite system the behaviour is diffusive but with a diffusion constant which diverges with system size.

These results are consistent with approximate calculations such as mode coupling [90, 91] for the spreading of excess mass fluctuations which predict that the drift speed is

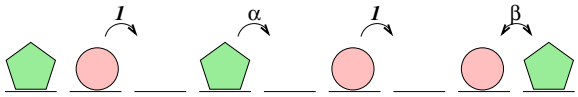


Figure 13. Dynamics of a system of first-class particles (circles) and defect particles (pentagons).

$1 - 2\rho$ as the spreading of density fluctuations around the drift grows as $t^{2/3}$ on an infinite system.

The dynamics of a second-class particle (2.62) is essentially passive in that it does not interfere with the dynamics of first-class particles. In this way the second-class particle can be though of as a 'passive scalar' for the TASEP. However, the density profile as seen from the second-class particle does have structure as explained earlier and gives insight into the structure of shocks. More recently generalizations of the second-class particle idea have been discussed in the context of passive scalars in other driven-diffusive systems [92, 93].

2.4.1. Defect particle. The dynamics of the second-class particle can easily be generalized to hop rates

$$1\,0 \xrightarrow{1} 0\,1 \qquad 2\,0 \xrightarrow{\alpha} 0\,2 \qquad 1\,2 \xrightarrow{\beta} 2\,1. \tag{2.66}$$

See also figure 13. In the general case $\beta \neq 1$ the dynamics of a two particle is no longer passive and does affect the dynamics of the normal particles. We will refer to the dynamics (2.66) as that of a defect particle. A related defect particle dynamics has been considered as an exactly solvable model of two-way traffic problem involving cars and trucks [94].

An extreme case is $\beta = 0$ where a normal particle cannot overtake the defect. In that case when $\alpha < 1$ the defect particle is slower than the other particles and a 'traffic jam' may build up behind the defect particle. Actually there is an interesting phase transition which occurs on increasing the density: for low density one has the 'traffic jam' phase where the density behind the defect particle is large and in front of the defect particle is small. For sufficiently high overall density there is a transition to a uniform density through the system. This jamming transition is mathematically analogous to Bose condensation, as we shall discuss in section 8.

For general α and β the steady state has been calculated using the matrix-product approach [95]. A phase diagram emerges for a large system with density ρ comprising four phases (see figure 22) characterized by different steady-state velocities and density profiles surrounding the defect particle. We shall discuss this phase diagram in detail in section 4.4.

2.5. The remainder of this review

In this section we have introduced the asymmetric simple exclusion process (ASEP) and some of its relatives, and through various approximate treatments (kinematic waves, mean-field theory and the extremal current principle) have got a feel for some of the interesting physics, such as shocks and phase transitions, that might emerge. We have also presented the basic ideas underlying the exact matrix-product approach that allows the existence of these phenomena to be confirmed. In the remainder of this review we flesh out these ideas considerably, presenting detailed accounts of the various methods by which matrix-product expressions can be evaluated and the physical properties of these systems fully revealed. As the review progresses, we will discuss all the main models that have so far been solved using the matrix-product method. Mostly these models are multispecies generalizations of

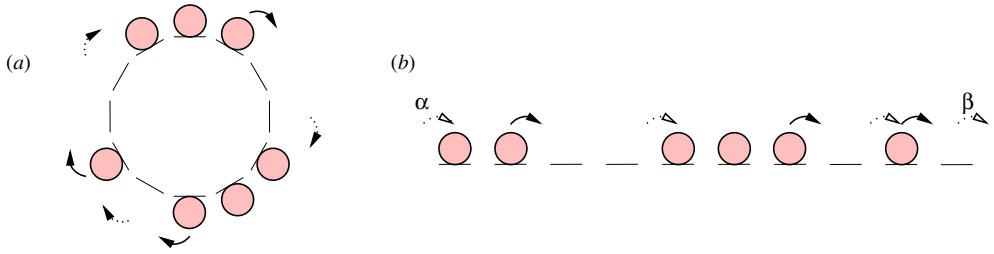


Figure 14. Transitions into (dotted arrows) and out of (solid arrows) a sample configuration for (*a*) the ASEP on a ring and (*b*) the ASEP with open boundaries. Transitions occur at unit rate unless otherwise indicated.

the ASEP in which the particle number is conserved by the dynamics (except, possibly, at the boundary sites) whilst a small class of models further admits non-conservation in the bulk. Meanwhile, some of these models can be found by means of a systematic search under certain assumptions on the existence of reduction relations for the matrices similar to (2.39)–(2.41). Others involve more complicated structures which allow, for example, the treatment of discrete-time updating schemes (such as parallel dynamics). Although in principle the matrix-product method can be used to solve a very wide class of nonequilibrium stochastic dynamical models in one dimension, solutions of many interesting systems have proved elusive. We therefore conclude our work by highlighting a few of these models in the hope that, after reading this review, the reader will be inspired to tackle these new and challenging problems.

For a full list of topics covered in sequence we refer the reader to the table of contents presented at the start of this review. We hope that this will also allow the reader to locate easily any sections that may be of particular interest.

3. Detailed analysis of the ASEP with open boundaries

Having introduced the ASEP with open boundaries (see figure 3) and the matrix-product expressions for the distribution (2.35), density (2.37) and current (2.38) in the steady state we now explain in detail how these are calculated using the matrix reduction relations (2.39)–(2.41). First, though, we must demonstrate that the distribution implied by (2.35) and (2.39)–(2.41) is, in fact, the stationary solution of the master equation for the process.

3.1. Proof of the reduction relations

3.1.1. Domain-based proof. As a first step towards understanding why the matrix-product solution of the ASEP with open boundaries given above is correct, we consider the simpler case of the ASEP on a ring. A typical configuration is shown in figure 14(*a*). Recall that the master equation (1.1) for the probability of being in configuration $\mathcal{C}$ has two contributions: a gain term from transitions into $\mathcal{C}$ and a loss term from transitions into another state. In the steady state, we require these loss and gain terms to balance.

As is evident from figure 14(*a*) there is for each domain of particles on the ring one way to enter a given configuration $\mathcal{C}$ (by a particle joining the end of the domain) and one way to exit (by a particle leaving the front of the domain). Therefore, we have for any given

configuration $\mathcal{C}$

$$\sum_{\mathcal{C}'} W(\mathcal{C}' \to \mathcal{C}) = \sum_{\mathcal{C}'} W(\mathcal{C} \to \mathcal{C}'). \tag{3.1}$$

The loss and gain terms in the master equation (1.1) thus cancel if $P(\mathcal{C}) = \text{const}\,\forall \mathcal{C}$. Putting this another way, every gain term $P(\mathcal{C}')W(\mathcal{C}' \to \mathcal{C})$ exactly balances one of the loss terms $P(\mathcal{C})W(\mathcal{C} \to \mathcal{C}'')$. Note this is similar, but not the same, as the cancellation that occurs when detailed balance (1.2) is satisfied since $\mathcal{C}' \neq \mathcal{C}''$. This more general cancellation scheme is sometimes referred to as *dynamic reversibility* [11] or *pairwise balance* [96].

When one has open boundaries, as in figure 3, the number of ways into a given configuration does not always equal the number of ways out. For example, when the left boundary site is occupied and the right boundary empty, the number of ways in exceeds the numbers out, and furthermore not all hops occur at the same rate (see figure 14(*b*)). The strategy to finding the steady-state solution of the ASEP is to have a *partial* cancellation between the terms corresponding to particles attaching to the rear or detaching from the front of domains in the bulk. The remainder is then cancelled with terms coming from the boundary interactions. As we now describe by reference to particular examples, it is this cancellation scheme that is implied by the matrix relations (2.39)–(2.41).

For our illustrative purposes, let us consider a particular six-site configuration $\_\circ\circ\_\circ\_$ for which the master equation reads

$$\begin{aligned}\frac{\partial}{\partial t} f(\_\circ\circ\_\circ\_) = {} & -\alpha f(\overset{\circ}{\curvearrowright}\_\circ\circ\_\circ\_) + f(\circ\overset{\curvearrowright}{\_}\circ\_\circ\_) - f(\_\circ\circ\overset{\curvearrowright}{\_}\circ\_) \\ & + f(\_\circ\circ\circ\overset{\curvearrowright}{\_}\_) - f(\_\circ\circ\_\circ\overset{\curvearrowright}{\_}) + \beta f(\_\circ\circ\_\circ\circ\overset{\curvearrowright}{\ }).\end{aligned} \tag{3.2}$$

In the gain terms (i.e., those with a positive sign), the arrows indicate the hop that leads to the configuration on the left-hand side; in the loss terms, the arrows indicate the transitions by which this configuration can be exited. On the right-hand side we insert now the matrix-product expressions:

$$\begin{aligned}\frac{\partial}{\partial t} f(\_\circ\circ\_\circ\_) = {} & -\alpha\langle W|\underline{E}DDEDE|V\rangle + \langle W|\underline{DE}DEDE|V\rangle - \langle W|ED\underline{DE}DE|V\rangle \\ & + \langle W|EDD\underline{DE}E|V\rangle - \langle W|EDDE\underline{DE}|V\rangle + \beta\langle W|EDDEDD|\underline{V}\rangle.\end{aligned} \tag{3.3}$$

Here we have underlined the mathematical objects (vectors or matrices) that correspond to the sites or boundary reservoirs involved the transitions indicated with arrows in (3.2). We note that all these underlined terms are of the form $\langle W|E$, DE or $D|V\rangle$ and so can be reduced using (2.39)–(2.41). Using these reduction relations we find

$$\begin{aligned}\frac{\partial}{\partial t} f(\_\circ\circ\_\circ\_) = {} & -\langle W|DDEDE|V\rangle \\ & + \langle W|\underline{D}DEDE|V\rangle + \langle W|\underline{E}DEDE|V\rangle - \langle W|ED\underline{E}DE|V\rangle \\ & - \langle W|ED\underline{D}DE|V\rangle + \langle W|EDD\underline{D}E|V\rangle + \langle W|EDD\underline{E}E|V\rangle \\ & - \langle W|EDDE\underline{E}|V\rangle - \langle W|EDDE\underline{D}|V\rangle + \langle W|EDDED|\underline{V}\rangle,\end{aligned} \tag{3.4}$$

where the underlined vectors and matrices show what remains after using the reduction relations. We see that there is a cancellation between the first and second term in each successive pair of terms. Hence the weight $f(\_\circ\circ\_\circ\_) = \langle W|EDDEDE|V\rangle$ is stationary, as claimed. If one has a configuration comprising n domains of particles, none of which contact either boundary, one finds a master equation of the form (3.2) with n pairs of terms like those in the second and third lines of (3.2), and the same cancellation mechanism goes through. For the case of the empty lattice, $n = 0$, one can verify that the term arising from the left boundary cancels with that from the right.

It only remains to check that weights corresponding to configurations with particles on the left, right or both boundary sites are also stationary. Let us consider a configuration that has a domain of particles at the left boundary. For such a configuration, the master equation begins with the terms

$$\frac{\partial}{\partial t} f(\circ\circ\circ\_\cdots) = \alpha f(\overset{\circ}{\_}\circ\circ\_\cdots) - f(\circ\circ\overset{\frown}{\circ}\_\cdots) + \cdots \tag{3.5}$$

$$= \alpha\langle W|\underline{E}DDE\cdots|V\rangle - \langle W|DD\underline{DE}\cdots|V\rangle + \cdots \tag{3.6}$$

$$= \langle\underline{W}|DDE\cdots|V\rangle - \langle W|DD\underline{E}\cdots|V\rangle - \langle W|DD\underline{D}\cdots|V\rangle + \cdots. \tag{3.7}$$

The first two terms on the last line cancel, leaving a single negative term. One can also verify that this term will cancel one coming from the next domain of particles on the lattice or from the right boundary. Similarly, if one has a domain of particles in contact with the right boundary, after using the reduction relations and performing a cancellation, a single positive term remains that cancels with one coming from the previous domain of particles on the lattice or from the left boundary. Finally, after checking that the matrix-product expression for the statistical weight of fully occupied lattice is also stationary, we can conclude that the matrix-product weights, defined by equation (2.35), in conjunction with the reduction relations (2.39)–(2.41), give the steady-state solution of the ASEP, as previously claimed.

3.1.2. Algebraic proof. We now show how this pictorial representation of the cancellation mechanism between domains can be encoded in an algebraic way. This will pave the way to understanding the general results of [97] that we review in section 9.

The master equation can be written as a sum of terms, each relating either to a pair of neighbouring sites or to one of the end sites, i.e., as

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\tau_1, \tau_2, \ldots, \tau_N) = \left[\hat{h}_L + \sum_{i=1}^{N-1} \hat{h}_{i,i+1} + \hat{h}_R\right] f(\tau_1, \tau_2, \ldots, \tau_N) \tag{3.8}$$

The operators $\hat{h}$ construct the required gain and loss terms in the master equation via

$$\hat{h}_{i,i+1} f(\cdots, \tau_i, \tau_{i+1}, \cdots) = W(\tau_{i+1}\tau_i \to \tau_i\tau_{i+1}) f(\cdots, \tau_{i+1}, \tau_i, \cdots) - W(\tau_i\tau_{i+1} \to \tau_{i+1}\tau_i) f(\cdots, \tau_i, \tau_{i+1}, \cdots) \tag{3.9}$$

$$\hat{h}_L f(\tau_1, \cdots) = W_1(1-\tau_1 \to \tau_1) f(1-\tau_1, \cdots) - W_1(\tau_1 \to 1-\tau_1) f(\tau_1, \cdots) \tag{3.10}$$

$$\hat{h}_R f(\cdots, \tau_N) = W_N(1-\tau_N \to \tau_N) f(\cdots, 1-\tau_N) - W_N(\tau_N \to 1-\tau_N) f(\cdots, \tau_N) \tag{3.11}$$

in which $W(\tau\tau' \to \tau'\tau)$ gives the rate at which a neighbouring pair of sites in configuration (τ, τ') exchange, $W_1(\tau_1 \to 1-\tau_1)$ gives the rate at which site 1 changes state and $W_N(\tau_N \to 1-\tau_N)$ gives the rate at which site N changes state. For the ASEP we have

$$W(10 \to 01) = 1 \tag{3.12}$$

$$W_1(0 \to 1) = \alpha \tag{3.13}$$

$$W_N(1 \to 0) = \beta \tag{3.14}$$

and all other rates zero.

In terms of matrix-product expressions (2.35) for the weights, (3.9)–(3.11) become

$$\hat{h}_{i,i+1}\langle W|\cdots X_{\tau_{i-1}}[X_{\tau_i}X_{\tau_{i+1}}]X_{\tau_{i+2}}\cdots|V\rangle = \langle W|\cdots X_{\tau_{i-1}}[W(\tau_{i+1}\tau_i \to \tau_i\tau_{i+1})X_{\tau_{i+1}}X_{\tau_i} - W(\tau_i\tau_{i+1} \to \tau_{i+1}\tau_i)X_{\tau_i}X_{\tau_{i+1}}]X_{\tau_{i+2}}\cdots|V\rangle \tag{3.15}$$

$$\hat{h}_L\langle W|X_{\tau_1}\cdots|V\rangle = \langle W|\left[W_1(1-\tau_1 \to \tau_1)X_{1-\tau_1} - W_1(\tau_1 \to 1-\tau_1)X_{\tau_1}\right]\cdots|V\rangle \tag{3.16}$$

$$\hat{h}_R\langle W|\cdots\left[W_N(1-\tau_N \to \tau_N)X_{1-\tau_N} - W_N(\tau_N \to 1-\tau_N)X_{\tau_N}\right]|V\rangle \tag{3.17}$$

where we have stretched the notation so that, for example, $\hat{h}_{i,i+1}$ now acts on the matrices $X_{\tau_i}, X_{\tau_{i+1}}$.

Meanwhile, we suppose there exists a set of auxiliary matrices $\tilde{X}_i$ such that additionally

$$\hat{h}_{i,i+1}\langle W|\cdots X_{\tau_i}X_{\tau_{i+1}}\cdots|V\rangle = \langle W|\cdots\left[\tilde{X}_{\tau_i}X_{\tau_{i+1}} - X_{\tau_i}\tilde{X}_{\tau_{i+1}}\right]\cdots|V\rangle \tag{3.18}$$

$$\hat{h}_L\langle W|X_{\tau_1}\cdots|V\rangle = \left[-\langle W|\tilde{X}_{\tau_1}\right]\cdots|V\rangle \tag{3.19}$$

$$\hat{h}_R\langle W|\cdots X_{\tau_N}|V\rangle = \langle W|\cdots\left[\tilde{X}_{\tau_N}|V\rangle\right]. \tag{3.20}$$

If this is the case then the sum in (3.8) telescopes to zero (i.e., there is a pairwise cancellation of terms). Therefore, if the right-hand sides of (3.15)–(3.17) and (3.18)–(3.20) can be shown to be equal, it follows that a stationary solution of the master equation for the process has been determined. Equating the terms in square brackets in (3.15)–(3.17) and (3.18)–(3.20) is sufficient to ensure overall equality of the two expressions. That is, if

$$W(\tau_{i+1}\tau_i \to \tau_i\tau_{i+1})X_{\tau_{i+1}}X_{\tau_i} - W(\tau_i\tau_{i+1} \to \tau_{i+1}\tau_i)X_{\tau_i}X_{\tau_{i+1}} = \tilde{X}_{\tau_i}X_{\tau_{i+1}} - X_{\tau_i}\tilde{X}_{\tau_{i+1}} \tag{3.21}$$

$$W_1(1-\tau_1 \to \tau_1)\langle W|X_{1-\tau_1} - W_1(\tau_1 \to 1-\tau_1)\langle W|X_{\tau_1} = -\langle W|\tilde{X}_{\tau_1} \tag{3.22}$$

$$W_N(1-\tau_N \to \tau_N)X_{1-\tau_N}|V\rangle - W_N(\tau_N \to 1-\tau_N)X_{\tau_N}|V\rangle = \tilde{X}_{\tau_N}|V\rangle. \tag{3.23}$$

Note that (3.21)–(3.23) amount to a total of eight conditions which become with the set of rates (3.12)–(3.14)

$$0 = \tilde{D}D - D\tilde{D} = \tilde{E}E - E\tilde{E} \tag{3.24}$$

$$DE = \tilde{E}D - E\tilde{D} = -\tilde{D}E + D\tilde{E} \tag{3.25}$$

$$\alpha\langle W|E = -\langle W|\tilde{D} = \langle W|\tilde{E} \tag{3.26}$$

$$\beta D|V\rangle = \tilde{E}|V\rangle = -\tilde{D}|V\rangle. \tag{3.27}$$

With a *scalar* choice for the auxiliaries, namely $\tilde{D} = -1$ and $\tilde{E} = 1$, these reduce to the three equations (2.39)–(2.41) as required.

The algebraic proof can be easily generalized to other models, for example those with more general bulk dynamics than exchanges, as shall be carried out in section 9.

3.2. Calculation of the nonequilibrium partition function Z_N

Our first calculational task is to compute Z_N given by (2.36), from which the current follows via (2.42). There are several approaches to the task and these are representative of the three main approaches to calculation which will be employed at different points in this review: direct matrix reordering, diagonalization of an explicit representation of the matrices and generating function techniques.

3.2.1. Direct calculation by matrix reordering. In section 2.3.1 we explained how the reduction relations can, in principle, be used repeatedly to bring any arbitrary matrix product U into a standard form (2.45) which can then be straightforwardly evaluated (2.46). We now put this principle into practice, and begin with the product $C^N = (D+E)^N$ required to calculate the normalization (2.36) and thence the current (2.38). By performing the reduction longhand for the first few N, one finds

$$C = D + E, \tag{3.28}$$

$$C^2 = DD + DE + ED + EE = DD + ED + EE + D + E, \tag{3.29}$$

$$\begin{aligned} C^3 &= DDD + DDE + DED + DEE + EDD + EDE + EED + EEE \\ &= DDD + EDD + EED + EEE \\ &\quad + DD + DE + DD + ED + DE + EE + ED + EE \\ &= DDD + EDD + EED + EEE + 2(DD + ED + EE) + 2(D + E). \end{aligned} \tag{3.30}$$

Eventually one is drawn to the general formula

$$C^N = \sum_{p=0}^{N} \frac{p(2N-p-1)!}{N!(N-p)!} \sum_{q=0}^{p} E^q D^{p-q}, \tag{3.31}$$

which can be proved by induction [62]. For this two preliminary identities are required. First,

$$D^N C = D + D^2 + \cdots + D^{N+1} + E, \tag{3.32}$$

which can be proved inductively by multiplying both sides from the left by D and using (2.39). The second identity involves the combinatorial factors

$$B_{N,p} = \begin{cases} \dfrac{p(2N-p-1)!}{N!(N-p)!} & 0 < p \leqslant N \\ 0 & \text{otherwise.} \end{cases} \tag{3.33}$$

It reads

$$B_{N,p} - B_{N,p+1} = B_{N-1,p-1} \qquad \forall\, N > 1, \quad p > 0 \tag{3.34}$$

which can be verified by substituting into the explicit expressions from (3.33).

We now assume that the identity (3.31) for C^N has been shown to hold for some N and derive an expression for C^{N+1} by multiplying from the right with C. Then, (3.32) implies

$$C^{N+1} = \sum_{p=0}^{N} B_{N,p} \sum_{q=0}^{p} [E^{q+1} + E^q (D + D^2 + \cdots + D^{p-q+1})]. \tag{3.35}$$

Now, using (3.34) we find

$$\begin{aligned} C^{N+1} &= \sum_{p=1}^{N+1} B_{N+1,p} \sum_{q=0}^{p-1} [E^{q+1} + E^q (D + D^2 + \cdots + D^{p-q})] \\ &\quad - \sum_{p=2}^{N+1} B_{N+1,p} \sum_{q=0}^{p-2} [E^{q+1} + E^q (D + D^2 + \cdots + D^{p-q-1})] \end{aligned} \tag{3.36}$$

$$= \sum_{p=1}^{n+1} B_{N+1,p} [E^p + E^{p-1} D + E^{p-2} D^2 + \cdots + D^p] \tag{3.37}$$

which is equal to the right-hand side of (3.31) with $N \to N+1$. Since (3.31) is obviously correct for $N = 1$ ($C = D + E$), it follows by induction that it is true for all N. Using (2.40), (2.41) and the convention $\langle W|V\rangle = 1$ we immediately find the exact expression

$$Z_N = \sum_{p=0}^{N} B_{N,p} \sum_{q=0}^{p} \left(\frac{1}{\alpha}\right)^q \left(\frac{1}{\beta}\right)^{p-q} = \sum_{p=0}^{N} B_{N,p} \frac{\left(\frac{1}{\alpha}\right)^{p+1} - \left(\frac{1}{\beta}\right)^{p+1}}{\frac{1}{\alpha} - \frac{1}{\beta}}. \tag{3.38}$$

3.2.2. Diagonalization of an explicit representation. So far we have not given any explicit representations of the matrices D, E and vectors $\langle W|$, $|V\rangle$ that satisfy (2.39)–(2.41). Several possible representations were proposed in [62]. For example, one can take

$$D = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 1 & 0 & \\ 0 & 0 & 1 & 1 & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \qquad E = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & \\ 0 & 1 & 1 & 0 & \\ 0 & 0 & 1 & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{3.39}$$

$$\langle W| = \kappa(1, \quad a, \quad a^2, \quad \dots) \qquad |V\rangle = \kappa \begin{pmatrix} 1 \\ b \\ b^2 \\ \vdots \end{pmatrix}, \tag{3.40}$$

where

$$a = \frac{1-\alpha}{\alpha} \qquad b = \frac{1-\beta}{\beta}. \tag{3.41}$$

$\kappa^2 = (\alpha + \beta - 1)/\alpha\beta$ is chosen to ensure that $\langle W|V\rangle = 1$. Note that for this we require $\alpha + \beta > 1$, but as we shall discuss later (in section 5.3) one can analytically continue results to $\alpha + \beta \leqslant 1$ in a straightforward manner. Some other representations are presented in appendix D.

Let us consider the eigenvectors of C, given in this representation by

$$C = \begin{pmatrix} 2 & 1 & 0 & 0 & \cdots \\ 1 & 2 & 1 & 0 & \\ 0 & 1 & 2 & 1 & \\ 0 & 0 & 1 & 2 & \\ \vdots & & & & \ddots \end{pmatrix}. \tag{3.42}$$

We parametrize the eigenvalues by $2(1 + \cos\theta)$ where $0 < \theta < 2\pi$ and denote the associated eigenvector by $|\cos\theta\rangle$ so that

$$C|\cos\theta\rangle = 2(1 + \cos\theta)|\cos\theta\rangle \tag{3.43}$$

where

$$|\cos\theta\rangle = \frac{1}{\sin\theta} \begin{pmatrix} \sin\theta \\ \sin 2\theta \\ \sin 3\theta \\ \vdots \end{pmatrix}. \tag{3.44}$$

Note that the elements of $|\cos\theta\rangle$ are precisely the Chebyshev polynomials of the second kind [98]. It is easy to show (by summing geometric series) that

$$\langle W|\cos\theta\rangle = \frac{\kappa}{\sin\theta}\sum_{n=0}^{\infty} a^n \sin(n+1)\theta = \frac{\kappa}{(1-a\,\mathrm{e}^{\mathrm{i}\theta})(1-a\,\mathrm{e}^{-\mathrm{i}\theta})} \tag{3.45}$$

and $\langle\cos\theta|V\rangle$ is given by a similar expression with a replaced by b. Also it is easy to show (using orthogonality of $\sin n\theta$ for different n) that

$$\mathbb{1} = \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta \sin^2\theta |\cos\theta\rangle\langle\cos\theta|. \tag{3.46}$$

Therefore Z_N can be evaluated as follows:

$$Z_N = \langle W|C^N|V\rangle \tag{3.47}$$

$$= \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\theta \sin^2\theta \langle W|C^N|\cos\theta\rangle\langle\cos\theta|V\rangle \tag{3.48}$$

$$= \frac{\kappa^2}{\pi}\int_0^{2\pi} \mathrm{d}\theta \frac{\sin^2\theta[2(1+\cos\theta)]^N}{(1-a\,\mathrm{e}^{\mathrm{i}\theta})(1-a\,\mathrm{e}^{-\mathrm{i}\theta})(1-b\,\mathrm{e}^{\mathrm{i}\theta})(1-b\,\mathrm{e}^{-\mathrm{i}\theta})}. \tag{3.49}$$

This integral representation of the normalization can be shown to be equivalent to (3.38) by using the calculus of residues (see appendix C).

3.2.3. Generating function approach. Let us define the generating function of the normalization as

$$\mathcal{Z}(z) = \sum_{N=0}^{\infty} z^N Z_N. \tag{3.50}$$

In order to determine this quantity in a fast way the idea is to work with formal power series involving matrices, i.e., expressions of the form

$$f(U) = \sum_{n=0}^{\infty} a_n U^n \tag{3.51}$$

where U is a matrix. The 'formal' component of this approach lies in the fact that we do not worry about convergence properties of these series. Rather, we simply use such expressions as a device for manipulating matrix-product expressions. For example, one might multiply both sides by some other formal series in a matrix V, bearing in mind that if U and V are noncommuting, one must multiply both sides from the same direction.

Consider the formal series

$$\frac{1}{1-zC} = \sum_{n=0}^{\infty} z^n C^n \tag{3.52}$$

so that the generating function of the normalization (3.50) can be written as

$$\mathcal{Z}(z) = \langle W|\frac{1}{1-zC}|V\rangle. \tag{3.53}$$

As was observed by Depken [99], equation (2.39) has the simple consequence that

$$(1-\eta D)(1-\eta E) = 1-\eta(D+E)+\eta^2 DE = 1-\eta(1-\eta)C. \tag{3.54}$$

Putting $z = \eta(1-\eta)$, and multiplying both sides from the left by $\frac{1}{1-\eta E}\frac{1}{1-\eta D}$ and from the right by $\frac{1}{1-zC}$ one learns that

$$\frac{1}{1-zC} = \frac{1}{1-\eta E}\frac{1}{1-\eta D}. \tag{3.55}$$

Using (2.40) and (2.41) one then finds that

$$\mathcal{Z}(z) = \left(1-\frac{\eta(z)}{\alpha}\right)^{-1}\left(1-\frac{\eta(z)}{\beta}\right)^{-1}. \tag{3.56}$$

Note that the condition $\mathcal{Z}(0) = 1$ implies $\eta(0) = 0$ which means that we must choose the root

$$\eta(z) = \tfrac{1}{2}(1-\sqrt{1-4z}). \tag{3.57}$$

The generating function approach provides a very fast route to the phase diagram and current via the asymptotic analysis of (3.56) that we perform in section 3.4. Furthermore, the exact expression (3.38) can also be recovered with very little work through an application of the Lagrange inversion formula for generating functions which we recap in appendix B. As indicated in appendix B, this formula can be applied because the combination $\eta(z)/z$ can be expressed in terms of $\eta(z)$ alone, i.e., its dependence on z enters only implicitly through the argument of $\eta(z)$. Then, the Lagrange inversion formula implies that the normalization Z_N is given by

$$\{z^N\}\mathcal{Z}[\eta(z)] = \frac{1}{N}\{\eta^{N-1}\}\left(\frac{\mathrm{d}}{\mathrm{d}\eta}\frac{1}{1-\frac{\eta}{\alpha}}\frac{1}{1-\frac{\eta}{\beta}}\right)\left(\frac{1}{1-\eta}\right)^N \tag{3.58}$$

$$= -\frac{1}{N}\left(\alpha\frac{\partial}{\partial\alpha}+\beta\frac{\partial}{\partial\beta}\right)\{\eta^N\}\frac{1}{1-\frac{\eta}{\alpha}}\frac{1}{1-\frac{\eta}{\beta}}\left(\frac{1}{1-\eta}\right)^N \tag{3.59}$$

$$= \sum_{p=0}^{N}\frac{p(2N-p-1)!}{N!(N-p)!}\sum_{q=0}^{p}\left(\frac{1}{\alpha}\right)^p\left(\frac{1}{\beta}\right)^{q-p} \tag{3.60}$$

where we introduce an important notation:

$\{x^n\}f(x)$ denotes the coefficient of x^n in the formal power series $f(x)$.

Expression (3.60) is in perfect agreement with (3.38), as required.

3.3. Exact density profile from direct matrix reordering

In order to calculate the density profile we use the approach of matrix reordering. This has the advantage of giving an exact finite sum expression (3.65) for finite systems. We start from (2.37), which contains the matrix product DC^j. Again, one can perform the reduction manually for small j (we shall omit this here) and deduce the identity

$$DC^j = \sum_{k=0}^{j-1} B_{k+1,1}C^{j-k} + \sum_{k=2}^{j+1} B_{j,k-1}D^k. \tag{3.61}$$

This was proved in [62], again using an inductive argument. Assuming (3.61) holds for some $j \geqslant 1$, we find this implies for $j+1$ that

$$(DC^j)C = \sum_{k=0}^{j-1} B_{k+1,1}C^{j+1-k} + \sum_{k=2}^{j+1} B_{j,k-1}D^kC \tag{3.62}$$

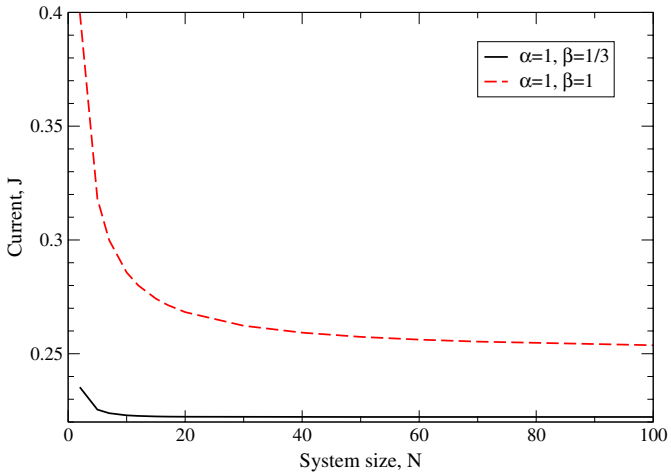


Figure 15. Particle current in the ASEP at finite system size for two different combinations of α and β.

$$= \sum_{k=0}^{j-1} B_{k+1,1} C^{j+1-k} + \sum_{k=2}^{j+1} (B_{j+1,k} - B_{j+1,k+1})(C + D^2 + D^3 + \cdots + D^{k+1}) \quad (3.63)$$

where both (3.32) and (3.34) have been used. Since, for $N \geqslant 1$, $B_{N+1,1} = B_{N+1,2}$, a little more algebra reveals that

$$DC^{j+1} = \sum_{k=0}^{(j+1)-1} B_{k+1,1} C^{(j+1)-k} + B_{j+1,1} D^2 + B_{j+1,2} D^3 + \cdots + B_{j+1,j+1} D^{j+2} \quad (3.64)$$

which is (3.61) with $j \to j+1$ as we desired. The case $j = 1$ is easily verified using the reduction relation (2.39), which completes the proof of (3.61) for all j. An expression for the density then follows by substituting (3.61) into (2.37). We find

$$\rho_i = \sum_{n=1}^{N-i} B_{n,1} \frac{Z_{N-n}}{Z_N} + \frac{Z_{i-1}}{Z_N} \sum_{p=1}^{N-i} B_{N-i,p} \frac{1}{\beta^{p+1}} \quad (3.65)$$

for $1 \leqslant i < N$. Note that the case $i = N$ is special: $\rho_N = Z_{N-1}/(\beta Z_N)$.

In the next subsection, we shall analyse the large-N behaviour of the current and density profiles. First, though, we remark that formulae (3.38) and (3.65) are of utility if one wishes to study *finite* systems, e.g., in the context of the biophysical applications of the ASEP where the thermodynamic limit is not appropriate. For example, the particle current as a function of system size is plotted in figure 15 for the combinations $\alpha = \beta = 1$ and $\alpha = 1, \beta = \frac{1}{3}$. The figure is suggestive that the decay to the asymptotic current with increasing system size is slower for the case $\alpha = \beta = 1$ compared to $\alpha = 1, \beta = \frac{1}{3}$. As one might guess from the mean-field treatment of the ASEP in section 2, these different decay forms arise from being in different phases of the model.

3.4. Asymptotic analysis of the current and density profiles

In order to obtain the exact phase diagram for the ASEP, it is necessary to determine the asymptotics (large-N forms) of the current and density profiles. There are a number of ways

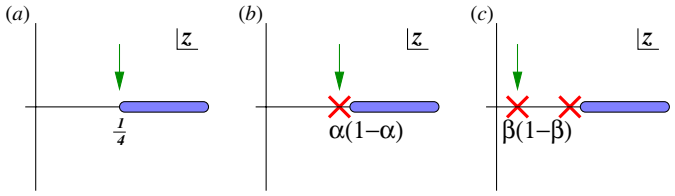


Figure 16. Poles (crosses) and branch cuts (solid lines) in the complex-z plane of the generating function $\mathcal{Z}(z;\alpha,\beta)$ at three different values of α and β. (*a*) $\alpha > \frac{1}{2}, \beta > \frac{1}{2}$; (*b*) $\alpha < \frac{1}{2}, \beta > \frac{1}{2}$ and (*c*) $\alpha < \frac{1}{2}, \beta < \frac{1}{2}$ and $\beta < \alpha$. In each case the singularity dominating the asymptotics of the normalization is the one closest to the origin and indicated with an arrow.

to achieve this. A method that is systematic and easy to apply in a wide range of cases is to construct the generating functions of the desired quantities. Then one can apply standard techniques for extracting the asymptotics of the coefficients from a generating function (see, e.g., [100]). For convenience, we have recapitulated these techniques in appendix B.

It will also be useful to have at our disposal some generating functions for common combinatorial quantities. We first note that $\eta(z)$ (3.57) that appears in (3.56) is actually (up to a factor of z) the generating function for Catalan numbers:

$$\eta(z) = \sum_{N=1}^{\infty} z^N \frac{1}{N} \binom{2(N-1)}{N-1}. \tag{3.66}$$

The Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ appear in a remarkably large number of combinatorial contexts [101] and will also be used many times in the course of this review article. Equation (3.66) can be seen by expanding the square root in (3.57) in powers of z. It is also instructive to consider the generating function for ballot numbers

$$G_p(z) = \sum_{N=p}^{\infty} z^N B_{N,p} \tag{3.67}$$

where the coefficients $B_{N,p}$ are defined in (3.33), and we have that $G_0(z) = 1$ and $G_1(z) = \eta(z)$. It turns out that $G_p(z)$ is simply the pth power of $\eta(z)$. To see this, note that (3.34) implies

$$G_p(z) = G_{p+1}(z) + zG_{p-1}(z) \tag{3.68}$$

and so that if $G_p(z) = [\eta(z)]^p$ one must have $\eta(z)[1-\eta(z)] = z$ which is indeed satisfied by (3.57).

As noted in appendix B, the asymptotics of the coefficients of a generating function (here, the normalization Z_N) are dominated by the singularity nearest the origin. Now, the function $\eta(z)$ has a square-root singularity at $z = \frac{1}{4}$ which is always present. Additionally there is a pole at $\eta(z) = \alpha$; however, this is present on the positive branch of the square root that appears in $\eta(z)$ *only* if $\alpha < \frac{1}{2}$. Similarly, there is a pole at $\eta(z) = \beta$ if $\beta < \frac{1}{2}$. As $\eta(z)$ increases monotonically from zero in the range $0 \leqslant z \leqslant \frac{1}{4}$, the dominant singularity is the square-root singularity at $z = \frac{1}{4}$ if both α and β exceed $\frac{1}{2}$. Otherwise, whichever of α or β is smaller supplies the dominant singularity—see figure 16. Note that the regions within which particular singularities dominate coincide with the phases previously established using a mean-field approach, as shown in figure 5. In other words, this analysis of the generating function reveals the phase diagram to be exact.

In the region identified as the low-density phase from the mean-field calculation, $\alpha < \beta, \alpha < \frac{1}{2}$, the pole in (3.56) at $z_0 = \alpha(1-\alpha)$ implies—via the results presented in appendix B—that

$$Z_N(\alpha,\beta) \sim \left[\lim_{z\to z_0}\left(1-\frac{z}{z_0}\right)\mathcal{Z}(z;\alpha,\beta)\right] z_0^{-N} \tag{3.69}$$

$$\sim \frac{\alpha\beta}{z_0\eta'(z_0)\,[\beta-\eta(z_0)]} z_0^{-N} \tag{3.70}$$

$$\sim \frac{\alpha\beta(1-2\alpha)}{(\beta-\alpha)}\,[\alpha(1-\alpha)]^{-N-1}\,. \tag{3.71}$$

Using (2.42), one finds that the current is $J = z_0 = \alpha(1-\alpha)$. In the high-density phase, $\beta < \alpha, \beta < \frac{1}{2}$ the pole at $z_0 = \beta(1-\beta)$ dominates and one obtains the same expression for the normalization but with α and β exchanged. This is in accordance with the particle–hole symmetry of the model. In the maximal-current phase, where both α and β are greater than $\frac{1}{2}$, we expand up the square root in (3.56) to find

$$\mathcal{Z}(z;\alpha,\beta) = \frac{4\alpha\beta}{(2\alpha-1)(2\beta-1)}\left(1-\left[\frac{1}{2\alpha-1}+\frac{1}{2\beta-1}\right]\sqrt{1-4z}+\mathrm{O}(1-4z)\right). \tag{3.72}$$

It then follows from the results of appendix B (equation (B.7)) that

$$Z_N(\alpha,\beta) \sim \frac{4\alpha\beta(\alpha+\beta-1)}{\sqrt{\pi}(2\alpha-1)^2(2\beta-1)^2}\frac{4^N}{N^{3/2}} \tag{3.73}$$

which can be shown to agree with the expression given in [62] after a little algebra. In this phase, the current $J = z_0 = \frac{1}{4}$, thus showing that the mean-field predictions for the asymptotic current in all three phases, as presented in table 1, are exact.

We now turn to the exact formula for the density profile (3.65) which for convenience we give again here:

$$\rho_i = \sum_{n=1}^{N-i} B_{n,1}\frac{Z_{N-n}}{Z_N} + \frac{Z_{i-1}}{Z_N}\sum_{p=1}^{N-i} B_{N-i,p}\frac{1}{\beta^{p+1}}. \tag{3.74}$$

Although formidable at a first glance, its analysis is in fact reasonably straightforward when both N and i are large. More precisely, we shall take $N \to \infty$ and $i \to \infty$ with the distance $j = N - i$ from the right boundary fixed. It is easily verified that in this limit one has in all phases

$$\lim_{N\to\infty}\frac{Z_{N-j}}{Z_N} = J^j \tag{3.75}$$

for any fixed finite j. Recall that J is the current in the thermodynamic limit, as determined above. Thus

$$\lim_{N\to\infty}\rho_{N-j} = \sum_{n=1}^{j} B_{n,1}J^n + \frac{J^{j+1}}{\beta}\sum_{p=1}^{j}\frac{B_{j,p}}{\beta^p}. \tag{3.76}$$

The second sum in this expression we have seen before: it is simply the normalization for the ASEP (2.48) in the limit $\alpha \to \infty$. Hence, for $j \gg 1$ we can use the asymptotic analyses from

above to determine that

$$\sum_{p=1}^{j} \frac{B_{j,p}}{\beta^p} \sim \begin{cases} \dfrac{\beta(1-2\beta)}{[\beta(1-\beta)]^{j+1}} & \beta < \dfrac{1}{2} \\ \dfrac{4^j}{\sqrt{\pi j}} & \beta = \dfrac{1}{2} \\ \dfrac{\beta 4^j}{\sqrt{\pi}(2\beta-1)^2 j^{3/2}} & \beta > \dfrac{1}{2}. \end{cases} \tag{3.77}$$

The case $\beta = \frac{1}{2}$ is special because the pole and square-root singularities in the generating function coincide, thus modifying the nature of the dominant singularity.

The first sum can be analysed through its generating function in a manner similar to that described for the normalization above. One ultimately finds

$$\sum_{n=1}^{j} B_{n,1} J^j = \{y^j\} \frac{\eta(yJ)}{1-y} \tag{3.78}$$

where we recall that $\{y^j\}$ picks out the coefficient of y^j from the generating function. In the high- and low-density phases, where $J < \frac{1}{4}$, the pole provides the dominant contribution to the sum for $j \to \infty$, with finite-j corrections supplied by the square-root singularity in $\eta(yJ)$ at $y = \frac{1}{4J}$. The result is

$$\sum_{n=1}^{j} B_{n,1} J^j \sim \eta(J) - \frac{J}{1-4J} \frac{(4J)^j}{\sqrt{\pi} j^{3/2}} + \cdots. \tag{3.79}$$

Note that $\eta(J) = \min\{\alpha, \beta\}$ in these phases. Meanwhile, in the maximal-current phase, the pole and square-root singularity coincide and then

$$\sum_{n=1}^{j} B_{n,1} J^j \sim \frac{1}{2}\left(1 - \frac{1}{\sqrt{\pi j}} + \mathrm{O}(j^{-3/2})\right). \tag{3.80}$$

3.5. Exact phase diagram

We can now put these results together to determine the density profile at a fixed (but large) distance j from the right boundary in the limit $N \to \infty$ using (3.65). We do not need to consider the left boundary explicitly because of a particle–hole symmetry in the model that implies

$$\rho_{N+1-i}(\alpha, \beta) = 1 - \rho_i(\beta, \alpha) \tag{3.81}$$

since holes enter the system at the right boundary at a rate β, hop with unit rate to the left and exit at rate α. There are a number of different cases to consider in obtaining the density profile, since not only does the first term in (3.74), and prefactor of the second, depend on which phase the system is in, but the second sum additionally takes two different functional forms at the right boundary in the low-density phase. The various distinct cases are as follows.

- *Maximal-current phase* $\beta > \frac{1}{2}, \alpha > \frac{1}{2}$. Here, the generating-function analysis of (3.65) involves only algebraic singularities rather than poles. Hence one finds a power-law decay of the density from the bulk value of $\frac{1}{2}$, namely,

$$\rho_{N-j} \sim \frac{1}{2}\left[1 - \frac{1}{\sqrt{\pi j}} + \mathrm{O}(j^{-3/2})\right]. \tag{3.82}$$

Although the mean-field theory predicts the bulk density correctly, the decay exponent is incorrectly predicted as 1 rather than $\frac{1}{2}$. Note also the universal (i.e., independent of α and β) prefactor of this decay which is also found at the left boundary as a consequence of the particle–hole symmetry.

- *Low-density phases* $\alpha < \frac{1}{2}, \alpha < \beta$. As noted above, the second term in (3.65) takes different forms depending on whether β is less than or greater than $\frac{1}{2}$. This gives rise to two sub-phases.
  - *LD-I* $\beta < \frac{1}{2}$. Here, the generating function for the density contains only poles, and so a purely exponential decay is obtained at the right boundary:
$$\rho_{N-j} \sim \alpha + (1-2\beta)\left(\frac{\alpha(1-\alpha)}{\beta(1-\beta)}\right)^{j+1}. \tag{3.83}$$
  - *LD-II* $\beta > \frac{1}{2}$. In this instance a pole and algebraic singularity combine to give an exponential decay modulated by a power law:
$$\rho_{N-j} \sim \alpha + \left[\frac{1}{(2\beta-1)^2} - \frac{1}{(2\alpha-1)^2}\right]\frac{1}{\sqrt{\pi}\, j^{3/2}} 4^j \left[\alpha(1-\alpha)\right]^{j+1}. \tag{3.84}$$
- *High-density phases* $\beta < \frac{1}{2}, \beta < \alpha$. In the high-density phases, all variations in the density are at the left boundary (which is out at infinity in the foregoing analysis). At the right boundary the density takes the constant value
$$\rho_{N-j} \sim 1 - \beta, \tag{3.85}$$
as was seen in the mean-field analysis.
- *High/low-density to maximal-current transition lines.* Along the boundaries between the high- or low-density and the maximal-current phases, the bulk density is $\frac{1}{2}$ with deviations of order $j^{-3/2}$ (the prefactor can be calculated by developing the asymptotic expansions given above: this is left as an exercise for the reader).
- *High- to low-density transition line.* Here, the situation is more interesting. Since $\alpha = \beta < \frac{1}{2}$, the generating function for the normalization has a double pole, and so $Z_N \sim (\text{const}) N J^{-N}$. This factor of N in the normalization gives rise to a linear density profile from the prefactor Z_{i-1}/Z_N of the second sum in (3.65). That is,
$$\rho_i = \alpha + (1-2\alpha)\frac{\mathrm{i}}{N}. \tag{3.86}$$
This density profile can be understood from the study of shock fronts discussed in section 2.4. A shock, at $x = 0$, is stationary if the density approaches $\rho_L < \frac{1}{2}$ as the spatial coordinate $x \to -\infty$ and a value $\rho_R = 1 - \rho_L > \rho_L$ as $x \to \infty$. One notes that this is the situation imposed by the boundary conditions on the open system in the $N \to \infty$ limit along this transition line: $\rho_L = \alpha, \rho_R = 1 - \alpha$. One has a *diffusing* shock front separating a low-density region to the left and a high-density region to the right. Given that this shock front performs a random walk, its position on the lattice is given by a flat distribution which in turn implies a linear *average* density profile of the form given above.

With these results we are now able to present the complete phase diagram for the ASEP. It is shown in figure 17 and differs from the mean-field version (figure 5) in that the high- and low-density phases have acquired two sub-phases in which the density decay near the boundary takes a different form. The lengthscale characterizing this density decay diverges as the maximal-current phase is approached from either the high- or low-density phases. This type of physics is suggestive of a continuous phase transition [63]. Meanwhile, the phase coexistence at the high- and low-density phase boundary is reminiscent of phase coexistence

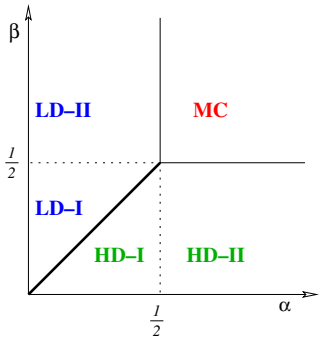


Figure 17. Exact phase diagram for the ASEP. The two high- and low-density sub-phases differ in the decay of the density profile to its bulk value from the boundaries.

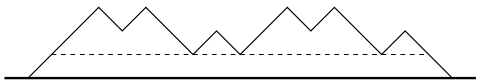


Figure 18. A Dyck path is a random walk on the rotated square lattice that returns to the origin for the first time at a fixed end-point. Between the initial upward and final downward steps, the walk comprises some numbers (possibly zero) of Dyck paths placed end-to-end.

at a first-order phase transition in an equilibrium system. Motivated by these observations, we now explore more deeply the relationship with equilibrium phase transitions.

3.6. Connection to equilibrium phase transitions

In section 3.4, we have seen that the generating function of Catalan numbers played a key part in the analysis of the ASEP's phase behaviour. This fact can be used to make contact with *equilibrium* phase transitions in two-dimensional surface models [102–104] since one of the many combinatorial objects counted by Catalan numbers are excursions on the rotated square lattice, known also as *Dyck paths* (see, e.g., [105] for further applications and references).

Specifically, a Dyck path of length $2N$ comprises N up-steps and N down-steps, so that it contacts the origin both at the start and at the end. It is further constrained never to contact the origin in the intermediate steps—see figure 18. The coefficient of z^N in the generating function $G(z)$ counts the number of paths of length $2N$. Referring to figure 18, we see that a Dyck path of a given length $2N$ begins with an up-step, ends on a down-step and comprises any number of excursions that together have length $2(N-1)$. Therefore, the generating function for Dyck paths obeys

$$G(z) = z[1 + G(z) + G^2(z) + G^3(z) + \cdots] \tag{3.87}$$

$$= \frac{z}{1 - G(z)}, \tag{3.88}$$

where on the rhs of (3.87) the factor z comes from the pair of initial and final up- and down-steps and the sum of terms $G^n(z)$ come from excursions which return exactly n times to site 1 in-between. We see then from (3.88) that $G(z) = \eta(z)$, the function defined in (3.57).

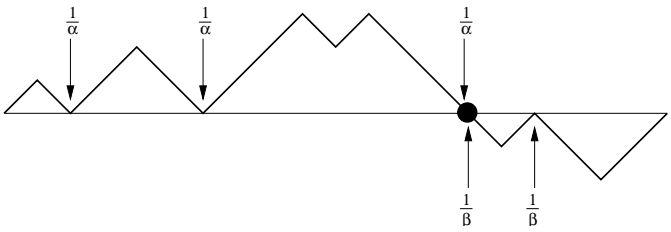


Figure 19. A one-transit walk whose statistical weight in the fixed-length ensemble depends on the density of contacts from above and below.

Similar reasoning implies that the coefficient of $z^N\alpha^{-n}\beta^{-m}$ in the generating function (3.56)

$$\mathcal{Z}(z) = \left(1 - \frac{\eta(z)}{\alpha}\right)^{-1}\left(1 - \frac{\eta(z)}{\beta}\right)^{-1} \tag{3.89}$$

counts the number of walks of length $2N$ that return to the origin a total of $n + m$ times (the last of those returns being the end of the path). It is convenient to have the first n of these excursions take place above the origin and the remaining m below. Then, the coefficient of z^N, which is the normalization of the ASEP Z_N for an N-site system, is also the partition function for the equilibrium ensemble of surfaces of length $2N$ that is constrained to start and end at the same height, crosses the origin at most once (and then only from above), and has a fugacity $1/\alpha$ associated with each contact with the origin from above, and $1/\beta$ with those from below—see figure 19. This construction has been dubbed a *one-transit walk* [103].

It is an elementary exercise in statistical mechanics to derive the free-energy per site in the thermodynamic limit as a function of the density of contacts ρ_a from above and ρ_b from below, these being the order parameters of the surface. This free energy is [103, 106]

$$\begin{aligned} f(\rho_a, \rho_b) = {} & \{\rho_a \ln\alpha + \rho_b \ln\beta\} \\ & - \{(2 - \rho_a - \rho_b)\ln(2 - \rho_a - \rho_b) - (1 - \rho_a - \rho_b)\ln(1 - \rho_a - \rho_b)\} \end{aligned} \tag{3.90}$$

in which the first term in curly brackets is an energy-like quantity that is lowered by the surface contacting the origin and the second an entropy-like quantity that is increased by the surface making large excursions away from the origin. The equilibrium state arising from this energy–entropy competition is found in the conventional manner, i.e., by minimizing the free energy with respect to the order parameters ρ_a and ρ_b.

The free energy is minimized by choosing ρ_a, ρ_b to lie somewhere on the boundary of the physical region $0 \leqslant \rho_a + \rho_b \leqslant 1$. When α and β are large, the entropy of walks that rarely contact the origin dominates and the surface is desorbed: $\rho_a = \rho_b = 0$. On the other hand, if one of α or β is smaller than the critical value of $\frac{1}{2}$, contacting the origin is energetically favourable. If $\alpha < \beta$, it is most favourable for all contacts to come from above, otherwise from below. Thus the low-density, high-density and maximal-current phases of the ASEP correspond to two adsorbed phases, one from above and one from below, and a desorbed phase respectively—see figure 20. We remark that a similar unbinding transition was explored using the mapping from ASEP to the disordered equilibrium problem of directed polymers in a random medium [107].

An interesting feature of this association of nonequilibrium and equilibrium phase behaviour is that the orders of the phase transitions must agree in both cases. In the equilibrium case, these are defined in the conventional way, i.e., through discontinuities in derivatives of the

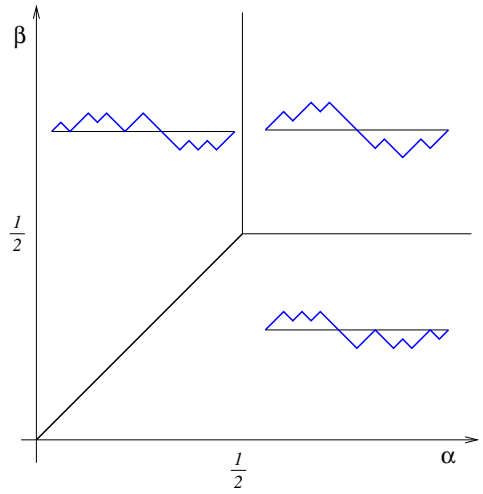


Figure 20. Phase diagram of the one-transit walk in the plane of equilibrium fugacities α and β.

equilibrium free energy with respect to the contact fugacities $1/\alpha$, $1/\beta$. In the nonequilibrium case, on the other hand, derivatives of the 'free energy' (the logarithm of the normalization—see equation (2.49)) with respect to the transition rates α and β have no obvious connection to moments of physical observables as they do for an equilibrium system. We remark that this phenomenon had previously been observed in a study of the Lee–Yang zeros of the ASEP normalization [19] and direct the interested reader to two recent thorough reviews of this subject [65, 108] for details.

3.7. Joint current-density distribution

Finally, in this section on the ASEP with open boundaries, we develop the generating function approach further to compute certain more complicated quantities. It is useful to introduce a pair of matrices

$$D' = (1 - x - y)\frac{D}{1 - xD} \tag{3.91}$$

$$E' = (1 - x - y)\frac{E}{1 - yE}. \tag{3.92}$$

Then, for example, the combination $\frac{x}{1-x-y}D'$ in a generating function constructs a domain of particles, the size of which is conjugate to the parameter (fugacity) x. These matrices have the special property that they satisfy the familiar relation

$$D'E' = D' + E' \tag{3.93}$$

as a consequence of (2.39). This means that results already obtained by, for example, direct ordering of D and E matrices can be carried across to expressions involving D' and E'. One quantity calculated in this way is a joint current-density distribution that was calculated in [109], the derivation of which we now briefly review.

The interest here is in the quantity

$$P_N(M, K) = \frac{1}{Z_N}\{x^M y^{N-M}\}\langle W|\frac{1}{1 - yE}\left(\frac{xD}{1 - xD}\frac{yE}{1 - yE}\right)^K \frac{1}{1 - xD}|V\rangle. \tag{3.94}$$

This gives the probability that a configuration of N sites containing M particles and K particle–hole domain walls is realized in the steady state of the ASEP. Such configurations support K units of flux in the bulk, and so $\lim_{N\to\infty} P_N(N\rho, NJ)$ gives the joint probability of observing a density ρ and current J in the thermodynamic limit.

To make use of earlier results for the ASEP, we note that

$$\left(\frac{xD}{1-xD}\frac{yE}{1-yE}\right)^K = \left[\frac{xy}{(1-x-y)^2}D'E'\right]^K = \frac{(xy)^K}{(1-x-y)^{2K}}(D'+E')^K. \tag{3.95}$$

Then, from (3.93) and the relations (2.39)–(2.41) we have

$$P_N(M,K) = \frac{1}{Z_N(\alpha,\beta)}\{x^{M-K}y^{N-M-K}\}\frac{1}{\left(1-\frac{y}{\alpha}\right)\left(1-\frac{x}{\beta}\right)}\frac{Z'_K}{(1-x-y)^{2K}} \tag{3.96}$$

in which Z'_N is the normalization for the ASEP as given in (3.38) but with the boundary rates α and β replaced by their primed versions

$$\alpha' = \frac{\alpha - y}{1-x-y} \tag{3.97}$$

$$\beta' = \frac{\beta - x}{1-x-y} \tag{3.98}$$

which arise from the requirement that $D'|V\rangle = (1/\beta')|V\rangle$ and $\langle W|E' = (1/\alpha')\langle W|$ in order for (3.96) to be correct. This expression is in agreement with equation (4) in [109] which was obtained using almost the same approach.

To analyse the joint current-density distribution $P_N(M,K)$ in the thermodynamic limit requires, once again, a careful analysis of the singularities in the generating function [109]. One anticipates expressions of the form $P_N(M,K) \approx \exp[-Nr(\rho,J)]$ for large N, with the function $r(\rho,J) = -\lim_{N\to\infty}\ln P_N(\rho N, JN)/N$ taking its minimum value of zero at the thermodynamic values of the current and density given in table 1. In the maximal-current phase one finds (up to additive constants relating to the normalization of P)

$$r(\rho,J) = 2J\ln J + (\rho - J)\ln(\rho - J) + (1-\rho-J)\ln(1-\rho-J) \tag{3.99}$$

and in the low-density phase

$$\begin{aligned} r(\rho,J) = {} & 2J\ln J + (\rho-J)\ln(\rho-J) + (1-\rho-J)\ln(1-\rho-J) \\ & + (1-\rho)\ln\alpha + \rho\ln(1-\alpha) - \rho\ln\rho - (1-\rho)\ln(1-\rho), \end{aligned} \tag{3.100}$$

from which the result for the high-density phase follows from the particle–hole symmetry. In each case, the thermodynamic values give the desired minimum, and one notes the presence of non-Gaussian fluctuations. In [109], such fluctuations are suggested to be a signature of the long-range correlations that are present in the model.

4. ASEP on a ring with two particle species

In section 2.4 we reviewed second-class particles and their various roles and introduced a generalization to two species of particles (see figure 21):

$$1\,0 \xrightarrow{1} 0\,1 \qquad 2\,0 \xrightarrow{\alpha} 0\,2 \qquad 1\,2 \xrightarrow{\beta} 2\,1. \tag{4.1}$$

The case $\alpha = \beta = 1$ corresponds to species 1 being first-class particles and species 2 being second-class particles. Let us stress here that we denote by a two-species exclusion process a system with two species of particles and in addition vacancies. In some works it has been convenient to consider the vacancies as a species of particles and thus sometimes the two-species system we study here has been referred to as a three-species system in the literature.

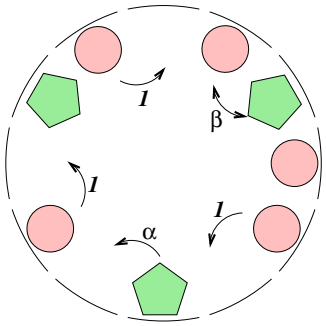


Figure 21. Asymmetric exclusion process on a ring with a second species of particle (shown as pentagons). The labels indicate the rates at which the various particle moves can occur.

4.1. Matrix-product expressions

A microscopic configuration is completely specified by the set of occupation numbers τ_i where $\tau_i = 0$ indicates that site i of the ring (relative to an arbitrary origin) is vacant and $\tau_i = 1, 2$ indicate occupancy by a particle of species 1 or species 2, respectively. We shall again deal with statistical weights (unnormalized probabilities) as in (2.34). On this occasion,

$$P(\tau_1, \tau_2, \ldots, \tau_N) = \frac{f(\tau_1, \tau_2, \ldots, \tau_N)}{Z_{N,M,P}} \tag{4.2}$$

where the normalization is the sum of the weights over all configurations with M species 1 and P species 2 particles on the N-site ring, since these quantities are all conserved by the dynamics. We extract a scalar from the matrix product appearing in (2.35) by the trace operation

$$f(\tau_1, \tau_2, \ldots, \tau_N) = \mathrm{tr}\left(X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N}\right) \tag{4.3}$$

since this reflects the rotational invariance of the system. As before we use the notation D and E for the matrices relating to species 1 particles and vacant sites, and introduce A for species 2 particles.

The reduction relations for these matrices are

$$DE = D + E \tag{4.4}$$

$$AE = \frac{1}{\alpha} A \tag{4.5}$$

$$DA = \frac{1}{\beta} A. \tag{4.6}$$

One notes the similarity with the relations (2.39)–(2.41) for the ASEP with open boundaries. Indeed, two sets of relations are equivalent if one chooses for A the outer product, or projector,

$$A = |V\rangle\langle W|. \tag{4.7}$$

In particular, if one has a single species 2 particle ($P = 1$), the weights become

$$f(\tau_1, \tau_2, \ldots, \tau_{N-1}) = \langle W | X_{\tau_1} X_{\tau_2} \cdots X_{\tau_{N-1}} | V \rangle \tag{4.8}$$

if one chooses the location of the origin as being that of the second-class particle. Although the weight is identical to the matrix-product expression (2.35) for the case of open boundaries,

the distributions (2.34) and (4.2) are, in principle, different because the ensembles, and hence normalizations, are not the same. We shall return to this important point again in subsections 4.3 and 4.3.1. First, we show that this set of matrix-product relations satisfies the stationary conditions for the system (4.1) (i.e., that illustrated in figure 21).

4.2. Proof of the reduction relations

We use a simple generalization of the algebraic approach of section 3.1.2 to prove the reduction relations. The master equation can be written as a sum of terms, each relating to a pair of sites, as

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\tau_1, \tau_2, \ldots, \tau_N) = \sum_{i=1}^{N} \hat{h}_{i,i+1} f(\tau_1, \tau_2, \ldots, \tau_N) \tag{4.9}$$

where in terms of matrix-product expressions (4.3) for the weights

$$\begin{aligned}\hat{h}_{i,i+1} \operatorname{tr}\left(\cdots X_{\tau_{i-1}}\left[X_{\tau_i} X_{\tau_{i+1}}\right] X_{\tau_{i+2}} \cdots\right) = \operatorname{tr}\Big(\cdots X_{\tau_{i-1}}\Big[&W(\tau_{i+1}\tau_i \to \tau_i\tau_{i+1}) X_{i+1} X_i \\ &- W(\tau_i\tau_{i+1} \to \tau_{i+1}\tau_i) X_{\tau_i} X_{\tau_{i+1}}\Big] X_{\tau_{i+2}} \cdots\Big),\end{aligned} \tag{4.10}$$

in which $W(\tau, \tau' \to \tau', \tau)$ gives the rate at which a neighbouring pair of sites in configuration (τ, τ') exchange. For two species of particles on a ring we have

$$W(10 \to 01) = 1 \tag{4.11}$$

$$W(20 \to 02) = \alpha \tag{4.12}$$

$$W(12 \to 21) = \beta \tag{4.13}$$

and all other rates zero.

Introducing auxiliary matrices, $\tilde{X}_\tau$, $\tau = 0, 1, 2$, to generate a pairwise cancellation, we arrive at a sufficient condition

$$W(\tau_{i+1}\tau_i \to \tau_i\tau_{i+1}) X_{\tau_{i+1}} X_{\tau_i} - W(\tau_i\tau_{i+1} \to \tau_{i+1}\tau_i) X_{\tau_i} X_{\tau_{i+1}} = \tilde{X}_{\tau_i} X_{\tau_{i+1}} - X_{\tau_i} \tilde{X}_{\tau_{i+1}}. \tag{4.14}$$

With a *scalar* choice for the auxiliaries and the set of rates (4.11)–(4.13) the nine equations in (4.14) reduce to three equations

$$DE = \tilde{E}D - \tilde{D}E \tag{4.15}$$

$$\alpha AE = \tilde{E}A - \tilde{A}E \tag{4.16}$$

$$\beta DA = \tilde{A}D - \tilde{D}A. \tag{4.17}$$

In this case, the choice $\tilde{D} = -1$, $\tilde{E} = 1$ and $\tilde{A} = 0$ yields (4.4)–(4.6) as required.

4.3. Normalization for a single defect particle

On a periodic system the numbers of each species of particle are fixed, which puts constraints on the set of allowed configurations. For example, when $P = 1$, so that there is a single species 2 particle (which we refer to as a defect), the normalization for an N-site ring populated by M particles is given by

$$Z_{N,M} = \sum_{\tau_1=0}^{1} \sum_{\tau_2=0}^{1} \cdots \sum_{\tau_{N-1}=0}^{1} \langle W | \prod_{j=1}^{N-1} [\tau_j E + (1 - \tau_j) D]^{\tau_1} | V \rangle \delta_{\sum_j \tau_j, M} \tag{4.18}$$

where $\delta_{i,j}$ is the usual Kronecker δ-symbol and enforces the fixed particle number constraint. Although this expression can be evaluated directly, the general expression is rather cumbersome and the derivation relies on choosing a particular representation of the matrices [110, 95]. We consider instead the generating function

$$F_N(u) = \sum_{M=0}^{N-1} Z_{N,M} u^M = \langle W|(uD + E)^{N-1}|V\rangle \tag{4.19}$$

which by virtue of being now rather similar to (2.36) can be readily evaluated. To do this, one transforms to a new pair of matrices D' and E' via

$$D = \frac{1}{\sqrt{u}}(D' - \mathbb{1}) + \mathbb{1} \qquad \text{and} \qquad E = \sqrt{u}(E' - \mathbb{1}) + \mathbb{1} \tag{4.20}$$

which, by (4.4), obey the relations $D'E' = D' + E'$, $D'|V\rangle = 1/\beta'|V\rangle$, $\langle W|E' = 1/\alpha'\langle W|$ with the boundary parameters

$$\frac{1}{\alpha'} = \left[1 + \frac{1}{\sqrt{u}}\left(\frac{1}{\alpha} - 1\right)\right] \qquad \text{and} \qquad \frac{1}{\beta'} = \left[1 + \sqrt{u}\left(\frac{1}{\beta} - 1\right)\right]. \tag{4.21}$$

This transformation is very similar to that described in section 3.7 to calculate the joint current-density distribution for the ASEP with open boundaries. Then, (4.19) may be written as

$$F_N(u) = \langle W|(\sqrt{u}[D' + E'] + [\sqrt{u} - 1]^2)^{N-1}|V\rangle. \tag{4.22}$$

We can now use the formal power series approach of section 3.2.3 to evaluate the generating function of $F_N(u)$. Consider the formal series

$$\begin{aligned}\mathcal{F}(u,z) &= \sum_{N=1}^{\infty} z^{N-1} F_N(u) = \sum_{N=0}^{\infty} \langle W|z^N(\sqrt{u}C' + (\sqrt{u} - 1)^2)^N|V\rangle \\ &= \frac{1}{1 - z(\sqrt{u} - 1)^2}\langle W|\left[1 - \frac{zu^{1/2}}{1 - z(\sqrt{u} - 1)^2}C'\right]^{-1}|V\rangle,\end{aligned} \tag{4.23}$$

where $C' = D' + E'$. Now, we may write

$$1 - \frac{zu^{1/2}}{1 - z(\sqrt{u} - 1)^2}C' = [1 - \omega(u,z)D'][1 - \omega(u,z)E'], \tag{4.24}$$

where

$$\omega(1 - \omega) = \frac{z\sqrt{u}}{1 - z(\sqrt{u} - 1)^2}. \tag{4.25}$$

Therefore,

$$\left[1 - \frac{zu^{1/2}}{1 - z(\sqrt{u} - 1)^2}C'\right]^{-1} = [1 - \omega(u,z)E']^{-1}[1 - \omega(u,z)D']^{-1} \tag{4.26}$$

and we deduce

$$\mathcal{F}(u,z) = \left[\left(1 - z(\sqrt{u} - 1)^2\right)\left(1 - \frac{\omega}{\alpha'}\right)\left(1 - \frac{\omega}{\beta'}\right)\right]^{-1}. \tag{4.27}$$

4.3.1. The grand-canonical ensemble. An explicit expression for $Z_{N,M}$ could be extracted from $F_N(u)$ (4.19) via, for example, the residue theorem

$$Z_{N,M} = \{u^M\}F_N(u) = \oint \frac{\mathrm{d}u}{2\pi \mathrm{i}} \frac{1}{u^{M+1}} F_N(u). \tag{4.28}$$

Equivalently we can work in the grand-canonical ensemble. To proceed we view (4.19) as a partition function for a system where the number of first-class particles can fluctuate. Then, u is a fugacity that controls the density of occupied sites. We remark that it is, in fact, possible to generate this grand-canonical ensemble dynamically by introducing certain specific dynamical rules which do not conserve the particle number as will be described in section 7.3.2.

In the grand-canonical ensemble the mean particle number ρ is given by

$$\rho = \lim_{N\to\infty} \frac{\langle M\rangle}{N} = \lim_{N\to\infty} \frac{1}{N} u \frac{\partial}{\partial u} \ln F_N(u). \tag{4.29}$$

From our understanding of the grand-canonical ensemble in equilibrium statistical physics, we anticipate that both the mean number of particles $\langle M\rangle$ and its variance $\langle M^2\rangle - \langle M\rangle^2$ in the ensemble will be proportional to the system size N in the thermodynamic limit. Thus, the particle number $\langle M\rangle$ will have fluctuations $O(N^{1/2})$ about the mean and the density will be sharply peaked about ρ. The grand-canonical and canonical ensembles should yield equivalent macroscopic properties. Since working at a fixed fugacity u is, in the thermodynamic limit, equivalent to working at fixed particle density ρ, there is no need to invert the generating function (4.19). Instead (for large N) one can substitute u for ρ via (4.29) into the expression for some physical quantity.

4.4. Phase diagram for a single defect particle

The phase diagram in the α–β plane for the model with a single species 2 particle was determined in [95] by considering the exact expressions for $Z_{N,M}$ for N large. An alternative approach was presented in [89] where the Bethe ansatz was used to develop contour integral representations of $Z_{N,M}$ which were then evaluated in the large-N limit. Here, we will derive the phase diagram from the asymptotic forms of the grand-canonical partition function (4.19). We therefore proceed to evaluate the asymptotics of (4.19) from (4.27).

We first note that the possible singularities of (4.27) as a function of z are (since u is positive) a square root at

$$z_0 = (1+\sqrt{u})^{-2}, \tag{4.30}$$

if $\alpha' < 1/2$ a pole at $\omega(z_\alpha) = \alpha'$ and if $\beta' < 1/2$ a pole at $\omega(z_\beta) = \beta'$. The locations of the two poles reduce to

$$z_\alpha = \frac{\alpha(1-\alpha)}{\alpha(u-1)+1} \tag{4.31}$$

$$z_\beta = \frac{\beta(1-\beta)}{u(1-\beta)+\beta}. \tag{4.32}$$

There is also an (irrelevant) pole at $z_1 = (1-\sqrt{u})^{-2}$.

Using the results of appendix B one can compute the following asymptotic contributions to $F_N(u)$:

$$\text{from} \quad z_0 \quad \frac{1}{z_0^N} \frac{1}{2\pi(1-z_0/z_1)^{3/2}} \frac{2\alpha'}{\alpha'-1} \frac{2\beta'}{\beta'-1} \left[\frac{2\alpha'}{\alpha'-1} + \frac{2\beta'}{\beta'-1}\right] \tag{4.33}$$

$$\text{from} \quad z_\alpha \quad \frac{1}{z_\alpha^N}\frac{[(1-\alpha)/\alpha - u\alpha/(1-\alpha)]}{[(1-\alpha)/\alpha - u(1-\beta)/\beta]} \tag{4.34}$$

$$\text{from} \quad z_\beta \quad \frac{1}{z_\beta^N}\frac{[\beta/(1-\beta) - u(1-\beta)/\beta]}{[(1-\alpha)/\alpha - u(1-\beta)/\beta]}. \tag{4.35}$$

Since the asymptotic form of F_N is always $\sim z^{-N}$, with z the dominant singularity, one can associate via (4.29) a density with each of the singularities

$$\rho_0 = -u\frac{\partial \ln z_0}{\partial u} = \frac{u^{1/2}}{1+u^{1/2}} \tag{4.36}$$

$$\rho_\alpha = -u\frac{\partial \ln z_\alpha}{\partial u} = \frac{u\alpha}{\alpha u + 1 - \alpha} \tag{4.37}$$

$$\rho_\beta = -u\frac{\partial \ln z_\beta}{\partial u} = \frac{u(1-\beta)}{u(1-\beta)+\beta}. \tag{4.38}$$

Note that since $z_\alpha < z_0$ and $z_\beta < z_0$ a contribution to $F_N(u)$ from z_α or z_β, if it exists, will dominate the contribution from z_0. For example, the condition for the contribution from z_α to exist is $u^{1/2} < (1-\alpha)/\alpha$. This becomes, using $\rho = \rho_\alpha$, (4.37) to give u as a function of ρ, $1-\alpha > \rho$. Similarly, the condition for the contribution from z_β to exist is $u^{1/2} > \beta/(1-\beta)$ and this becomes using (4.38) $\beta > \rho$.

Finally, we note that in both the z_α and z_β contributions to $F_N(u)$, equations (4.34) and (4.35), there is a pole at

$$u_p = \frac{\beta(1-\alpha)}{(1-\beta)\alpha}. \tag{4.39}$$

The presence of this pole means that as the parameters α and β change, u cannot cross u_p. For the contribution coming from z_α, one finds that as β decreases to ρ, u increases to u_p. Similarly, for the contribution coming from z_β, as α decreases to $1-\rho$, u decreases to u_p. Therefore in the region $\alpha < 1-\rho$, $\beta < \rho$, u is fixed at $u = u_p$. Actually, this holds strictly in the thermodynamic limit; to correctly achieve the density $\langle\rho\rangle = \rho$ one should choose $u = u_p + \gamma(\alpha,\beta,\rho)/N$, where the function γ is some function to be determined, then let $N \to \infty$.

In the thermodynamic limit, we can identify four phases as follows:

(i) $1-\alpha < \rho < \beta$ $\quad z = z_0 = (1-\rho)^2$ and $u = \frac{\rho}{(1-\rho)^2}$,

(ii) $1-\alpha > \rho$ and $\rho < \beta$ $\quad z = z_\alpha = \alpha(1-\rho)$ and $u = \frac{\rho(1-\alpha)}{(1-\rho)\beta}$,

(iii) $1-\alpha < \rho$ and $\rho > \beta$ $\quad z = z_\beta = (1-\beta)(1-\rho)$ and $u = \frac{\rho\beta}{(1-\rho)(1-\beta)}$,

(iv) $1-\alpha > \rho$ and $\rho > \beta$ $\quad z = z_\alpha = z_\beta = \alpha(1-\beta)$ and $u = \frac{(1-\alpha)\beta}{\alpha(1-\beta)}$.

We now calculate the velocity of the defect particle. In the canonical ensemble this is given by $\alpha(1-\rho_+)$, where ρ_+ is the probability of a hole in front of the defect minus $\beta\rho_-$, where ρ_- is the probability of a particle behind the defect particle: using (4.5) and (4.4)

$$v = \frac{Z_{N-1,M} - Z_{N-1,M-1}}{Z_{N,M}}. \tag{4.40}$$

In the grand-canonical ensemble the corresponding expression is

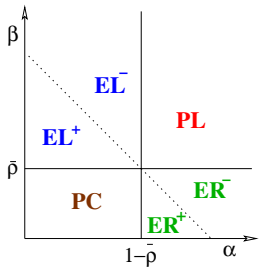


Figure 22. Phase diagram for the ASEP with a single second-class particle on a ring.

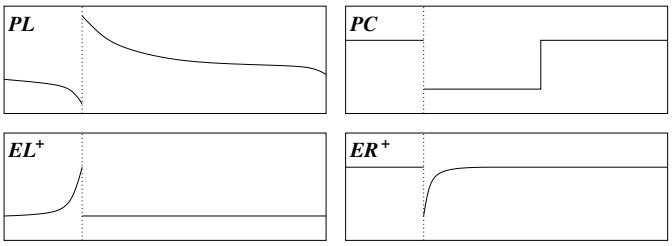


Figure 23. Density profiles as seen from the defect particle (shown as a vertical dotted line) in the four phases of the ASEP on a ring with a defect particle: power law (PL), phase coexistence (PC), exponential decay to the right (ER) and left (EL) of the defect particle.

Table 2. Velocity, density in front of the defect and density behind the defect in the various phases of the single defect system.

Region	Phase	v	ρ_+	ρ_-
(i)	PL	$1-2\rho$	$1-(1-\rho)^2/\alpha$	ρ^2/β
(ii)	ER	$\alpha-\rho$	ρ	$\rho(1-\alpha)/\beta$
(iii)	EL	$1-\beta-\rho$	$1-(1-\rho)(1-\beta)/\alpha$	ρ
(iv)	PC	$\alpha-\beta$	β	$1-\alpha$

$$v = \frac{\alpha\langle\alpha|E(uD+E)^{N-1}|\beta\rangle - \beta\langle\alpha|(uD+E)^{N-1}uD|\beta\rangle}{F_N(u)} \tag{4.41}$$

$$= \frac{F_{N-1}(u) - uF_{N-1}(u)}{F_N(u)}. \tag{4.42}$$

We find using the form of the asymptotic behaviour, $F_N(u) \sim z^{-N}$, that

$$v \to z(1-u) \quad \text{for } N \text{ large.} \tag{4.43}$$

The different phases are characterized by the values in table 2. The phase diagram which results is shown in figure 22, and the corresponding density profiles in figure 23. Most of the phases are recognizable from the ASEP with open boundaries. First, the region $\beta > \rho > 1-\alpha$, which includes the case of a second-class particle ($\alpha = \beta = 1$), has $v = 1-2\rho$ and is marked PL (power law) in figure 22. The effect of the defect is to cause a disturbance which results in an increase in the density in front of the defect and a decrease behind. The density profile as seen from the defect particle decays as a power law with distance from the defect towards its asymptotic value ρ:

$$\rho_{+j} \simeq \bar{\rho} + \left(\frac{\bar{\rho}(1-\bar{\rho})}{\pi j}\right)^{1/2}, \tag{4.44}$$

$$\rho_{-j} \simeq \bar{\rho} - \left(\frac{\bar{\rho}(1-\bar{\rho})}{\pi j}\right)^{1/2}. \tag{4.45}$$

We note that this is essentially the same profile as that obtained for the ASEP with open boundaries in the maximal-current phase.

Since the bulk density far away from the defect is fixed at ρ in the periodic system, we do not have low- and high-density phases, as such, in this case. However, the characteristic density profiles of these phases, in which one has exponential decays to the left and right of the defect particle respectively, are present here in the phases marked $\mathrm{EL}^{\pm}$ and $\mathrm{ER}^{\pm}$, respectively. In the ER phase the defect only causes a disturbance in front of it and the density profile decays exponentially towards the asymptotic value. Similarly, in the EL phase the defect only causes a disturbance behind it. In the figure the + or − superscript indicates whether the density increases or decreases as one goes clockwise around the ring. The dashed line separating these regions with positive and negative density gradients has $\alpha + \beta = 1$ and, in common with the ASEP with open boundaries, along this line the density profile is completely flat.

The main novelty arising from the particle-number conservation is the PC (phase coexistence) phase, where $\alpha < 1 - \rho, \beta < \rho$. Here, there is a coexistence between a region of low-density β in front of the defect particle and high-density $1-\alpha$ behind the defect particle. At the point where the two regions meet there is a shock in the density. The shock is localized at a position a distance xN in front of the defect particle, where x is given by $\rho = \beta x + (1-\alpha)x$ [111, 95]. The interesting point is that the shock exists in an entire region of the α–β plane as opposed to the open-boundary condition case where phase coexistence only occurs along the line $\alpha = \beta < 1/2$.

To understand the PC phase from the analysis, we note that there are two competing contributions to $F_N(u)$ from the pole in (4.27) at z_α and the pole at z_β. Using $u = u_p$ and (4.37), (4.38) these correspond to densities β and $1-\alpha$, respectively. Finally, we remark that in all phases the current of first-class particles $J = \rho(1-\rho)$ *except* the PC phase where $J = \rho(\alpha - \beta) + \beta(1-\alpha)$. We also mention that other defect particle dynamics have been studied [112–117] and we will discuss these models further in section 7.

4.5. Some properties of the second-class particle case

The second-class particle case $\alpha = \beta = 1$ was subject of a detailed study by Derrida, Janowsky, Lebowitz and Speer [83]. In addition to the case of a single second-class particle, results were obtained for several second-class particles and a finite density of second-class particles (as far as we are aware these computations have not yet been extended to the case of general α and β). Here we briefly summarize some of their findings.

For the case of two second-class particles the probability of finding the particles a distance r apart was computed and shown to decay as a $1/r^{3/2}$ power law. Thus there is an effective attractive interaction and the two particles form a weak bound-state in which, despite the attractive interaction, the mean separation of the two particles diverges. The computations of [83] also considered a finite density, ρ_2, of second-class particles and showed that in the limit $\rho_2 \to 0$ the environment around a second-class particle is different from the case of a single second-class particle. Although in this limit the global density is zero there must be a clustering of the second-class particles into a bound state.

A finite system with first- and second-class particles cannot support two regions of different densities. However via an astute construction [83, 86] one can study shocks on such a system.

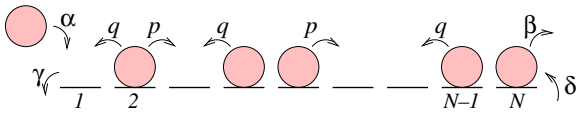


Figure 24. The full parameter space of the partially asymmetric exclusion process with open boundaries. Particles can hop in either direction on the lattice, and can enter and leave the system at both ends, all at different rates.

The idea is that a system with finite densities ρ_1, ρ_2 of first- and second-class particles can be viewed as a one-species system in two different ways. Treating the second-class particles as holes gives a TASEP with density ρ_1 whereas ignoring the difference between first- and second-class particles, so that exchanges between them are immaterial, yields a TASEP with density $\rho_1 + \rho_2$. Then one can construct shock profiles as follows: consider the density profile as seen from some second-class particle where in front of this second-class particle both first- and second-class particles are treated as particles and behind the second-class particle the second-class particles are treated as holes. Then letting the system size go to infinity yields a profile where ρ_i tends to $\rho_1 + \rho_2$ for i large and ρ_{-i} tends to ρ_1 for i large. Using this approach shock profiles were obtained. Later these profiles were recovered from a calculation directly on the infinite system (see section 7.2.3).

5. The partially asymmetric exclusion process (PASEP) with open boundaries

In this section we consider a generalization of the ASEP to the situation where particles can hop both to the left and to the right, with different rates. We refer to this variant as the partially asymmetric exclusion process (PASEP). The dynamics are as shown in figure 24, and one has particles on the lattice hopping to the left and right with rates p and q, respectively, as long as the hard-core exclusion constraint is maintained at all times. Furthermore, particles enter the system at the left boundary at rate α, as before, but can also leave from the left boundary at rate γ. Likewise, particles leave the right boundary at rate β, as before, but also enter there at rate δ. We shall once again define the unit of time by setting $p = 1$.

Physical interest in this more general model comes from at least two quarters. First, there is a mapping between exclusion models and certain nonequilibrium surface growth models [118] (note these are distinct from the equilibrium surface models discussed in section 3.6). At large lengthscales, the shape of the interface is governed by a stochastic partial differential equation, the KPZ equation (see, e.g., [33] for a review), that has a nonlinear term proportional to $1 - q$. The presence or otherwise of this nonlinearity governs in turn the universal properties of the interface. Thus, by taking $q \to 1$ in the PASEP we should learn something of the crossover between universality classes [119]. The second source of interest lies in the special case where $\gamma = \delta = 0$ and $q > 1$. Then, the bias in the stochastic motion of the particles opposes the current imposed by the boundary conditions. This gives rise to a new *reverse-bias* phase not seen in the ASEP. Furthermore, the matrix algebra that is used to solve the PASEP also arises in the context of ballistic reaction dynamics and has been used to calculate the late-time density decay in the system [120].

In section 3.2, we discussed the three main approaches to calculation with matrix products. For the PASEP with open boundaries the most natural approach turns out to be diagonalization of an explicit representation. Indeed, to our knowledge, the results have thus far been obtained only through this approach [121–125].

5.1. Quadratic algebra

As for the ASEP, the stationary distribution is given through products of two matrices D and E that represent particles and holes, respectively. Again, the product is contracted with a pair of vectors $\langle W|$ and $|V\rangle$ to obtain the desired scalar weight, as in (2.35). The relations that allow the matrix products to be reduced are

$$DE - qED = D + E \tag{5.1}$$

$$(\beta D - \delta E)|V\rangle = |V\rangle \tag{5.2}$$

$$\langle W|(\alpha E - \gamma D) = \langle W|. \tag{5.3}$$

These can be proved to solve the stationary master equation either by considering terms that cancel between domains, as in section 3.1, or using a slight modification of the purely algebraic approach of section 3.1.2. We shall run through the latter briefly here.

As in section 3.1.2, we decompose the terms in the master equation into a sum over terms relating to interactions at the boundaries and between pairs of sites in the bulk. That is,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\tau_1, \ldots, \tau_N) = \left(\hat{h}_L + \sum_{i=1}^{N-1} \hat{h}_{i,i+1} + \hat{h}_R \right) f(\tau_1, \ldots, \tau_N) \tag{5.4}$$

where $\hat{h}_{i,i+1}$ generate the terms arising from particle hops in the bulk, as in (3.9), and $\hat{h}_L$ and $\hat{h}_R$ perform the corresponding duties for interactions at the left and right boundaries as in (3.10), (3.11).

As in section 4.2 we seek to telescope this sum, and we do so by assuming that the terms in the master equation generated by the $\hat{h}$ operators can be written using auxiliary matrices as

$$\hat{h}_{i,i+1}\langle W| \cdots X_{\tau_i} X_{\tau_{i+1}} \cdots |V\rangle = \langle W| \cdots [\tilde{X}_{\tau_i} X_{\tau_{i+1}} - X_{\tau_i} \tilde{X}_{\tau_{i+1}}] \cdots |V\rangle \tag{5.5}$$

$$\hat{h}_L\langle W|X_{\tau_1} \cdots |V\rangle = -\langle W|\tilde{X}_{\tau_1} \cdots |V\rangle \tag{5.6}$$

$$\hat{h}_R\langle W| \cdots X_{\tau_N}|V\rangle = \langle W| \cdots \tilde{X}_{\tau_N}|V\rangle. \tag{5.7}$$

Matching these right-hand sides up with the terms generated by the respective operators in the master equation yields

$$0 = \tilde{D}D - D\tilde{D} = \tilde{E}E - E\tilde{E} \tag{5.8}$$

$$DE - qED = \tilde{E}D - E\tilde{D} = -\tilde{D}E + D\tilde{E} \tag{5.9}$$

$$\langle W|(\gamma D - \alpha E) = -\langle W|\tilde{E} = \langle W|\tilde{D} \tag{5.10}$$

$$(\delta E - \beta D) = \tilde{D}|V\rangle = -\tilde{E}|V\rangle. \tag{5.11}$$

Again, the scalar choices $\tilde{D} = -1$, $\tilde{E} = 1$ yield the consistent set of relations (5.1)–(5.3).

It turns out that the mathematical structure of the solution is much the same for the case with nonzero γ and δ as that with $\gamma = \delta = 0$. However, in the former case the various boundary rates combine with one another in a complicated way, so we shall take $\gamma = \delta = 0$ in the analysis that follows for ease of presentation. At the end of this section we shall explain how nonzero γ and δ were catered for in the works [124, 125].

5.2. Calculation of Z_N using an explicit representation

As we observed in section 2.3.1 of the introduction, an argument was presented in [62] for the D and E matrices to have, in general, infinite dimension, at least for the case $q = 0$. This is also the case for general q, and one can find a number of different representations that satisfy the relations (5.1)–(5.3). The first of these that we shall consider has

$$D = \frac{1}{1-q}\begin{pmatrix} 1+b & \sqrt{c_0} & 0 & 0 & \cdots \\ 0 & 1+bq & \sqrt{c_1} & 0 & \\ 0 & 0 & 1+bq^2 & \sqrt{c_2} & \\ 0 & 0 & 0 & 1+bq^3 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{5.12}$$

$$E = \frac{1}{1-q}\begin{pmatrix} 1+a & 0 & 0 & 0 & \cdots \\ \sqrt{c_0} & 1+aq & 0 & 0 & \\ 0 & \sqrt{c_1} & 1+aq^2 & 0 & \\ 0 & 0 & \sqrt{c_2} & 1+aq^3 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{5.13}$$

where the parameters a, b, c_n are given by

$$a = \frac{1-q}{\alpha} - 1, \qquad b = \frac{1-q}{\beta} - 1 \tag{5.14}$$

and

$$c_n = (1-q^{n+1})(1-abq^n). \tag{5.15}$$

In this case, the appropriate choice for the boundary vectors is

$$\langle W_1| = \langle 0| \qquad \text{and} \qquad |V_1\rangle = |0\rangle \tag{5.16}$$

where we have introduced a set of basis vectors

$$|n\rangle = (\underbrace{0, \ldots, 0}_{n}, 1, 0, \ldots, 0)^T \tag{5.17}$$

for $n = 0, 1, 2, \ldots$ and their corresponding duals $\langle n|$ by transposition. One can check explicitly that the reduction relations (5.1)–(5.3) are satisfied.

One way to evaluate the normalization using this representation is to diagonalize the matrix $C = D + E$, as was done in section 3.2.2 for the case $q = 0$. In analogy with (3.43) we parametrize the eigenvalues of C by $2(1+\cos\theta)/(1-q)$ where $0 < \theta < 2\pi$ so that

$$C|\cos\theta\rangle = \frac{2(1+\cos\theta)}{1-q}|\cos\theta\rangle. \tag{5.18}$$

Writing

$$|\cos\theta\rangle = \sum_{n=0}^{\infty} f_n(\cos\theta)|n\rangle \tag{5.19}$$

implies from (5.12) and (5.13) that

$$\frac{2(1+\cos\theta)}{1-q} f_n(\cos\theta) = \frac{1}{1-q} \times \{\sqrt{c_{n-1}} f_{n-1}(\cos\theta) + [2 + (a+b)q^n] f_n(\cos\theta) + \sqrt{c_n} f_{n+1}(\cos\theta)\}. \tag{5.20}$$

Putting

$$f_n(\cos\theta) = \frac{P_n(\cos\theta)}{\prod_{k=0}^{n-1}\sqrt{c_k}}, \tag{5.21}$$

this three-term recursion simplifies to

$$2\cos\theta\, P_n(\cos\theta) = (1-q^n)(1-abq^{n-1})P_{n-1}(\cos\theta) + (a+b)q^n P_n(\cos\theta) + P_{n+1}(\cos\theta). \tag{5.22}$$

Given the boundary conditions $P_{-1}(\cos\theta) = 0$, $P_0(\cos\theta) = 1$, this recursion defines the set of Al-Salam–Chihara polynomials [126]. The relevance of these polynomials to the PASEP was noticed by Sasamoto [121].

Although these polynomials are rather complicated, we need to know very little about them to evaluate the normalization $Z_N = \langle W|C^N|V\rangle$, other than that they satisfy an orthogonality relation [121, 126]

$$\frac{1}{2\pi}\frac{(q,ab;q)_\infty}{(q,ab;q)_n}\int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta},\mathrm{e}^{-2\mathrm{i}\theta};q)_\infty}{(a\,\mathrm{e}^{\mathrm{i}\theta},a\,\mathrm{e}^{-\mathrm{i}\theta},b\,\mathrm{e}^{\mathrm{i}\theta},b\,\mathrm{e}^{-\mathrm{i}\theta};q)_\infty} P_n(\cos\theta)P_m(\cos\theta) = \delta_{n,m}, \tag{5.23}$$

which holds when $|a|<1$, $|b|<1$ and $|q|<1$. Here we have invoked the notation

$$(a_1,a_2,\ldots,a_m;q)_n = \prod_{k=0}^{n-1}(1-a_1q^k)(1-a_2q^k)\cdots(1-a_mq^k) \tag{5.24}$$

since the quantity on the right-hand side is ubiquitous in the theory of q-series, of which the Al-Salam–Chihara polynomials are an example [126]. Using this orthogonality relation, we can construct a representation of the identity

$$\mathbb{1} = \frac{(q,ab;q)_\infty}{2\pi}\int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta},\mathrm{e}^{-2\mathrm{i}\theta};q)_\infty}{(a\,\mathrm{e}^{\mathrm{i}\theta},a\,\mathrm{e}^{-\mathrm{i}\theta},b\,\mathrm{e}^{\mathrm{i}\theta},b\,\mathrm{e}^{-\mathrm{i}\theta};q)_\infty}|\cos\theta\rangle\langle\cos\theta| \tag{5.25}$$

where $\langle\cos\theta|$ is simply the transpose of $|\cos\theta\rangle$.

An integral representation for the normalization Z is then obtained as follows. First, one incorporates this identity into the definition of Z via $Z_N = \langle W|C^N\mathbb{1}|V\rangle$. This yields

$$Z_N = \frac{(q,ab;q)_\infty}{2\pi}\int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta},\mathrm{e}^{-2\mathrm{i}\theta};q)_\infty}{(a\,\mathrm{e}^{\mathrm{i}\theta},a\,\mathrm{e}^{-\mathrm{i}\theta},b\,\mathrm{e}^{\mathrm{i}\theta},b\,\mathrm{e}^{-\mathrm{i}\theta};q)_\infty}\langle W|C^N|\cos\theta\rangle\langle\cos\theta|V\rangle. \tag{5.26}$$

We then use the fact that, by construction, $C|\cos\theta\rangle = 2(1+\cos\theta)/(1-q)$ to obtain

$$Z = \frac{(q,ab;q)_\infty}{2\pi}\int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta},\mathrm{e}^{-2\mathrm{i}\theta};q)_\infty}{(a\mathrm{e}^{\mathrm{i}\theta},a\mathrm{e}^{-\mathrm{i}\theta},b\mathrm{e}^{\mathrm{i}\theta},b\mathrm{e}^{-\mathrm{i}\theta};q)_\infty}\langle W|\cos\theta\rangle\left[\frac{2(1+\cos\theta)}{1-q}\right]^N\langle\cos\theta|V\rangle. \tag{5.27}$$

Finally, we note that $\langle W|\cos\theta\rangle = \langle 0|\cos\theta\rangle = 1$ and likewise that $\langle\cos\theta|V\rangle = 1$. We find then an integral representation of the normalization

$$Z_N = \frac{(q,ab;q)_\infty}{2\pi}\int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta},\mathrm{e}^{-2\mathrm{i}\theta};q)_\infty}{(a\mathrm{e}^{\mathrm{i}\theta},a\mathrm{e}^{-\mathrm{i}\theta},b\mathrm{e}^{\mathrm{i}\theta},b\mathrm{e}^{-\mathrm{i}\theta};q)_\infty}\left[\frac{2(1+\cos\theta)}{1-q}\right]^N. \tag{5.28}$$

In the limit $q\to 0$ this can be written as

$$Z_N = \frac{1-ab}{\pi}\int_{-\pi}^\pi \mathrm{d}\theta \frac{\sin^2\theta}{(1+a^2-2a\cos\theta)(1+b^2-2b\cos\theta)}[2(1+\cos\theta)]^N \tag{5.29}$$

which can be shown to agree with the finite sum (3.38) as it should—see appendix C. In [122] it was also shown that (5.28) is equivalent to a finite sum expression for general q. However, the resulting expression is rather cumbersome, and so in section 5.3 we will concentrate

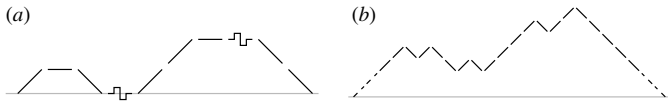


Figure 25. (*a*) A bicoloured Motzkin path. Each step may increase or decrease the height of the surface by one unit or leave it unchanged. These latter horizontal segments come in two *colours*, denoted here by the flat and crenellated segments. (*b*) The same path mapped onto a Dyck path, wherein each Motzkin segment becomes two Dyck segments. Two additional segments (shown dotted) are needed at both ends to prevent the Dyck path from crossing the origin.

on the procedure to extract the asymptotics of the normalization directly from the integral representation (5.28).

5.2.1. Relation to Motzkin paths. The representation given as (5.12) and (5.13) has an interpretation in terms of random walks and surfaces, similar to that described in section 3.6. To see this, let the vector $|n\rangle$ represent a height n above some origin. A matrix element $\langle m|U|n\rangle$, where U is a product of ℓ D and E matrices, can then be interpreted as the weight of a set of paths of horizontal length ℓ connecting a point at height m (at the left-hand end of the surface) to a point at height n (at the right-hand end of the surface). The particular choice of boundary vectors $\langle W_1| = \langle 0|$ and $|V_1\rangle = |0\rangle$ thus corresponds to a set of paths that begin and end at the origin.

The bidiagonal forms of the D and E matrices imply that a step of the surface must be flat, a diagonal up-step or a diagonal down-step. Furthermore, the fact that the matrices are semi-infinite means that the paths may never go below the origin. Finally, the matrix product $Z_N = \langle 0|(D+E)^N|0\rangle$ describes the ensemble of all possible such paths of length N, and has up- and down-step pairs connecting height n and $n+1$ weighted by c_n, and two different types (or 'colours') of horizontal segments, coming in with weights $1+aq^n$ and $1+bq^n$, where again n is the height of the segment above the origin, see figure 25. These paths sometimes appear in the combinatorial literature as *bicoloured Motzkin paths*.

We remark that each bicoloured Motzkin path can be mapped uniquely to a Dyck path, which has only diagonal up- and down-steps as described in section 3.6. This is achieved by mapping each segment of a Motzkin path to *two* consecutive segments of a Dyck path. Specifically, diagonal up- and down-steps are mapped to a pair of up- and down-steps, respectively, whilst the two colours of horizontal steps are mapped to up-down and down-up pairs. In order that the resulting path contacts the origin only at its two ends (i.e., *is* a Dyck path) two additional up-steps are needed at the left end and similarly two down-steps at the right, see figure 25. Thus a Motzkin path of length N corresponds to a Dyck path of length $2(N+2)$. Although a similar treatment of the PASEP thermodynamics to that described for the ASEP in section 3.6 should in principle be possible using this approach, to our knowledge it has not yet been achieved.

5.3. Asymptotics of the normalization from the integral representation

To analyse the asymptotics of the normalization from its integral representation (5.28) we first rewrite it as the contour integral

$$Z_N = \frac{(q,ab;q)_\infty}{4\pi\mathrm{i}} \int_K \frac{\mathrm{d}z}{z} \frac{(z^2,z^{-2};q)_\infty}{(az,a/z,bz,b/z;q)_\infty} \left[\frac{2+z+z^{-1}}{1-q}\right]^N. \tag{5.30}$$

where K is the positively oriented unit circle centred on the origin. Recall that the orthogonality relation (5.23) was valid only when $|a|<1$ and $|b|<1$. In this case the contour K encloses

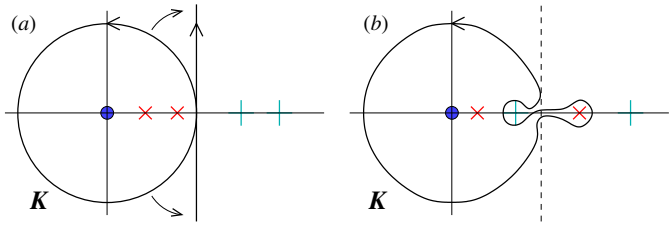


Figure 26. Integration contour K for the representation (5.30). The ($\times$) symbols show the first of each of the sequences of poles $z = a, qa, q^2a, \ldots$ and $z = a, qb, q^b, \ldots$ that are required to be inside K, and the (+) symbols those that must be outside. (a) When $|a| < 1$ and $|b| < 1$, K can be continuously deformed into the contour for the saddle-point integration (the vertical line) as shown. (b) When $|a| > 1$, this deformation is no longer possible, and it is necessary to add to the saddle-point result the difference between the residues at $z = a$ and $z = 1/a$.

two sequences of poles at $z = a, qa, q^2b, \ldots$ and $z = b, qb, q^b, \ldots$ along with a number of singularities at the origin, as shown in figure 26(a). We can continue the function Z_N to $a > 1$ and/or $b > 1$ by ensuring the contour K is deformed still to enclose these poles, and further exclude those at $z = 1/a, q/a, q^2/a, \ldots$ and $z = 1/b, q/b, q^2/b, \ldots$, as shown in figure 26($b$).

In the former case, where $|a| < 1$ and $|b| < 1$, the contour K can be continuously deformed without passing through any poles to coincide with the line that passes parallel to the imaginary axis through the saddle point in the integrand at $z = 1$, see figure 26(a). Then, the normalization is well approximated for large N by the expression obtained from applying the saddle-point method. This is [121, 122]

$$Z_N \sim \frac{4}{\sqrt{\pi}} \frac{(q;q)_\infty^3 (ab;q)_\infty}{(a;q)_\infty^2 (b;q)_\infty^2} \frac{4^N}{N^{3/2}}. \tag{5.31}$$

Let us now consider the case $a > 1 > qa$, $|b| < 1$ where the deformed contour shown in figure 26(b) cannot be continuously deformed into that used in the saddle-point method. One must therefore correct the result by applying the residue theorem to the poles at $z = a$ and $z = 1/a$ so that the former is included in the integration contour (as required) and the latter excluded. This procedure gives a correction

$$Z_N \sim \frac{(1/a^2;q)_\infty}{(b/a,q)_\infty} (2 + a + 1/a)^N \tag{5.32}$$

which, for sufficiently large N, overwhelms the contribution from saddle-point expression (5.31) because $2 + a + 1/a > 4$ when $a > 1$. Therefore, the normalization is well approximated by the contribution (5.32) from the pair of poles at $z = a$ and $z = 1/a$. As a is further increased, more and more poles need—in principle—to be accounted for in the deformation of the contour. However, one can verify that all further corrections grow less fast than $(2 + a + 1/a)^N$ and so (5.32) remains the leading contribution to the normalization.

When $|a| < 1$ but $b > 1$, we arrive at the same expression (5.32) for the normalization but with a and b exchanged. When both a and b are larger than 1, both of these contributions are present. Here, (5.32) dominates when $a > b$, and the corresponding expression with $a \leftrightarrow b$ when $a < b$. Thus, for the PASEP we have the same three phases as for the ASEP. In each of these the current can be determined using the matrix expression

$$J = \frac{\langle W|C^{i-1}(DE - qED)C^{N-i-1}|V\rangle}{Z_N} = \frac{Z_{N-1}}{Z_N} \tag{5.33}$$

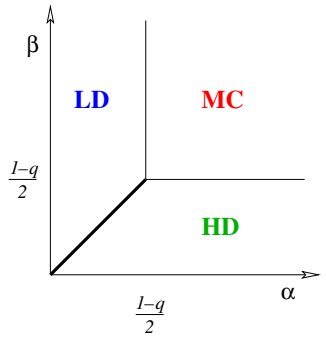


Figure 27. Phases of the PASEP over which the particle current is analytic.

Table 3. Asymptotic currents in the various phases of the PASEP.

Region	Phase	Current J
$\alpha < \beta, \alpha < \frac{1-q}{2}$	Low-density (LD)	$\frac{\alpha(1-q-\alpha)}{1-q}$
$\beta < \alpha, \beta < \frac{1-q}{2}$	High-density (HD)	$\frac{\beta(1-q-\beta)}{1-q}$
$\alpha > \frac{1}{2}, \beta > \frac{1-q}{2}$	Maximal-current (MC)	$\frac{1-q}{4}$

where the reduction relation (5.1) has been used to yield the same ratio of normalizations that gave the current for the ASEP. Rewriting the inequalities on a and b given above in terms of α and β we find the phase diagram shown in figure 27 with the corresponding currents given in table 3. Note the similarities with figure 5 and table 1.

5.4. Density profiles from the q-deformed harmonic oscillator algebra

Whilst the matrix representation given in (5.12) and (5.13) provides a quick route to the normalization (5.28), it is less suited to the task of calculating the density profiles. Here, progress is made by using a representation that involves the q-deformed harmonic oscillator algebra, a well-studied mathematical object for which many results are known [127, 128]. The relevance of the q-deformed harmonic oscillator algebra to the PASEP was first noted in [129].

Central to this approach is a pair of creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ that satisfy a q-commutation relation

$$\hat{a}\hat{a}^\dagger - q\hat{a}^\dagger\hat{a} = 1 - q \tag{5.34}$$

and act on basis vectors according to

$$\hat{a}^\dagger|n\rangle = \sqrt{1-q^{n+1}}|n+1\rangle \tag{5.35}$$

$$\hat{a}|n\rangle = \sqrt{1-q^{n}}|n-1\rangle. \tag{5.36}$$

If we take

$$D = \frac{1}{1-q}(\mathbb{1} + \hat{a}) \qquad \text{and} \qquad E = \frac{1}{1-q}(\mathbb{1} + \hat{a}^\dagger) \tag{5.37}$$

we find from (5.34) that the reduction relation (5.1) is satisfied. Thus this representation looks like

$$D = \frac{1}{1-q}\begin{pmatrix} 1 & \sqrt{1-q} & 0 & 0 & \cdots \\ 0 & 1 & \sqrt{1-q^2} & 0 & \\ 0 & 0 & 1 & \sqrt{1-q^3} & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{5.38}$$

$$E = \frac{1}{1-q}\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ \sqrt{1-q} & 1 & 0 & 0 & \\ 0 & \sqrt{1-q^2} & 1 & 0 & \\ 0 & 0 & \sqrt{1-q^3} & 1 & \\ \vdots & & & & \ddots \end{pmatrix}. \tag{5.39}$$

In order that (5.2) and (5.3) are also respected (bearing in mind $\gamma = \delta = 0$) we require

$$\langle W|n\rangle = \kappa \frac{a^n}{\sqrt{(q;q)_n}} \qquad \text{and} \qquad \langle n|V\rangle = \kappa \frac{b^n}{\sqrt{(q;q)_n}} \tag{5.40}$$

where the parameters a and b are as previously given in (5.14). Thus, this representation is the generalization to $q \neq 0$ of that used in section 3.2.2. The constant κ appearing in (5.40) is fixed by the convention that $\langle V|W\rangle = 1$. That is, we require

$$\frac{1}{\kappa^2} = \sum_{n=0}^{\infty} \frac{(ab)^n}{(q;q)_n}. \tag{5.41}$$

This sum, which converges when $|ab| < 1$, features prominently in the literature on q-series as it is the q-analogue of the exponential function [126]. When it converges, it can be expressed as an infinite product and $\kappa^2 = (ab;q)_\infty$.

Diagonalization of the matrix $C = D + E$ proceeds in much the same way as described in subsection 5.2. Here, the eigenfunctions are a vector of functions

$$|\cos\theta\rangle = \sum_{n=0}^{\infty} \frac{H_n(\cos\theta)}{\sqrt{(q;q)_n}}|n\rangle \tag{5.42}$$

with eigenvalue $\lambda(\cos\theta) = 2(1+\cos\theta)/(1-q)$ as before. The polynomials $H_n(\cos\theta)$ are found to satisfy the three-term recurrence

$$2\cos\theta H_n(\cos\theta) = (1-q^n)H_{n-1}(\cos\theta) + H_{n+1}(\cos\theta) \qquad n \geqslant 0 \tag{5.43}$$

using the same approach as previously. Taking $H_0 = 1$ allows the functions $H_n(\cos\theta)$ to be identified as q-Hermite polynomials, the properties of which are discussed in [98, 126, 9]. Of particular importance is the q-exponential generating function of these polynomials, since this allows the computation of the scalar products $\langle W|\cos\theta\rangle$ and $\langle\cos\theta|V\rangle$ that appear during the calculation of the normalization [121, 122]. One finds for the former

$$\langle W|\cos\theta\rangle = \kappa \sum_{n=0}^{\infty} \frac{a^n H_n(\cos\theta)}{(q;q)_n} = \frac{1}{(a\,\mathrm{e}^{\mathrm{i}\theta}, a\,\mathrm{e}^{-\mathrm{i}\theta};q)_\infty} \tag{5.44}$$

and the corresponding expression with $a \to b$ for the latter. With the knowledge of the orthogonality relation satisfied by the polynomials [98, 126] one can write the identity

$$\mathbb{1} = \frac{(q;q)_\infty}{2\pi} \int_0^\pi \mathrm{d}\theta\,(\mathrm{e}^{2\mathrm{i}\theta}, \mathrm{e}^{-2\mathrm{i}\theta};q)_\infty |\cos\theta\rangle\langle\cos\theta| \tag{5.45}$$

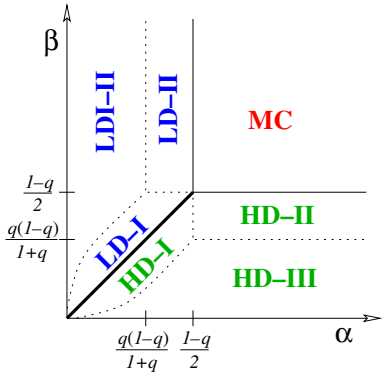


Figure 28. Full phase diagram of the PASEP showing additional sub-phases that have different density decay forms.

which is the final part of the jigsaw that is an expression for the normalization. One can verify that the same integral (5.28) is obtained using this alternative representation.

The key benefit offered by this approach is that expressions for the difference in density between neighbouring pairs of sites

$$\Delta_i \equiv \rho_i - \rho_{i+1} = \frac{\langle W|C^{i-1}[DC - CD]C^{N-i-2}|V\rangle}{Z_N} \tag{5.46}$$

involve the matrix $DC - CD = DE - ED$ which is diagonal in this representation. This one sees from (5.37) and the definitions (5.35) and (5.36). Specifically,

$$\langle n|[DE - ED]|m\rangle = \frac{\langle n|[\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}]|m\rangle}{1-q} \tag{5.47}$$

$$= \frac{[(1-q^{n+1}) - (1-q^n)]}{1-q}\delta_{n,m} = q^n\delta_{n,m}. \tag{5.48}$$

To express the density gradient (5.46) in terms of integrals over the q-Hermite polynomials, one inserts two of the identities (5.45), one in front of the combination $[DC - ED]$ and one behind. This gives rise to a double integral which is analysed for large N and i in much the same way as discussed in subsection 5.3 but with additional complexity stemming from the fact that one has the deformation of two contours to contend with.

The detailed analysis of this double integral was conducted in [123]. As with the ASEP, one finds additional sub-phases relating to different boundary profiles within the high-density and low-density phases. The full phase diagram is as shown in figure 28 and the density profiles one obtains are summarized as follows:

- *Maximal-current phase* $\alpha > \frac{1-q}{2}$, $\beta > \frac{1-q}{2}$. Once again, one has the universal power-law density profile seen both in the ASEP with open boundaries and on a ring for sufficiently large α and β. That is, near the right boundary one has the decay

$$\rho_{N+1-j} \sim \frac{1}{2}\left(1 - \frac{1}{\sqrt{\pi j}}\right) \tag{5.49}$$

and a similar decay from the left obtained by invoking the particle hole symmetry that was also present in the ASEP, namely, $\rho_i(\alpha, \beta) = 1 - \rho_{N+1-i}(\beta, \alpha)$.

- *Low-density phases* $\alpha < \frac{1-q}{2}, \alpha < \beta$. In all of the low-density phases there is a region of constant density $\rho = \frac{\alpha}{1-q} < \frac{1}{2}$ that extends from the left boundary into the bulk. Near the right boundary the decay takes three different forms which define three sub-phases.
  - *LD-I* $\alpha < \beta < \frac{\alpha}{q+a}, \beta < \frac{1-q}{2}$. Here, there is an exponential decay with a characteristic length ξ given by
$$\xi^{-1} = \ln\left[\frac{\beta(1-q-\beta)}{\alpha(1-q-\alpha)}\right]. \tag{5.50}$$
    This phase is present in the totally symmetric case $q = 0$ (see figure 17).
  - *LD-II* $\frac{q(1-q)}{1+q} < \alpha < \frac{1-q}{2}, \beta > \frac{1-q}{2}$. This phase, which is also present when $q = 0$, has an exponential decay characterized by
$$\xi^{-1} = \ln\left[\frac{(1-q)^2}{4\alpha(1-q-\alpha)}\right] \tag{5.51}$$
    which is additionally modulated by a power-law decay with exponent $\frac{3}{2}$, as in the totally asymmetric case.
  - *LD-III* $\beta > \frac{\alpha}{q+a}, \alpha < \frac{q(1-q)}{1+q}$. When $q > 0$, a new phase appears that has a purely exponential decay with a characteristic length
$$\xi^{-1} = \ln\left[\frac{q}{(q+\alpha)^2}\right]. \tag{5.52}$$
- *High-density phases* $\beta < \frac{1-q}{2}, \beta < \alpha$. The properties of the high-density phases are obtained from the particle–hole symmetry. These all have an exponential decay at the left boundary to a constant bulk density that extends to the right boundary.

5.5. *Reverse-bias phase*

The foregoing analysis applies to the case $q < 1$, since this is a restriction on the representations of the identity (5.25) and (5.45) that were required in the analysis. As we previously noted, the case where $q > 1$ is of interest since then the particles in the bulk prefer to hop to the left, even though the boundary conditions enforce a current that flows in the opposite direction.

Although the q-Hermite polynomials have an extension to $q > 1$ [130] it is unclear how to handle the resulting integral representation of the normalization and other matrix-product expressions [122]. An alternative approach, followed in [122], is to note that if one were actually to evaluate the integral (5.28) by performing, say, a Laurent expansion, one would end up with a sum containing a finite number of terms. This is because this integral is another way to express the result of a direct reordering using the reduction relations (5.1)–(5.3), a procedure that is known to terminate. Such an expression is then valid for any q, not just $q < 1$ as was required to derive the integral representation in the first place. In other words, we can access the asymptotics for $q > 1$ by first fixing N at some finite value with $q < 1$, put $q > 1$ in the result and only then at the very end take $N \to \infty$.

It is possible, using a clutch of identities from the theory of q-series, to find the finite sum implied by (5.28) [122]. We do not quote this formula—a complicated mixture of ordinary and q-combinatorial factors—here. Rather, we simply state that the leading term in the normalization for $q > 1$ is, for large N [122],

$$Z_N \sim A(q,\alpha,\beta)\left(\frac{1}{q-1}\sqrt{\frac{(q-1+\alpha)(q-1+\beta)}{\alpha\beta}}\right)^N q^{\frac{1}{4}N^2} \tag{5.53}$$

where $A(q, \alpha, \beta)$ is a function that is independent of N. Note that this result holds for all values of α and β: there is only one reverse-bias phase. For (5.33) we see that in this phase, the current vanishes asymptotically as

$$J \sim \left(\frac{\alpha\beta(q-1)^2}{(q-1+\alpha)(q-1+\beta)} \right)^{\frac{1}{2}} q^{\frac{1}{4}-\frac{1}{2}N}. \tag{5.54}$$

Although the density profile in the reverse-bias phase has not been calculated explicitly, one can conjecture its form using a simple argument. First we expect most of the particles to be in a domain by the left boundary and most of the vacant sites by the right. Now, if the interface between these particle- and hole-rich regions were closer to the right boundary than the left, a particle at the interface will escape from the right boundary in a shorter time than a hole from the left. Therefore, in this instance the interface will tend to drift to the left. On the other hand, if the interface were closer to the left boundary it would, by the same argument, drift to the right. Therefore, the lattice will typically be half full in the steady state, a feature noticed in computer simulations for small systems.

5.6. Entry and exit at both boundaries

We finally briefly mention the case of the PASEP where particles may enter and leave the system at both boundaries, as shown in figure 24. A representation similar to that used in subsection 5.2 was established in [124], where it was also shown that the functions appearing in the eigenvector of the matrix C are a complete set of Askey–Wilson polynomials [131, 126]. These are a set of very general polynomials that contain both the Al-Salam–Chihara and q-Hermite polynomials encountered above as special cases. They have four parameters related to the transition rates α, β, γ and δ as follows:

$$a_\pm = \frac{1}{2\alpha}[(1-q-\alpha+\gamma) \pm \sqrt{(1-q-\alpha+\gamma)^2+4\alpha\gamma}] \tag{5.55}$$

$$b_\pm = \frac{1}{2\beta}[(1-q-\beta+\delta) \pm \sqrt{(1-q-\beta+\delta)^2+4\beta\delta}]. \tag{5.56}$$

As in subsection 5.2, the boundary vectors are $\langle W| = \langle 0|$ and $|V\rangle = |0\rangle$, and the polynomials satisfy an orthogonality relation similar to (5.23), but with additional factors appearing to accommodate the new parameters [131, 126]. These carry through to the expression one eventually obtains for the normalization

$$Z_N = \frac{(q, a_+a_-, b_+b_-, a_+b_+, a_+b_-, a_-b_+, a_-b_-; q)_\infty}{2\pi(a_+b_+a_-b_-, q)_\infty} \times \int_0^\pi \mathrm{d}\theta \frac{(\mathrm{e}^{2\mathrm{i}\theta}, \mathrm{e}^{-2\mathrm{i}\theta}; q)_\infty}{(a_+\mathrm{e}^{\mathrm{i}\theta}, a_+\mathrm{e}^{-\mathrm{i}\theta}, b_+\mathrm{e}^{\mathrm{i}\theta}, b_+\mathrm{e}^{-\mathrm{i}\theta}, a_-\mathrm{e}^{\mathrm{i}\theta}, a_-\mathrm{e}^{-\mathrm{i}\theta}, b_-\mathrm{e}^{\mathrm{i}\theta}, b_-\mathrm{e}^{-\mathrm{i}\theta}; q)_\infty} \left[\frac{2(1+\cos\theta)}{1-q} \right]^N. \tag{5.57}$$

One recognizes a structure similar to (5.28) and the asymptotic analysis via a contour integral detailed in [124] is basically the same as described in subsection 5.3. When $q < 1$, one finds the three familiar phases: a maximal-current phase when $a_+ > 1$, $b_+ > 1$; a low-density phase when $a_+ > 1$ and $a_+ > b_+$; and a high-density phase when $b_+ > 1$ and $b_+ > a_+$. When $q > 1$ the same three phases are manifested as the model is invariant under the exchanges $(\alpha, \beta, \gamma, \delta, q) \leftrightarrow (q^{-1}\delta, q^{-1}\gamma, q^{-1}\beta, q^{-1}\alpha, q^{-1})$, at least when all four boundary rates are nonzero. Since the physical properties of these phases are very similar to their counterparts in the ASEP with open boundaries, we do not discuss them further here. Further details,

including an analysis of the bulk densities through the introduction of a fugacity as described in subsection 4.3.1 are given in [124]. As further shown in [125], the eigenvalue spectrum of the matrix C can be used to determine the density profiles and correlation functions. These show once again the universal power-law decay in the maximal-current phase, and various exponential decays in the high- and low-density phases.

6. Macroscopic density profiles for open-boundary ASEP

In this review, we have so far mostly concentrated on how the current and density profiles in driven-diffusive systems are calculated from the matrix-product expressions for the steady state. Since one knows the entire stationary distribution of microstates, it is in principle possible to calculate any macroscopic quantity of interest. For example, the existence of a matrix-product solution has recently admitted the calculation of the Gibbs–Shannon entropy in the symmetric exclusion process with open boundaries [132]. In this section we take a closer look at a quantity that has received a great deal of recent attention in the context of open driven-diffusive systems, namely free-energy functionals which characterize the probability that a deviation from the most likely density profile is seen in a macroscopically large (but finite) system.

To be more precise, we consider macrostates corresponding to a density profile $\rho(x)$ where x lies in the (one-dimensional) interval [0, 1] which contains some large number N of lattice sites. As we shall see explicitly below, each macroscopic profile $\rho(x)$ can be realized in number of ways that grows exponentially with the number of microscopic degrees of freedom N. This exponential growth leads to the key property of statistical mechanical systems that in the thermodynamic limit, $N \to \infty$, one particular macrostate—let us call this $\bar{\rho}(x)$—is very much more likely to be realized than any other (except possibly at a phase transition). That is, the probability $P[\rho]$ of seeing a particular macrostate behaves for large N as

$$P[\rho] \sim \exp\left(-N\mathcal{F}[\rho]\right) \tag{6.1}$$

where $\mathcal{F}[\rho]$ is a functional of the density profile that vanishes when $\rho(x) = \bar{\rho}(x)$. In an equilibrium system, $\mathcal{F}[\rho]$ is essentially the free energy of the macrostate, relative to the most likely profile. Of course, in a nonequilibrium steady state, the probability of seeing a deviation from the most likely profile can be measured, and so by analogy with the equilibrium case we refer to the object $\mathcal{F}[\rho]$ that results as a free-energy functional. Progress in calculating such free-energy functionals and related quantities has been summarized in the lecture notes of Derrida [133].

6.1. Free-energy functional for the symmetric simple exclusion process (SSEP)

In this subsection, we will sketch the derivation of the free-energy functional for the nonequilibrium symmetric simple exclusion process (SSEP) with open boundaries that has been achieved through the matrix-product solution for the steady-state distribution [134, 135]. This version of the model has symmetric hopping of a single species at unit rate in the bulk, entry at the left and right boundaries at rates α and δ, respectively, and exit at the left and right boundaries at rates β and γ. In other words, we have the dynamics illustrated in figure 24 of section 5 with the bias $q = 1$.

In order to illustrate the procedure by which the free-energy functional was first calculated in [134, 135] from the matrix-product expressions, we shall follow it here for the much simpler case in which the model parameters are chosen so as to give rise to an *equilibrium* steady state. To find this set of parameters, we employ a Kolmogorov criterion [11, 136]. This states that

if for every loop $\mathcal{C}_1 \to \mathcal{C}_2 \to \cdots \to \mathcal{C}_n \to \mathcal{C}_1$ in configuration space, the equality

$$W(\mathcal{C}_1 \to \mathcal{C}_2)W(\mathcal{C}_2 \to \mathcal{C}_3)\cdots W(\mathcal{C}_{n-1} \to \mathcal{C}_n)W(\mathcal{C}_n \to \mathcal{C}_1)$$
$$= W(\mathcal{C}_n \to \mathcal{C}_{n-1})W(\mathcal{C}_{n-1} \to \mathcal{C}_{n-2})\cdots W(\mathcal{C}_2 \to \mathcal{C}_1)W(\mathcal{C}_1 \to \mathcal{C}_n) \tag{6.2}$$

holds, the steady state of the system is an equilibrium state for which the detailed balance relation

$$\frac{f(\mathcal{C})}{f(\mathcal{C}')} = \frac{W(\mathcal{C}' \to \mathcal{C})}{W(\mathcal{C} \to \mathcal{C}')} \tag{6.3}$$

also holds for every pair of configurations $\mathcal{C}$ and $\mathcal{C}'$ between which transitions occur at nonzero rate. Here we find that (6.2) is automatically satisfied unless, in total, some nonzero number m particles exits at one boundary and re-enters at the other to return to the starting configuration going one way round the loop. Then, the ratio of the two products in (6.2) is $(\alpha\beta/\gamma\delta)^m$. Therefore, to realize an equilibrium steady state one must have $\alpha\beta = \gamma\delta$.

Since the bulk hopping occurs at the same rate in both directions, the detailed balance relation (6.3) implies that all configurations that have M particles on the lattice are equally likely in the steady state. By considering a pair of configurations that differ by the addition or removal of a single particle at the left boundary, one finds that if f_M is the steady-state weight of any one of the M-particle configurations,

$$f_{M+1} = \frac{\alpha}{\gamma} f_M. \tag{6.4}$$

Hence, the probability of seeing a particular configuration $\tau_1, \tau_2, \ldots, \tau_N$ (where, as previously, $\tau_i = 1$ if site i is occupied and zero otherwise) is given by the product

$$P(\tau_1, \tau_2, \ldots, \tau_N) = \prod_{i=1}^{N} \frac{\left(\frac{\alpha}{\gamma}\right)^{\tau_i}}{1 + \frac{\alpha}{\gamma}}. \tag{6.5}$$

Once the underlying distribution of microstates has been established, the free-energy functional is obtained by taking an appropriate combination of thermodynamic and continuum limits. To this end, we divide the system into a number n of boxes, box j containing $N_j = y_j N$ sites and $M_j = \rho_j N_j$ particles. The j pairs (y_j, ρ_j) can thus be used to specify a macrostate of the system in the limit $N \to \infty$. To find the probability of this macrostate we note that

$$P(N_1, M_1; N_2, M_2; \cdots; N_j, M_j) = \prod_{i=1}^{j} P_{\text{box}}(N_i, M_i) \tag{6.6}$$

where $P_{\text{box}}(N, M)$ is the probability that N particles are in a box of size M in the steady state. This is given by

$$P_{\text{box}}(N, M) = \binom{N}{M} \frac{\left(\frac{\alpha}{\gamma}\right)^M}{\left(1 + \frac{\alpha}{\gamma}\right)^N} \tag{6.7}$$

$$\sim \exp\left(-N\left[\ln\left(1 + \frac{\alpha}{\gamma}\right) - \frac{M}{N}\ln\frac{\alpha}{\gamma} + \frac{M}{N}\ln\frac{M}{N} + \left(1 - \frac{M}{N}\right)\ln\left(1 - \frac{M}{N}\right)\right]\right) \tag{6.8}$$

where the second expression, valid for large N, has been obtained using Stirling's approximation. Inserting this into (6.6), and taking the thermodynamic limit, one finds that

$$-\lim_{N\to\infty} \frac{1}{N} \ln P(y_1, \rho_1; \cdots; y_j, \rho_j) = \sum_{i=1}^{j} y_i \left[\rho_i \ln \frac{\rho_i}{\frac{\alpha}{\gamma+\alpha}} + (1 - \rho_i)\ln\frac{1 - \rho_i}{1 - \frac{\alpha}{\gamma+\alpha}}\right] \tag{6.9}$$

in terms of the intensive box sizes y_j and densities ρ_j. The continuum limit is now straightforward: we take $y_i = \delta x_i \to 0$ and $j \to \infty$ where δx_i is some portion of the interval $x \in [0, 1]$ such that $\sum_i \delta x_i = 1$ and the density profile $\rho(x)$ becomes a function of x. This procedure yields

$$\mathcal{F}^{\mathrm{eq}}[\rho] = \int_0^1 \mathrm{d}x \left[\rho(x) \ln \frac{\rho(x)}{\bar{\rho}} + (1 - \rho(x)) \ln \frac{1 - \rho(x)}{1 - \bar{\rho}} \right] \tag{6.10}$$

for the free-energy functional, where we have introduced the parameter $\bar{\rho} = \frac{\alpha}{\alpha+\gamma}$. It is easily verified that $\mathcal{F}[\rho(x)]$ vanishes if $\rho(x) = \bar{\rho}\, \forall x \in [0, 1]$, indicating that the flat profile is the most likely profile in the thermodynamic limit as one would expect at equilibrium. We also note that the free-energy functional is local in the density field $\rho(x)$; that is, one can express (6.10) as an integral over a free-energy density $f(x)$ that depends only on the local properties of the density field $\rho(x)$ at the point x (and nowhere else).

When the equilibrium constraint on the boundary rates $\alpha\beta = \gamma\delta$ is relaxed, this local property of the free-energy density no longer holds. The reason for this is that when the system is divided into boxes, the joint distribution of box sizes and occupancy does not factorize as in (6.6). That is, there are long-range correlations in the density profiles that are implied by the matrix-product structure of the steady-state weights. This makes evaluating the free-energy functional by taking the appropriate combination of thermodynamic and continuum limits a much more complicated exercise. We therefore omit the technical details of the calculation here, referring the reader to [135] for a full account.

The way in which the free-energy functional is obtained in [134, 135] is to introduce a pair of fugacities for each box, one of which controls its mean size y_i and the other its density ρ_i. This procedure is similar to the use of the grand-canonical ensemble described in section 4.3.1, albeit with a much larger number of fugacities, and yields the generating function of the joint distribution that extends to the nonequilibrium case the expression (6.6). The values of the fugacities are then fixed by finding the saddle point of this generating function. It turns out that the resulting expression for the free-energy functional,

$$\mathcal{F}[\rho] = \int_0^1 \mathrm{d}x \left[\rho(x) \ln \frac{\rho(x)}{\sigma(x)} + (1 - \rho(x)) \ln \frac{1 - \rho(x)}{1 - \sigma(x)} + \ln \frac{\sigma'(x)}{\sigma(1) - \sigma(0)} \right], \tag{6.11}$$

involves a companion function $\sigma(x)$ which is related to the fugacities described above, and that is fixed by the saddle-point analysis as the function that maximizes the functional $\mathcal{F}[\rho]$ for a given profile $\rho(x)$. One can verify, by asking for a vanishing functional derivative $\frac{\delta\mathcal{F}}{\delta\sigma(x)}$, that this function must satisfy the nonlinear differential equation

$$\sigma(x) + \sigma(x)(1 - \sigma(x)) \frac{\sigma''(x)}{\sigma'(x)^2} = \rho(x). \tag{6.12}$$

Additionally, the detailed analysis presented in [134, 135] shows that $\sigma(x)$ must further be a monotonic function and satisfy the boundary conditions

$$\sigma(0) = \frac{\alpha}{\alpha + \gamma} \qquad \text{and} \qquad \sigma(1) = \frac{\delta}{\delta + \beta}. \tag{6.13}$$

We note that the monotonicity requirement on $\sigma(x)$ ensures that the argument of the final logarithm appearing in (6.11) is positive and further that the value of the companion function $\sigma(x)$ at some point x will in general depend on the *entire* profile $\rho(x)$ through the differential equation (6.12)—the free-energy density is non-local, as previously claimed.

6.2. Effective local thermal equilibrium and additivity principle

The companion function $\sigma(x)$ that appears in the nonequilibrium free-energy functional (6.11) is somewhat mysterious at a first encounter. However, a physical meaning can be ascribed to

this function, starting from the observation that the values imposed at the boundaries coincide with the mean densities of the left and right boundary sites in the steady state, and hence also the densities of the particle reservoirs at each end of the system. In the bulk, one finds that an infinitesimal segment of the system located at a point $x \in (0, 1)$ can be thought of as being in local equilibrium with a reservoir at density $\sigma(x)$.

To see this, it is convenient first to generalize the functional (6.11) to a density profile $\rho(x)$ defined on an arbitrary interval $x \in [a, b]$ in such a way that if the profile is translated and scaled so that it covers a different interval $x \in [a', b']$, the free-energy per unit length is unchanged. The functional that has this property and coincides with (6.11) for the case $a = 0, b = 1$ is

$$\mathcal{F}_{[a,b]}([\rho]; \sigma(a), \sigma(b)) = \int_a^b \mathrm{d}x \left[\rho(x) \ln \frac{\rho(x)}{\sigma(x)} + (1 - \rho(x)) \ln \frac{1 - \rho(x)}{1 - \sigma(x)} + \ln \frac{(b-a)\sigma'(x)}{\sigma(b) - \sigma(a)} \right]. \tag{6.14}$$

In the limit of an infinitesimal interval, $b \to a$, this expression approaches the corresponding expression for the equilibrium system

$$\mathcal{F}^{\mathrm{eq}}_{[a,b]}([\rho]; \bar{\rho}) = \int_a^b \mathrm{d}x \left[\rho(x) \ln \frac{\rho(x)}{\bar{\rho}} + (1 - \rho(x)) \ln \frac{1 - \rho(x)}{1 - \bar{\rho}} \right] \tag{6.15}$$

if we identify $\sigma(a)$ with $\bar{\rho}$, the density of the reservoirs coupled to the equilibrium system.

The observation that an effective local thermal equilibrium applies does not itself allow one to recover the full expression (6.14) for the free-energy functional since *a priori* one does not know what densities the effective intermediate reservoirs should have. However, it was noticed in [135] that one can construct the free-energy functional through an additivity principle that involves the modified free energy

$$\mathcal{H}_{[a,b]}([\rho]; \sigma(a), \sigma(b)) = \mathcal{F}_{[a,b]}([\rho]; \rho_a, \rho_b) + (b - a) \ln J(\sigma(a), \sigma(b)) \tag{6.16}$$

where $J(\sigma(a), \sigma(b))$ is the steady-state current through a system of length $b - a$ coupled to a reservoir of density $\sigma(a)$ at the left boundary and of density $\sigma(b)$ at the right. For the SSEP, this current is

$$J = \frac{\sigma(b) - \sigma(a)}{b - a}, \tag{6.17}$$

where we have assumed that $\sigma(b) > \sigma(a)$. (Since the dynamics are symmetric, the case $\sigma(a) > \sigma(b)$ can be treated by making the replacement $x \to 1 - x$.) The additivity principle given in [135] relates the free-energy functional for a system to that of two subsystems created by inserting a reservoir at a point $y \in (a, b)$. It reads

$$\mathcal{H}_{[a,b]}([\rho]; \sigma(a), \sigma(b)) = \max_{\sigma(y)} \{\mathcal{H}_{[a,y]}([\rho]; \sigma(a), \sigma(y)) + \mathcal{H}_{[y,b]}([\rho]; \sigma(y), \sigma(b))\}. \tag{6.18}$$

That is, $\sigma(y)$ is the density one should choose for the reservoir placed at the point y so that the combined free energy of the two subsystems is maximized. One can verify by recursively subdividing the subsystems created by inserting intermediate reservoirs, and by assuming that each infinitesimal segment created in this way is in a local equilibrium with its boundary reservoirs, that the expression (6.11) results. The fact that one is looking for a maximum in (6.18) further implies that the reservoir densities $\sigma(x)$ must satisfy the differential equation (6.12). Thus the additivity principle (6.18) in tandem with the effective local-equilibrium property and the free-energy functional given by (6.12) are equivalent.

6.3. *Properties and applications of the free-energy functional for the SSEP*

It is worth reiterating a few properties exhibited by the free-energy functional (6.11) noted in [135]. First, as is required, the equilibrium version (6.10) is recovered in the limit where the right boundary reservoir density $\sigma(1)$ approaches that of the left $\sigma(0)$. To see this, we write (without loss of generality) for the effective reservoir density in the bulk

$$\sigma(x) = \sigma(0) + [\sigma(1) - \sigma(0)]x + \delta\sigma(x). \tag{6.19}$$

If one substitutes this expression into the differential equation (6.12) that governs $\sigma(x)$, one finds that the function $\delta\sigma(x)$ must be proportional to $[\sigma(1) - \sigma(0)]^2$ if $\sigma(x)$ is to remain bounded as $\sigma(1) \to \sigma(0)$. In turn, this implies that $\sigma(x) \to \sigma(0) = \sigma(1)$ in this limit, yielding the desired equilibrium free-energy functional, with $\bar{\rho}$ equal to both boundary reservoir densities.

It is also straightforward to show that the most probable profile coincides with the stationary profile, which for the SSEP is simply the linear function $\bar{\rho}(x) = \sigma(0) + [\sigma(1) - \sigma(0)]x$. This is achieved by asking for the functional derivative of (6.11) with respect to $\rho(x)$ to vanish. That is,

$$\frac{\delta\mathcal{F}[\rho]}{\delta\rho(x)} = \ln\left[\frac{\rho(x)}{\sigma(x)}\frac{1-\sigma(x)}{1-\rho(x)}\right] = 0, \tag{6.20}$$

and hence we must have $\sigma(x) = \rho(x)$. The differential equation (6.12) then implies that the second derivative of $\sigma(x)$ must vanish, which (along with the boundary conditions) results in the optimal $\sigma(x)$ and $\rho(x)$ coinciding with the linear stationary profile. It is easy to see that then (6.11) is zero, indicating that this profile appears with probability 1 in the thermodynamic limit; it can also be shown [135] that any other macroscopic profile appears with a probability exponentially small in the number of lattice sites.

It is interesting to compare the magnitude of the free-energy functional for a given profile $\rho(x)$ with that for an equilibrium system that has the same stationary profile. This can be realized by coupling a one-dimensional chain of sites along its length to a series of particle reservoirs that have their chemical potentials tuned in such a way that the desired equilibrium density $\bar{\rho}(x)$ in a small interval $x \in [a, b]$ is attained. Since the equilibrium free-energy density is local, the free-energy functional for the whole system is obtained by summing (6.15) over all these small intervals. This gives

$$\mathcal{F}^{\mathrm{eq}}[\rho] = \int_0^1 \mathrm{d}x \left[\rho(x)\ln\frac{\rho(x)}{\bar{\rho}(x)} + (1-\rho(x))\ln\frac{1-\rho(x)}{1-\bar{\rho}(x)}\right]. \tag{6.21}$$

Now (6.11) reduces to the equilibrium expression (6.21) when $\sigma(x) = \bar{\rho}(x)$, the equilibrium profile (which is linear). However, the nonequilibrium free energy is obtained by maximizing the expression (6.11) with respect to σ and generally the maximum is not realized by $\sigma(x) = \bar{\rho}(x)$. Therefore, $\mathcal{F}[\rho] \geqslant \mathcal{F}^{\mathrm{eq}}[\rho]$ which means that the probability of witnessing a particular fluctuation away from the optimal profile for a system driven out of equilibrium is suppressed compared to that for an equilibrium system with the same optimal profile. This is not always the case, however: a similar analysis for the asymmetric simple exclusion process (see below) shows that fluctuations can also be enhanced relative to the equilibrium state.

A final further application of the free-energy functional is to explore the optimal profiles in the nonequilibrium system after imposing additional global constraints. For example, one can ask for the most likely profile given that the overall density $\bar{\rho} = \int_0^1 \mathrm{d}x\, \rho(x)$ is fixed. It turns out [135] that the most likely profile has an exponential form and, unless the overall imposed density happens to be equal to the equilibrium density (in which limit the linear profile is recovered), the density is discontinuous at the boundaries $x = 0, 1$.

6.4. *Free-energy functional for the partially asymmetric simple exclusion process (PASEP)*

Free-energy functionals have also been derived for the partially asymmetric exclusion process (PASEP) in the case where the entry and exit of particles occur at the left and right boundaries, respectively, and the bulk bias is to the right (i.e., the forward-bias regime, $q < 1$) [137, 138]. In principle, one could follow through the procedure outlined in subsection 6.1, where one starts with the full stationary distribution of microstates and takes the continuum limit. It turns out to be more straightforward instead to use the microscopic distribution to extend the additivity principle discussed in subsection 6.2. Then, by the effective local-equilibrium property, one can construct the full free-energy functional as before.

When the bulk dynamics are asymmetric, it turns out that two versions of the additivity principle come into play. One applies when the left boundary reservoir density $\sigma(0)$ is less than that at the right; the other when the reverse is true. In both cases, the additivity relation involves the modified free energy defined by equation (6.16), but where now the current is that found (for example) by an application of the extremal current principle discussed in section 2. That is,

$$J(\sigma,\sigma') = \begin{cases} \min\limits_{\rho\in[\sigma,\sigma']} \rho(1-\rho) & \sigma \leqslant \sigma' \\ \max\limits_{\rho\in[\sigma',\sigma]} \rho(1-\rho) & \sigma \geqslant \sigma'. \end{cases} \tag{6.22}$$

When $\sigma(0) \geqslant \sigma(1)$, the additivity formula is the same as for the symmetric case, equation (6.18) (see [138] for the details of the calculation). Using the local-equilibrium property, one then finds the free-energy functional to be

$$\mathcal{F}_{[a,b]}([\rho];\sigma(a),\sigma(b)) = -(b-a)\ln J(\sigma(0),\sigma(1)) + \max_{\sigma(x)} \int_a^b \mathrm{d}x\,[\rho(x)\ln(\rho(x)[1-\sigma(x)]) + (1-\rho(x))\ln[(1-\rho(x))\sigma(x)]]\,. \tag{6.23}$$

Again, the companion function $\sigma(x)$ that gives the effective reservoir densities in the bulk must match the actual reservoir densities at the boundaries and must also be a nonincreasing function.

When $\sigma(0) \leqslant \sigma(1)$, the additivity formula takes on the different form

$$\mathcal{H}_{[a,b]}([\rho];\sigma(a),\sigma(b)) = \min_{\rho_y\in\{\sigma(a),\sigma(b)\}} \{\mathcal{H}_{[a,y]}([\rho];\sigma(a),\sigma(y)) + \mathcal{H}_{[y,b]}([\rho];\sigma(y),\sigma(b))\}, \tag{6.24}$$

where again we refer the reader to [138] for the details of the calculation. The fact that the intermediate reservoir inserted at the point y always takes on a density that is equal to that of one of the boundary reservoirs implies that the effective reservoir density is constant on one side of the point y. Further subdivision on that side then has no effect. On the other side the situation is the same as for the previous subdivision: on one side of the division, the density will be constant and equal to that of the appropriate boundary reservoir, whereas on the other further subdivisions will be required. The upshot of this is that the reservoir density function $\sigma(x)$ will be a step function, taking the value $\sigma(x) = \sigma(a)$ for $a < x < y$ and $\sigma(x) = \sigma(b)$ for $y < x < b$ for some value y. This value will be such that the overall free energy is minimized, that is

$$\mathcal{F}_{[a,b]}([\rho];\sigma(a),\sigma(b)) = -(b-a)\ln J(\sigma(a),\sigma(b)) + \min_{a\leqslant y\leqslant b} \left\{ \int_a^y \mathrm{d}x\,[\rho(x)\ln(\rho(x)[1-\sigma(a)]) + (1-\rho(x))\ln[(1-\rho(x))\sigma(a)]] + \int_y^b \mathrm{d}x\,[\rho(x)\ln(\rho(x)[1-\sigma(b)]) + (1-\rho(x))\ln[(1-\rho(x))\sigma(b)]] \right\}. \tag{6.25}$$

Note that despite considerable cosmetic differences between this formula and (6.23), the two expressions are in fact very similar: in both cases, one needs to find the set of reservoir densities $\sigma(x)$ that leads to an extremum of the same joint functional of $\rho(x)$ and $\sigma(x)$. The fact that in one case the additivity principle involves a maximum and in another a minimum has its origins in the nature of the saddle point which provides the relevant asymptotics [133].

We also note that the bias parameter q does not enter explicitly into these equations, only implicitly through the relationship between the boundary reservoir densities $\sigma(a)$ and $\sigma(b)$, and the microscopic transition rates α, β, γ, δ and q. Specifically

$$\sigma(a) = \frac{\alpha}{1-q} \qquad \text{and} \qquad \sigma(b) = 1 - \frac{\beta}{1-q}, \tag{6.26}$$

and so q affects phase boundaries in the α–β plane, as shown in section 5. The q-independence of the free-energy functionals does not contradict the rich structure seen in the density profiles in section 5, since the deviations from the bulk density at the boundaries vanish under rescaling in the $N \to \infty$ limit. This lack of q-dependence also implies that (6.11) is not obtained as $q \to 1$: in other words, the limits $N \to \infty$ and $q \to 1$ do not commute (as we have already seen). The free-energy functional in the weakly asymmetric limit $1 - q = \lambda/N$ has been computed explicitly using the matrix-product approach and on this scale a q (or λ) dependence is apparent [139].

Closer scrutiny of these free-energy functionals for the PASEP [138] shows the similar properties to those seen for the SSEP in subsection 6.3. For example, one finds that the stationary profile (here, a constant profile after rescaling in the thermodynamical limit) is also the most likely profile, except along the boundary between the high- and low-density phases along which—as has been discussed before—there is a superposition of shocks with the shock location distributed uniformly across the system. Then, any one of these shock profiles is found to minimize the free-energy functional. Again, one can compare the relative size of a fluctuation away from the most likely profile in the nonequilibrium system and an equilibrium system coupled to a reservoir with a spatially-varying chemical potential. In the case where $\sigma(a) > \sigma(b)$, it is found that (as for the SSEP) such fluctuations are suppressed; however, when $\sigma(a) < \sigma(b)$ the opposition between the boundary densities and the bulk bias results in these fluctuations being enhanced.

A final interesting feature of the PASEP is that density fluctuations in the maximal-current phase are non-Gaussian. Evidence for this is provided by the existence of a discontinuity at a density of $\rho = \frac{1}{2}$ in the probability of witnessing a density ρ in a box located somewhere in the bulk of the system. That the distribution is indeed non-Gaussian is confirmed by explicit calculations for the totally asymmetric case ($q = 0$) [138, 140] that exploit the relationship to equilibrium surface models outlined in subsection 3.6.

6.5. Remarks

In this section we have seen that once one has found a function that is additive when two subsystems are connected together via a reservoir with an appropriate density, the distribution of density profile macrostates can be constructed given the existence of a local-equilibrium property. This is an appealing approach, and one hopes that it can be applied to a much wider range of systems than exclusion processes. The difficulty is, however, that one does not know *a priori* what form the additivity principle should take: here, we have relied on the complete knowledge of the underlying distribution of microstates to construct it. Nevertheless, it is worth remarking that the free-energy functional for the symmetric exclusion process (6.11) has been obtained in a purely macroscopic formalism [4]. In this approach it is assumed that, in the combined thermodynamic and continuum limit, the macroscopic density profile evolves

deterministically as a consequence of the law of large numbers. Then, it can be shown that the probability of witnessing a deviation $\rho(x)$ from the most likely profile $\bar{\rho}(x)$ in the steady state is given by a functional of the most likely trajectory through phase space that begins at time $t = -\infty$ at stationarity $\bar{\rho}(x)$ and is constrained to reach $\rho(x)$ at $t = 0$ [3].

7. Two-species models with quadratic algebra

So far we have seen exact matrix-product solutions for the steady state of two classes of model systems with nonequilibrium dynamics: the open-boundary ASEP and PASEP—various aspects of which were discussed in detail in sections 2, 3, 5 and 6—which had a single species of particles hopping on a one-dimensional lattice with open boundaries; a two-species model with periodic boundary conditions, which we examined in section 4. In this section we are going to search for two-species exclusion models that can be solved using the matrix-product approach. For clarity, we reiterate that in this work the number of species n relates to the number of particle species, excluding vacancies.

It is not known how to perform an exhaustive search of *all* possible matrix-product steady states, but it is possible to search a restricted subset where the matrix products involved can be systematically *reduced* using expressions such as (2.39)–(2.41) and (4.4)–(4.6). To this end, let us recall the proof of the reduction relations that applied for exclusion models on the ring geometry given in subsection 4.2. This proof concerned stationary weights that were given by a trace of matrices

$$f(\tau_1, \tau_2, \ldots, \tau_N) = \mathrm{tr}\left(X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N}\right) \tag{7.1}$$

where the variable $\tau_i = 0, 1, 2, \ldots, n$ indicates the species of particle occupying site i ($\tau_i = 0$ denotes a vacancy) and X_{τ_i} is the corresponding matrix whose form is to be determined. We showed that when $W(\tau_i \tau_{i+1} \to \tau_{i+1} \tau_i)$ is the rate at which particles on sites i and $i+1$ exchange places, if one can find auxiliary matrices $\tilde{X}_i$ such that

$$W(\tau_{i+1}\tau_i \to \tau_i \tau_{i+1}) X_{i+1} X_i - W(\tau_i \tau_{i+1} \to \tau_{i+1} \tau_i) X_i X_{i+1} = \tilde{X}_i X_{i+1} - X_i \tilde{X}_{i+1} \tag{7.2}$$

is satisfied, the weights given by (7.1) are stationary.

If we are to obtain matrix reduction relations, we require that the auxiliaries $\tilde{X}_i$ are scalars (rather than matrices). We must also insist that these reduction relations describe an associative algebra, i.e., that no matter what order the reduction relations are applied, one always ends up with the same sum of irreducible strings. The various ways in which this can be achieved were formalized, classified and catalogued by Isaev, Pyatov and Rittenberg [141]. In this section, we briefly outline the classification scheme for the case of two particle species, $n = 2$, and show what physical dynamics the various possibilities correspond to. We then move on to discuss geometries other than the ring.

7.1. Classification of Isaev, Pyatov and Rittenberg for two-species models

In this section we will mostly use the notation established above, where vacancies are denoted by 0 and particles by 1 and 2. Occasionally, it will be helpful to state model dynamics using a natural 'charge representation', where particles of species 1 and vacancies are relabelled as positive and negative charges and particles of species 2 as vacancies:

$$1 \mapsto + \qquad 2 \mapsto 0 \qquad 0 \mapsto -. \tag{7.3}$$

Note that our notation differs from that used by [141]. We will also use a shorthand for the transition rates

$$w_{\tau\tau'} \equiv W(\tau\tau' \to \tau'\tau) \tag{7.4}$$

and, since we are assuming that the auxiliary matrices are scalars, we will write them as $\tilde{X}_\tau = x_\tau$.

We first restrict ourselves to the case $w_{12} > 0, w_{20} > 0$ and $w_{12} > 0$, i.e., there are exchanges in at least one direction between particles and vacancies and between the two species of particles. Then, the relations (7.2) become

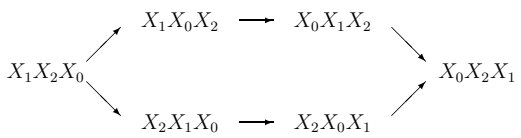
$$\begin{aligned} w_{10}X_1X_0 - w_{01}X_0X_1 &= x_0X_1 - x_1X_0 \\ w_{20}X_2X_0 - w_{02}X_0X_2 &= x_0X_2 - x_2X_0 \\ w_{12}X_1X_2 - w_{21}X_2X_1 &= x_2X_1 - x_1X_2. \end{aligned} \tag{7.5}$$

It remains to check whether the relations (7.5) are consistent. The approach of Isaev, Pyatov and Rittenberg [141] (see also that of Karimipour [142]) is to generalize the reduction process leading to (2.45). That is, one seeks to use (7.5) to reduce any product of matrices to a sum of irreducible strings. First we fix the order of irreducible strings as $X_0^l X_2^m X_1^n$. Then we require that reducing an arbitrary product U of the matrices X_i to a sum of irreducible strings

$$U = \sum_{l,m,n} a_{l,m,n} X_0^l X_2^m X_1^n \tag{7.6}$$

is independent of the order in which the reduction is carried out using the elementary rules (7.5). Reduction rules which satisfy this requirement are referred to by Isaev, Pyatov and Rittenberg as *PBW-type algebras* [143].

For the two-species case the requirement is that the two possible reductions of the product $X_1X_2X_0$, illustrated schematically below, give equivalent results.



Calculating explicitly using (7.5) yields [141]

$$\begin{aligned} &x_1w_{02}(w_{12} - w_{21} - w_{10} + w_{01})X_0X_2 + x_2w_{01}(w_{12} - w_{21} + w_{20} - w_{02})X_0X_1 \\ &\quad + x_0w_{21}(w_{20} - w_{02} - w_{10} + w_{01})X_2X_1 + x_1x_2(w_{12} - w_{21} + w_{20} - w_{10})X_0 \\ &\quad + x_1x_0(w_{21} - w_{02})X_2 + x_2x_0(w_{10} - w_{12} - w_{20} + w_{02})X_1 = 0, \end{aligned} \tag{7.7}$$

which results in the following six conditions on the hopping rates w_{ij} from the requirement that each of the six terms in the above equation vanish:

$$x_1w_{02}(w_{12} - w_{21} - w_{10} + w_{01}) = 0, \tag{7.8}$$

$$x_2w_{01}(w_{12} - w_{21} + w_{20} - w_{02}) = 0, \tag{7.9}$$

$$x_0w_{21}(w_{20} - w_{02} - w_{10} + w_{01}) = 0, \tag{7.10}$$

$$x_1x_2(w_{12} - w_{21} + w_{20} - w_{10}) = 0, \tag{7.11}$$

$$x_1x_0(w_{21} - w_{02}) = 0, \tag{7.12}$$

$$x_2x_0(w_{10} - w_{12} - w_{20} + w_{02}) = 0. \tag{7.13}$$

The solutions of these equations were classified in [141]. Here, we summarize the various nontrivial solutions which are physically relevant and the corresponding models which have been previously studied in the literature. The classification hinges on how many of x_i are zero.

7.1.1. Class A (includes Karimipour's model of overtaking). The first class of solutions has none of the $x_i = 0$. There are then two possible solutions to (7.8)–(7.13).

- *Solution AI* $w_{ij} = w\ \forall i, j$. Physically, this corresponds to symmetric exclusion with two species which, although labelled distinctly, have identical dynamics. Matrices are easily found by setting $X_2 = (x_2/x_1)X_1$ which reduces (7.5) to the single species condition $w(X_1X_0 - X_0X_1) = x_0X_1 - x_1X_0$. On the ring this system has a simple steady state where all allowed configurations are equally likely.
- *Solution AII* $w_{ij} = v_i - v_j > 0$ for $i < j$ and $w_{ij} = 0$ for $i > j$. This corresponds to a model where the hop rates are totally asymmetric and of the form

$$1\,0 \xrightarrow{v_1} 0\,1 \qquad 2\,0 \xrightarrow{v_2} 0\,2 \qquad 1\,2 \xrightarrow{v_1 - v_2} 2\,1. \tag{7.14}$$

  That is, since $v_1 > v_2$, a particle of species 1 is faster than a particle of species 2 and overtakes with rate $v_1 - v_2$. The corresponding matrix algebra is

$$\begin{aligned} v_1X_1X_0 &= x_0X_1 - x_1X_0 \\ v_2X_2X_0 &= x_0X_2 - x_2X_0 \\ (v_1 - v_2)X_1X_2 &= x_2X_1 - x_1X_2. \end{aligned} \tag{7.15}$$

  This solution of (7.8)–(7.13) was first noted and studied by Karimipour [142] and followed up in [144]. As it happens, this model can be generalized to more than two species, and as such will be discussed in its full generality in section 8.2. Representations are given in appendix D, equation (D.19).

7.1.2. Class B (includes asymmetric second-class particle dynamics). This second class of solutions has one of $x_i = 0$. The first two families of solutions within this class are obtained by taking $x_2 = 0$, and then choosing the hop rates to satisfy (7.8), (7.10) and (6.12) in two different ways.

- *Solution BI.* $w_{12} = w_{20} = p_2, w_{21} = w_{02} = q_2$, $w_{10} = p_1$ and $w_{01} = q_1$ with $p_2 - q_2 = p_1 - q_1$. This corresponds to a model with the hop rates

$$1\,0 \underset{q_1}{\overset{p_1}{\rightleftharpoons}} 0\,1 \qquad 2\,0 \underset{q_2}{\overset{p_2}{\rightleftharpoons}} 0\,2 \qquad 1\,2 \underset{q_2}{\overset{p_2}{\rightleftharpoons}} 2\,1 \tag{7.16}$$

  or, in the charge representation,

$$+\,- \underset{q_1}{\overset{p_1}{\rightleftharpoons}} -\,+ \qquad 0\,- \underset{q_2}{\overset{p_2}{\rightleftharpoons}} -0 \qquad +\,0 \underset{q_2}{\overset{p_2}{\rightleftharpoons}} 0\,+\,. \tag{7.17}$$

  This model is thus an asymmetric generalization of the second-class particle of section 4 and was first used by [83] for the special case $q_2 = q_1$, $p_2 = p_1$. The corresponding matrix algebra is

$$\begin{aligned} p_1X_1X_0 - q_1X_0X_1 &= x_0X_1 - x_1X_0 \\ p_2X_2X_0 - q_2X_0X_2 &= x_0X_2 \\ p_2X_1X_2 - q_2X_2X_1 &= -x_1X_2. \end{aligned} \tag{7.18}$$

  A representation of this algebra is given in appendix D.
- *Solution BII.* $w_{02} = w_{21} = 0, w_{12} = \alpha, w_{10} = p, w_{01} = q$ and $w_{12} = \beta$. The corresponding model has hop rates

$$1\,0 \underset{q}{\overset{p}{\rightleftharpoons}} 0\,1 \qquad 2\,0 \xrightarrow{\alpha} 0\,2 \qquad 1\,2 \xrightarrow{\beta} 2\,1 \tag{7.19}$$

or, in the charge representation,

$$+-\underset{q}{\overset{p}{\rightleftharpoons}}-+ \qquad 0-\xrightarrow{\alpha}-0 \qquad +0\xrightarrow{\beta}0+. \tag{7.20}$$

This model is another generalization of the second-class problem first studied with $p = 1, q = 0$ for a single species 2 particle [95] and later generalized in [112]. The matrix algebra is

$$\begin{aligned} pX_1X_0 - qX_0X_1 &= x_0X_1 - x_1X_0 \\ \alpha X_2X_0 &= x_0X_2 \\ \beta X_1X_2 &= -x_1X_2. \end{aligned} \tag{7.21}$$

This algebra can be mapped onto that used to solve the PASEP with open boundaries, (5.1)–(5.3), if one takes

$$X_2 = |V\rangle\langle W| \tag{7.22}$$

and $x_0 = 1, x_1 = -1$. Thus one can make use of the results from section 5, along with techniques for models on the ring described in section 4 to analyse various cases of this model [112–117].

Choosing x_1 or x_3 to be zero results in the same solution as BII under relabelling particles and one additional solution when $x_1 = 0$. This is

- *Solution BIII.* $w_{01} = w_{21} = 0$, $w_{10} = \alpha$, $w_{21} = \beta$, $w_{20} = p$, $w_{02} = q$ and $p - q = \alpha - \beta$. The corresponding model has hop rates

$$1\,0\xrightarrow{\alpha}0\,1 \qquad 2\,0\underset{q}{\overset{p}{\rightleftharpoons}}0\,2 \qquad 1\,2\xrightarrow{\beta}2\,1 \tag{7.23}$$

and the matrix algebra is

$$\begin{aligned} \alpha X_1X_0 &= x_0X_1 \\ pX_2X_0 - qX_0X_2 &= x_0X_2 - x_2X_0 \\ \beta X_1X_2 &= x_2X_1. \end{aligned} \tag{7.24}$$

However on the ring this results in a steady state in which all allowed configurations are equally likely: as can be seen from (7.24) in any periodic string of matrices, all X_2 and X_0 can be eliminated through the first and third relations.

7.1.3. Class C (includes non-overtaking two-species dynamics). The third class of solutions is obtained by taking two of the scalar quantities x_i equal to zero. Equations (7.11)–(7.13) and two of (7.8)–(7.10) are then automatically satisfied. The equation that remains has a structure that is independent of which of x_i are taken to be nonzero, so we take $x_1 = x_2 = 0$ which leaves us to satisfy (7.10). This can be done in two ways.

- *Solution CI.* $w_{10} = p_1$, $w_{01} = q_1$, $w_{20} = p_2$, $w_{02} = q_2$ and $p_1 - q_1 = p_2 - q_2$. This corresponds to a model in which all six exchanges may take place

$$1\,0\underset{q_1}{\overset{p_1}{\rightleftharpoons}}0\,1 \qquad 2\,0\underset{q_2}{\overset{p_2}{\rightleftharpoons}}0\,2 \qquad 1\,2\underset{w_{21}}{\overset{w_{12}}{\rightleftharpoons}}2\,1 \tag{7.25}$$

and that has the matrix algebra

$$\begin{aligned} p_1X_1X_0 - q_1X_0X_1 &= x_0X_1 \\ p_2X_2X_0 - q_2X_0X_2 &= x_0X_2 \\ w_{12}X_1X_2 - w_{21}X_2X_1 &= 0. \end{aligned} \tag{7.26}$$

Unfortunately this algebra is not useful in describing a physical system with periodic boundary conditions. To see this one can take for example a representation where $X_0 = \mathbb{1}$ and $x_0 = p_1 - q_1 = p_2 - q_2$ which satisfies the first two relations of (7.26). However the third relation of (7.26) has the form of a vanishing deformed commutator

$$X_1 X_2 - r X_2 X_1 = 0. \tag{7.27}$$

Thus if we commute say a matrix X_1 all the way around the ring we will end up with the same matrix product multiplied by a factor r^{N_2} where N_2 is the number of species 2 particles on the ring and $r = w_{21}/w_{12}$. Thus, only for $w_{21} = w_{12}$ ($r = 1$) can we use this algebra on a periodic system; alternatively, it can be used for general r on a closed segment as we discuss in section 7.2.2.

- *Solution CII.* Equation (7.10) can also be satisfied by taking $w_{12} = w_{21} = 0$ and unrestricted choices for the remaining rates $w_{10} = p_1, w_{01} = q_1, w_{20} = p_2, w_{02} = q_2$. This has similar dynamics to the previous model, but with no exchange of species 1 and 2:

$$1\,0 \underset{q_1}{\overset{p_1}{\rightleftharpoons}} 0\,1 \qquad 2\,0 \underset{q_2}{\overset{p_2}{\rightleftharpoons}} 0\,2. \tag{7.28}$$

The matrix algebra is similar to (7.26) but with the third relation absent. In contrast to the previous case, this algebra can be used to describe the above dynamics on a ring [145]. This model and its generalization to multiple species will be discussed further in section 8.

7.1.4. Class D (includes the cyclically symmetric ABC model). In this case, we take all $x_i = 0$ which leaves all six hop rates free:

$$1\,0 \underset{w_{01}}{\overset{w_{10}}{\rightleftharpoons}} 0\,1 \qquad 2\,0 \underset{w_{02}}{\overset{w_{20}}{\rightleftharpoons}} 0\,2 \qquad 1\,2 \underset{w_{21}}{\overset{w_{12}}{\rightleftharpoons}} 2\,1. \tag{7.29}$$

The algebra is a set of deformed commutators

$$w_{10} X_1 X_0 - w_{01} X_0 X_1 = 0 \tag{7.30}$$

$$w_{20} X_2 X_0 - w_{02} X_0 X_2 = 0 \tag{7.31}$$

$$w_{12} X_1 X_2 - w_{21} X_2 X_1 = 0. \tag{7.32}$$

Explicit forms for X_0, X_1 and X_2 are discussed in section 7.2.2. On a periodic system, as with case CI discussed above, we require that a matrix product is left invariant after commuting one of the matrices once around the ring. The condition for this is that

$$\left[\frac{w_{01}}{w_{10}}\right]^{n_0}\left[\frac{w_{21}}{w_{12}}\right]^{n_2} = \left[\frac{w_{02}}{w_{20}}\right]^{n_0}\left[\frac{w_{12}}{w_{21}}\right]^{n_1} = \left[\frac{w_{10}}{w_{01}}\right]^{n_1}\left[\frac{w_{20}}{w_{02}}\right]^{n_2} = 1 \tag{7.33}$$

where n_i is the number of particles of species i. These conditions are satisfied for example by the ABC model [146, 147] in the special case where all particle numbers are equal $n_0 = n_1 = n_2$. The fact that the right hand sides of (7.30)–(7.32) are all zero implies that to use these algebraic relations detailed balance must hold in the steady state. Therefore, the condition (7.33) is actually the condition for detailed balance to hold. The corresponding energy turns out to be an interesting long-range function [147].

In the ABC model, the particles and vacancies are relabelled as

$$1 \mapsto A \qquad 2 \mapsto B \qquad 0 \mapsto C \tag{7.34}$$

and have dynamics

$$AB \underset{1}{\overset{q}{\rightleftharpoons}} BA \qquad BC \underset{1}{\overset{q}{\rightleftharpoons}} CB \qquad CA \underset{1}{\overset{q}{\rightleftharpoons}} AC \tag{7.35}$$

where we take $q < 1$. This corresponds to the choice of rates $w_{12} = w_{20} = w_{01} = q$ and $w_{21} = w_{02} = w_{10} = 1$. Note that the dynamics are invariant under cyclic permutation of the three particle types and that A particles tend to move to the left of B particles, B particles tend to move to the left of C particles and C particles tend to move to the left of A particles. In the steady state of this model this results in a strong phase separation into three domains of A, B and C in the order ABC. In the limit $N \to \infty$ with $q < 1$ fixed, the domains are pure in that far away from the boundary of the domain the probability of finding a particle of a different species (to that of the domain) tends to zero. That is, particles from a neighbouring domain may penetrate only a finite distance. In the case where we have equal numbers of each species condition (7.33) is satisfied and one can use the matrix product to calculate the steady state exactly [147]. This model also exhibits an interesting phase transition in the weakly asymmetric limit where q varies with system size N as $q = \exp(-\beta/N)$, thus approaching unity in the thermodynamic limit. Then, according to the value of β, the steady state can either order into three domains which are rich in A, B and C (but are not pure domains) or a disordered phase where the particles are typically in a disordered configuration [148].

7.2. Geometries other than the ring

In the previous subsection, we summarized a classification of all the possible sets of dynamics for which a matrix-product state with scalar auxiliaries (i.e., one that has a set of reduction relations) exists with a unique decomposition into irreducible strings of matrices. We also assessed whether these matrix-product states could be used for models on the ring. We now extend this enquiry to other one-dimensional geometries that can be constructed.

7.2.1. Open-boundary conditions. We first consider open-boundary conditions, i.e., those where particles can enter and leave at the boundaries. The most general dynamics is to have rates α_{ij} at which a particle of species i (or a vacancy if $i = 0$) is transformed into a particle of species j at site 1 and rates β_{ij} at which a particle of species i is transformed into a particle of species j (or a vacancy if $j = 0$) at site N. If we are to have statistical weights of the form (2.35), i.e.,

$$f(\tau_1, \ldots, \tau_N) = \langle W | X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N} | V \rangle \tag{7.36}$$

we then obtain additional conditions involving the matrices and vectors $\langle W|$ and $|V\rangle$. These may be written as

$$\sum_i \langle W | \alpha_{ij} X_i = -\langle W | x_j \tag{7.37}$$

$$\sum_i \beta_{ij} X_i | W \rangle = x_j | W \rangle \tag{7.38}$$

where $\alpha_{ii} = -\sum_{j \neq i} \alpha_{ij}$ and $\beta_{ii} = -\sum_{j \neq i} \beta_{ij}$. Since $\sum_j \alpha_{ij} = \sum_j \beta_{ij} = 0$ we require

$$\sum_i x_i = 0. \tag{7.39}$$

In the open-boundary case, typically, in addition to the conditions for the solution classes of section 7.1, one has further constraints on the boundary rates to allow conditions (7.37), (7.38) to be satisfied. We do not attempt to catalogue all solutions here, rather we point out which of the solution classes of section 7.1 may, in principle, be used in an open system and refer some examples. Some general considerations of the open-boundary conditions for

which one has a matrix-product state using the quadratic algebra (7.5) are discussed in the two-species cases in [149].

- *Class A.* Solution AI, which corresponds to symmetric exclusion of two differently labelled species with identical bulk dynamics, can be used in the open-boundary case where the two species are injected and extracted with different rates, under certain conditions. Solution AII can be combined with certain open-boundary conditions that were found by Karimipour [144]. We shall specify these boundary interactions and discuss this model more thoroughly in section 8.
- *Class B.* Both solutions BI and BII can be used in models with open-boundary conditions. One case that has been studied is the dynamics implied by solution BII with $q = 0$ [150, 151, 20, 152]. This is sometimes referred to as the bridge model because one has two particle species with opposite velocities that slow down to exchange places when they meet, rather like cars on a narrow bridge. Unfortunately, the full phase diagram of the bridge model—and in particular an interesting broken symmetry region—cannot be described by a quadratic algebra (i.e., one where the auxiliaries are scalars and reduction relations are implied on the steady-state weights). The BII $q \neq 0$ case has also been studied [153, 154].
- *Class C.* Recall that class C comprises solutions where only one of x_i are nonzero. Hence, it is not possible to satisfy that sum rule (7.39), and hence the dynamics implied by these solutions cannot be solved in an open system using a simple matrix-product algebra.
- *Class D.* Algebras within class D can only be used in models with open boundaries under the special conditions of detailed balance, which generally does not hold. On the other hand, as we discuss below, detailed balance *does* apply in conservative models with closed boundaries and thus such algebras are relevant there—see below.

7.2.2. Closed segment. A closed segment is a one-dimensional lattice of size N with *closed* boundary conditions. That is, particles cannot enter or leave at the left boundary or right boundaries (sites 1 and N). For the moment, let us consider conserving models where particles are neither created nor destroyed in the bulk: the only moves that are allowed are those where particles or vacancies on neighbouring sites exchange places. It is straightforward to show using a Kolmogorov criterion (6.2) that detailed balance is satisfied in the steady state of any such model. The reason for this is that in order to create a loop in configuration space, every exchange of a pair must be accompanied by an exchange in the opposite direction (as long as the dynamics is not totally asymmetric). Hence the forward loop contains the same set of exchanges as the reverse, and (6.2) is satisfied. Some quadratic algebras for closed segments and segments closed at one end and open at the other were discussed in [149].

The simplest example of this type is the partially asymmetric exclusion process (PASEP) on the closed segment, i.e., just one species of particle [155]. The appropriate matrix algebra in this case is

$$DE - qED = 0. \tag{7.40}$$

Taking $q < 1$ the configuration with the highest weight will be the configuration with all particles stacked up to the right end of the lattice as shown in figure 29. The weight of any configuration can then be obtained by moving particles to the left from the maximal weight configuration and multiplying the weight by a factor q at each move as implied by (7.40). Then the statistical weight of any configuration $\tau_1, \tau_2, \ldots, \tau_N$ will be given by

$$f(\tau_1, \tau_2, \ldots, \tau_N) = q^{-\sum_k k\tau_k}. \tag{7.41}$$

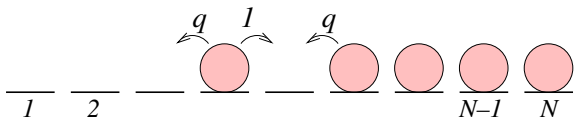


Figure 29. Typical configuration in the steady state of the partially asymmetric exclusion process on a closed segment of N sites with $q < 1$.

and the maximal weight is

$$q^{-\sum_{k=N-M+1}^{N} k} = q^{-\binom{N+1}{2}+\binom{N-M+1}{2}}.$$

The steady-state weight (7.41) is in fact a Boltzmann weight with an energy function $\mathcal{H} \propto |\ln q| \sum_k k\tau_k$ and the dynamics respects detailed balance with respect to this weight.

Let us now turn to class D models, where the matrix algebra takes the form of three deformed commutators (7.32). An explicit form for the operators X_1, X_2, X_3 involves tensor products of matrices which obey deformed commutation relations

$$X_1 = D \otimes E(\omega_{21}/\omega_{12}) \otimes \mathbb{1} \tag{7.42}$$

$$X_2 = \mathbb{1} \otimes D \otimes E(\omega_{02}/\omega_{20}) \tag{7.43}$$

$$X_0 = E(\omega_{10}/\omega_{01}) \otimes \mathbb{1} \otimes D \tag{7.44}$$

where

$$DE(r) - rE(r)D = 0. \tag{7.45}$$

So, for example,

$$\begin{aligned} &w_{10}X_1X_0 - w_{01}X_0X_1 \\ &\quad = [w_{10}DE(\omega_{10}/\omega_{01}) - w_{01}E(\omega_{10}/\omega_{01})D] \otimes E(\omega_{21}/\omega_{12}) \otimes D = 0. \end{aligned} \tag{7.46}$$

To obtain statistical weights here, we need to prescribe a suitable contraction operation, i.e., appropriate boundary vectors. Using the representation of the deformed commutator algebra in appendix D (D.29) we see that if there are n_1 species 1 particles, n_2 species 2 particles and n_0 vacancies then

$$\langle W| = \langle 0| \otimes \langle 0| \otimes \langle 0| \qquad |V\rangle = |n_1\rangle \otimes |n_2\rangle \otimes |n_0\rangle \tag{7.47}$$

will give a nonzero contraction for those configurations with the correct number of each species. In particular for the reference configuration the weight is

$$\langle W|X_0^{n_0}X_2^{n_2}X_1^{n_1}|V\rangle = \left(\frac{\omega_{21}}{\omega_{12}}\right)^{n_2}. \tag{7.48}$$

7.2.3. A single defect on an infinite system. In order to investigate the structure of shocks, Derrida, Lebowitz and Speer studied a single second-class particle on an infinite system [156]. The idea was to describe the stationary measure as seen from the second-class particle by a matrix product. The model considered was the partially asymmetric generalization with dynamics

$$1\,0 \underset{q}{\overset{p}{\rightleftharpoons}} 0\,1 \qquad 2\,0 \underset{q}{\overset{p}{\rightleftharpoons}} 0\,2 \qquad 1\,2 \underset{q}{\overset{p}{\rightleftharpoons}} 2\,1. \tag{7.49}$$

Taking the second-class particle to be at the origin and considering a window of m sites to the left and n sites to the right of the second-class, the stationary probabilities (note, not weights) are written as

$$P(\tau_{-m},\ldots,\tau_{-1},\tau_1,\ldots,\tau_n) = \langle W| \left[\prod_{i=-m}^{-1} X_{\tau_i}\right] A \left[\prod_{j=1}^{n} X_{\tau_j}\right] |V\rangle \tag{7.50}$$

where A is once more the matrix corresponding to a second-class particle. In order for this form to hold for a window of arbitrary size, i.e. all values of $n, m = 0,\ldots,\infty$ one requires that

$$\langle W|C = \langle W| \qquad C|V\rangle = |V\rangle \tag{7.51}$$

where $C = D + E$. To ensure that the probabilities are correctly normalized we further require

$$\langle W|A|V\rangle = 1. \tag{7.52}$$

The bulk dynamics falls within the solution class BI, and so the matrix algebra (7.18) should be used with $p_1 = p_2 = p$ and $q_1 = q_2 = p$. This particular application has the unusual property that the choice of x_0 and x_1 has physical consequences. This is because x_0 and x_1 cannot be scaled out of equations (7.51). The correct choice is made by insisting that the desired asymptotic densities ρ_+ far to the right of the second-class particle ($n \to \infty$) and ρ_- far to the left ($m \to \infty$) are obtained. It turns out that one should take

$$x_0 = (p-q)(1-\rho_-)(1-\rho_+) \tag{7.53}$$

$$x_1 = -(p-q)\rho_-\rho_+. \tag{7.54}$$

Choosing $\rho_+ > \rho_-$ corresponds to a shock profile as soon from the second-class particle with the second-class particle tracking the position of the shock. The structure of the shock was analysed, in particular the decay to the asymptotic values is exponential. An interesting result is that the characteristic length of the decay becomes independent of the asymmetry when

$$\frac{p}{q} > \left(\frac{\rho_+(1-\rho_-)}{\rho_-(1-\rho_+)}\right)^{1/2}. \tag{7.55}$$

To obtain a representation of the required matrices and vectors, one can use (D.21)–(D.23) given in appendix D for solution class BI. Then, using (7.51), one can construct $\langle W|, |V\rangle$ to be eigenvectors $D + E$ with eigenvalue 1. The expressions for the boundary vectors are, however, quite complicated in this representation. An alternative approach, used in [156], exploits an infinite-dimensional (rather than semi-infinite) representation. Finite-dimensional representations along special curves were studied in [157].

7.3. Non-conserving two-species models with quadratic algebra

In the foregoing, we have assumed that the bulk dynamics conserve particle numbers. In this section we show how the matrix algebras which were encountered can be adapted to models that have non-conservative dynamics. The basic idea is to augment the dynamics with additional processes that move between different sectors (i.e., configurations with a particular number of particles). These moves will be constructed in such a way that detailed balance between sectors is satisfied, even though detailed balance does not hold within a sector. As such, one can realize any prescribed distribution of sectors: we choose weights that correspond to the grand-canonical ensemble that was introduced in section 4 as a means to study systems with fixed particle numbers. We first discuss this method in detail and then give two concrete examples.

7.3.1. Non-conserving dynamics which generate a grand-canonical ensemble. As we have just described, the aim is to construct some dynamics such that the statistical weight of a configuration with M particles (one a designated species) on an N-site ring is

$$f(\tau_1, \tau_2, \ldots, \tau_N; M) = u^M \operatorname{tr}\left(X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N}\right) \tag{7.56}$$

where X_τ are the matrices that give the steady state for some conservative process and u is a fugacity that controls the mean particle number M. Let us denote transition rates for this conservative process as $W_0(\mathcal{C} \to \mathcal{C}')$, and the additional rates that serve to increase or decrease M by one as $W_\pm(\mathcal{C} \to \mathcal{C}')$. The statistical weights must then satisfy the master equation

$$\begin{aligned}\sum_{\mathcal{C}'} \left[f(\mathcal{C}'; M) W_0(\mathcal{C}' \to \mathcal{C}) - f(\mathcal{C}; M) W_0(\mathcal{C} \to \mathcal{C}')\right] \\ + \sum_{\mathcal{C}'} \left[f(\mathcal{C}'; M-1) W_+(\mathcal{C}' \to \mathcal{C}) - f(\mathcal{C}; M) W_-(\mathcal{C} \to \mathcal{C}')\right] \\ + \sum_{\mathcal{C}'} \left[f(\mathcal{C}'; M+1) W_-(\mathcal{C}' \to \mathcal{C}) - f(\mathcal{C}; M) W_+(\mathcal{C} \to \mathcal{C}')\right] = 0.\end{aligned} \tag{7.57}$$

Now, the first summation is what appears in the master equation for the conservative process, and since all the weights for fixed M are proportional to u^M, this first summation vanishes. The second and third summations can be made to vanish if we choose

$$\frac{W_-(\mathcal{C} \to \mathcal{C}')}{W_+(\mathcal{C}' \to \mathcal{C})} = \frac{f(\mathcal{C}'; M-1)}{f(\mathcal{C}; M)} = \frac{1}{u} \frac{\operatorname{tr}\left(X_{\tau_1'} X_{\tau_2'} \cdots X_{\tau_N'}\right)}{\operatorname{tr}\left(X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N}\right)} \tag{7.58}$$

where τ_i and τ_i' are the site occupation variables specifying to configurations $\mathcal{C}$ and $\mathcal{C}'$ that have M and $M-1$ particles, respectively. As we can see, this takes the form of a detailed balance relation (6.3). The two examples that follow serve to demonstrate.

7.3.2. Grand-canonical dynamics for the ASEP with a second-class particle. Recall that in section 4 we considered a version of the ASEP with a second-class particle, i.e., a model with dynamics

$$1\,0 \xrightarrow{1} 0\,1 \qquad 2\,0 \xrightarrow{\alpha} 0\,2 \qquad 1\,2 \xrightarrow{\beta} 2\,1. \tag{7.59}$$

This falls within the family of solutions denoted as BII above. In the analysis of this model we introduced a fugacity u as a trick to expedite the calculation of statistical weights, tuning its value to yield the desired mean particle density on the ring in the canonical ensemble.

This ensemble can also be generated using physical dynamics as we have just described. Let M count the number of positive charges (species 1 particles). We note that the algebra (7.21) implies

$$\frac{\operatorname{tr}(\cdots X_2 X_0 \cdots)}{\operatorname{tr}(\cdots X_1 X_2 \cdots)} = -\frac{\beta}{\alpha} \frac{x_0}{x_1}. \tag{7.60}$$

Hence from (7.58), we see that the grand-canonical ensemble is realized physically if we introduce two non-conserving processes

$$+\,0 \underset{\gamma}{\overset{\delta}{\rightleftharpoons}} 0\,- \tag{7.61}$$

where the rates satisfy

$$\frac{\delta}{\gamma} = -\frac{\beta}{\alpha u} \frac{x_0}{x_1}. \tag{7.62}$$

We remark that δ and γ as free parameters that enforce a particular fugacity u in the grand-canonical ensemble generated by these non-conservative dynamics. This particular

implementation of a model with non-conserving dynamics was used in [158] to generate a grand-canonical ensemble dynamically. A partially asymmetric variation of this model was studied in [159].

7.3.3. A model in which no particle numbers are conserved. A second non-conserving generalization of the ASEP on a ring with a second-class particle is best described using the charge representation. For simplicity, we take the dynamics of BII with $p = 1$ and $q = 0$ and take $\alpha = \beta = 1$. The case $q \neq 0$ was studied in [160]. Thus, the conserving dynamics are

$$+\,0 \xrightarrow{1} 0\,+ \qquad 0\,- \xrightarrow{1} -0 \qquad +\,- \xrightarrow{1} -\,+\,. \tag{7.63}$$

To these dynamics we seek to add the following non-conserving moves:

$$+\,0 \underset{1}{\overset{w}{\rightleftharpoons}} 0\,0 \qquad 0\,- \underset{1}{\overset{w}{\rightleftharpoons}} 0\,0. \tag{7.64}$$

in such a way that the BII matrix algebra (7.21) can be exploited. To achieve this, we require matrices such that (7.58) is satisfied for the non-conserving moves. That is,

$$w = \frac{1}{u}\frac{\mathrm{tr}(\cdots X_0 X_0 \cdots)}{\mathrm{tr}(\cdots X_+ X_0 \cdots)} = \frac{1}{u}\frac{\mathrm{tr}(\cdots X_0 X_0 \cdots)}{\mathrm{tr}(\cdots X_0 X_- \cdots)} \tag{7.65}$$

where u is a fugacity counting the total number of positive and negative charges. To make contact with more familiar problems, let us put $X_+ = D$, $X_- = E$ and $X_0 = A$. We see that the previous equation is satisfied if

$$u = \frac{1}{w} \qquad \text{and} \qquad A^2 = A, \tag{7.66}$$

i.e., we require A to be a projector. This is achieved using the matrices for the second-class particle problem of section 4, (4.4)–(4.7), i.e., $A = |V\rangle\langle W|$, where $D|V\rangle = |V\rangle$, $\langle W|E = \langle W|$ and $\langle W|V\rangle = 1$. The statistical weight of a configuration is then given by

$$f(\tau_1, \tau_2, \ldots, \tau_N) = w^M \,\mathrm{tr}(X_{\tau_1} X_{\tau_2} \cdots \tau_N) \tag{7.67}$$

where M is the number of vacancies in the configuration.

In [160] it was shown that an interesting phase transition arises as w is varied. For $w < 1$ there is a tendency to create positive and negative particles and this ultimately leads to a vanishing steady-state density of vacancies. On the other hand for $w > 1$ there is a tendency to eliminate positive and negative charges at the left and right boundaries respectively of domains of vacancies. Thus domains of vacancies tend to grow and in the steady state this results in a finite density of vacancies. The transition is remarkable in that it is a phase transition in a periodic one dimensional system with local dynamics without an absorbing state. Here, we shall use the generating function technique of section 3.2.3 to quickly obtain the asymptotics of the partition function and hence demonstrate the phase transition.

First, note from (7.64) that configurations with no vacancies are dynamically inaccessible. Therefore, all configurations have at least one vacancy and we write the partition function as

$$Z_N = \mathrm{tr}[wA(D + E + wA)^{N-1}] = w\langle W|(D + E + wA)^{N-1}|V\rangle \tag{7.68}$$

(An unimportant subtlety is that we have ignored the degeneracy factor in placing a given configuration of particles on the periodic lattice; this factor is bounded from above by N and hence does not contribute to the exponential part of Z_N which controls the macroscopic physics, as will be seen below.)

We introduce the generating function

$$F(z) = \sum_{N=1}^{\infty} z^N Z_N = wz\langle W|\frac{1}{1 - zG}|V\rangle \tag{7.69}$$

where we define $G = D + E + wA$. Now, we may write

$$1 - zG = (1 - \eta D - \nu A)(1 - \eta E - \nu A) \tag{7.70}$$

where

$$z = \eta(1 - \eta) \qquad zw = 2\nu(1 - \eta) - \nu^2. \tag{7.71}$$

Then using the generating function technique and (7.71) we obtain

$$F(z) = wz\langle W|\frac{1}{(1 - \eta E - \nu A)}\frac{1}{(1 - \eta D - \nu A)}|V\rangle \tag{7.72}$$

$$= \frac{wz}{(1 - \eta - \nu)^2} = \frac{wz}{(1 - \eta)^2 - wz}. \tag{7.73}$$

The singularities of $F(z)$ are a square-root singularity at $z = 1/4$ (coming from $\eta(z) = (1 - \sqrt{1 - 4z})/2$) and a pole at $z = w/(1 + w)^2$ which only exists if $w > 1$. Using the results of appendix B one quickly obtains the asymptotic behaviour of Z_N

$$\text{if} \quad w < 1 \qquad Z_N \sim \frac{1}{\pi^{1/2}}\frac{w}{(1 - w)^2}\frac{4^N}{N^{3/2}} \tag{7.74}$$

$$\text{if} \quad w > 1 \qquad Z_N \sim \frac{w - 1}{w + 1}\frac{(w + 1)^{2N}}{w^N}. \tag{7.75}$$

Due to the form (7.67) of the weights, θ, the density of vacancies is given by

$$\theta \equiv \lim_{N\to\infty}\frac{\overline{M}}{N} = \lim_{N\to\infty}\frac{w}{N}\frac{\partial \ln \mathcal{Z}_N}{\partial w}, \tag{7.76}$$

where $\overline{M}$ is the average number of vacancies in the system. Therefore we obtain

$$\theta = 0 \qquad \text{for} \quad w < 1 \tag{7.77}$$

$$\theta = \frac{w - 1}{w + 1} \qquad \text{for} \quad w > 1 \tag{7.78}$$

demonstrating the phase transition in the density of vacancies at $w = 1$.

8. Multispecies models with quadratic algebra

In the previous section we identified all possible two-species dynamics that have a steady state of matrix-product form with a quadratic algebra, i.e., obeying relations of the type given by (7.5). We deferred the discussion of two families of dynamics, as both can be generalized in a straightforward way to arbitrarily many species of particles. We now examine these cases in more detail. Further discussions of multispecies quadratic algebras are given in [141, 161–164].

8.1. ASEP with disordered hopping rates

We first consider a variant of the ASEP on the ring in which each particle has a different forward and reverse hop rate (and there is no overtaking) [145, 165]. This is a generalization of the two-species model, case CII of section 7. There are M particles and $N - M$ vacancies

on the ring of size N. We give each of the M particles an index μ, and introduce for each μ a pair of rates p_μ and q_μ so that the dynamics of particle μ are

$$\mu 0 \underset{q_\mu}{\overset{p_\mu}{\rightleftharpoons}} 0\mu. \tag{8.1}$$

The corresponding matrix algebra is a generalization of class CII of section 7 to M particle species. It takes the form

$$p_\mu D_\mu E - q_\mu E D_\mu = D_\mu \tag{8.2}$$

where we have adopted the notation $X_0 = E$ and $X_\mu = D_\mu$. One possible representation of these matrices is given in appendix D as equation (D.14). However, it is simplest to perform calculations not with a representation, but directly from the algebra (8.2).

On a periodic system let the steady-state weight of a configuration specified by $\{n_\mu\}$, where n_μ is the number of empty sites in front of particle μ, be $f_N(n_1, \ldots, n_M)$. The matrix-product form for f_N is

$$f_N(n_1, n_2, \ldots, n_M) = \mathrm{Tr}\left[D_1 E^{n_1} D_2 E^{n_2} \ldots D_M E^{n_M}\right]. \tag{8.3}$$

Using (8.2) for $\mu = M$ in (8.3) gives the relation

$$f_N(n_1, \ldots, n_M) - \frac{q_M}{p_M} f_N(n_1, \ldots, n_{M-1} + 1, n_M - 1) = \frac{1}{p_M} f_{N-1}(n_1, \ldots, n_{M-1}, n_M - 1).$$

The procedure is continued using, in sequence, (8.2) with $\mu = M-1, M-2, \ldots, 1$ to obtain

$$\left[1 - \prod_{i=0}^{M-1} \frac{q_{M-i}}{p_{M-i}}\right] f_N(n_1, \ldots, n_M) = \left[\sum_{i=0}^{M-1} \frac{1}{p_{M-i}} \prod_{j=M+1-i}^{M} \frac{q_j}{p_j}\right] f_{N-1}(n_1, \ldots, n_M - 1).$$

The effect of this manipulation has been to commute a hole initially in front of the particle M backwards one full turn around the ring. The result is that the weight of a configuration of size N is expressed as a multiple of the weight of a configuration of size $N-1$ with one hole fewer in front of particle M. Repeating the commutation procedure for a hole initially in front of particles $M-1, M-2, \ldots, 1$ implies that the weights are

$$f_N(\{n_\mu\}) = \prod_{\mu=1}^{M} g_\mu^{n_\mu} f_M(0, \ldots, 0) \qquad \text{where}$$

$$g_\mu = \left[\sum_{i=0}^{M-1} \frac{1}{p_{\mu-i}} \prod_{j=\mu+1-i}^{\mu} \frac{q_j}{p_j}\right] \left[1 - \prod_{k=1}^{M} \frac{q_k}{p_k}\right]^{-1} \tag{8.4}$$

and we can take $f_M(0, \ldots, 0) = 1$.

8.1.1. Bose–Einstein condensation. An interesting phenomenon that may occur in this system is that of condensation. In that case the steady state is dominated by configurations where one particle has an extensive value for n_μ [145, 166]. This is easiest to understand when $q_\mu = 0$, i.e., we consider only forward hops. Then (8.4) reduces to $g_\mu = 1/p_\mu$ and we may write

$$f_N(\{n_\mu\}) = \exp\left(-\sum_{\mu=1}^{M} n_\mu \epsilon_\mu\right) \qquad \text{where} \qquad \epsilon_\mu = \ln p_\mu. \tag{8.5}$$

The weight (8.5) clearly has the form of the weight of an ideal Bose gas with $kT = 1$. Here the 'bosons' are vacancies and the Bose states correspond to particles with the energy of the

state determined by the hop rate of the particle. The equivalent of the density of states for the Bose system is the distribution of particle hop rates which we denote as $\sigma(p)$. If there is a minimum hop rate $p_{\min}$ this will correspond to the ground state of the Bose system. Then, if for low hop rates the distribution of particle hop rates follows

$$\sigma(p) \sim (p - p_{\min})^{\gamma} \qquad \text{with} \qquad \gamma > 0, \tag{8.6}$$

Bose condensation will occur for high enough vacancy density or equivalently *low* enough particle density. When condensation occurs the slowest particle (the one with the minimum hop rate $p_{\min}$) has an extensive number of vacancies in front of it whereas the rest of the particles will have gaps in front of them comprising some finite number of vacancies. Thus a 'traffic jam' forms behind the slowest particle. In [14] it is shown that the ASEP on a periodic system may be mapped onto a zero-range process and the condensation transition is fully discussed in the context of the zero-range process.

Recently, further progress has been made in understanding the case where $q_\mu \neq 0$ by using extreme values statistics and renormalization arguments [167].

8.2. *Karimipour's model of overtaking dynamics*

This is the generalization of the two-species case AII, which had the two particle species hopping with different speeds, and the faster ones overtaking the slower. The many-species generalization allows an arbitrary number of particle species labelled $\mu = 1, \ldots, P$. A particle of species μ has a forward hop rate or 'velocity' v_μ, and adjacent particles μ and ν exchange places with rate $v_\mu - v_\nu$ if $v_\mu > v_\nu$, i.e.,

$$\begin{aligned} &\mu 0 \overset{v_\mu}{\to} 0\mu \\ &\mu\nu \overset{v_\mu - v_\nu}{\to} \nu\mu \qquad \text{if} \quad v_\mu > v_\nu. \end{aligned} \tag{8.7}$$

The model provides an elegant generalization of the open-boundary TASEP and its algebra. The generalization of (7.15) to the multispecies case is

$$\begin{aligned} &v_\mu X_\mu X_0 = x_0 X_\mu - x_\mu X_0 \qquad \mu > 0 \\ &(v_\mu - v_\nu) X_\mu X_\nu = x_\nu X_\mu - x_\mu X_\nu \qquad v_\mu > v_\nu. \end{aligned} \tag{8.8}$$

Writing $X_0 = E$, $X_\mu = D_\mu / P$ and making the convenient choice $x_0 = 1, x_\mu = -v_\mu / P$ for $\mu > 0$ yields

$$\begin{aligned} &D_\mu E = \frac{1}{v_\mu} D_\mu + E \\ &D_\mu D_\nu = \frac{v_\mu}{v_\mu - v_\nu} D_\nu - \frac{v_\nu}{v_\mu - v_\nu} D_\mu \qquad v_\mu > v_\nu. \end{aligned} \tag{8.9}$$

On a periodic lattice the relations (8.9) are the only ones which need be satisfied and this can be done by choosing a scalar representation $E = 1/\epsilon$ and $D_\mu = \frac{v_\mu}{v_\mu - \epsilon}$ where ϵ is chosen to be less than the lowest hop rate. Thus on a ring all allowed configurations are equally likely.

More interestingly, the relations (8.2) may be used for an open system with suitable boundary conditions. As discussed in section 7.2.1, we require $\sum_{\mu=0}^{P} x_\mu = 0$ which with the above choice of x_0 and x_μ implies

$$\frac{1}{P} \sum_{\mu=1}^{P} v_\mu = 1. \tag{8.10}$$

At the left-hand boundary a particle of species μ is injected with rate α_μ if the first site is empty. The choice $\alpha_\mu = \alpha v_\mu / P$ reduces (7.37) to

$$\langle W|E = \frac{1}{\alpha}\langle W|. \tag{8.11}$$

At the left-hand boundary a particle of species μ leaves the system with rate β_μ. The choice $\beta_\mu = v_\mu + \beta - 1$ ensures that explicit representations of the matrices can be constructed (see appendix D) and (7.38) becomes

$$D_\mu|V\rangle = \frac{v_\mu}{v_\mu - (1-\beta)}|V\rangle. \tag{8.12}$$

Relations (8.9), (8.12) and (8.11) provide an elegant generalization of the usual ASEP relations (2.39)–(2.41) which are recovered when we have just one species of particle with hop rate $v_1 = 1$. In general, an infinite-dimensional representation of (8.9) is needed to satisfy the boundary conditions (8.12); see, for example, (D.19) in appendix D.

The fully disordered system is realized when each particle that enters has a velocity drawn from some distribution $\sigma(v)$, with support $[v_{\min}, \infty]$, so that the lowest velocity is $v_{\min}$, and (8.10) is replaced by

$$\int_{v_{\min}}^{\infty} \mathrm{d}v\, \sigma(v) v = 1. \tag{8.13}$$

Thus, when site one is vacant a particle enters with rate α and its velocity v is assigned according to the distribution $\sigma(v)v$ giving an effective rate $\alpha(v) = \sigma(v)v\alpha$. Therefore we can think of a reservoir of particles with velocities distributed according to $\sigma(v)$ each attempting to enter with rate $v\alpha$. A particle reaching site N exits with rate $\beta(v) = v + \beta - 1$. Then we can replace X_μ in the discrete case by $X(v)$ in (8.8). Letting $X_0 = E$ and $X(v) = \sigma(v)D(v)$, $x_0 = 1$ and $x(v) = -\sigma(v)v$ recovers the same algebra (8.9), (8.12) and (8.11) with D_v replaced by $D(v)$, e.g.,

$$D(v)D(v') = \frac{v}{v - v'}D(v') - \frac{v'}{v - v'}D(v) \qquad \text{for} \quad v > v'. \tag{8.14}$$

Karimipour [168, 144] studied, amongst other things, the phase diagram of the open-boundary model and found that the structure depends on the distribution $\sigma(v)$ through the parameter

$$l[\sigma] = \frac{1}{v_{\min}^2} - \left\langle \frac{v}{(v - v_{\min})^2} \right\rangle \tag{8.15}$$

which characterizes the form of $\sigma(v)$ near the lowest hop rate $v_{\min}$. If $l[\sigma] < 0$ the three phases (low-density, high-density and maximal-current) of the open-boundary ASEP remain although the phase boundaries depend on parameters such as $l[\sigma]$ and $v_{\min}$. However if $l[\sigma] \geqslant 0$ the high-density phase is suppressed and only the low-density and maximal-current phases exist.

9. More complicated matrix-product states

Recall that in section 3.1.2 we set out a general cancellation scheme for matrices that guarantees a stationary solution of a master equation for a process on the ring. This involved a set of ordinary matrices X_τ and their auxiliaries $\tilde{X}_\tau$. By restricting to the case of scalar auxiliaries, we showed, in the case of two particle species in section 7, that it was possible to determine all possible sets of dynamics that give rise to a matrix-product steady state. In this section, we shall

consider the more complicated case where the auxiliaries are matrices or more complicated operators.

Here it is not possible—as far as we are aware—to perform an exhaustive search for solutions. However, it has been shown that if one has a one-dimensional lattice with n particle species, open-boundary conditions and arbitrary nearest-neighbour dynamics in the bulk, there does exist a matrix-product solution, involving auxiliaries that in general will not be scalars, obeying the algebraic relations set out in section 3.1.2. In the next subsection we review this existence proof which is due to Krebs and Sandow [97]. Unfortunately, this proof does not lead to any convenient reduction relations, like (2.39)–(2.41) for the ASEP, that might be in operation, Also, the proof is not constructive in that it requires the steady state to already be known in order to construct explicit matrices. Rather, the proof demonstrates that there are no internal inconsistencies in the cancellation mechanism of section 3.1.2.

In the last two subsections we discuss two models of which we are aware that give examples of a matrix-product state with matrix auxiliaries.

9.1. Existence of a matrix-product solution for models with open boundaries

The existence proof of a matrix-product state [97] (see also [169] for the discrete-time case) applies to models with open-boundary conditions and arbitrary nearest-neighbour interactions. Let us restate the master equation for a general process in the form introduced in section 3.1.2. It reads

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\tau_1, \ldots, \tau_N) \equiv \hat{H} f(\tau_1, \ldots, \tau_N) = \left(\hat{h}_L + \sum_{i=1}^{N-1} \hat{h}_{i,i+1} + \hat{h}_R \right) f(\tau_1, \ldots, \tau_N), \tag{9.1}$$

where the operator $\hat{h}_{i,i+1}$ applied to the function of N state variables τ_i generates the gain and loss terms arising from interactions between neighbouring pairs of sites in the bulk, and $\hat{h}_L$ and $\hat{h}_R$ do the same at the boundaries. Here we shall consider models that have n distinct particle species in addition to vacancies, i.e., the occupation variables take the values $\tau_i = 0, 1, \ldots, n$. The matrix-product expressions

$$f(\tau_1, \tau_2, \cdots, \tau_N) = \langle W | X_{\tau_1} X_{\tau_2} \cdots X_{\tau_N} | V \rangle \tag{9.2}$$

provide the steady-state solution to (9.1) for some set of matrices X_τ and vectors $\langle W|$ and $|V\rangle$ if one can find a second set of auxiliary matrices $\tilde{X}_\tau$ such that the relations

$$\hat{h}_{i,i+1} \langle W | \cdots X_{\tau_i} X_{\tau_{i+1}} \cdots | V \rangle = \langle W | \cdots [\tilde{X}_{\tau_i} X_{\tau_{i+1}} - X_{\tau_i} \tilde{X}_{\tau_{i+1}}] \cdots | V \rangle \tag{9.3}$$

$$\hat{h}_L \langle W | X_{\tau_1} \cdots | V \rangle = -\langle W | \tilde{X}_{\tau_1} \cdots | V \rangle \tag{9.4}$$

$$\hat{h}_R \langle W | \cdots X_{\tau_N} | V \rangle = \langle W | \cdots \tilde{X}_{\tau_N} | V \rangle \tag{9.5}$$

hold. Then, one obtains a zero right-hand side of the master equation (9.1) via a pairwise cancellation of terms coming from (9.2)–(9.5). Thus, relations (3.21)–(3.23) for the ASEP generalize to

$$\hat{h}_{i,i+1} X_{\tau_i} X_{\tau_{i+1}} = \tilde{X}_{\tau_i} X_{\tau_{i+1}} - X_{\tau_i} \tilde{X}_{\tau_{i+1}} \tag{9.6}$$

$$\hat{h}_L \langle W | X_{\tau_1} = -\langle W | \tilde{X}_{\tau_1} \tag{9.7}$$

$$\hat{h}_R X_{\tau_N} | V \rangle = \tilde{X}_{\tau_N} | V \rangle. \tag{9.8}$$

Relations (9.6)–(9.8) give the general form for a cancellation mechanism involving matrix (or possibly tensor) auxiliaries. Such a cancellation scheme has been proposed in various contexts

[28, 170, 171] including a generalization to longer (but finite) range interactions [172]. Some consequences of this scheme have been explored in [173].

The existence of the matrices X_τ and $\tilde{X}_\tau$ and vectors $\langle W|$ and $|V\rangle$ appearing in (9.6)–(9.8) is proved by constructing them within an explicit representation. This representation has basis vectors that correspond to configurations of the rightmost $N-k+1$ sites of the N-site lattice. These we denote as $|\tau_k\tau_{k+1}\cdots\tau_N\rangle$. The vector $|V\rangle$ is then ascribed the role of a vacuum state, and the matrices X_τ that of creation operators in such a way that

$$X_\tau|V\rangle = |\tau\rangle \qquad \text{and} \qquad X_\tau|\tau_{k+1}\tau_{k+2}\cdots\tau_N\rangle = |\tau\tau_{k+1}\tau_{k+2}\cdots\tau_N\rangle. \tag{9.9}$$

One then defines the vector $\langle W|$ via the scalar products

$$\langle W|\tau_1\tau_2\ldots\tau_N\rangle = f(\tau_1,\tau_2,\cdots,\tau_N) \tag{9.10}$$

so that (9.2) gives the desired statistical weights.

The auxiliary matrices $\tilde{X}_\tau$ are defined as

$$\tilde{X}_\tau = \left(\sum_{j=1}^{N-1}\hat{h}_{j,j+1} + \hat{h}_R\right) X_\tau \tag{9.11}$$

where the operators $\hat{h}_{j,j+1}$ and $\hat{h}_R$ are extended to the full space of all sub-configurations $|\tau_k\tau_k+1\cdots\tau_N\rangle$ as follows. First, the bulk operator is defined in terms of the microscopic transition rates $W(\mathcal{C}\to\mathcal{C}')$ as

$$\langle\tau_k\tau_{k+1}\cdots\tau_i'\tau_{i+1}'\cdots\tau_N|\hat{h}_{i,i+1}|\tau_k\tau_{k+1}\cdots\tau_i\tau_{i+1}\cdots\tau_N\rangle = W(\tau_i\tau_{i+1}\to\tau_i'\tau_{i+1}'), \tag{9.12}$$

when $i\geqslant k$ and $(\tau_i',\tau_{i+1}')\neq(\tau_i,\tau_{i+1})$, and

$$\langle\tau_k\tau_{k+1}\cdots\tau_i\tau_{i+1}\cdots\tau_N|\hat{h}_{i,i+1}|\tau_k\tau_{k+1}\cdots\tau_i\tau_{i+1}\cdots\tau_N\rangle = -\sum_{\tau_i',\tau_{i+1}'} W(\tau_i\tau_{i+1}\to\tau_i'\tau_{i+1}') \tag{9.13}$$

under the same condition that $i\geqslant k$. All other elements of these operators are zero. Similarly, at the right boundary we have

$$\langle\tau_k\cdots\tau_N'|\hat{h}_R|\tau_k\cdots\tau_N\rangle = W(\tau_N\to\tau_N'), \tag{9.14}$$

when $\tau_N\neq\tau_N'$, and

$$\langle\tau_k\cdots\tau_N|\hat{h}_R|\tau_k\cdots\tau_N\rangle = -\sum_{\tau_N'} W(\tau_N\to\tau_N'). \tag{9.15}$$

With these definitions established, the relations (9.3)–(9.5) follow after some straightforward manipulations [97].

Let us take stock of this construction. We have seen that not only does the existence of a set of vectors and matrices that satisfy (9.3)–(9.5) imply a stationary solution of matrix-product form of the master equation (9.1), but also that such a set of vectors and matrices can always be constructed *once* one knows the weights (9.10). As we have previously seen in section 5 for the PASEP, one can find a number of different representations of the matrices D and E (which correspond to X_1 and X_0 in the more general setting) even when the auxiliary matrices $\tilde{D}$ and $\tilde{E}$ are the same. In the representation detailed above the auxiliaries are not scalars but rather complicated objects. Hence we see that (i) there may be many choices of both the matrices X_τ and their auxiliaries that correspond to the stationary solution of a single master equation; (ii) matrix reduction relations, like those found for the ASEP (2.39)–(2.41), constitute only sufficient conditions on X_τ, $\langle W|$ and $|V\rangle$, since valid representations where these relations do not hold can be found (that used in this section provides an example) and (iii) the construction of the auxiliary matrices through the generators of the stochastic process in (9.11) does not necessarily imply the existence of any convenient reduction relations for the X matrices that allow, for example, efficient computation of statistical properties.

9.1.1. Existence of a matrix-product state for models on the ring. As we have seen in section 4 the matrix-product state on the ring involves using a trace operation. The algebraic relations to be satisfied are given in the general case by (9.6). However, as we have seen in section 7, sometimes it happens that, although the algebraic relations are consistent, the rotational invariance of the trace operation leads to global constraints on particle numbers. For example, in the ABC model discussed in 7.1.4 we found that the matrix-product state was consistent only for models in which the numbers of each particle species was the same. It is essentially this property of models with periodic boundary conditions which has precluded the development of an existence proof for a matrix-product state parallel to that given above. These difficulties are discussed in more detail in [174].

9.2. Non-conserving models with finite-dimensional representations

We now discuss a specific reaction–diffusion model whose steady state can be represented in matrix-product form, but where the auxiliaries $\tilde{X}_\tau$ are matrices. This model contains one species of particle and the stochastic processes are diffusion, coagulation and decoagulation. These last two updates involve the annihilation or creation of a particle adjacent to another particle. If the bulk rates are chosen in the following way [175]

$$01 \underset{q^{-1}}{\overset{q}{\rightleftharpoons}} 10 \qquad 11 \underset{\Delta q^{-1}}{\overset{q}{\rightleftharpoons}} 10 \qquad 11 \underset{\Delta q}{\overset{q^{-1}}{\rightleftharpoons}} 01$$

then for suitable boundary conditions the steady state may be written in matrix-product form.

Originally in [175] a closed segment was considered. Then, one requires $\hat{D}|V\rangle = \hat{E}|V\rangle = \langle W|\hat{D} = \langle W|\hat{E} = 0$, the conditions to be satisfied come from the bulk algebra

$$\begin{aligned}
\hat{E}E - E\hat{E} &= 0 \\
\hat{E}D - E\hat{D} &= q^{-1}DE + q^{-1}DD - (\Delta+1)qED \\
\hat{D}E - D\hat{E} &= qED + qDD - (\Delta+1)q^{-1}ED \\
\hat{D}D - D\hat{D} &= \Delta qED + \Delta q^{-1}DD - (q+q^{-1})ED.
\end{aligned}$$

A four-dimensional representation of the matrices was found

$$E = \begin{pmatrix} q^{-2} & q^{-2} & 0 & 0 \\ 0 & (\Delta+1)^{-1} & (\Delta+1)^{-1} & 0 \\ 0 & 0 & 1 & q^2 \\ 0 & 0 & 0 & q^2 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \Delta/(\Delta+1) & \Delta/(\Delta+1) & 0 \\ 0 & 0 & \Delta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\hat{E} = \begin{pmatrix} 0 & 0 & q^{-1} & -(q-q^{-1})^{-1} \\ 0 & 0 & (q-q^{-1}) & -q \\ 0 & 0 & \Delta(q-q^{-1}) & -\Delta q \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\hat{D} = \begin{pmatrix} 0 & -\Delta q^{-1} & 0 & 0 \\ 0 & -\Delta(q-q^{-1}) & 0 & 0 \\ 0 & 0 & -\Delta(q-q^{-1}) & \Delta q \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

with boundary vectors

$$\langle W| = (1-q^2, 1, 0, a) \qquad |V\rangle = (b, 0, q^2, q^2-1)^T \tag{9.16}$$

where $a \neq b$ so that $\langle W|V\rangle \neq 0$.

However there is a subtlety with this model on a closed segment which is that the empty lattice is dynamically inaccessible and itself comprises an inactive steady state. If we wish to exclude this configuration we should choose a and b so that $\langle W|E^N|V\rangle$ vanishes. This can be accomplished but in doing so renders a and b N-dependent.

A first-order phase transition occurs at $q^2 = 1+\Delta$. For $q^2 < 1+\Delta$ the system is in a low-density phase whereas for $q^2 > 1+\Delta$ the system is in a high-density phase. Within the matrix-product form of the steady state the transition can be understood as the crossing of eigenvalues of C. This may occur here not because the matrices are of infinite dimension but because they have negative elements (cf the transfer matrices that arise in equilibrium statistical mechanics, as described in section 2.3.3).

Various other non-conserving models with finite-dimensional matrix-product states have been identified. For example, the bulk dynamics just described has also been studied in the case of an open left boundary, where particles enter with rate α and leave with rate γ, and a closed right boundary by Jafarpour [176]. Then, one requires $\hat{D}|V\rangle = \hat{E}|V\rangle = 0$ and

$$-\langle W|\hat{D} = \alpha\langle W|\hat{E} - \gamma\langle W|\hat{D} \tag{9.17}$$

$$-\langle W|\hat{E} = \gamma\langle W|\hat{E} - \alpha\langle W|\hat{D}. \tag{9.18}$$

If

$$\alpha = \left(\frac{1}{q} - q + \beta\right)\Delta, \tag{9.19}$$

one can find two-dimensional representations of the algebra. Further generalizations of this class of model have been shown to have finite-dimensional matrix-product states [177]. A general systematic approach to determine a necessary condition for the existence of a finite-dimensional matrix-product state has been put forward by Hieida and Sasamoto [178] in which more examples are given.

9.3. Several classes of particles

Second-class particles were discussed in section 2.4 and the matrix-product solution to the second-class particle problem was detailed in section 4. The matrix algebras for two-species generalizations of the second-class particle arose in the BI and BII solution classes of the classification scheme of section 7. We now consider a natural generalization of the second-class particle to many classes of particle, labelled $\tau = 1, 2, \ldots, n$, with dynamics

$$\tau 0 \xrightarrow{1} 0\tau \quad \forall\tau \qquad \text{and} \qquad \tau\tau' \xrightarrow{1} \tau'\tau \quad \text{if} \quad \tau < \tau'. \tag{9.20}$$

In this case it turns out one requires more complicated operators and auxiliaries than matrices.

For $n = 3$, the case of three particle classes, the algebraic relations to be satisfied, (9.6), becomes

$$\tilde{X}_i X_i - X_i \tilde{X}_i = 0, \qquad i = 0, 1, \ldots 3 \tag{9.21}$$

and

$$\begin{aligned}
X_\tau X_0 &= \tilde{X}_0 X_\tau - X_0 \tilde{X}_\tau = -\tilde{X}_\tau X_0 + X_\tau \tilde{X}_0, \qquad \tau > 0 \\
X_1 X_2 &= \tilde{X}_2 X_1 - X_2 \tilde{X}_1 = -\tilde{X}_1 X_2 + X_1 \tilde{X}_2 \\
X_1 X_3 &= \tilde{X}_3 X_1 - X_3 \tilde{X}_1 = -\tilde{X}_1 X_3 + X_1 \tilde{X}_3 \\
X_2 X_3 &= \tilde{X}_3 X_2 - X_3 \tilde{X}_2 = -\tilde{X}_2 X_3 - X_2 \tilde{X}_3.
\end{aligned} \tag{9.22}$$

The two-class problem, where the last lines of (9.22) are absent, can be solved (as we have seen in section 4.2) with scalar choices $\tilde{X}_1 = -1$, $\tilde{X}_0 = 1$ and $\tilde{X}_2 = 0$. It turns out that the three-class problem is not a simple generalization of the two-class case, but instead a much more complication solution is needed [179]. Matrices which satisfy (9.21), (9.22) have been given as

$$X_1 = \begin{pmatrix} D & 0 & E & 0 & \cdots \\ 0 & D & 0 & E & \\ 0 & 0 & D & 0 & \\ 0 & 0 & 0 & D & \\ \vdots & & & & \ddots \end{pmatrix}, \qquad X_2 = \begin{pmatrix} D & -E & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix}$$
$$X_3 = \begin{pmatrix} E & 0 & 0 & 0 & \cdots \\ D & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix}, \qquad X_0 = \begin{pmatrix} E & 0 & 0 & 0 & \cdots \\ 0 & E & 0 & 0 & \\ D & 0 & E & 0 & \\ 0 & D & 0 & E & \\ \vdots & & & & \ddots \end{pmatrix}. \tag{9.23}$$

The objects D and E that appear as elements of these infinite-dimensional matrices are themselves infinite-dimensional matrices D and E which satisfy the familiar relation $DE = D + E$. In other words X_i are rank four tensors, as indeed are the auxiliaries in this representation. These were also given by [179] and take the form

$$\tilde{X}_1 = \begin{pmatrix} D/2 + \mathbb{1} & 0 & E/2 - \mathbb{1} & 0 & \cdots \\ 0 & D/2 + \mathbb{1} & 0 & E/2 - \mathbb{1} & \\ 0 & 0 & D/2 + \mathbb{1} & 0 & \\ 0 & 0 & 0 & D/2 + \mathbb{1} & \\ \vdots & & & & \ddots \end{pmatrix},$$
$$\tilde{X}_2 = \begin{pmatrix} \mathbb{1} - D/2 & \mathbb{1} - E/2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix},$$
$$\tilde{X}_3 = \begin{pmatrix} E/2 - \mathbb{1} & 0 & 0 & 0 & \cdots \\ \mathbb{1} - D/2 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix}, \tag{9.24}$$
$$\tilde{X}_0 = -\begin{pmatrix} E/2 + \mathbb{1} & 0 & 0 & 0 & \cdots \\ 0 & E/2 + \mathbb{1} & 0 & 0 & \\ D/2 - \mathbb{1} & 0 & E/2 + \mathbb{1} & 0 & \\ 0 & D/2 - \mathbb{1} & 0 & E/2 + \mathbb{1} & \\ \vdots & & & & \ddots \end{pmatrix}.$$

The tensor, or matrix within matrix, form of (9.23) is closely related to that of the operators used to determine the steady-state current fluctuations in an open system [170].

In [179], some algebraic relations involving *only* the matrices X_τ (i.e., not the auxiliaries) were quoted. These included relations which had triples and quartets of matrices reducing to pairs, and others which transformed triples and pairs to other triples and pairs. As noted in [179], the relations given comprise only a subset of those that would actually be required to perform a complete reduction of any string of matrices to an irreducible form. Hence, a concise statement of the steady state of this system, like that encoded by the reduction relations (2.39)–(2.41) for the ASEP, is not easy. As such, no calculations beyond simple cases [179] have yet been attempted for the stationary properties of this model.

However recently, generalizing a construction by Angel [68] for the two-class problem, Ferrari and Martin [180] have shown how to generate the steady state of the multi-class problem. This construction can then be used to determine operators and auxiliaries to solve the steady state for more than three classes [181].

10. Discrete-time updating schemes

So far we have reviewed interacting particle systems involving continuous time or, equivalently, random-sequential dynamics as discussed in section 2. However, this is not necessarily the most natural choice of dynamics with which to model some physical systems of interest. For example, in traffic flow and pedestrian modelling [35, 53] it is desirable that the microscopic constituents are able to move simultaneously—this often originates from the existence of a smallest relevant timescale, e.g., reaction times in traffic. Therefore, in these systems, parallel dynamics in which all particles are updated in one discrete timestep may be used. In this section we consider the ASEP under three types of discrete-time dynamics: sublattice, ordered sequential and fully parallel.

In all these discrete-time cases the steady-state weights are given by the eigenvector with eigenvalue one of the transfer matrix of the dynamics, which we write in a schematic notation as

$$\hat{T} f(\tau_1, \ldots, \tau_N) = f(\tau_1, \ldots, \tau_N). \tag{10.1}$$

The transfer matrix $\hat{T}$ applied to the weight $f(\tau_1, \ldots, \tau_N)$ generates a sum of terms, each being the weight of a configuration multiplied by the transition *probability* in one discrete timestep from that configuration to $\{\tau_1, \ldots, \tau_N\}$.

10.1. Sublattice parallel dynamics

In sublattice updating the timestep is split into two halves: in the first half site 1, the even bonds $(2, 3), (4, 5), \ldots, (2L-2, 2L-1)$ and site $2L$ are updated simultaneously; in the second half-timestep the odd bonds $(1, 2), (3, 4), \ldots, (2L-1, 2L)$ are updated simultaneously. Note that we require the total number of sites $N = 2L$ to be even.

In figure 30, we show how this type of updating is applied to partially asymmetric exclusion dynamics with open-boundary conditions. In an update of site 1, if site 1 is empty a particle enters with probability α. In an update of a bond $i, i+1$, the two possible transitions are: if site i is occupied and site $i+1$ is empty the particle moves forward with probability p or else if site $i+1$ is occupied and site i is empty the particle moves backward with probability q. In an update of site N, if site N is occupied the particle leaves the system with probability β. We note that this two-step process avoids the possibility of a conflict occurring (e.g., two particles attempting to hop into the same site simultaneously).

The ASEP with sublattice dynamics was first studied in the case of deterministic bulk dynamics ($p = 1, q = 0$) by Schütz [182]. A matrix-product solution for this case was found

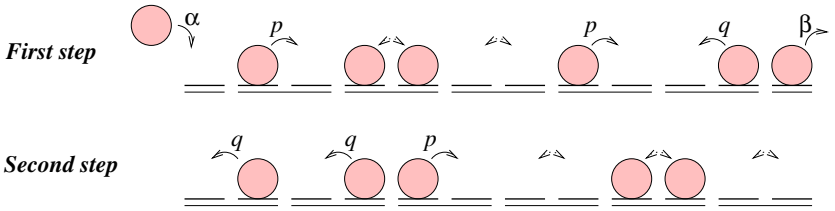


Figure 30. Typical configuration in the steady state of the partially asymmetric exclusion process on a closed segment of N sites with $q < 1$.

by Hinrichsen [183] and a matrix-product solution for the general case of stochastic bulk dynamics was found by Rajewsky *et al.* [169, 81]. We outline this general case here.

One proceeds by constructing the transfer matrix for the whole timestep as a product of operators for each half-timestep

$$\hat{T} = \hat{T}_2 \hat{T}_1 \tag{10.2}$$

where

$$T_1 = \hat{L}\hat{T}_{23}\hat{T}_{45}\ldots\hat{T}_{2L-2\,2L-1}\hat{R} \tag{10.3}$$

$$T_2 = \hat{T}_{12}\hat{T}_{34}\ldots\hat{T}_{2L-1\,2L} \tag{10.4}$$

$\hat{T}_{ii+1}$, $\hat{L}$, $\hat{R}$ when acting on $f(\tau_1\cdots\tau_N)$ generate the weights of configurations from which $\tau_1\cdots\tau_N$ could have been reached by a transition associated with that bond or boundary site, multiplied by the transition probability.

We may write

$$\hat{T}_{ii+1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-q & p & 0 \\ 0 & q & 1-p & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10.5}$$

$$\hat{L} = \begin{pmatrix} 1-\alpha & 0 \\ \alpha & 1 \end{pmatrix} \qquad \hat{R} = \begin{pmatrix} 1 & \beta \\ 0 & 1-\beta \end{pmatrix} \tag{10.6}$$

where the basis for $\hat{T}$ is 00, 01, 10, 11 and for $\hat{L}$ and $\hat{R}$ is 0, 1. So, for example,

$$\begin{aligned} &\hat{T}_{ii+1} f(\cdots\tau_{i-1}10\tau_{i+1}\cdots) \\ &\quad = pf(\cdots\tau_{i-1}01\tau_{i+1}\cdots) + (1-p)f(\cdots\tau_{i-1}10\tau_{i+1}\cdots). \end{aligned} \tag{10.7}$$

The matrix-product state for sublattice updating takes the form

$$f_N(\tau_1, \tau_2, \ldots, \tau_N) = \langle W|\hat{X}_{\tau_1}X_{\tau_2}\cdots\hat{X}_{\tau_{2L-1}}X_{\tau_{2L}}|V\rangle \tag{10.8}$$

in which, for i odd, X_{τ_i} is a matrix D if site i is occupied by a particle ($\tau_i = 1$) or a matrix E otherwise; and, for i even, $\hat{X}_{\tau_i}$ is a matrix $\hat{D}$ if site i is occupied by a particle or a matrix $\hat{E}$ otherwise. Note that different matrices, hatted and unhatted, are used for the odd and even sublattices.

The cancellation mechanism is as follows:

$$\hat{T}_{ii+1}\left[X_{\tau_i}\hat{X}_{\tau_{i+1}}\right] = \hat{X}_{\tau_i}X_{\tau_{i+1}} \tag{10.9}$$

$$\langle W|\hat{L}\hat{X}_{\tau_1} = \langle W|X_{\tau_1} \qquad \hat{R}X_{\tau_{2L}}|V\rangle = \hat{X}_{\tau_{2L}}|V\rangle. \tag{10.10}$$

Note that the action of the first half-timestep transfer matrix $\hat{T}_1$ is to put hatted matrices at the even sites and unhatted matrices on the odd sites, then the action of $\hat{T}_2$ is to restore them to their original sites. Thus, the matrix-product state (10.8) is indeed an eigenvector with eigenvalue one of the full transfer matrix (10.1). Using the form (10.5), (10.6) of the operators, the mechanism (10.9), (10.10) implies the following algebraic relations

$$\begin{aligned} &[E, \hat{E}] = [D, \hat{D}] = 0 \\ &\hat{E}D = (1-q)E\hat{D} + pD\hat{E} \\ &\hat{D}E = (1-p)D\hat{E} + qE\hat{D} \end{aligned} \tag{10.11}$$

and the boundary conditions

$$\begin{aligned} &\langle W|\hat{E}(1-\alpha) = \langle W|E \qquad (1-\beta)D|V\rangle = \hat{D}|V\rangle \\ &\langle W|(\alpha\hat{E} + \hat{D}) = \langle W|D \qquad (E + \beta D)|V\rangle = \hat{E}|V\rangle. \end{aligned} \tag{10.12}$$

Remarkably, the same matrices and vectors D, E as in the random-sequential case can be used to solve these relations. To see this we make the ansatz [169]

$$\hat{E} = E + \lambda\mathbb{1}; \qquad \hat{D} = D - \lambda\mathbb{1}, \tag{10.13}$$

where λ is a scalar to be fixed, and we find that (10.11), (10.12) reduce to

$$pDE - qED = \lambda[(1-p)D + (1-q)E] \tag{10.14}$$

$$D|V\rangle = \frac{\lambda}{\beta}|V\rangle; \qquad \langle W|E = \langle W|\frac{\lambda(1-\alpha)}{\alpha}. \tag{10.15}$$

Finally, we may set

$$\lambda^2 = \frac{1}{(1-q)(1-p)} \tag{10.16}$$

and define

$$E' = \lambda(1-q)E; \qquad D' = \lambda(1-p)D \tag{10.17}$$

which then satisfy

$$pD'E' - qE'D' = D' + E' \tag{10.18}$$

$$D'|V\rangle = \frac{1}{(1-q)\beta}|V\rangle; \qquad \langle W|E' = \langle W|\frac{(1-\alpha)}{\alpha(1-p)}. \tag{10.19}$$

Thus, we have rewritten (10.14), (10.15) in terms of matrices obeying the usual algebra for the PASEP (5.1)–(5.3) with redefined α and β. From this rewriting, the phase diagram can, in principle, be deduced using the general results of [120].

10.2. Ordered sequential dynamics

Another discrete-time updating scheme is to update each site in a fixed sequence in each timestep. Two particularly obvious choices of sequence are as follows:

Forward-ordered update. Here, the timestep begins by updating site 1 wherein if site 1 is empty a particle enters with probability α. Next, the bonds 1, 2 is updated such that if site 1 is occupied and site 2 is empty the particle moves forward with probability p or else if site 2 is occupied

and site 1 is empty the particle moves backward with probability q. Then the bonds $i, i+1$ for $i = 2 \cdots N-1$ are similarly updated in order. The timestep concludes with site N being updated wherein if site N is occupied the particle leaves the system with probability β. Note that in the forward sequence it is possible for a particle to move several steps forward in one timestep.

Backward-ordered update. In this case, the timestep begins by updating site N, then bonds $N-1, N$ to $1, 2$ in backward sequence and finally site 1, where all the individual updates follow the usual rules, as described above. Note that due to particle–hole symmetry the dynamics of the vacancies in the backward order is the same as the updating of particles in the forward order. Therefore, there is a symmetry between the steady states for the forward and backward orders.

Let us focus on the backward-ordered dynamics. We construct the transfer matrix for the whole timestep as a product of operators for each update of the sequence

$$\hat{T} = \hat{L}\hat{T}_{12}\hat{T}_{23}\ldots\hat{T}_{N-1N}\hat{R} \tag{10.20}$$

where $\hat{L}$, $\hat{T}_{ii+1}$, $\hat{R}$ are as in (10.5), (10.6).

The matrix-product solution is of the usual form

$$f_N(\tau_1, \tau_2, \ldots, \tau_N) = \langle W|X_{\tau_1}X_{\tau_2}\ldots X_{\tau_N}|V\rangle. \tag{10.21}$$

The cancellation mechanism is precisely the same as for the sublattice parallel case (10.9), (10.10). In the ordered case the hat matrices do not appear in the steady-state weights at the end of the timestep rather they are auxiliary matrices which appear during the update procedure and move from right to left through the lattice.

We conclude that backward-ordered updating has the exact same phase diagram as sublattice updating, although exact expressions for correlation functions do depend on the details of the updating [81, 184]. Finally, the phase diagram for forward-ordered updating can be obtained from the particle–hole symmetry mentioned above. Current and density profiles have been calculated in [185, 186].

10.3. Fully parallel dynamics

In parallel dynamics (sometimes referred to as fully parallel to distinguish from the sublattice updating) *all* bonds and boundary sites are updated simultaneously. This dynamics is considered the most natural for modelling traffic flow [35]. In an update at the bond $(i, i+1)$, if site i is occupied and site $i+1$ is empty the particle moves forward with probability p. Under this type of updating scheme one cannot include reverse hopping without introducing the possibility of conflicts occurring or of a particle hopping to two places at once. Also note that it is the occupancies at the beginning of the timestep which determine the dynamical events.

In this case, the matrix-product solution is of the usual form (10.21) but the algebraic relations have a more complicated structure which was first elucidated in [184]. There it was found that recursion relations between systems of different sizes were higher than first order: for example the relation

$$f_N(\ldots0100\ldots) = (1-p)f_{N-1}(\ldots010\ldots) + f_{N-1}(\ldots000\ldots) + pf_{N-2}(\ldots00\ldots), \tag{10.22}$$

which relates the weights of size N to those of size $N-1$ and size $N-2$, was discovered to hold. This in turn implies that algebraic relations between the operators are *quartic* rather than quadratic. For example, the rules in the bulk are

$$EDEE = (1-p)EDE + EEE + pEE \tag{10.23}$$

$$EDED = EDD + EED + pED \tag{10.24}$$

$$DDEE = (1-p)DDE + (1-p)DEE + p(1-p)DE \tag{10.25}$$

$$DDED = DDD + (1-p)DED + pDD \tag{10.26}$$

and there are other rules which we do not quote here for matrices near to the boundary (see [184]). These rules were proved using a domain approach related to, but more complicated than, that presented in section 3.1 [184]. Subsequently, an algebraic proof similar in spirit to the cancellation mechanism for the ordered sequential updating case was found [187] but it is still too involved to present here.

The quartic algebra, (10.23)–(10.26), as well as the other conditions mentioned above, can be reduced to quadratic rules by making a convenient choice for the operators involved. The trick is to write

$$D = \begin{pmatrix} D_1 & 0 \\ D_2 & 0 \end{pmatrix}, \qquad E = \begin{pmatrix} E_1 & E_2 \\ 0 & 0 \end{pmatrix}, \tag{10.27}$$

where D_1, D_2, E_1 and E_2 are matrices of, in general, infinite dimension; that is, D and E are written as rank four tensors with two indices of (possibly) infinite dimension and the other two indices of dimension two. Correspondingly, we write $\langle W|$ and $|V\rangle$ in the form

$$\langle W| = (\langle W_1|, \langle W_2|), \qquad |W\rangle = \begin{pmatrix} |V_1\rangle \\ |V_2\rangle \end{pmatrix}, \tag{10.28}$$

where $\langle W_1|$, $\langle W_2|$, $|V_1\rangle$ and $|V_2\rangle$ are vectors of the same dimension as D_1 and E_1. D_1, E_1, $\langle W_1|$ and $|V_1\rangle$ satisfy the quadratic relations

$$D_1E_1 = (1-p)\,[D_1 + E_1 + p]\,, \tag{10.29}$$

$$D_1|V_1\rangle = \frac{p(1-\beta)}{\beta}|V_1\rangle, \qquad \langle W_1|E_1 = \langle W_1|\frac{p(1-\alpha)}{\alpha}, \tag{10.30}$$

and D_2, E_2, $\langle W_2|$ and $|V_2\rangle$ satisfy

$$E_2D_2 = p\,[D_1 + E_1 + p]\,, \tag{10.31}$$

$$E_2|V_2\rangle = p|V_1\rangle, \qquad \langle W_2|D_2 = \langle W_1|p, \tag{10.32}$$

and $\langle W_1|$ satisfying (10.29), (10.30) are presented in [184]. In addition, one can choose $D_2 \propto E_1$, $E_2 \propto D_1$, $\langle W_2| \propto \langle W_1|$, $|V_2\rangle \propto |V_1\rangle$ so that (10.31), (10.32) reduce to (10.29), (10.30). Along the curve

$$(1-\alpha)(1-\beta) = 1-p, \tag{10.33}$$

scalar representations of D_1, E_1, $|V_1\rangle$ and $\langle W_1|$ can be found and D, E are 2×2 operators.

The continuous time limit is recovered by setting $p = \mathrm{d}t$, replacing $\alpha \to \alpha\,\mathrm{d}t$, $\beta \to \beta\,\mathrm{d}t$ and letting the timestep $\mathrm{d}t \to 0$. Then (10.29), (10.30) reduce to the usual ASEP quadratic algebra (2.39)–(2.41).

The phase diagram is presented in figure 31. It appears similar to the familiar continuous time phase diagram except that the transition lines from low density to maximal current and high density to maximal current are at $\alpha = 1 - \sqrt{1-p}$ and $\beta = 1 - \sqrt{1-p}$, respectively. In the continuous time limit described above one recovers $\alpha = 1/2$ and $\beta = 1/2$. Also, in the limit of deterministic bulk dynamics, $p \to 1$ [188], the maximal-current phase disappears from the phase diagram as is also the case in other updating schemes in this limit.

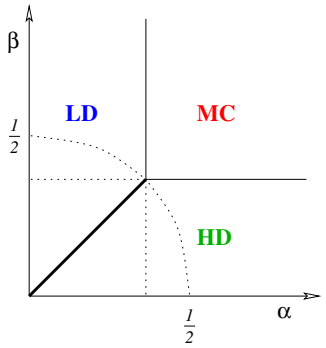


Figure 31. Phase diagram for the ASEP with parallel update. Note that $\alpha \leqslant 1, \beta \leqslant 1$. MC is the maximum-current phase, LD and HD are the low- and high-density phases, respectively. The boundaries between the low-density and maximal-current phase and the high-density and maximal-current phase are at $\alpha = 1 - \sqrt{1-p}$ and $\beta = 1 - \sqrt{1-p}$, respectively. The curved dashed line is the line given by (10.33) and intersects the line $\alpha = \beta$ at $\alpha = \beta = 1 - \sqrt{1-p} = q$.

11. Summary and outstanding challenges

In this work we have reviewed the physical properties of systems that are driven out of equilibrium, with particular reference to those one-dimensional models that can be solved exactly using a matrix-product method. The most prominent physical phenomena exhibited by these systems are phase transitions which in open systems are induced by changes in the interactions with the boundaries and in periodic systems by the addition of particle species that have different dynamics. Additionally, we have seen in a number of cases the formation of shocks in these one-dimensional systems.

As we have also discussed, the matrix-product approach can be used to determine the steady-state statistics of models with a variety of microscopic dynamics. The majority of models that have matrix-product states involves diffusing particles with hard-core interactions and whose numbers are conserved, except possibly at boundary sites. Under such conditions, certain generalizations to multispecies models have been found. Furthermore, there are a few cases in which non-conserving particle reactions can occur that also admit a convenient solution in terms of matrix products.

However, it has to be said that the dream scenario, of being able to systematically construct any nonequilibrium steady state in matrix-product form starting from the microscopic dynamics of the system, appears a remote goal. Although the existence proofs discussed in section 9 tell us that a matrix-product formulation should nearly always be possible, we are in practice still feeling our way with particular examples.

There remain a number of simple physical systems that exhibit nontrivial out-of-equilibrium behaviour, but that have nevertheless been unyielding to exact solution by matrix products or any other means. It is perhaps appropriate here to highlight some of these cases as challenges for future research.

11.1. Particlewise and sitewise disorder in the ASEP

As was discussed in section 8.1, it is possible to solve for the steady state of the ASEP on the ring when particles have different hop rates, as would occur when each particle is randomly assigned a hop rate from some distribution. As far as we aware, the corresponding dynamics

with open boundaries has not been solved apart from Karimipour's model of overtaking which we reviewed in section 8.2. Furthermore, if the disordered hop rates are associated with sites rather than particles, even the model on the ring has so far eluded a complete exact solution. Indeed, even if only a single site has a different hop rate to the rest, the steady state does not appear to have a manageable form [189, 190]. On the other hand, some nontrivial symmetries in the disordered case have been identified [191, 192].

One phenomenon that can emerge when disorder is present is phase separation. One sees this most clearly for the periodic system with particlewise disorder, where as we noted in section 8.1.1 a condensation of particles behind the slowest may occur [145]. With sitewise disorder [191], a flattening of the macroscopic current-density relation $J(\rho)$ is seen (recall that it is parabolic $J(\rho) = \rho(1 - \rho)$ in the ASEP without disorder), and the densities at which transitions into a maximal-current phase are correspondingly lowered [193]; it has been suggested that a linear portion in the $J(\rho)$ curve may also provide a mechanism for phase separation [194]. Another interesting observation is that under sitewise disorder the location of the first-order transition for the ASEP becomes sample dependent (i.e., the free-energy-like quantity for the nonequilibrium system is not self-averaging) [195].

11.2. Driven n-mers with open boundaries

In all the models described in this work, particles occupy a single site of a lattice. However, from the earliest work on the ASEP as a model of biopolymerization [38], there has been interest in variants that have particles occupying multiple sites. If one has extended objects on a ring, the dynamics are unaffected. With open boundaries, however, the situation changes and even if there is a single particle species and all particles are the same length, a matrix-product solution for the steady state has proved elusive. Although it is quite easy to arrange for cancellation of terms in the master equation corresponding to particles joining and leaving domains (see section 3.1), the necessary cancellation with boundary terms is much harder to arrange: it is not even clear what the most convenient choice of boundary conditions should be (e.g., extended particles sliding on and off site-by-site or appearing and disappearing in their entirety or a mixture of the two). A phenomenological diffusion equation for the macroscopic density profile in the open system with extended objects appearing and disappearing at both boundaries has been presented [196]. Furthermore, it was conjectured in that work to be the equation that would result in the continuum limit (along the lines of that taken in section 2.2), if an exact matrix-product solution of this model were to be found. Whether this should turn out to be the case or not, an application of the extremal current principle (see section 2.2.4) gives predictions for the phase diagram that are in good agreement with simulation data: the same three phases that emerge in the ASEP are seen, but with shifted transition points. Meanwhile an approach to this problem based on a local-equilibrium approximation [197] has shown excellent agreement with Monte Carlo results and have also been conjectured to be exact. The case of the open system with extended objects and a single spatial inhomogeneity in the hop rates has also been considered [198].

11.3. Bridge model

The bridge model [150, 151] comprises oppositely-charged particles in the sense of section 7. Positive charges enter at the left at rate α, hop with unit rate to the right and exit at the right boundary at rate β. The dynamics of negative charges is exactly the same, but in the opposite direction. When two opposing particles meet they may exchange with some rate q. Thus, this model has both a parity and charge-conjugation symmetry. In certain parameter ranges, this

symmetry is reflected in the steady state that results: both positive and negative charges flow freely in their preferred directions. In particular, the limit $\alpha \to \infty$ at $\beta = 1$, for which an exact solution in terms of matrix products exists, is included within this range. On the other hand, when β is small the model's symmetries are found to be broken. This has been proved in the $\beta \to 0$ limit [199] but there is no known exact solution for general β. In a finite system, the current flips between the positive and negative direction, but as the system size increases the time between flips increases exponentially, and so in the thermodynamic limit only the state with positive or negative current state is seen. It is unclear whether the structure of the matrix-product solution allows for the description of such symmetry-broken states. This may also be related to whether the boundary conditions correspond to particle reservoirs with well-defined densities [200]. Meanwhile, progress has been made on understanding spontaneous symmetry breaking in the sublattice bridge model [201].

11.4. ABC model

The ABC model [146, 147] is another deceptively simple model that has yet to be completely solved. This model has a ring that is fully occupied by particles from three species, which are labelled A, B and C and exhibit a symmetry under cyclic permutation of the labels. Specifically, the rates at which neighbouring particles exchange are

$$AB \underset{1}{\overset{q}{\rightleftharpoons}} BA \qquad BC \underset{1}{\overset{q}{\rightleftharpoons}} CB \qquad CA \underset{1}{\overset{q}{\rightleftharpoons}} AC. \tag{11.1}$$

That is, if $q < 1$, A particles prefer to be to the left of Bs, Bs to the left of Cs and Cs to the left of A s. This model was discussed in section 7.1.4 and was shown to be exactly solvable when the numbers of A, B and C particles are all equal as a consequence of detailed balance being satisfied. Here a phase separation is seen when $q \neq 1$. If $q = 1$, on the other hand, all configurations become equivalent and the stationary distribution is uniform. When there is even the slightest imbalance and particle numbers *and* hop rates, a small current starts to flow and nothing is known about the properties of the stationary distribution.

The model is of particular importance because in the special case of equal numbers of each species it is one of few models along with the SSEP and KMP model [202] where a free-energy functional for the density profile is known.

11.5. Fixed random sequence and shuffled dynamics

We discussed in section 10 different updating schemes for which matrix-product steady states have been determined. Here, we mention some other updating schemes which have not been solved so far. First an ordered discrete-time sequential updating scheme could have the order as a randomly chosen sequence. The forward and backward schemes in section 10 correspond to special sequences. It would be of interest to construct a matrix-product steady state for an arbitrary fixed sequence.

Recently a *shuffled* updating scheme has been considered [206, 207], in which a new random sequence is chosen at each timestep. It has been argued that this scheme is relevant to the modelling of pedestrian dynamics. The shuffled updating scheme guarantees that each site or bond is updated exactly once in each timestep but is more stochastic than an ordered sequential scheme. Again it would be of interest to see if a matrix product could be used to describe the steady state.

11.6. KLS model

We end by mentioning the original paradigm of a system driven out of equilibrium, the model due to Katz, Lebowitz and Spohn [32, 203, 204]. This model has Ising-like interactions between pairs of neighbouring spins and evolves by neighbouring spins exchanging places (Kawasaki dynamics) but with a symmetry-breaking external field that favours hops of up-spins along a particular axis. When the system is open, or has periodic boundaries, a nonequilibrium steady state ensues. In the limit where spins become non-interacting and there is only one spatial direction, the partially asymmetric exclusion process (PASEP) described in section 5 is recovered. When the spins interact, an Ising-like steady state that involves products of 2×2 matrices (as described in section 2.3.3) can be found, albeit only if certain conditions on the hop rates are satisfied. (These conditions do admit nonequilibrium steady states in which a current flows, however.) An exact solution for this model along a special line in parameter space has also been given in [205]. In two dimensions, one has, of course, a finite critical temperature in the absence of a driving force, and there has been particular interest in the nature of the low-temperature ordered phase in the presence of an external drive [32].

Acknowledgments

We thank all the colleagues with whom it has been a pleasure to discuss nonequilibrium matrix-product states over the years. In particular, we acknowledge our collaborators and other authors whose work we have summarized. For their useful discussions and insightful comments during the preparation of this manuscript we would wish to thank Bernard Derrida, Rosemary Harris, Des Johnston, Joachim Krug, Kirone Mallick, Andreas Schadschneider, Gunter Schütz and Robin Stinchcombe. RAB further acknowledges the Royal Society of Edinburgh for the award of a Research Fellowship.

Appendix A. Method of characteristics

The continuity equation (2.3) is a particularly simple case of a first-order quasi-linear differential equation for the density

$$a(x,t,\rho)\frac{\partial \rho}{\partial t} + b(x,t,\rho)\frac{\partial \rho}{\partial x} = c(x,t,\rho). \tag{A.1}$$

The term on the right-hand side would represent a source or sink of particles and is relevant for example to the case where creation and annihilation processes exist. Such an equation can generally be solved by the method of characteristics [208] which involves identifying the characteristic curves along which information from the boundary or initial condition propagates through the spacetime domain. The characteristic curves satisfy

$$\frac{\mathrm{d}t}{a} = \frac{\mathrm{d}x}{b} = \frac{\mathrm{d}\rho}{c} \tag{A.2}$$

of which the two independent conditions may be written as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{b}{a} \qquad \frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{c}{b} \tag{A.3}$$

These equations generally have two families of solution

$$\phi(x,t,\rho(x,t)) = C_1 \qquad \psi(x,t,\rho(x,t)) = C_2. \tag{A.4}$$

Then the solution to (A.1) is of the form

$$F(\phi,\psi) = 0 \tag{A.5}$$

where the function F is fixed by the initial data.

For the case (2.3), i.e. $c = 0$, the characteristics in the x–t plane are straight lines with slope $v_g(\rho)$ and ρ is constant along these lines, as can be seen from (A.3). More generally the characteristics are curves in the x–t plane and the density varies along the characteristic. This can be seen by integrating (A.3)

$$G(\rho) - G(\rho_0) = t \qquad \text{where} \qquad G(\rho) = \int^{\rho} \mathrm{d}\rho \frac{b(\rho)}{c(\rho)} \tag{A.6}$$

which gives ρ implicitly as a function of t, then the trajectory of the characteristic is given by

$$x = x_0 + \int_0^t b(\rho(t'))\,\mathrm{d}t'. \tag{A.7}$$

If we consider a patch of the initial density profile of density ρ_0 at x_0, its density evolves according to (A.6) and (A.7) gives the position of the patch. These more complicated characteristics can produce a richer variety of boundary-induced phase transitions [46, 209–211] including stationary shocks.

Appendix B. Generating functions and asymptotics

In this appendix, we collect together general formulae for obtaining both exact and asymptotic expressions for the coefficients appearing in a generating function if the latter is known. We refer the reader to [100] for proofs and more detailed discussion. For brevity, we introduce the notation $\{x^n\}f(x)$ to mean 'the coefficient of x^n in the (formal) power series $f(x)$'.

Lagrange inversion formula. If a generating function $f(x)$ satisfies the functional relation

$$f(x) = x\Phi(f) \tag{B.1}$$

where $\Phi(f)$ satisfies $\Phi(0) = 1$, one has an expression for x as a function of f which can be inverted to find f as a function of x. This procedure then allows one to determine with relative ease the coefficient of x^n in the expansion of any *arbitrary* function $F(f)$. The formula for doing so reads

$$\{x^n\}F(f) = \frac{1}{n}\{f^{n-1}\}[F'(f)\Phi(f)^n]. \tag{B.2}$$

Asymptotics. If a generating function $f(x)$ has a singularity at a point x_0 in the complex plane, it contributes a term $Ax_0^{-n}n^{-\mu}$ to the coefficient $\{x^n\}f(x)$, where A and μ are constants to be determined. For sufficiently large n, this coefficient is dominated by the singularity closest to the origin. To be more precise, we say that two sequences a_n and b_n are asymptotically equivalent, denoted as $a_n \sim b_n$, if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1. \tag{B.3}$$

Then, we have that $\{x^n\}f(x) \sim Ax_0^{-n}n^{-\mu}$.

Let us decompose the generating function $f(x)$ into its regular and singular parts around x_0, denoting the former by $f_r(x)$. That is,

$$f(x) = f_r(x)\left(1 - \frac{x}{x_0}\right)^{-\nu}. \tag{B.4}$$

The cases that interest us are poles $\nu = 1, 2, \ldots$, and algebraic singularities (noninteger ν). For the former case one finds

$$\{x^n\}f(x) \sim \binom{\nu + n - 1}{n} f_r(x_0)x_0^{-n}. \tag{B.5}$$

For sufficiently large n, the binomial coefficient behaves as

$$\binom{\nu+n-1}{n} \sim \frac{n^{\nu-1}}{(\nu-1)!}. \tag{B.6}$$

Replacing the factorial with a Gamma function gives the extension of (B.5) to algebraic singularities

$$\{x^n\} f(x) \sim \frac{f_r(x_0)}{\Gamma(\nu)} x_0^{-n} n^{\nu-1}. \tag{B.7}$$

The prefactor $f_r(x_0)$ in this expression can be obtained by taking the limit

$$f_r(x_0) = \lim_{x \to x_0} \left(1 - \frac{x}{x_0}\right)^{\nu} f(x). \tag{B.8}$$

Appendix C. Equivalence of the integral and sum representations of the ASEP normalization

In order to show that the integral representation of the normalization (3.49) is equivalent to the summation (3.38), the simplest approach is to show that the generating functions for Z_N computed from the two expressions are equal. As we have seen in section 3.2.3, the generating function computed from the finite sum (3.38) yields (3.56) which can be developed further to give

$$\mathcal{Z}(z) \equiv \sum_{N=0}^{\infty} z^N Z_N \tag{C.1}$$

$$= \frac{\alpha\beta[2\alpha\beta - 2z - (\alpha+\beta-1)(1+\sqrt{1-4z})]}{2\,[z-\alpha(1-\alpha)]\,[z-\beta(1-\beta)]}. \tag{C.2}$$

We now turn to the integral representation (3.49) which may be recast as a contour integral by change of variable $u = \mathrm{e}^{\mathrm{i}\theta}$:

$$Z_N = \frac{(1-ab)}{4\pi\mathrm{i}} \oint_K \frac{\mathrm{d}u}{u} \frac{(2-u^2-u^{-2})(2+u+u^{-1})^N}{(1-au)(1-au^{-1})(1-bu)(1-bu^{-1})} \tag{C.3}$$

where K is the (positively oriented) unit circle in the complex u plane.

We compute the generating function (C.1) using (C.3). Summing a geometric series (which converges for $|z| < 1/4$) we find

$$\mathcal{Z}(z) = \frac{(1-ab)}{4\pi\mathrm{i}} \oint_K \frac{\mathrm{d}u}{u} \frac{(2-u^2-u^{-2})}{(1-au)(1-au^{-1})(1-bu)(1-bu^{-1})(1-z(2+u+u^{-1}))}. \tag{C.4}$$

We now use the residue theorem to evaluate this integral. We take a and $b < 1$ (although we get the same final result for other ranges of a and b). In this case, the singularities within the unit circle are simple poles at $u = a, u = b$ and $u = u_-$ where we define

$$u_\pm = 1 - \frac{1}{2z}(1 \pm \sqrt{1-4z}). \tag{C.5}$$

The residues from the three poles yield

$$\mathcal{Z}(z) = \frac{(1-a^2)}{2(a-b)} \frac{\alpha^2}{[z-\alpha(1-\alpha)]} - \frac{(1-b^2)}{2(a-b)} \frac{\beta^2}{[z-\beta(1-\beta)]} \tag{C.6}$$

$$-\frac{(1-ab)(2-u_-^2-u_-^{-2})}{2z(1+a^2-a(u_-+u_-^{-1}))(1+b^2-b(u_-+u_-^{-1}))(u_--u_+)}. \tag{C.7}$$

Some simple algebra then shows that this expression is equivalent to (C.2).

Appendix D. Matrix representations

Here, we collect together some representations of the various matrix-product algebras that have been discussed in this work.

D.1. Representations of PASEP algebra

We begin by recapping representations for the PASEP algebra in the case of injection at the left boundary and extraction at the right boundary:

$$DE - qED = D + E \tag{D.1}$$

$$\beta D|V\rangle = |V\rangle \tag{D.2}$$

$$\alpha\langle W|E = \langle W|. \tag{D.3}$$

These representations are generalizations of the three representations for the totally asymmetric case first given in [62].

From section 5.2, we have

$$D = \frac{1}{1-q}\begin{pmatrix} 1+b & \sqrt{c_0} & 0 & 0 & \cdots \\ 0 & 1+bq & \sqrt{c_1} & 0 & \\ 0 & 0 & 1+bq^2 & \sqrt{c_2} & \\ 0 & 0 & 0 & 1+bq^3 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.4}$$

$$E = \frac{1}{1-q}\begin{pmatrix} 1+a & 0 & 0 & 0 & \cdots \\ \sqrt{c_0} & 1+aq & 0 & 0 & \\ 0 & \sqrt{c_1} & 1+aq^2 & 0 & \\ 0 & 0 & \sqrt{c_2} & 1+aq^3 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.5}$$

where

$$a = \frac{1-q}{\alpha} - 1, \qquad b = \frac{1-q}{\beta} - 1, \qquad c_n = (1-q^{n+1})(1-abq^n), \tag{D.6}$$

and the boundary vectors are

$$\langle W| = \langle 0| \qquad \text{and} \qquad |V\rangle = |0\rangle. \tag{D.7}$$

From this representation one sees that for certain parameter curves, namely

$$1 - abq^n = 0, \tag{D.8}$$

the representations become finite dimensional since $c_n = 0$ and the upper left corner of the matrices become disconnected from the rest. Curves of this type were first noted in the more general case $\gamma, \delta \neq 0$ in [212] and finite-dimensional representations were catalogued in [213].

In the limit $q \to 1$ (5.12) and (5.12) have well-defined limits

$$D = \begin{pmatrix} 1/\beta & \sqrt{g} & 0 & 0 & \cdots \\ 0 & 1+1/\beta & \sqrt{2g} & 0 & \\ 0 & 0 & 2+1/\beta & \sqrt{3g} & \\ 0 & 0 & 0 & 3+1/\beta & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.9}$$

$$E = \begin{pmatrix} 1/\alpha & 0 & 0 & 0 & \cdots \\ \sqrt{g} & 1+1/\alpha & 0 & 0 & \\ 0 & \sqrt{2g} & 2+1/\alpha & 0 & \\ 0 & 0 & \sqrt{3g} & 3+1/\alpha & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.10}$$

where $g = \frac{1}{\alpha} + \frac{1}{\beta}$.

A second representation of (D.1)–(D.3) given in section 5.3 is

$$D = \frac{1}{1-q} \begin{pmatrix} 1 & \sqrt{1-q} & 0 & 0 & \cdots \\ 0 & 1 & \sqrt{1-q^2} & 0 & \\ 0 & 0 & 1 & \sqrt{1-q^3} & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.11}$$

$$E = \frac{1}{1-q} \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ \sqrt{1-q} & 1 & 0 & 0 & \\ 0 & \sqrt{1-q^2} & 1 & 0 & \\ 0 & 0 & \sqrt{1-q^3} & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.12}$$

with

$$\langle W|n\rangle = \kappa \frac{a^n}{\sqrt{(q;q)_n}} \qquad \text{and} \qquad \langle n|V\rangle = \kappa \frac{b^n}{\sqrt{(q;q)_n}} \tag{D.13}$$

where the parameters a and b are given in (D.6), and κ is a constant usually chosen so that $\langle V|W\rangle = 1$. For these matrices, the $q \to 1$ limit is not defined.

A third representation, which we have not yet encountered, is

$$D = \begin{pmatrix} 1/\beta & 1/\beta & 1/\beta & 1/\beta & \cdots \\ 0 & d^{(11)} & d^{(12)} & d^{(13)} & \cdots \\ 0 & 0 & d^{(22)} & d^{(23)} & \cdots \\ 0 & 0 & 0 & d^{(33)} & \\ \vdots & & & & \ddots \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.14}$$

$$\langle W|n\rangle = (1/\alpha)^n \quad \langle n|V\rangle = \delta_{n,0} \tag{D.15}$$

where

$$d^{(ij)} = \sum_{m=0}^{i-1} q^m \binom{j-i+m}{m} + \frac{q^i}{\beta} \binom{j}{i}. \tag{D.16}$$

In the $q \to 1$ limit the elements of D reduce to

$$d^{(ij)} = \binom{j}{i-1} + \frac{1}{\beta}\binom{j}{i}. \tag{D.17}$$

D.2. Representations of two-species and multispecies algebras

We now give representations corresponding to the physically relevant and nontrivial algebras classified in section 7 and the multispecies generalizations of section 8.

Solution AII and multispecies generalization. This is a multispecies model with matrices E and $D(v)$, with an algebra

$$\begin{aligned} D(v)E &= \frac{1}{v}D(v) + E \\ D(v)D(v') &= \frac{v}{v-v'}D(v') - \frac{v'}{v-v'}D(v) \qquad v > v' \\ D(v)|V\rangle &= \frac{1}{\beta(v)}|V\rangle \\ \langle W|E &= \frac{1}{\alpha}, \end{aligned} \tag{D.18}$$

where $\beta(v) = \frac{v-(1-\beta)}{v}$. One possible representation is

$$D(v) = \begin{pmatrix} 1/\beta(v) & 1/(\beta(v)v) & 1/(\beta(v)v^2) & 1/(\beta(v)v^3) & \cdots \\ 0 & 1 & 1/v & 1/v^2 & \cdots \\ 0 & 0 & 1 & 1/v & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & & & & \ddots \end{pmatrix}$$

$$E = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ \vdots & & & & \ddots \end{pmatrix}$$

$$\langle W|n\rangle = (1/\alpha)^n \quad \langle n|V\rangle = \delta_{n,0}. \tag{D.19}$$

This is the generalization of the third PASEP representation (D.15) in the case $q = 0$.

Solution BI and BII. With the algebra (7.18),

$$\begin{aligned} p_1X_1X_0 - q_1X_0X_1 &= x_0X_1 - x_1X_0 \\ p_2X_2X_0 - q_2X_0X_2 &= x_0X_2 \\ p_2X_1X_2 - q_2X_2X_1 &= -x_1X_2, \end{aligned} \tag{D.20}$$

one has a representation

$$X_1 = -\frac{x_1}{(p_2-q_2)} \begin{pmatrix} 1 & \sqrt{c_0} & 0 & 0 & \cdots \\ 0 & 1 & \sqrt{c_1} & 0 & \\ 0 & 0 & 1 & \sqrt{c_2} & \\ 0 & 0 & 0 & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.21}$$

$$X_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & (q_2/p_2) & 0 & 0 & \\ 0 & 0 & (q_2/p_2)^2 & 0 & \\ 0 & 0 & 0 & (q_2/p_2)^3 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.22}$$

$$X_0 = \frac{x_0}{(p_2 - q_2)} \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ \sqrt{c_0} & 1 & 0 & 0 & \\ 0 & \sqrt{c_1} & 1 & 0 & \\ 0 & 0 & \sqrt{c_2} & 1 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.23}$$

where

$$c_n = (1 - (q_1/p_1)^{n+1}). \tag{D.24}$$

As noted in the main text a representation for solution BII is to use a representation of the PASEP algebra and write $X_2 = |V\rangle\langle W|$, i.e. a projector.

Solution CII and multispecies generalization. This is the case of the ASEP with disordered hopping rates. The matrix algebra (8.2)

$$p_\mu D_\mu E - q_\mu E D_\mu = D_\mu \tag{D.25}$$

can be represented via

$$D_\mu = \begin{pmatrix} d_\mu^{(00)} & d_\mu^{(01)} & d_\mu^{(02)} & d_\mu^{(03)} & \cdots \\ 0 & d_\mu^{(11)} & d_\mu^{(12)} & d_\mu^{(13)} & \cdots \\ 0 & 0 & d_\mu^{(22)} & d_\mu^{(23)} & \cdots \\ 0 & 0 & 0 & d_\mu^{(33)} & \\ \vdots & & & & \ddots \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ \vdots & & & & \ddots \end{pmatrix} \tag{D.26}$$

where

$$d_\mu^{(ij)} = \binom{i+j}{i} \frac{q_\mu^i}{p_\mu^{i+j}}. \tag{D.27}$$

Solution D and deformed commutators. As noted in section 7.2.2 the solution class D operators can be constructed from tensor products of operators obeying deformed commutator relations. The deformed commutator is defined as

$$X_1 X_2 - r X_2 X_1 = 0 \tag{D.28}$$

where r is a deformation parameter: $r = 1$ corresponds to the usual commutator $[X_1, X_2] = 0$. Deformed commutators appear in the analysis of the PASEP and in some of the two-species models of section 7. One possible representation of these matrices is

$$X_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 & \\ \vdots & & & & \ddots \end{pmatrix} \qquad X_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & r & 0 & 0 & \\ 0 & 0 & r^2 & 0 & \\ 0 & 0 & 0 & r^3 & \\ \vdots & & & & \ddots \end{pmatrix}. \tag{D.29}$$

We note that both X_1 and any product of X_1 and X_2 including at least one X_1 are traceless.

References

[1] Onsager L and Machlup S 1953 Fluctuations and irreversible processes *Phys. Rev.* **91** 1505
[2] Onsager L and Machlup S 1953 Fluctuations and irreversible processes: II. Systems with kinetic energy *Phys. Rev.* **91** 1512
[3] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2001 Fluctuations in stationary nonequilibrium states of irreversible processes *Phys. Rev. Lett.* **87** 040601
[4] Bertini L, De Sole A, Gabrielli D, Jona-Lasinio G and Landim C 2002 Macroscopic fluctuation theory for stationary non-equilibrium states *J. Stat. Phys.* **107** 635
[5] Kurchan J 2007 Non-equilibrium work relations *J. Stat. Mech.: Theor. Exp.* P07005
[6] Harris R J and Schütz G M 2007 Fluctuation theorems for stochastic dynamics *J. Stat. Mech: Theor. Exp.* P07020
[7] Taniguchi T and Cohen E G D 2006 Onsager–Machlup theory for nonequilibrium steady states and fluctuation theorems *J. Stat. Phys.* **126** 1–41
[8] Bray D 2000 *Cell Movements* (Garland) 2nd edn
[9] Evans M R and Blythe R A 2002 Nonequilibrium dynamics in low-dimensional systems *Physica* A **313** 110
[10] Durrett R 2001 *Essentials of Stochastic Processes* (New York: Springer)
[11] Kelly F P 1979 *Reversibility and Stochastic Networks* (New York: Wiley)
[12] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: Elsevier)
[13] Evans M R, Majumdar S N and Zia R K P 2004 Factorized steady states in mass transport models *J. Phys. A: Math. Gen.* **37** L275
[14] Evans M R and Hanney T 2005 Nonequilibrium statistical mechanics of the zero-range process and related models *J. Phys. A: Math. Gen.* **38** R195–240
[15] Blythe R A 2001 Nonequilibrium steady states and dynamical scaling regimes *PhD Thesis* University of Edinburgh
[16] Zia R K P and Schmittmann B 2007 Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states *J. Stat. Mech.: Theor. Exp.* P07012
[17] Arndt P F 2000 Yang–Lee theory for a nonequilibrium phase transition *Phys. Rev. Lett.* **84** 814
[18] Dammer S M, Dahmen S R and Hinrichsen H 2002 Yang–Lee zeros for a nonequilibrium phase transition *J. Phys. A: Math. Gen.* **35** 4527
[19] Blythe R A and Evans M R 2002 Lee–Yang zeros and phase transitions in nonequilibrium steady states *Phys. Rev. Lett.* **89** 080601
[20] Derrida B and Evans M R 1996 The asymmetric exclusion model: exact results through a matrix approach *Nonequilibrium Statistical Mechanics in One Dimension* (Cambridge: Cambridge University Press)
[21] Derrida B 1998 An exactly soluble non-equilibrium system: the asymmetric simple exclusion process *Phys. Rep.* **301** 65
[22] Gwa L H and Spohn H 1992 Bethe solution for the dynamical-scaling exponent of the noisy Burgers equation *Phys. Rev.* A **46** 844
[23] Kim D 1995 Bethe ansatz solution for crossover scaling functions of the asymmetric XXZ chain and the Karder–Parisi–Zhang-type growth model *Phys. Rev.* E **52** 3512
[24] de Gier J and Essler F H L 2005 Bethe ansatz solution of the asymmetric exclusion process with open boundaries *Phys. Rev. Lett.* **95** 240601
[25] Schütz G M 1997 Exact solution of the master equation for the asymmetric exclusion process *J. Stat. Phys.* **88** 427–45
[26] Prähofer M and Spohn H 2000 Universal distributions for growth processes in 1 + 1 dimensions and random matrices *Phys. Rev. Lett.* **84** 4882–5
[27] Nagao T and Sasamoto T 2004 Asymmetric simple exclusion process and modified random matrix ensembles *Nucl. Phys.* B **699** 487
[28] Stinchcombe R B and Schütz G M 1995 Application of operator algebras to stochastic dynamics and the Heisenberg chain *Phys. Rev. Lett.* **75** 140–3
[29] Schütz G M 1998 Dynamic matrix ansatz for integrable reaction–diffusion processes *Eur. Phys. J.* B **5** 589–97
[30] Golinelli O and Mallick K 2006 The asymmetric simple exclusion process: an integrable model for non-equilibrium statistical mechanics *J. Phys. A: Math. Gen.* **39** 12679–705
[31] Sasamoto T 2007 Fluctuations of the one-dimensional asymmetric exclusion process using random matrix techniques *J Stat. Mech.: Theor. Exp.* P07007
[32] Schmittmann B and Zia R K P 1995 *Statistical Mechanics of Driven Diffusive Systems (Phase Transitions and Critical Phenomena* vol 17*)* (London: Academic)
[33] Krug J 1997 Origin of scale invariance in growth processes *Adv. Phys.* **46** 139

[34] Schütz G M 2000 *Exactly Solvable Models for Many-Body Systems Far From Equilibrium (Phase Transitions and Critical Phenomena* vol 19*)* (London: Academic)

[35] Chowdhury D, Santen L and Schadschneider A 2000 Statistical physics of vehicular traffic and some related systems *Phys. Rep.* **329** 199

[36] Stinchcombe R B 2001 Stochastic non-equilibrium systems *Adv. Phys.* **50** 431–96

[37] Schütz G M 2003 Critical phenomena and universal dynamics in one-dimensional driven diffusive systems with two species of particles *J. Phys. A: Math. Gen.* **36** R339–79

[38] MacDonald C T, Gibbs J H and Pipkin A C 1968 Kinetics of biopolymerization on nucleic acid templates *Biopolymers* **6** 1

[39] Heckmann K 1972 Single file diffusion *Passive Permeability of Cell Membranes* ed F Kreuzer and J F G Slegers (New York: Plenum) (Biomembranes) 3 127

[40] Harris T E 1965 Diffusion with 'collisions' between particles *J. Appl. Probab.* **2** 323–38

[41] Spitzer F 1970 Interaction of Markov processes *Adv. Math.* **5** 2

[42] Schadschneider A 2001 Statistical mechanics of traffic flow *Physica* A **285** 101

[43] Aghababaie Y, Menon G I and Plischke M 1999 Universal properties of interacting Brownian motors *Phys. Rev.* E **59** 2578

[44] Chou T and Lohse D 1999 Entropy-driven pumping in zeolites and biological channels *Phys. Rev. Lett.* **82** 3552

[45] Klumpp S and Lipowsky R 2003 Traffic of molecular motors through tube-like components *J. Stat. Phys.* **113** 233

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

[47] Leduc C, Campàs O, Zeldovich K B, Roux A, Jolimaitre P, Bourel-Bonnet L, Goud B, Joanny J F, Bassereau P and Prost J 2004 Cooperative extraction of membrane nanotubes by molecular motors *Proc. Natl Acad. Sci.* **101** 17096

[48] Chowdhury D, Schadschneider A and Nishinari K 2005 Physics of transport and traffic phenomena in biology: from molecular motors and cells to organisms *Phys. Life Rev.* **2** 318

[49] Liggett T M 1985 Interacting particle systems *Grundlehren der Mathematischen Wissenschaften* vol 276 (New York: Springer)

[50] Liggett T M 1999 *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* (Berlin: Springer)

[51] Golinelli O and Mallick K 2005 Spectral gap of the totally asymmetric exclusion process at arbitrary filling *J. Phys. A: Math. Gen.* **38** 1419

[52] Golinelli O and Mallick K 2005 Spectral degeneracies in the totally asymmetric exclusion process *J. Stat. Phys.* **120** 779

[53] Helbing D 2001 Traffic and related self-driven many-particle systems *Rev. Mod. Phys.* **73** 1067–141

[54] Lighthill M J and Whitham G B 1955 On kinematic waves: II. A theory of traffic flow on long crowded roads *Proc. R. Soc.* A **229** 317

[55] Kolomeisky A B, Schütz G M, Kolomeisky E B and Straley J P 1998 Phase diagram of one-dimensional driven lattice gases with open boundaries *J. Phys. A: Math. Gen.* **31** 6911

[56] Popkov V and Schütz G 1999 Steady-state selection in driven diffusive systems with open boundaries *Europhys. Lett.* **48** 257

[57] Hager J S, Krug J, Popkov V and Schütz G M 2001 Minimal current phase and universal boundary layers in driven diffusive systems *Phys. Rev.* E **63** 056110

[58] Santen L and Appert C 2002 The asymmetric exclusion process revisited: fluctuations and dynamics in the domain wall picture *J. Stat. Phys.* **106** 187–99

[59] Plischke M and Bergersen B 2005 *Equilibrium Statistical Physics* 3rd edn (Singapore: World Scientific)

[60] Derrida B, Domany E and Mukamel D 1992 An exact solution of a one-dimensional asymmetric exclusion model with open boundaries *J. Stat. Phys.* **69** 667

[61] Krug J 1991 Boundary-induced phase transitions in driven diffusive systems *Phys. Rev. Lett.* **67** 1882

[62] Derrida B, Evans M R, Hakim V and Pasquier V 1993 Exact solution of a 1d asymmetric exclusion model using a matrix formulation *J. Phys. A: Math. Gen.* **26** 1493

[63] Schütz G and Domany E 1993 Phase transitions in an exactly soluable one-dimensional exclusion process *J. Stat. Phys.* **72** 277

[64] Derrida B and Evans M R 1993 Exact correlation functions in an asymmetric exclusion model with open boundaries *J. Physique* I **3** 311

[65] Blythe R A and Evans M R 2003 The Lee–Yang theory of equilibrium and nonequilibrium phase transitions *Braz. J. Phys.* **33** 464

[66] Comtet L 1974 *Advanced Combinatorics* (Boston: Reidel)

[67] Duchi E and Schaeffer G 2005 A combinatorial approach to jumping particles *J. Comb. Theory* A **110** 1–29

[68] Angel O 2006 The stationary measure of a 2-type totally asymmetric exclusion process *J. Comb. Theory* A **113** 625–35
[69] Brak R, Corteel S, Essam J, Parviainen R and Rechnitzer A 2006 A combinatorial derivation of the PASEP stationary state *Electron. J. Combin.* **13** R108
[70] Corteel S and Williams L 2006 Permutation tableaux and the asymmetric exclusion process *Preprint* math/06002109
[71] Gantmacher F R 1959 *Matrix Theory* vol II (Chelsea, London) chapter XIII
[72] Seneta E 1981 *Non-Negative Matrices and Markov Chains* (Berlin: Springer)
[73] Karlin S and Taylor H M 1975 *A First Course in Stochastic Processes* (New York: Academic)
[74] Evans M R 2000 Phase transitions in one-dimensional nonequilibrium systems *Braz. J. Phys.* **30** 42–57
[75] Klümper A, Schadschneider A and Zittartz J 1991 Equivalence and solution of anisotropic spin-1 models and generalized *t-J* fermion models in one dimension *J. Phys. A: Math. Gen.* **24** L955–9
[76] Fannes M, Nachtergaele B and Werner R F 1992 Finitely correlated states on quantum spin chains *Commun. Math. Phys.* **144** 443–90
[77] White S R 1998 Strongly correlated electron systems and the density matrix renormalization group *Phys. Rep.* **301** 187–204
[78] Carlon E, Henkel M and Schollwöck U 1999 Density matrix renormalization group and reaction–diffusion processes *Eur. Phys. J.* B **12** 99–114
[79] Schollwock U 2005 The density-matrix renormalization group *Rev. Mod. Phys.* **77** 259
[80] Verstraete F and Cirac J I 2006 Matrix product states represent ground states faithfully *Phys. Rev.* B **73** 094423
[81] Rajewsky N, Santen L, Schreckenberg M and Schadschneider A 1998 The asymmetric exclusion process: comparison of update procedures *J. Stat. Phys.* **92** 151–94
[82] Klümper A, Schadschneider A and Zittartz J 1993 Matrix product ground states for one-dimensional spin-1 quantum antiferromagnets *Europhys. Lett.* **24** 293–7
[83] Derrida B, Janowsky S A, Lebowitz J L and Speer E R 1993 Exact solution of the totally asymmetric simple exclusion process: shock profiles *J. Stat. Phys.* **73** 813
[84] Ferrari P A 1992 Shock fluctuations in asymmetric simple exclusion *Probab. Theor. Rel. Fields* **91** 81
[85] Andjel E D, Bramson M and Liggett T M 1988 Shocks in the asymmetric simple exclusion process *Probab. Theor. Rel. Fields* **78** 231
[86] Ferrari P A, Kipnis C and Saada E 1991 Microscopic structure of travelling waves for asymmetric simple exclusion process *Ann. Probab.* **19** 226
[87] Boldrighini C, Cosimi G, Frigio S and Nunes M G 1989 Computer simulation of shock waves in the completely asymmetric simple exclusion process *J. Stat. Phys.* **55** 611
[88] Spohn H 1991 *Large Scale Dynamics of Interacting Particles* (Berlin: Springer)
[89] Derrida B and Evans M R 1999 Bethe ansatz solution for a defect particle in the asymmetric exclusion model *J. Phys. A: Math. Gen.* **32** 4833
[90] van Beijeren H, Kutner R and Spohn H 1985 Excess noise for driven diffusive systems *Phys. Rev. Lett.* **54** 2026
[91] van Beijeren H 1991 Fluctuations in the motions of mass and of patterns in one-dimensional driven diffusive systems *J. Stat. Phys.* **63** 47
[92] Levine E, Mukamel D and Schütz G M 2005 Long-range attraction between probe particles mediated by a driven fluid *Europhys. Lett.* **70** 565
[93] Chatterjee S and Barma M 2007 Dynamics of shock probes in driven diffusive systems *J. Stat. Mech.: Theor. Exp.* L01004
[94] Lee H W, Popkov V and Kim D 1997 Two-way traffic flow: exactly solvable model of traffic jam *J. Phys. A: Math. Gen.* **30** 8497
[95] Mallick K 1996 Shocks in the asymmetry exclusion model with an impurity *J. Phys. A: Math. Gen.* **29** 5375
[96] Schütz G, Ramaswamy R and Barma M 1996 Pairwise balance and invariant measures for generalized exclusion processes *J. Phys. A: Math. Gen.* **29** 837
[97] Krebs K and Sandow S 1997 Matrix product eigenstates for one-dimensional stochastic models and quantum spin chains *J. Phys. A: Math. Gen.* **30** 3165
[98] Andrews G E, Askey R and Roy R 2000 *Special Functions (Encyclopaedia of Mathematics and its Applications* vol 71*)* (Cambridge: Cambridge University Press)
[99] Depken M 2003 Models of nonequilibrium systems *PhD Thesis* University of Oxford
[100] Wilf H S 1994 *Generatingfunctionology* (San Diego, CA: Academic)
[101] Stanley R P 2001 *Enumerative Combinatorics* vol II (Cambridge: Cambridge University Press)
[102] Brak R and Essam J W 2004 Asymmetric exclusion model and weighted lattice paths *J. Phys. A: Math. Gen.* **37** 4183

[103] Brak R, de Gier J and Rittenberg V 2004 Nonequilibrium stationary states and equilibrium models with long range interactions *J. Phys. A: Math. Gen.* **37** 4303

[104] Blythe R A, Johnston D A, Janke W and Kenna R 2004 The grand-canonical asymmetric exclusion process and the one-transit walk *J. Stat. Mech.: Theor. Exp.* P06001

[105] van Rensburg E J J 2000 *The Statistical Mechanics of Interacting Walks, Polygons, Animals and Vesicles* (Oxford: Oxford University Press)

[106] Blythe R A, Johnston D A, Janke W and Kenna R 2004 Dyck paths, Motzkin paths and traffic jams *J. Stat. Mech.: Theor. Exp.* P10007

[107] Krug J and Tang L H 1994 Disorder-induced unbinding in confined geometries *Phys. Rev.* E **50** 104–15

[108] Bena I, Droz M and Lipowski A 2005 Statistical mechanics of equilibrium and nonequilibrium phase transitions: The Yang–Lee formalism *Int. J. Mod. Phys.* B **19** 4269–329

[109] Depken M and Stinchcombe R B 2004 Exact joint density-current probability function for the asymmetric exclusion process *Phys. Rev. Lett.* **93** 040602

[110] Derrida B, Evans M R and Mukamel D 1993 Exact diffusion constant for the one-dimensional asymmetric exclusion models *J. Phys. A: Math. Gen.* **36** 4911–8

[111] Derrida B 1996 Phase transitions in nonequilibrium systems *Statphys-19 (19th IUPAP International Conference on Statistical Physics)* ed B L Hao (Singapore: World Scientific)

[112] Arndt P F, Heinzel T and Rittenberg V 1998 Spontaneous breaking of translational invariance in one-dimensional stationary states on a ring *J. Phys. A: Math. Gen.* **31** L45–51

[113] Arndt P F, Heinzel T and Rittenberg V 1999 Spontaneous breaking of translational invariance and spatial condensation in stationary states on a ring: I. The neutral system *J. Stat. Phys.* **97** 1–65

[114] Jafarpour F H 2000 Partially asymmetric simple exclusion model in the presence of an impurity on a ring *J. Phys. A: Math. Gen.* **33** 1797–808

[115] Jafarpour F H 2000 Exact solution of an exclusion model in the presence of a moving impurity on a ring *J. Phys. A: Math. Gen.* **33** 8673–80

[116] Sasamoto T 2000 One-dimensional partially asymmetric simple exclusion process on a ring with a defect particle *Phys. Rev.* E **61** 4980

[117] Rajewsky N, Sasamoto T and Speer E R 2000 Spatial particle condensation for an exclusion process on a ring *Physica* A **279** 123

[118] Meakin P, Ramanlal P, Sander L M and Ball R C 1986 Ballistic deposition on surfaces *Phys. Rev.* A **34** 5091

[119] Derrida B and Mallick K 1997 Exact diffusion constant for the one-dimensional partially asymmetric exclusion model *J. Phys. A: Math. Gen.* **30** 1031–46

[120] Blythe R A, Evans M R and Kafri Y 2000 Stochastic ballistic annihilation and coalescence *Phys. Rev. Lett.* **85** 3750–3

[121] Sasamoto T 1999 One-dimensional partially asymmetric simple exclusion process with open boundaries: orthogonal polynomials approach *J. Phys. A: Math. Gen.* **32** 7109

[122] Blythe R A, Evans M R, Colaiori F and Essler F H L 2000 Exact solution of a partially asymmetric exclusion model using a deformed oscillator algebra *J. Phys. A: Math. Gen.* **33** 2313

[123] Sasamoto T 2000 Density profile of the one-dimensional partially asymmetric simple exclusion process with open boundaries *J. Phys. Soc. Japan* **69** 1055

[124] Uchiyama M, Sasamoto T and Wadati M 2004 Asymmetric simple exclusion process with open boundaries and Askey–Wilson polynomials *J. Phys. A: Math. Gen.* **37** 4985

[125] Uchiyama M and Wadati M 2005 Correlation function of asymmetric simple exclusion process with open boundaries *J. Nonlinear Math. Phys.* **12** 676

[126] Gasper G and Rahman M 2004 Basic hypergeometric series *Encyclopæ dia of Mathematics and its Applications* 2nd edn (Cambridge: Cambridge University Press)

[127] Biedenhahn L C 1989 The quantum group $SU_q(2)$ and a q-analogue of the boson operators *J. Phys. A: Math. Gen.* **22** L873

[128] MacFarlane A J 1989 On q-analogues of the quantum harmonic oscillator and the quantum group $SU(2)_q$ *J. Phys. A: Math. Gen.* **22** 4581

[129] Sandow S 1994 Partially asymmetric exclusion process with open boundaries *Phys. Rev.* E **50** 2660

[130] Askey R 1989 Continuous q-Hermite polynomials when $q > 1$ *q-Series and Partitions* (New York: Springer)

[131] Askey R and Wilson J 1985 Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials *Mem. Am. Math. Soc.* **54** 1

[132] Derrida B, Lebowitz J L and Speer E R 2007 Entropy of open lattice systems *J. Stat. Phys.* **126** 1083–108

[133] Derrida B 2007 Non equilibrium steady states: fluctuations and large deviations of the density and of the current *J. Stat. Mech.: Theor. Exp.* P07023

[134] Derrida B, Lebowitz J L and Speer E R 2001 Free energy functional for nonequilibrium systems: an exactly solvable case *Phys. Rev. Lett.* **87** 150601
[135] Derrida B, Lebowitz J L and Speer E R 2002 Large deviation of the density profile in the steady state of the symmetric simple exclusion process *J. Stat. Phys.* **107** 599
[136] Mukamel D 2000 Phase transitions in nonequilibrium systems *Soft and Fragile Matter: Metastability and Flow* ed M E Cates and M R Evans (Bristol: Institute of Physics Publishing)
[137] Derrida B, Lebowitz J L and Speer E R 2002 Exact free energy functional for a driven diffusive open stationary nonequilibrium system *Phys. Rev. Lett.* **89** 30601
[138] Derrida B, Lebowitz J L and Speer E R 2003 Exact large deviation functional of a stationary open driven diffusive system: the asymmetric exclusion process *J. Stat. Phys.* **110** 775
[139] Enaud C and Derrida B 2004 Large deviation functional of the weakly asymmetric exclusion process *J. Stat. Phys.* **114** 537–62
[140] Derrida B, Enaud C and Lebowitz J L 2004 The asymmetric exclusion process and Brownian excursions *J. Stat. Phys.* **115** 365
[141] Isaev A P, Pyatov P N and Rittenberg V 2001 Diffusion algebras *J. Phys. A: Math. Gen.* **34** 5815–34
[142] Karimipour V 1999 Multispecies asymmetric simple exclusion process and its relation to traffic flow *Phys. Rev.* E **59** 205–12
[143] Bergman G M 1978 The diamond lemma for ring theory *Adv. Math.* **29** 178–218
[144] Khorrami M and Karimipour V 2000 Exact determination of the phase structure of a multi-species asymmetric exclusion process *J. Stat. Phys.* **100** 999–1030
[145] Evans M R 1996 Bose–Einstein condensation in disordered exclusion models and relation to traffic flow *Europhys. Lett.* **36** 13–8
[146] Evans M R, Kafri Y, Koduvely H M and Mukamel D 1998 Phase separation in one-dimensional driven diffusive systems *Phys. Rev. Lett.* **80** 425–9
[147] Evans M R, Kafri Y, Koduvely H M and Mukamel D 1998 Phase separation and coarsening in one-dimensional driven diffusive systems: local dynamics leading to long-range Hamiltonians *Phys. Rev.* E **58** 2764–78
[148] Clincy M, Derrida B and Evans M R 2003 Phase transition in the ABC model *Phys. Rev.* E **67** 066115
[149] Alcaraz F C, Dasmahapatra S and Rittenberg V 1998 *n*-species stochastic models with boundaries and quadratic algebras *J. Phys. A: Math. Gen.* **31** 845–78
[150] Evans M R, Foster D P, Godreche C and Mukamel D 1995 Spontaneous symmetry-breaking in a one-dimensional driven diffusive system *Phys. Rev. Lett.* **74** 208–11
[151] Evans M R, Foster D P, Godreche C and Mukamel D 1995 Asymmetric exclusion model with 2 species—spontaneous symmetry-breaking *J. Stat. Phys.* **80** 69–102
[152] Arita C 2006 Phase transitions in the two-species totally asymmetric exclusion process with open boundary conditions *J. Stat. Mech.: Theor. Exp.* P12008
[153] Kolomeisky A B 1997 Exact solutions for a partially asymmetric exclusion model with two species *Physica* A **245** 523
[154] Uchiyama M 2007 Two-species asymmetric simple exclusion process with open boundaries *Preprint* cond-mat/0703660
[155] Sandow S and Schütz G 1994 On $U(Q)[SU(2)]$-symmetrical driven diffusion *Europhys. Lett.* **26** 7–12
[156] Derrida B, Lebowitz J L and Speer E R 1997 Shock profiles for the asymmetric simple exclusion process in one dimension *J. Stat. Phys* **89** 135–67
[157] Speer E R 1997 Finite-dimensional representations of a shock algebra *J. Stat. Phys* **89** 169–75
[158] Evans M R, Levine E, Mohanty P K and Mukamel D 2004 Modelling one-dimensional driven diffusive systems by the zero-range process *Eur. Phys. J. B.* **41** 223–30
[159] Jafarpour F H and Khaki P 2007 A family of exactly-solvable driven-diffusive systems in one-dimension *Preprint* 0706.1826
[160] Evans M R, Kafri Y, Levine E and Mukamel D 2002 Phase transition in a non-conserving driven system *J. Phys. A: Math. Gen.* **35** L433–8
[161] Aneva B 2002 The noncommutative space of stochastic diffusion systems *J. Phys. A: Math. Gen.* **35** 859–77
[162] Aneva B 2003 Deformed coherent and squeezed states of multiparticle processes *Eur. Phys. J.* C **31** 403–14
[163] Pyatov P N and Twarock R 2002 Construction of diffusion algebras *J. Math. Phys.* **43** 3268–79
[164] Twarock R 2003 Quadratic algebras in traffic flow models *Rep. Math. Phys.* **51** 381–9
[165] Benjamini I, Ferrari P A and Landim C 1996 Asymmetric conservative processes with random rates *Stoch. Proc. Appl.* **61** 181–204
[166] Krug J and Ferrari P A 1996 Phase transitions in driven diffusive systems with random rates *J. Phys. A: Math. Gen.* **29** 35

[167] Juhàsz R, Santen L and Igloi F 2005 Partially asymmetric exclusion models with quenched disorder *Phys. Rev. Lett.* **94** 010601
[168] Karimipour V 1999 A multi-species asymmetric exclusion process, steady state and correlation functions on a periodic lattice *Europhys. Lett.* **47** 304–10
[169] Rajewsky N and Schreckenberg M 1997 Exact results for one-dimensional cellular automata with different types of updates *Physica* A **245** 139–44
[170] Derrida B, Evans M R and Mallick K 1995 Exact diffusion constant of a one-dimensional asymmetric exclusion model with open boundaries *J. Stat. Phys.* **79** 833–74
[171] Schütz G M 1996 Experimental realizations of integrable reaction–diffusion processes in biological and chemical systems *Proc. 7th Nankai Workshop on Symmetry, Statistical Mechanics Models, and Applications* ed F Y Wu and M L Ge (Singapore: World Scientific)
[172] Klauck K and Schadschneider A 1999 On the ubiquity of matrix-product states in one-dimensional stochastic processes with boundary interaction *Physica* A **271** 102–17
[173] Karimipour V 2000 Some general features of matrix product states in stochastic systems *J. Phys. A: Math. Gen.* **33** 709–19
[174] Krebs K 2000 On matrix product states for periodic boundary conditions *J. Phys. A: Math. Gen.* **33** L149
[175] Hinrichsen H, Sandow S and Peschel I 1996 On matrix product ground states for reaction–diffusion models *J. Phys. A: Math. Gen.* **29** 2643–9
[176] Jafarpour F H 2003 First-order phase transition in a reaction–diffusion model with open boundary: the Yang–Lee theory approach *J. Phys. A: Math. Gen.* **36** 7497
[177] Jafarpour F H 2004 Matrix product states of three families of one-dimensional interacting particle systems *Physica* A **339** 369–84
[178] Hieida Y and Sasamoto T 2004 Construction of a matrix product stationary state from solutions of a finite-size system *J. Phys. A: Math. Gen.* **37** 9873–89
[179] Mallick K, Mallick S and Rajewsky N 1999 Exact solution of an exclusion process with three classes of particles and vacancies *J. Phys. A: Math. Gen.* **32** 8399–410
[180] Ferrari P A and Martin J B 2007 Stationary distributions of multi-type totally asymmetric exclusion processes *Ann. Probab.* **35** 807–32
[181] Evans M R, Ferrari P A and Mallick K 2007 unpublished
[182] Schütz G 1993 Time-dependent correlation functions in a one-dimensional asymmetric exclusion process *Phys. Rev.* E **47** 4265–77
[183] Hinrichsen H 1996 Matrix product ground states for exclusion processes with parallel dynamics *J. Phys. A: Math. Gen.* **29** 3659–67
[184] Evans M R, Rajewsky N and Speer E R 1999 Exact solution of a cellular automaton for traffic *J. Stat. Phys* **95** 45–96
[185] Brankov J, Pesheva N and Valkov N 2000 Exact results for a fully asymmetric exclusion process with sequential dynamics and open boundaries *Phys. Rev.* E **61** 2300–18
[186] Brankov J and Pesheva N 2001 Exact density profiles for the fully asymmetric exclusion process with discrete-time dynamics on semi-infinite chains *Phys. Rev.* E **63** 046111
[187] de Gier J and Nienhuis B 1999 Exact stationary state for an asymmetric exclusion process with fully parallel dynamics *Phys. Rev.* E **59** 4899–911
[188] Tilstra L G and Ernst M H 1998 Synchronous asymmetric exclusion processes *J. Phys. A: Math. Gen.* **31** 5033–63
[189] Janowsky S A and Lebowitz J L 1992 Finite-size effects and shock fluctuations in the asymmetric simple-exclusion process *Phys. Rev.* A **45** 618–25
[190] Janowsky S A and Lebowitz J L 1994 Exact results for the asymmetric simple exclusion process with a blockage *J. Stat. Phys.* **77** 35–51
[191] Tripathy G and Barma M 1998 Driven lattice gases with quenched disorder: exact results and different macroscopic regimes *Phys. Rev.* E **58** 1911
[192] Kolwankar K M and Punnoose A 2000 Disordered totally asymmetric simple exclusion process: Exact results *Phys. Rev.* E **61** 2453
[193] Harris R J and Stinchcombe R B 2004 Disordered asymmetric simple exclusion process: mean-field treatment *Phys. Rev.* E **70** 016108
[194] Krug J 2000 Phase separation in disordered exclusion models *Braz. J. Phys.* **30** 97
[195] Enaud C and Derrida B 2004 Sample-dependent phase transitions in disordered exclusion models *Europhys. Lett.* **66** 83
[196] Shaw L B, Zia R K P and Lee K H 2003 Totally asymmetric exclusion process with extended objects: a model for protein synthesis *Phys. Rev.* E **68** 021910

[197] Lakatos G and Chou T 2003 Totally asymmetric exclusion processes with particles of arbitrary size *J. Phys. A: Math. Gen.* **36** 2027–41
[198] Shaw L B, Kolomeisky A B and Lee K H 2004 Local inhomogeneity in asymmetric simple exclusion processes with extended objects *J. Phys. A: Math. Gen.* **37** 2105–13
[199] Godreche C, Luck J M, Evans M R, Mukamel D, Sandow S and Speer E R 1995 Spontaneous symmetry breaking: exact results for a biased random walk model of an exclusion process *J. Phys. A: Math. Gen.* **28** 6039
[200] Popkov V and Schütz G M 2004 Why spontaneous symmetry breaking disappears in a bridge system with PDE-friendly boundaries *J. Stat. Mech.: Theor. Exp.* P12004
[201] Großkinsky S, Schütz G M and Willmann R D 2007 Rigorous results on spontaneous symmetry breaking in a one-dimensional driven particle system *J. Stat. Phys.* **128** 587
[202] Bertini L, Gabrielli D and Lebowitz J L 2005 Large deviation for a stochastic model of heat flow *J. Stat. Phys* **121** 843
[203] Katz S, Lebowitz J L and Spohn H 1983 Phase transitions in stationary nonequilibrium states of model lattice systems *Phys. Rev.* B **28** 1655
[204] Katz S, Lebowitz J L and Spohn H 1984 Nonequilibrium steady states of stochastic lattice gas models of fast ionic conductors *J Stat. Phys.* **34** 497–537
[205] Antal T and Schütz G M 2000 Asymmetric exclusion process with next-nearest-neighbor interaction: some comments on traffic flow and a nonequilibrium reentrance transition *Phys. Rev.* E **62** 83
[206] Wolki M, Schadschneider A and Schreckenberg M 2006 Asymmetric exclusion processes with shuffled dynamics *J. Phys. A: Math. Gen.* **39** 33–44
[207] Smith D A and Wilson R E 2007 Dynamical pair approximation for cellular automata with shuffle update *J. Phys. A: Math. Theor.* **40** 2651
[208] Debnath L 1997 *Nonlinear Partial Differential Equations for Scientists and Engineers* (Boston: Birkhäuser)
[209] Evans M R, Juhàsz R and Santen L 2003 Shock formation in an exclusion process with creation and annihilation *Phys. Rev.* E **68** 026117
[210] Popkov V, Rákos A, Willmann R, Kolomeisky A and Schütz G 2003 Localization of shocks in driven diffusive systems without particle number conservation *Phys. Rev.* E **67** 66117
[211] Mukherji S and Bhattacharjee S M 2005 Nonequilibrium criticality at shock formation in steady states *J. Phys. A: Math. Gen.* **38** L285
[212] Essler F H L and Rittenberg V 1996 Representations of the quadratic algebra and partially asymmetric diffusion with open boundaries *J. Phys. A: Math. Gen.* **29** 3375–407
[213] Mallick K and Sandow S 1997 Finite-dimensional representations of the quadratic algebra: applications to the exclusion process *J. Phys. A: Math. Gen.* **30** 4513