Continued fractions and the partially asymmetric exclusion process

2009 J. Phys. A: Math. Theor. 42 325002

(http://iopscience.iop.org/1751-8121/42/32/325002)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 325002 (21pp) doi:10.1088/1751-8113/42/32/325002

Continued fractions and the partially asymmetric exclusion process

R A Blythe[1], W Janke[2], D A Johnston[3] and R Kenna[4]

[1] SUPA, School of Physics and Astronomy, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, UK
[2] Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig, Postfach 100 920, 04009 Leipzig, Germany
[3] Department of Mathematics and the Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK
[4] Applied Mathematics Research Centre, Coventry University, Coventry, CV1 5FB, UK

Received 24 April 2009, in final form 22 June 2009
Published 27 July 2009
Online at stacks.iop.org/JPhysA/42/325002

Abstract
We note that a tridiagonal matrix representation of the algebra of the partially asymmetric exclusion process (PASEP) lends itself to interpretation as the transfer matrix for weighted Motzkin lattice paths. A continued-fraction ('J fraction') representation of the lattice-path-generating function is particularly well suited to discussing the PASEP, for which the paths have height-dependent weights. We show that this not only allows a succinct derivation of the normalization and correlation lengths of the PASEP, but also reveals how finite-dimensional representations of the PASEP algebra, valid only along special lines in the phase diagram, relate to the general solution that requires an infinite-dimensional representation.

PACS numbers: 64.60.De, 64.60.Ht, 64.70.qd

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

1. Introduction

Although the asymmetric exclusion process (ASEP)—a model in which hard-core particles hop in a preferred direction along a one-dimensional lattice—has been reinvented in various different guises over the years, it is only relatively recently that exact solutions for the steady state(s) of the model have been available. The solution of the ASEP with open boundary conditions in [1] using a matrix product ansatz was a landmark in the study of driven diffusive systems.

As discussed in a recent review of the matrix product approach to solving for the steady state of nonequilibrium Markov processes [2], there are a range of different methods for analysing the thermodynamic phase behaviour of the simplest versions of the ASEP. By

contrast, more general models—collectively known as the partially asymmetric exclusion process (PASEP)—that admit particles to hop in both directions in the bulk, and even more generally to enter and exit at both left and right boundaries, have so far been studied only through a diagonalization of the matrices appearing in the formalism [3–5]. In this work, we extend a technique that previously admitted an extremely quick derivation of the ASEP phase behaviour under various updating schemes [6, 7] to these more general models.

The idea is to consider the behaviour of a 'grand-canonical partition function' for the model. More precisely, we examine the generating function of the normalization of the nonequilibrium steady-state distribution over an ensemble of different lattice lengths whose mean is controlled by a fugacity. The thermodynamic phase behaviour can then be read off from the singularities of this generating function. Whilst obtaining this generating function is straightforward for the ASEP [6, 7], a convenient closed form for the PASEP has remained elusive.

Our aim here is to demonstrate that a representation of the generating function that allows the thermodynamic phase behaviour to be determined with relative ease takes the form of an infinite continued fraction. This we arrive at through an interpretation of the PASEP normalization as the (equilibrium) partition function of lattice paths, which we discuss in section 3 after recalling the model definition and its basic properties in the following section. In section 4, we show how to analyse the singularities embedded in the continued-fraction representation. The results we obtain are, of course, equivalent to those obtained within other approaches [3–5]. However, given that continued fractions are not frequently encountered in statistical mechanical contexts, we feel there is some value in using the PASEP as an illustrative example of how to handle them.

We find that the analysis is intimately related to an approach based on finite-dimensional matrix representations [8, 9], exact along special lines in the phase diagram, and that the continued fraction shows how these particular solutions and the general solution are related. We further show in section 5 that the continued-fraction approach extends to the most general version of the PASEP, solved in [5], and that one can access both currents and correlation lengths through it (section 6). In all of these cases, the steady state of the system is one in which the preferred hopping direction acts in the same direction as the chemical potential imposed by the boundaries. There are, however, two other possibilities: (i) where the chemical potential gradient *opposes* the bulk driving; and (ii) where the boundaries impose a gradient, but hopping in the bulk is unbiased. We consider these cases in sections 7 and 8, respectively. We find that the nonequilibrium phase transition identified in [4] has a rather appealing interpretation in terms of an equilibrium lattice path transition. Finally, in section 9, we observe that the continued-fraction analysis introduced in this work reveals a further special line in the phase diagram which, to our knowledge, has not previously been considered in the context of the PASEP.

2. Model definition and basic properties

The dynamics of the PASEP take place on a finite one-dimensional lattice with open boundaries. In its simplest form, the microscopic dynamics of the PASEP are specified by four rates, one of which can be set to unity by an overall scaling. For a rate λ associated with a particular event, the probability that the event happens in an infinitesimal time interval Δt is $\lambda \Delta t$. Moves that would lead to two particles occupying a single lattice site at any one time are prohibited due to the hard-core repulsion between them.

In the PASEP, particles are inserted onto the left boundary site (when empty) at a rate α and removed from the right boundary site at a rate β, see figure 1. Once on the lattice a

Figure 1. Typical particle configuration and allowed moves in the PASEP model.

particle hops by one site to the right at rate 1 or by one site to the left at a rate q when sites are available (i.e. empty). It is possible to expand this set of moves to allow particles to enter at the right at a rate δ and exit at the left at a rate γ, while still retaining the solvability of the model [5].

In all the models we consider, we want to calculate Z, which normalizes the statistical weight, $f(\mathcal{C})$, of a lattice configuration, $\mathcal{C}$, in the steady state. This is given by

$$Z = \sum_{\mathcal{C}} f(\mathcal{C}), \tag{2.1}$$

so the normalized probability of being in state $\mathcal{C}$ is $P(\mathcal{C}) = f(\mathcal{C})/Z$.

The weights themselves are obtained through the stationarity condition on the transition rates $W(\mathcal{C} \to \mathcal{C}')$,

$$\sum_{\mathcal{C}' \neq \mathcal{C}} [f(\mathcal{C}')W(\mathcal{C}' \to \mathcal{C}) - f(\mathcal{C})W(\mathcal{C} \to \mathcal{C}')] = 0, \tag{2.2}$$

where $W(\mathcal{C} \to \mathcal{C}')$ is the probability of making the transition from configuration $\mathcal{C}$ to $\mathcal{C}'$ in a single time step. This is less restrictive than the detailed balance condition for equilibrium states, which is obtained when the sum in equation (2.2) vanishes term by term.

The solution of the ASEP in [1] and the PASEP in [3–5] made use of a matrix product ansatz [2]. In this the steady-state probability $P(\mathcal{C})$ of a configuration of particles $\mathcal{C}$ on a chain of length N is represented by an ordered product of matrices $X_1 X_2 \cdots X_N$, where $X_i = D$ if site i is occupied and $X_i = E$ if it is empty. We expect $P(\mathcal{C})$ to be a function of both the number and position of particles on the lattice, which suggests the choice of non-commuting objects, matrices, for the ansatz. To obtain a scalar probability value from this matrix product, it is sandwiched between two vectors $\langle W|$ and $|V\rangle$:

$$P(\mathcal{C}) = \frac{\langle W|X_1 X_2 \cdots X_N|V\rangle}{Z_N}. \tag{2.3}$$

The factor Z_N is included to ensure that $P(\mathcal{C})$ is properly normalized. This latter quantity plays the role of a partition function in equilibrium problems,

$$Z_N = \langle W|(D+E)^N|V\rangle = \langle W|C^N|V\rangle, \tag{2.4}$$

where we have defined $C = D + E$. Indeed, we shall see in what follows that Z_N *is* the partition function for an equivalent two-dimensional lattice path problem.

The algebraic properties of the matrices D and E can be deduced from the master equation for the dynamics of the ASEP, PASEP and various other related models [2]. For the variant of the PASEP discussed above, sufficient conditions for equation (2.3) to hold are

$$DE - qED = D + E, \tag{2.5}$$

$$\alpha\langle W|E = \langle W|, \tag{2.6}$$

$$\beta D|V\rangle = |V\rangle. \tag{2.7}$$

These relations allow one to calculate Z_N and other quantities of physical interest by a range of methods, such as 'normal ordering' of the matrices, or through use of explicit representations [2].

In this work, we focus on an approach based around the generating function of Z_N, namely $\mathcal{Z}(z) = \sum_N Z_N z^N$, which can be thought of as a 'grand-canonical' normalization. As is well known [10], the large-N form of the 'canonical' normalization (Z_N) can be determined from the dominant singularity $z_{\rm cr}$ of $\mathcal{Z}(z)$. Typically, $Z_N \sim z_{\rm cr}^{-N} N^{-\nu}$, where the exponent $\nu \geqslant 0$ depends on the nature of the singularity. Then, by defining a function f in analogy to the free-energy density of an equilibrium system via

$$f = -\lim_{N\to\infty} \frac{1}{N} \ln Z_N, \tag{2.8}$$

we find $f = \ln z_{\rm cr}$. Nonanalyticities in f can then be associated with phase transitions in the physical system [2, 6, 7, 11].

For orientation, let us recall the results for the ASEP, which has $q = 0$. The canonical normalization can be shown by direct matrix reordering [1] to be

$$Z_N = \sum_{p=1}^{N} \frac{p(2N-1-p)!}{N!(N-p)!} \frac{(1/\beta)^{p+1} - (1/\alpha)^{p+1}}{(1/\beta) - (1/\alpha)}. \tag{2.9}$$

Performing the summation [6] gives the grand-canonical normalization

$$\mathcal{Z}(z) = \frac{\alpha\beta}{(\alpha - x(z))(\beta - x(z))}, \tag{2.10}$$

where $x(z) = (1 - \sqrt{1-4z})/2$.

This function has a pole at $x(z) = \alpha$ when $\alpha < 1/2$ and similarly at $x(z) = \beta$ when $\beta < 1/2$. These correspond to $z_{\rm cr} = \alpha(1-\alpha)$ and $z_{\rm cr} = \beta(1-\beta)$. When neither of these poles contribute, all that remains is the square-root singularity at $z_{\rm cr} = 1/4$. This allows one to very quickly establish the behaviour of f as a function of α and β:

$$f = \begin{cases} \ln\left[\frac{1}{4}\right] & \text{for} \quad \alpha, \beta > 1/2 \\ \ln\left[\alpha(1-\alpha)\right] & \text{for} \quad \beta > \alpha, \alpha < 1/2 \\ \ln\left[\beta(1-\beta)\right] & \text{for} \quad \alpha > \beta, \beta < 1/2. \end{cases} \tag{2.11}$$

It turns out that for the ASEP, $z_{\rm cr}$ corresponds to the particle current, and $x(z_{\rm cr})$ is the bulk density in the thermodynamic limit. Hence, the phase diagram for the model, figure 2, is quickly recovered using this generating-function (or grand-canonical) analysis. Further details of these methods as applied to the ASEP can be found in [2, 6, 7]. In the remainder of this work, we show how to elicit the structure of the grand-canonical normalization of the PASEP, where direct summation of the canonical normalization, given explicitly in [4], does not lead to a compact expression like (2.10).

3. Explicit matrix representation and lattice path interpretation

A useful route to the grand-canonical normalization for the PASEP is via the generating function for an ensemble of lattice paths, which in turn can be read off from an explicit representation of the matrices and vectors appearing in equations (2.5)–(2.7) that define the matrix algebra. A number of representations are known, see [1–3, 5]; the one that is of use here is that for which the vectors $\langle W|$ and $|V\rangle$ have nonzero entries only in their first element:

$$\langle W_q| = \langle W| = h_0^{1/2}(1,0,0,\ldots) \qquad |V_q\rangle = |V\rangle = h_0^{1/2}(1,0,0,\ldots)^T, \tag{3.1}$$

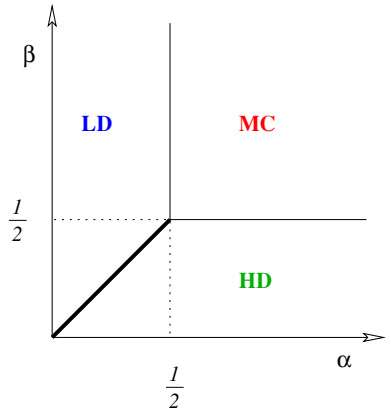

Figure 2. The phase diagram of the ASEP. Here HD, LD and MC denote the high-density, low-density and maximal-current phases, respectively.

where h_0 is a constant to be given shortly. One can verify that, with this choice of boundary vectors, the following tridiagonal representations of D and E satisfy (2.5)–(2.7):

$$D_q = \frac{1}{1-q}\begin{pmatrix} 1+\tilde{\beta} & \sqrt{c_1} & 0 & \cdots \\ 0 & 1+\tilde{\beta}q & \sqrt{c_2} & \\ 0 & 0 & 1+\tilde{\beta}q^2 & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix},$$
$$E_q = \frac{1}{1-q}\begin{pmatrix} 1+\tilde{\alpha} & 0 & 0 & \cdots \\ \sqrt{c_1} & 1+\tilde{\alpha}q & 0 & \\ 0 & \sqrt{c_2} & 1+\tilde{\alpha}q^2 & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}. \tag{3.2}$$

The various parameters that appear are

$$\tilde{\alpha} = \frac{1-q}{\alpha} - 1, \tag{3.3}$$

$$\tilde{\beta} = \frac{1-q}{\beta} - 1, \tag{3.4}$$

$$c_n = (1-q^n)(1-\tilde{\alpha}\tilde{\beta}q^{n-1}), \tag{3.5}$$

$$h_0 = \frac{1}{(\tilde{\alpha}\tilde{\beta};q)_\infty} = \sum_{n=0}^{\infty} \frac{(\tilde{\alpha}\tilde{\beta})^n}{(q;q)_n} = \langle W|V\rangle, \tag{3.6}$$

in which we have used the standard notation for (shifted) q factorials

$$\begin{aligned} (a;q)_n &= \prod_{j=0}^{n-1}(1-aq^j), \\ (a;q)_0 &= 1, \\ (a,b,\ldots,c;q)_n &= (a;q)_n(b;q)_n\cdots(c;q)_n. \end{aligned} \tag{3.7}$$

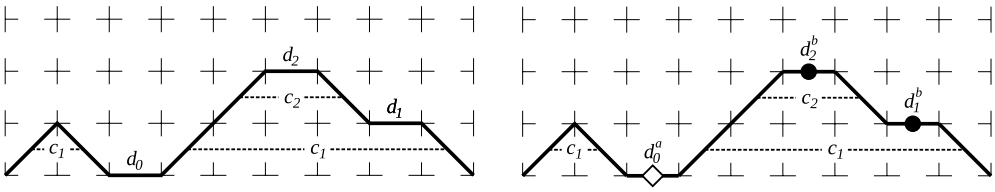

Figure 3. Two Motzkin path transliterations of the tridiagonal matrix representation of the PASEP matrices. In both cases, up-step down-step pairs contribute a weight c_n, where n is the height above the origin of the upper end of the steps. Left: there is a single type of horizontal step that contributes a weight d_n, where n is the height of the segment above the origin. Right: an equivalent interpretation, in which there are two types of horizontal steps, shown with diamonds and circles at the their midpoints. These contribute weights d_n^a and d_n^b, respectively.

There are various ways to arrive at an interpretation in terms of lattice paths from a matrix representation. One was suggested by Brak and Essam [12, 13], who used the fact that $D + E = DE$ for the ASEP to interpret the D and E as odd–even and even–odd height transfer matrices separately. In later works [6, 7], the path interpretation was inferred from the grand-canonical normalization once this had been obtained by another means. Here, we shall take the most direct approach, which is to associate a height $n \geqslant 0$ above the origin with the vector $|n\rangle = (00\cdots010\cdots)^T$ (i.e. n is the number of zero entries that appear before the single nonzero entry). We then interpret $\langle m|X|n\rangle$, where X is some combination of D and E matrices, as the weight of paths connecting a point at height n to another point at height m.

Of particular importance is the matrix $X = C^N$, which appears in equation (2.4) for the normalization. From the above expressions we have that

$$C_q = D_q + E_q = \frac{1}{1-q}\begin{pmatrix} 2+\tilde{\alpha}+\tilde{\beta} & \sqrt{c_1} & 0 & \cdots \\ \sqrt{c_1} & 2+(\tilde{\alpha}+\tilde{\beta})q & \sqrt{c_2} & \\ 0 & \sqrt{c_2} & 2+(\tilde{\alpha}+\tilde{\beta})q^2 & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}. \tag{3.8}$$

The matrix element $\langle m|C^N|n\rangle$ then gives the combined weight of paths that begin at height m, end at height n and contain N steps, each of which may raise or lower the height by one unit, or leave the height unchanged. Since the height is a nonnegative quantity, $n \geqslant 0$, these paths may never descend beneath the origin. Additionally, since $\langle W| \propto \langle 0|$ and $|V\rangle \propto |0\rangle$, the paths that contribute to Z_N begin and end at the origin. Paths with these properties are known as *Motzkin paths*.

The weight of various path components can now be obtained by inspecting the form of C_q. For the path to begin and end at the origin, every up-step must be accompanied by a down-step; each up-step–down-step pair connecting height $n-1$ to height n contributes a weight c_n to the path. Each horizontal step at height n contributes a weight $d_n = 2+(\tilde{\alpha}+\tilde{\beta})q^n$ to the path. An alternative interpretation has two types (or 'colours') of horizontal path segments, one of which contributes a weight $d_n^a = 1+\tilde{\alpha}q^n$ and the other $d_n^b = 1+\tilde{\beta}q^n$; see figure 3. To arrive at the canonical normalization Z_N, we sum over all the paths of length N, and multiply by the factor $h_0/(1-q)^N$. The grand-canonical normalization is constructed by summing paths of all lengths, weighting each segment by z and finally multiplying by h_0. In the following section, we shall see an equivalent recursive construction which can be expressed as a continued fraction.

As a check of the path representation, one can compare the expression for Z_N calculated from enumerating the weighted Motzkin paths with that which emerges from directly evaluating the matrix product expression for the stationary state. This was done in [4] by using the matrix product approach to write down an integral representation of the normalization. By recasting trigonometric functions appearing in the integrand in terms of a set of 'q-deformed' polynomials of the harmonic oscillator algebra, one could then exploit their orthogonality properties to convert the integral representation into a sum that contains a finite number of terms when $N < \infty$. The resulting expression is not particularly simple, but we quote it here for completeness:

$$Z_N = \langle W|V\rangle \left(\frac{1}{1-q}\right)^N \sum_{n=0}^{N} R_{N,n}(q) B_n(\tilde{\alpha}, \tilde{\beta}; q), \tag{3.9}$$

where

$$R_{N,n}(q) = \sum_{k=0}^{\lfloor \frac{N-n}{2} \rfloor} (-1)^k \left[\binom{2N}{N-n-2k} - \binom{2N}{N-n-2k-2} \right] q^{\binom{k+1}{2}} \begin{bmatrix} n+k \\ k \end{bmatrix}, \tag{3.10}$$

and

$$B_n(\tilde{\alpha}, \tilde{\beta}; q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} \tilde{\alpha}^{n-k} \tilde{\beta}^k. \tag{3.11}$$

In the above, we have used the standard notation for the q-binomial coefficient

$$\begin{bmatrix} n \\ k \end{bmatrix} = \frac{(q;q)_n}{(q;q)_{n-k}(q;q)_k}. \tag{3.12}$$

For instance, taking the simple case of Z_2 a direct calculation using the above formulae gives

$$Z_2 = \frac{5 - q + \tilde{\alpha}^2 + \tilde{\alpha}\tilde{\beta} + \tilde{\beta}^2 + q\tilde{\alpha}\tilde{\beta} + 4\tilde{\alpha} + 4\tilde{\beta}}{(1-q)^2}, \tag{3.13}$$

where we have dropped the overall normalization $\langle W|V\rangle$. The $(1-q)^2$ denominator disappears when this is written in terms of the original rate parameters α and β,

$$Z_2 = \frac{\alpha\beta q + \alpha^2 + \beta^2\alpha + \alpha\beta + \alpha^2\beta + \beta^2}{\alpha^2\beta^2}. \tag{3.14}$$

The expression for Z_2 in (3.13) can be seen to be the sum of weights for an up–down step pair from level 0 to 1, given by $(1-q)(1-\tilde{\alpha}\tilde{\beta})/(1-q)^2$ and all four possible combinations of two horizontal steps at the level zero of either 'colour', given by $(2+\tilde{\alpha}+\tilde{\beta})(2+\tilde{\alpha}+\tilde{\beta})/(1-q)^2$ (where we have again dropped the overall normalization $\langle W|V\rangle$). While the diagrammatics becomes increasingly complicated for larger Z_N, the principle remains the same.

4. The continued-fraction representation of the path-generating function

Despite the availability of an exact expression for the canonical normalization, equation (3.9) from [4], we have not been able to find a convenient expression for its grand-canonical counterpart due to the q-dependence of the PASEP weights. A more fruitful route is to represent the lattice-path-generating function as a continued fraction, a procedure first expounded by Flajolet [14]. As we now show, this representation can be read off more-or-less directly from the matrix representation (3.8), and the form that emerges is particularly well adapted to the discussion of the PASEP.

First, for convenience, we subsume the prefactor $1/(1-q)$ appearing in (3.8) into the parameters

$$\begin{aligned}\tilde{d}_n &= \frac{2+(\tilde{\alpha}+\tilde{\beta})q^n}{1-q},\\ \tilde{c}_n &= \frac{(1-q^n)(1-\tilde{\alpha}\tilde{\beta}q^{n-1})}{(1-q)^2},\end{aligned} \tag{4.1}$$

so that then

$$C_q = \begin{pmatrix} \tilde{d}_0 & \sqrt{\tilde{c}_1} & 0 & \cdots \\ \sqrt{\tilde{c}_1} & \tilde{d}_1 & \sqrt{\tilde{c}_2} & \\ 0 & \sqrt{\tilde{c}_2} & \tilde{d}_2 & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}. \tag{4.2}$$

In the lattice-path language, this means that horizontal steps at height n are weighted by $\tilde{d}_n$ and up–down step pairs between heights n and $n+1$ are weighted by $\tilde{c}_n$.

Let now $\mathcal{M}_n(z)$ be the generating function of the weighted Motzkin paths that start and end at height n, never go below this height and have their lengths $N \geqslant 0$ counted by powers of z. That is, the coefficient of z^N in $\mathcal{M}_n(z)$ is the weight of such paths of length N. The grand-canonical normalization for the PASEP is then given by $\mathcal{Z}(z) = h_0\mathcal{M}_0(z)$. Let us suppose that $\mathcal{M}_{n+1}(z)$ is known for some $n \geqslant 0$. Then, we can construct $\mathcal{M}_n(z)$ by concatenating contiguous components of two types: (i) sequences of horizontal segments of arbitrary (possibly zero) length at height n; and (ii) Motzkin paths starting at height $n+1$ enclosed by an up–down pair. Denoting these components schematically as $\_\_$ and $/\mathcal{M}_{n+1}\backslash$ respectively, we can write the recursion

$$\mathcal{M}_n = \_\_ + \_\_/\mathcal{M}_{n+1}\backslash\_\_ + \_\_/\mathcal{M}_{n+1}\backslash\_\_/\mathcal{M}_{n+1}\backslash\_\_ + \cdots \tag{4.3}$$

$$= \_\_(1 + [/\mathcal{M}_{n+1}\backslash\_\_] + [/\mathcal{M}_{n+1}\backslash\_\_]^2 + \cdots) \tag{4.4}$$

$$= \frac{\_\_}{1 - /\mathcal{M}_{n+1}\backslash\_\_}. \tag{4.5}$$

The generating function for a (possibly empty) sequence of horizontal segments, each weighted by $\tilde{d}_n$, is simply $(1-\tilde{d}_n z)^{-1}$. An up–down pair $/\cdots\backslash$ from height n to $n+1$ contributes the weight $\tilde{c}_{n+1}z^2$. We thus arrive at the generating-function recursion

$$\mathcal{M}_n(z) = \frac{(1-\tilde{d}_n z)^{-1}}{1-\tilde{c}_n z^2 \mathcal{M}_{n+1}(z)(1-\tilde{d}_n z)^{-1}} = \frac{1}{1-\tilde{d}_n z - \tilde{c}_{n+1}z^2\mathcal{M}_{n+1}(z)}. \tag{4.6}$$

Starting at $n=0$ and iterating, we find that the generating function $\mathcal{Z}(\tilde{\alpha},\tilde{\beta},q,z)$ for the Motzkin paths of arbitrary length (and hence the grand-canonical PASEP normalization) is given by the infinite continued fraction

$$\mathcal{Z}(\tilde{\alpha},\tilde{\beta},q,z) = \cfrac{1}{1-\tilde{d}_0 z - \cfrac{\tilde{c}_1 z^2}{1-\tilde{d}_1 z - \cfrac{\tilde{c}_2 z^2}{1-\tilde{d}_2 z - \cfrac{\tilde{c}_3 z^2}{\cdots}}}}, \tag{4.7}$$

where we have dropped the factor h_0 since this does not contribute to any physical quantities. Such a continued fraction containing both z and z^2 terms is usually denoted a *Jacobi continued fraction*, or 'J fraction' for short [15].

Before considering the case of general q let us take $q = 0$ and see how the expression for the grand-canonical normalization, equation (2.10), is recovered. We have

$$\mathcal{Z}(\tilde{\alpha}, \tilde{\beta}, 0, z) = \cfrac{1}{1 - \tilde{d}_0 z - \cfrac{\tilde{c}_1 z^2}{1 - 2z - \cfrac{z^2}{1 - 2z - \cfrac{z^2}{1 - 2z - \cfrac{z^2}{\cdots}}}}}, \tag{4.8}$$

where $\tilde{d}_0 = 1/\alpha + 1/\beta$ and $\tilde{c}_1 = 1/\alpha + 1/\beta - 1/(\alpha\beta) = \kappa^2$ when $q = 0$. Note that in this case, the continued fraction is periodic after the first level. That is, in the above notation

$$\mathcal{M}_n(z) = \frac{1}{1 - 2z - z^2 \mathcal{M}_{n+1}(z)} \qquad \forall n \geqslant 1. \tag{4.9}$$

Hence, we must have that $\mathcal{M}_1(z) = \mathcal{M}_2(z) = \cdots$ and hence

$$\mathcal{M}_1(z)[1 - 2z - z^2 \mathcal{M}_1(z)] = 1 \tag{4.10}$$

or

$$\mathcal{M}_1(z) = \frac{1 - 2z - \sqrt{1 - 4z}}{2z^2}, \tag{4.11}$$

which is the generating function familiar from many Catalan counting problems. For $n = 0$, we have

$$\mathcal{Z}(\tilde{\alpha}, \tilde{\beta}, 0, z) = \mathcal{M}_0(z) = \frac{1}{1 - \tilde{d}_0 z - \tilde{c}_1 z^2 \mathcal{M}_1(z)}, \tag{4.12}$$

which coincides with (2.10) when both are expanded and rationalized.

At $q = 0$, the luxury of being able to sum the continued fraction to get equation (4.12) makes the phase structure, which (as previously discussed) is determined by the singularities of $\mathcal{Z}(\tilde{\alpha}, \tilde{\beta}, 0, z)$ in z, immediately apparent. If we are *not* able to easily sum explicitly the continued fraction, as is the case for the PASEP, we can use more indirect methods to determine the singularities. We focus here on the 'forward-bias' regime, $q < 1$, in which the continued fraction has a finite radius of convergence in the complex-z plane. In the reverse-bias regime, $q > 1$, the continued fraction is unconditionally divergent, a fact we will interpret in section 7.

We first appeal to Worpitzky's theorem on the convergence of continued fractions [15, 16] which states that a continued fraction of the form

$$\cfrac{1}{1 + \cfrac{a_2}{1 + \cfrac{a_3}{1 + \cfrac{a_4}{\cdots}}}}$$

converges if the partial numerators a_p satisfy

$$|a_p| < 1/4, \qquad p = 2, 3, 4, \ldots. \tag{4.13}$$

For $\mathcal{Z}(z)$ given by (4.7) this translates to a radius of convergence z_{cr} given by

$$\frac{4\tilde{c}_n z_{\mathrm{cr}}^2}{(1 - \tilde{d}_{n-1} z_{\mathrm{cr}})(1 - \tilde{d}_n z_{\mathrm{cr}})} = 1 \ \forall n \geqslant 1 \tag{4.14}$$

and shows that

$$z_{\mathrm{cr}} \to (1 - q)/4 \tag{4.15}$$

as $n \to \infty$. To decide if this is the dominant singularity, one must also divine the location of any poles in the complex-z plane from the continued fraction (4.7).

The strategy is to examine the nth *convergent* of the continued fraction, that is, the expression obtained by truncating the continued fraction at the nth level (counting from zero). Denoting this as K_n, we have that

$$K_0 = \frac{1}{1 - \tilde{d}_0 z}, \tag{4.16}$$

$$K_1 = \frac{1}{1 - \tilde{d}_0 z - \dfrac{\tilde{c}_1 z^2}{1 - \tilde{d}_1 z}}, \tag{4.17}$$

and so on. We observe that the continued fraction (4.8) is given *exactly* by the convergent K_n if $\tilde{c}_{n+1} = 0$. Inspection of (4.1) reveals that this occurs on the special line in the phase diagram given by

$$\tilde{\alpha}\tilde{\beta} = q^{-n}. \tag{4.18}$$

We observe that on such special lines, the matrix C_q decomposes into two blocks: the first is $(n + 1)$ dimensional, whilst the second does not contribute to the normalization because only the first elements of the boundary vectors are nonzero. Finite-dimensional representations of the matrix algebra for the PASEP were used prior to the advent of the full solution to study the model along these lines, and to conjecture the phase behaviour elsewhere [8, 9].

To locate the poles of K_n in the complex-z plane, we use the fact [17] that the reciprocal of each convergent K_n can be expressed as

$$(K_n)^{-1} = \frac{A_n}{B_n}, \tag{4.19}$$

where A_n and B_n both satisfy the same three-term recurrence

$$A_n(z) = (1 - \tilde{d}_n z)A_{n-1}(z) - \tilde{c}_n z^2 A_{n-2}(z), \tag{4.20}$$

$$B_n(z) = (1 - \tilde{d}_n z)B_{n-1}(z) - \tilde{c}_n z^2 B_{n-2}(z), \tag{4.21}$$

but with slightly different initial conditions: $A_{-2} = B_{-1} = 0$ and $A_{-1} = B_0 = 1$. Hence, K_n has poles that coincide with the zeros of $A_n(z)$. These relations for $A_n(z)$ and $B_n(z)$ hold for arbitrary choices of $\tilde{\alpha}$ and $\tilde{\beta}$. Let us now define the function $\mathcal{A}_n(z)$ that is given by $A_n(z)$ for the particular choice $\tilde{\beta} = 1/(q^n\tilde{\alpha})$. That is, the zeros of $\mathcal{A}_n(z)$ give the poles of the PASEP grand-canonical generating function along the line in the phase diagram along which the $(n + 1)$-dimensional matrix representation is exact.

As a consequence of the recurrence (4.20) the $A_n(z)$ are closely related to the al-Salam–Chihara polynomials [18] and may be written explicitly as

$$A_n(z) = \frac{(\tilde{\alpha}\tilde{\beta}; q)_{n+1} z^{n+1}}{(1 - q)^{n+1}\tilde{\alpha}^{n+1}} \sum_{k=0}^{n+1} \frac{(q^{-(n+1)}; q)_k (\tilde{\alpha}\, \mathrm{e}^{\mathrm{i}\theta}; q)_k (\tilde{\alpha}\, \mathrm{e}^{-\mathrm{i}\theta}; q)_k}{(\tilde{\alpha}\tilde{\beta}; q)_k (q; q)_k} q^k, \tag{4.22}$$

where $\cos(\theta) = (1 - q)/(2z) - 1$. If we take $\tilde{\beta} = 1/(q^n\tilde{\alpha})$ to obtain $\mathcal{A}_n(z)$ all the terms in the sum in (4.22) will be zero apart from the $(n + 1)$ th, since they contain a $(1 - \tilde{\alpha}\tilde{\beta}q^n)$ factor in the numerator, which is cancelled only in the last term. Using

$$(\tilde{\alpha}\, \mathrm{e}^{\mathrm{i}\theta}; q)_{n+1} (\tilde{\alpha}\, \mathrm{e}^{-\mathrm{i}\theta}; q)_{n+1} = \prod_{j=0}^{n} (1 - 2\tilde{\alpha}q^j \cos(\theta) + \tilde{\alpha}^2 q^{2j}), \tag{4.23}$$

we thus find

$$\mathcal{A}_n(z) = \prod_{k=0}^{n} \left(1 - \frac{z}{z_k}\right), \tag{4.24}$$

where

$$z_k = \frac{1-q}{(1+\tilde{\alpha}q^k)\left(1+\frac{1}{\tilde{\alpha}q^k}\right)}. \tag{4.25}$$

Equally, we could have taken $\tilde{\alpha} = 1/(q^n\tilde{\beta})$ and obtained the same result with $\tilde{\alpha} \leftrightarrow \tilde{\beta}$. These roots correspond to the reciprocal of the eigenvalues of the finite-dimensional matrices C_q examined in [8, 9] (note, however, that the representation used in those works differs from that used here, but is related by a similarity transformation [3]). This fact can be understood by making the ansatz $A_n(z) = z^n\left|C_q^{(n)} - \frac{1}{z}\right|$, where $C_q^{(n)}$ is the submatrix formed by the first n rows and columns of C_q. After substituting into (4.20), one obtains the expression that results from a co-factor expansion along the bottom row in the determinant $\left|C_q^n - \lambda\right|$, as long as λ is identified with $\frac{1}{z}$.

We see, then, that the sequence of generating functions $\mathcal{A}_n(z)$ has a truly remarkable property: all poles of $\mathcal{A}_n(z)$ are poles of $\mathcal{A}_m(z)$, where $m > n$. In principle, a completely different set of singularities could have been obtained after truncating at a deeper level. Furthermore, when $q < 1$, $|z_n| < |z_m|$ if $n < m$, and hence all convergents have the same dominant pole, z_0. Writing this in terms of the original rates α and β gives

$$z_{\mathrm{cr}} = \frac{1-q}{2+\tilde{\alpha}+1/\tilde{\alpha}} = \frac{\alpha(1-q-\alpha)}{1-q} \tag{4.26}$$

or

$$z_{\mathrm{cr}} = \frac{1-q}{2+\tilde{\beta}+1/\tilde{\beta}} = \frac{\beta(1-q-\beta)}{1-q}. \tag{4.27}$$

Taking now the number of levels (or, equivalently, the dimensionality of the matrix representation) to infinity, we find that the dominant singularity will always be the smallest of (4.15), (4.26) or (4.27). This observation allows us to recover f for the PASEP previously calculated in [3, 4], and identify the regions of the phase diagram within which they apply

$$\begin{aligned}
f &= \ln\left[\frac{1-q}{4}\right], && \text{for} \quad \alpha, \beta > (1-q)/2, \\
f &= \ln\left[\frac{\alpha(1-q-\alpha)}{1-q}\right], && \text{for} \quad \beta > \alpha, \alpha < (1-q)/2, \\
f &= \ln\left[\frac{\beta(1-q-\beta)}{1-q}\right], && \text{for} \quad \alpha > \beta, \beta < (1-q)/2.
\end{aligned} \tag{4.28}$$

As for the ASEP, this phase structure comprises high- and low-density phases and a maximal current phase. The key differences lie in the q-dependent position of the second-order transition line, $\alpha = \beta = (1-q)/2$, and in the subextensive corrections to the free energy, which manifest themselves in the density profile at finite distances from the boundaries [20].

As noted above, this analysis of the PASEP phase behaviour is similar to that based on finite-dimensional matrix representations [8, 9]. The main additional benefit of the present approach is that the relationship between the general expression for the normalization, given through the non-terminating continued fraction (4.7), and the versions that apply only on special lines is clearer. For example, the singularity that arises from non-convergence of the continued fraction, and which governs the maximal-current behaviour, is not detectable by inspection of the finite-dimensional matrices. Furthermore, some light is shed on the

'continuity arguments' expounded in [8, 9] that lead one to extend physical properties found along special lines to regions of the phase diagram, in that each additional level of the continued fraction yields a singularity that is subdominant to all of those present at the previous level. We remark that the fact that the asymptotically dominant singularity is present even in the one-dimensional representation is suggestive of an explanation as to why the phase diagram and currents obtained from mean-field theory coincide with the exact results.

5. Lattice paths for non-zero 'wrong direction' removal and injection rates γ and δ

As noted in the introduction it is possible to generalize the PASEP boundary conditions to allow 'wrong direction' injection and removal rates. In addition to injection at a rate α on the left boundary particles may also be removed there at a rate γ. Similarly, in addition to removal at a rate β on the right boundary particles may also be injected there at a rate δ.

A tridiagonal representation of D and E with these generalized boundary conditions still exists for the case of nonzero γ and δ [5] and which can be written as

$$\hat{D}_q = \frac{1}{1-q}\begin{pmatrix} 1+d_0^{\natural} & d_0^{\sharp} & 0 & \cdots \\ d_0^{\flat} & 1+d_1^{\natural} & d_1^{\sharp} & \\ 0 & d_1^{\flat} & 1+d_2^{\natural} & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix},$$
$$\hat{E}_q = \frac{1}{1-q}\begin{pmatrix} 1+e_0^{\natural} & e_0^{\sharp} & 0 & \cdots \\ e_0^{\flat} & 1+e_1^{\natural} & e_1^{\sharp} & \\ 0 & e_1^{\flat} & 1+e_2^{\natural} & \ddots \\ \vdots & & \ddots & \ddots \end{pmatrix}, \tag{5.1}$$

$$\langle \tilde{W}_q| = h_0^{1/2}(1,0,0,\ldots), \qquad |\tilde{V}_q\rangle = h_0^{1/2}(1,0,0,\ldots)^T, \tag{5.2}$$

where the expressions for $d_n^{\sharp}$, $d_n^{\natural}$, $e_n^{\natural}$, $e_n^{\flat}$ (a notation borrowed from [5]) and h_0 are given in the appendix.

We can define $\hat{C}_q = \hat{D}_q + \hat{E}_q$ in an analogous manner to the earlier discussion which leads to a generating function of the form

$$\mathcal{Z}(a,b,c,d,q,z) = \cfrac{1}{1-\hat{d}_0 z - \cfrac{\hat{c}_1 z^2}{1-\hat{d}_1 z - \cfrac{\hat{c}_2 z^2}{1-\hat{d}_2 z - \cfrac{\hat{c}_3 z^2}{\cdots}}}}, \tag{5.3}$$

where the coefficients are now

$$\hat{d}_n = \frac{2+d_n^{\natural}+e_n^{\natural}}{1-q}, \qquad \hat{c}_{n+1} = \frac{\left(d_n^{\sharp}+e_n^{\sharp}\right)\left(d_n^{\flat}+e_n^{\flat}\right)}{(1-q)^2}. \tag{5.4}$$

Even with these more complicated coefficients, the singularity structure of the generating function is identical to the $\gamma = \delta = 0$ case when $q < 1$. We can still extract poles in order to discern the high- and low-density phases. Explicitly,

$$\hat{c}_{n+1} = \frac{(1-q^{n-1}abcd)(1-q^{n+1})(1-q^n ab)(1-q^n ac)(1-q^n ad)(1-q^n bc)(1-q^n bd)(1-q^n cd)}{(1-q^{2n-1}abcd)(1-q^{2n}abcd)^2(1-q^{2n+1}abcd)(1-q)^2} \tag{5.5}$$

(in which the parameters a, b, c and d are also given in the appendix) so that we can pick up the various poles depending on the zeros of the numerator. For nonzero γ and δ Askey–Wilson polynomials [19] play the role of the al-Salam–Chihara polynomials so in this case the denominators of the convergents are given by

$$A_n(z) = \frac{(ab, ac, ad; q)_{n+1} z^{n+1}}{(1-q)^{n+1} a^{n+1}} \sum_{k=0}^{n+1} \frac{(q^{-(n+1)}; q)_k (q^n abcd)(a\,\mathrm{e}^{\mathrm{i}\theta}; q)_k (a\,\mathrm{e}^{-\mathrm{i}\theta}; q)_k}{(ab; q)_k (ac; q)_k (ad; q)_k (q; q)_k} q^k, \tag{5.6}$$

where $\cos(\theta) = (1-q)/(2z) - 1$ as before. If we now choose $q^n ab = 1$, for example, to enforce $\hat{c}_{n+1} = 0$, we find that only the final term in the sum contributes again, resulting in the dominant poles

$$z_{\mathrm{cr}} = \frac{1-q}{2+a+1/a} \tag{5.7}$$

or

$$z_{\mathrm{cr}} = \frac{1-q}{2+b+1/b} \tag{5.8}$$

depending on the relative size of a and b. The maximal current phase is still observed in a similar manner to the $\gamma = \delta = 0$ case via Worpitzky's theorem, but a major difference is that the reverse bias phase is absent when $q > 1$, since the particles can now escape at both ends. Algebraically, this is manifested in the reflection symmetry relating the parameters for $q < 1$ and $q > 1$ when γ and δ are nonzero:

$$a, b, c, d, q \to b^{-1}, a^{-1}, d^{-1}, c^{-1}, q^{-1}. \tag{5.9}$$

We remark that this continued-fraction representation provides a very quick route to the identification of the dominant singularities in the generating function, and therewith the extensive part of the free energy.

6. Correlation lengths

We have so far restricted our discussion to the normalization itself, but it is also possible to extract the correlation lengths, which determine the various subphases within the PASEP phase diagram, in a rather straightforward manner. We have seen that the leading singularities of the grand-canonical normalization $\mathcal{Z}$ determine the phase structure of the PASEP and that these can be extracted from the continued-fraction representation either directly as poles when the fraction terminates (the high- and low-density phases), or using Worpitzky's theorem (the maximal current phase).

For a chain of length N, the one- and two-point density correlation functions for the PASEP may be obtained in terms of the matrices D_q and C_q as

$$\begin{aligned} \langle \tau_i \rangle &= \frac{1}{Z_N} \langle W | C_q^{i-1} D_q C_q^{N-i} | V \rangle, \\ \langle \tau_i \tau_j \rangle &= \frac{1}{Z_N} \langle W | C_q^{i-1} D_q C_q^{j-i-1} D_q C_q^{N-j} | V \rangle. \end{aligned} \tag{6.1}$$

We have already evaluated the thermodynamic limit $N \to \infty$ of Z_N by picking up the leading singularity in the partial fraction representation. If we switch back to the transfer matrix

picture of C_q, the leading singularity of the generating function $z_{\rm cr}$ is just the inverse of the largest eigenvalue of C_q, $\lambda_{\rm cr} = 1/z_{\rm cr}$, which dominates the normalization as

$$\begin{aligned}\langle W|C_q^N|V\rangle &= \sum_\lambda \langle W|C_q^N|\lambda\rangle\langle\lambda|V\rangle \\ &= \sum_\lambda \langle W|\lambda\rangle\lambda^N\langle\lambda|V\rangle \\ &\simeq \text{const}\ \lambda_{\rm cr}^N. \end{aligned} \tag{6.2}$$

For the two-point function, there is a similar expansion [21]

$$\begin{aligned}\langle W|C_q^{i-1}D_qC_q^{j-i-1}D_qC_q^{N-j}|V\rangle &\simeq \lambda_{\rm cr}^{N-2}\langle W|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|V\rangle \\ &+ \lambda_{\rm cr}^{N-2}\left(\frac{\lambda_{\rm sub}}{\lambda_{\rm cr}}\right)^{i-1}\langle W|\lambda_{\rm sub}\rangle\langle\lambda_{\rm sub}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|V\rangle \\ &+ \lambda_{\rm cr}^{N-2}\left(\frac{\lambda_{\rm sub}}{\lambda_{\rm cr}}\right)^{j-i-1}\langle W|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm sub}\rangle\langle\lambda_{\rm sub}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|V\rangle \\ &+ \lambda_{\rm cr}^{N-2}\left(\frac{\lambda_{\rm sub}}{\lambda_{\rm cr}}\right)^{N-j}\langle W|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm cr}\rangle\langle\lambda_{\rm cr}|D_q|\lambda_{\rm sub}\rangle\langle\lambda_{\rm sub}|V\rangle, \end{aligned} \tag{6.3}$$

where the second largest eigenvalue $\lambda_{\rm sub}$ is the inverse of the subdominant singularity $z_{\rm sub}$ in the generating function $\mathcal{Z}$. From this, it is clear that the ratio $\lambda_{\rm sub}/\lambda_{\rm cr}$, or alternatively $z_{\rm cr}/z_{\rm sub}$, determines the correlation length. We have assumed that it is possible to define a second largest eigenvalue in the above, which will be the case in the high- and low-density phases. The spectrum, however, in the maximal current phase is continuous as can be seen from the nature of the respective singularities in the continued fraction, where the high- and low-density phases are associated with the poles, and the maximal current singularity emerges from Worpitzky's theorem.

If we consider first the high- and low-density phases the subdominant singularity is separated from the dominant one and both can be read off directly from the continued-fraction representation of $\mathcal{Z}$. In this case, we descend two levels in the continued fraction in order to extract both the dominant and subdominant singularities, which means that we impose $\tilde{c}_2 = 0$, i.e. $\tilde{\beta} = 1/(\tilde{\alpha}q)$ or $\tilde{\alpha} = 1/(\tilde{\beta}q)$ depending on the values of the parameters. The generating function then truncates to

$$\begin{aligned}\mathcal{Z}_2(\tilde{\alpha},\tilde{\beta},q,z) &= \frac{1}{1-\tilde{d}_0z-\dfrac{\tilde{c}_1z^2}{1-\tilde{d}_1z}} \\ &= \frac{1-\tilde{d}_1z}{1-(\tilde{d}_0+\tilde{d}_1)z+(\tilde{d}_0\tilde{d}_1-\tilde{c}_1)z^2}, \end{aligned} \tag{6.4}$$

which allows us to extract both the dominant singularity $z_{\rm cr}$ and subdominant singularity $z_{\rm sub}$ for various ranges of the parameters $\tilde{\alpha}, \tilde{\beta}, q$. As we have already noted in extracting the normalization from the continued fraction, descending to deeper levels still picks up $z_{\rm cr}$ and $z_{\rm sub}$ as the dominant and subdominant poles at every deeper level, in addition to the further subleading singularities.

This in turn allows us to classify various subphases within the high- and low-density phases which have different behaviours of the correlation length. For example, the low-density phase may be divided into the three subphases found in [21]:

(A_1) $\tilde{\alpha}q > 1, \tilde{\alpha}q > \tilde{\beta}$:

$$z_{\rm cr} = \frac{1-q}{(1+\tilde{\alpha})(1+\tilde{\alpha}^{-1})}, \qquad z_{\rm sub} = \frac{1-q}{(1+\tilde{\alpha}q)(1+(\tilde{\alpha}q)^{-1})}; \tag{6.5}$$

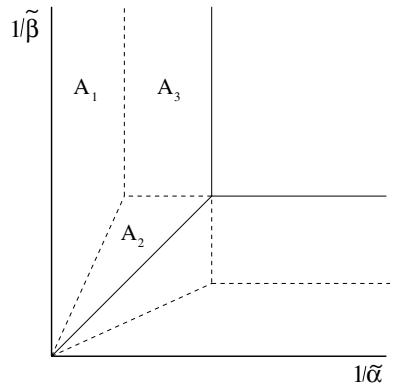

Figure 4. The various subphases $A_1, A_2, \ldots$ of the PASEP defined by the behaviour of the correlation length are delineated.

(A_2) $\tilde{\alpha} > \tilde{\beta} > \tilde{\alpha}q, \tilde{\beta} > 1$:

$$z_{\rm cr} = \frac{1-q}{(1+\tilde{\alpha})(1+\tilde{\alpha}^{-1})}, \qquad z_{\rm sub} = \frac{1-q}{(1+\tilde{\beta})(1+\tilde{\beta}^{-1})}; \tag{6.6}$$

(A_3) $\tilde{\alpha} > 1 > \tilde{\alpha}q, \tilde{\beta} < 1$:

$$z_{\rm cr} = \frac{1-q}{(1+\tilde{\alpha})(1+\tilde{\alpha}^{-1})}, \qquad z_{\rm sub} = \frac{1-q}{4}. \tag{6.7}$$

The $z_{\rm sub}$ values which play a role in the A_1 and A_2 subphases are the (subdominant) poles of the truncated $\mathcal{Z}_2(\tilde{\alpha}, \tilde{\beta}, q, z)$, whereas in the A_3 subphase it is the singularity determined using Worpitzky's theorem for the full, untruncated $\mathcal{Z}$. The high-density phase splits in a similar fashion with $\tilde{\alpha}$ and $\tilde{\beta}$ interchanged in the various expressions for $z_{\rm cr}$ and $z_{\rm sub}$ above. In all these subphases, the correlation functions decay exponentially and the correlation length is given by the log of the ratio of the poles $\xi = (\ln(z_{\rm sub}/z_{\rm cr}))^{-1}$. The phase structure deduced in this manner by considering the singularities of $\mathcal{Z}$ agrees, as it should, with that obtained from considering the integral representation of Z_N in [21] and is shown in figure 4.

It is also clear from the continued-fraction analysis that descending to deeper levels will pick up further subdominant poles which can endow the higher correlation functions with a yet finer phase structure inside $A_1, A_2, \ldots$ as the parameters are varied and the relative positions of the poles change, a possibility already remarked in [3].

7. (Motzkin) path transitions

In our earlier work on Dyck path representations of the ASEP normalization, it was found that the high- and low-density phases of the ASEP were reflected in bound phases for the lattice paths, whereas the maximal current phase corresponded to an unbound phase. The second-order transition lines of the ASEP corresponded to unbinding transition lines for one of the two lattice paths which described the ASEP normalization, and the first-order transition between the high- and low-density phases of the ASEP corresponded to a cooperative transition involving both paths.

The situation for the $q < 1$ phase structure of the PASEP is analogous. In the region of the phase diagram corresponding to the high- and low-density phases of the PASEP predominantly, one type of horizontal step is bound to the x-axis, reflecting the majority of particles or holes

in the high- and low-density phases of the PASEP. In the maximal current phase, the entropy dominates and paths are unbound. For instance, in the low-density A phases, indicated in figure 4, the horizontal on-axis steps weighted by $1+\tilde{\alpha}$ will make a greater contribution to the partition function than those weighted by $1+\tilde{\beta}$ and so will dominate the sum. The reverse is true in the high-density phase where the steps weighted by $1+\tilde{\beta}$ dominate.

The non-trivial terms in the weights of steps at height n are proportional to q^n so these steps do not contribute to the energy as $N \to \infty$ when $q < 1$, but only entropically. It is therefore not surprising that for $q < 1$ the PASEP phase diagram is essentially identical to the ASEP, with only the position of the transition lines changing as q is varied. It is interesting to inquire about the nature of the paths contributing to the partition function for $q > 1$. The expression for Z_N in this reverse bias phase was calculated in [4],

$$Z_N \sim A(\tilde{\alpha}, \tilde{\beta}; q)(q^{-1}\tilde{\alpha}\tilde{\beta}, 1/\tilde{\alpha}\tilde{\beta}; q^{-1})_\infty \left(\frac{\sqrt{\tilde{\alpha}\tilde{\beta}}}{q-1}\right)^N q^{\frac{1}{4}N^2}, \tag{7.1}$$

where

$$A(\tilde{\alpha}, \tilde{\beta}; q) = \sqrt{\frac{\pi}{\ln q}} \exp\left[M(q) + \frac{(\ln \tilde{\beta}/\tilde{\alpha})^2}{4 \ln q}\right], \tag{7.2}$$

and

$$M(q) \simeq -\sum_{k=1}^{\infty} \frac{1}{k} \frac{1}{q^k - 1}. \tag{7.3}$$

The most interesting feature of (7.1) is that one now sees an area-like behaviour emerging, whereby $Z_N \sim q^{\frac{1}{4}N^2}$. This contrasts earlier naive expectations that the normalization should grow exponentially with the system size, cf equations (2.8) and (6.2). Implicit in these analyses is the assumption that the grand-canonical normalization has a finite radius of convergence around the origin in the complex-z plane. It turns out that when $q > 1$, the series is divergent.

The mechanism for the appearance of the $q^{\frac{1}{4}N^2}$ factor is similar to the case of a model of square-lattice polygons with both edge and area fugacities [22], where the generating function may be written as

$$G(x, q) = \sum_{M,N} c_{M,N} x^N q^M, \tag{7.4}$$

with $c_{M,N}$ being the number of polygons with area M and perimeter $2N$. We can sum over the area to get a one-parameter generating function

$$Q_N(q) = \sum_{M} c_{M,N} q^M \tag{7.5}$$

and define an associated critical coupling

$$x_c(q) = \lim_{N\to\infty} (Q_N(q))^{-1/N} \tag{7.6}$$

from which we can observe that

$$\begin{aligned} &Q_N(1) \sim \mu_{\text{saw}}^{2N}, \\ &Q_N(q > 1) \sim q^{N^2/4}, \\ &x_c(1) \sim \mu_{\text{saw}}^{-2}, \\ &x_c(q > 1) = 0, \end{aligned} \tag{7.7}$$

where μ_{saw} is the connectivity constant for the $2D$ self-avoiding walk on a square lattice. We thus see that the square-lattice polygon model undergoes a transition from 'one-dimensional'

behaviour $Q_N(1) \sim \mu_{\rm saw}^{2N}$ to 'two-dimensional' behaviour $Q_N(q > 1) \sim q^{N^2/4}$. For $q > 1$, the dominant configurations are inflated, convex polygons.

To see how such behaviour emerges in our model of one-dimensional Motzkin paths when $q > 1$ consider a tent-shaped path contributing to Z_N with $N/2$ upward steps followed by $N/2$ downward steps. This has a weight

$$\frac{(\tilde{\alpha}\tilde{\beta};q)_{N/2}(q;q)_{N/2}}{(q-1)^N}, \tag{7.8}$$

which has a leading term of the form

$$\frac{(\tilde{\alpha}\tilde{\beta}-1)(\tilde{\alpha}\tilde{\beta})^{\frac{N}{2}-1}q^{\frac{N^2}{4}}}{(q-1)^N}. \tag{7.9}$$

The reverse bias expression for Z_N in (7.1) displays exactly this behaviour. With a slight rewriting of the second factor,

$$Z_N \sim A(\tilde{\alpha},\tilde{\beta};q)(q^{-1}\tilde{\alpha}\tilde{\beta},1/q\tilde{\alpha}\tilde{\beta};q^{-1})_\infty \left(\frac{\tilde{\alpha}\tilde{\beta}-1}{\tilde{\alpha}\tilde{\beta}}\right)\left(\frac{\sqrt{\tilde{\alpha}\tilde{\beta}}}{q-1}\right)^N q^{\frac{1}{4}N^2}, \tag{7.10}$$

which matches the behaviour of (7.9) coming from a 'tent' path. Heuristically, this makes sense since paths which gain in altitude as quickly as possible will donate the highest powers of q to Z_N and dominate when $q > 1$.

This change in behaviour at $q > 1$ to an area-like scaling is thus interpreted in the lattice path picture as an inflation transition, where the fluctuating path becomes the boundary of an inflated vesicle (with particular edge weights favouring a speedy ascent) pinned at its ends to the horizontal axis. The matrices D_q and E_q initially denoted the presence or absence of particles on a particular site in the original matrix product ansatz. We can associate an up-step and one colour of horizontal step in a path with, say, the presence of a particle in the original PASEP steady state. Looking at the form of the leading contribution to Z_N when $q > 1$, namely $N/2$ up-steps followed by $N/2$ down-steps, we can see that this corresponds to a half-filled lattice, which has also been observed in simulations of the reverse bias phase and heuristically derived using particle/hole symmetry arguments [4].

8. The case $q = 1$

The expression for Z_N may also be evaluated in the case of the symmetric simple exclusion process (SSEP) when $q = 1$. If we introduce some more standard notation, this time for q-numbers,

$$[n]_q = \frac{1-q^n}{1-q} = 1 + q + \cdots q^{n-1}, \tag{8.1}$$

(with $[n]_q = 0$ for $n \leqslant 0$) we can rewrite the Motzkin path weights as

$$\begin{aligned} \tilde{d}_n &= 2[n]_q + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)q^n, \\ \tilde{c}_n &= [n]_q\left([n-1]_q + \left(\frac{1}{\alpha}+\frac{1}{\beta}\right)q^{n-1} - \frac{(1-q)q^{n-1}}{\alpha\beta}\right), \end{aligned} \tag{8.2}$$

or, splitting the horizontal weights into two types of steps again in order to make the correspondence with bicoloured Motzkin paths explicit,

$$\tilde{d}_{n,1} = [n]_q + \frac{q^n}{\alpha}, \qquad \tilde{d}_{n,2} = [n]_q + \frac{q^n}{\beta}. \tag{8.3}$$

The q-integers become normal integers when $q = 1$, so the weights at $q = 1$ are given by

$$\begin{aligned}\tilde{d}_{n,1} &= n + \frac{1}{\alpha},\\ \tilde{d}_{n,2} &= n + \frac{1}{\beta},\\ \tilde{c}_n &= n\left(n - 1 + \frac{1}{\alpha} + \frac{1}{\beta}\right),\end{aligned} \tag{8.4}$$

which simplify even further when $\alpha = \beta = 1$ to $n + 1$ and $n(n + 1)$. This gives

$$\mathcal{Z}(1, 1, 1, z) = \cfrac{1}{1 - 2z - \cfrac{2z^2}{1 - 4z - \cfrac{2 \times 3z^2}{1 - 6z - \cfrac{3 \times 4z^2}{\cdots}}}}, \tag{8.5}$$

which can be seen to be a continued-fraction expansion for the divergent power series

$$\mathcal{Z} = \sum_{N=0}^{\infty} (N + 1)! z^N, \tag{8.6}$$

so $Z_N = (N + 1)!$ in this case. A similar expansion exists for general α, and β and gives Z_N as the ratio of two Gamma functions

$$Z_N = \frac{\Gamma(\lambda + N + 1)}{\Gamma(\lambda + 1)}, \tag{8.7}$$

where $\lambda = \frac{1}{\alpha} + \frac{1}{\beta} - 1$. This agrees with a direct calculation in [5]. Using Stirling's approximation for the Gamma functions, and remembering that λ is of $O(1)$ we see that $\ln Z_N \sim N \ln N$ so $Z_N \sim \mathrm{e}^{N \ln N}$ at $q = 1$, intermediate between the linear behaviour for $q < 1$ and the area law behaviour when $q > 1$.

9. The case $q^{n+1} = 1$

The continued-fraction representation of $\mathcal{Z}$ will also terminate when $q^{n+1} = 1$. These values of q have recently proved of interest for the Bethe ansatz approach to the open XXZ spin chain with non-diagonal boundary conditions [23–25]. The Hamiltonian of the XXZ spin chain is equivalent (up to a unitary transformation) to the transfer matrix of the PASEP. If we denote the expression for $A_n(z)$ when $q^{n+1} = 1$ by $\tilde{\mathcal{A}}_n(z)$ the sum in (4.22) truncates to two terms

$$\tilde{\mathcal{A}}_n(z) = \frac{z^{n+1}}{(1 - q)^{n+1}\tilde{\alpha}^{n+1}}((\tilde{\alpha}\tilde{\beta}; q)_{n+1} - (\tilde{\alpha}\,\mathrm{e}^{\mathrm{i}\theta}; q)_{n+1}(\tilde{\alpha}\,\mathrm{e}^{-\mathrm{i}\theta}; q)_{n+1}), \tag{9.1}$$

where we have used a limiting procedure

$$\lim_{\epsilon \to 0} \frac{(q^{-n+1}; q)_{n+1}}{(q; q)_{n+1}} = \lim_{\epsilon \to 0} (-1)^{n+1} q^{-(n+1)(n+2)/2} = -1 \tag{9.2}$$

with $q = \exp(\epsilon + 2\pi i m/(n + 1))$ to handle the indeterminate final term [26]. Noting that $(\tilde{\alpha}\,\mathrm{e}^{\mathrm{i}\theta}, q)_{n+1} = (1 - \tilde{\alpha}^{n+1}\,\mathrm{e}^{\mathrm{i}(n+1)\theta})$ when $q^{n+1} = 1$, this can be written as

$$\tilde{\mathcal{A}}_n(z) = \frac{z^{n+1}}{(1 - q)^{n+1}}(\mathrm{e}^{\mathrm{i}(n+1)\theta} + \mathrm{e}^{-\mathrm{i}(n+1)\theta} - (\tilde{\alpha}^{n+1} + \tilde{\beta}^{n+1})), \tag{9.3}$$

and we find that the zeros of $\tilde{\mathcal{A}}_n(z)$ are given by

$$z_k = \frac{1-q}{(1+rq^k)\left(1+\frac{1}{rq^k}\right)}, \tag{9.4}$$

where r is the root with a smaller argument of

$$r^{n+1} = \frac{(\tilde{\alpha}^{n+1}+\tilde{\beta}^{n+1})}{2} + \sqrt{\frac{(\tilde{\alpha}^{n+1}+\tilde{\beta}^{n+1})^2}{4}-1} \tag{9.5}$$

or

$$r^{-(n+1)} = \frac{(\tilde{\alpha}^{n+1}+\tilde{\beta}^{n+1})}{2} - \sqrt{\frac{(\tilde{\alpha}^{n+1}+\tilde{\beta}^{n+1})^2}{4}-1}. \tag{9.6}$$

It is interesting that the overall structure of the roots remains similar to that observed when $\tilde{\alpha} = 1/(q^n\tilde{\beta})$, but one no longer has a dominant pole appearing at first order and remaining.

10. Conclusions

We have shown that the matrix $C_q = D_q + E_q$ in a particular tridiagonal representation of the PASEP algebra can be interpreted as the transfer matrix for weighted Motzkin paths, with two colours of horizontal steps. Writing the generating function for these paths as a continued fraction allowed a succinct derivation of the thermodynamic limit of the normalization of the PASEP in its various phases. In particular, it allowed calculations without explicitly summing the generating function in a closed form, which was possible (or at any rate, easy to perform) only in the $q = 0$ case. Consideration of the subleading singularities in the continued fraction also allowed the determination of the correlation length in the high- and low-density phases.

A further interesting feature of the continued-fraction representation of the generating function was that it made clear how the finite-dimensional representations of the PASEP algebra, valid only along special lines in the phase diagram, related to the general, infinite-dimensional solution via truncation. The unusual structure of the poles of the convergents of the continued fraction along these special lines was highlighted, with lower order poles remaining present as the order of the convergent increased and the dominant pole (for the high- and low-density phases) already being present at zeroth order. The presence of the maximal current phase, on the other hand, was deduced from the continued-fraction representation by using Worpitzky's theorem.

The phase transitions of the PASEP can be identified with transitions in the lattice path model. When $q < 1$, there are similar unbinding transitions to those seen in the Dyck path model for the ASEP, with bound states corresponding to the high- and low-density phases of the PASEP and an unbound phase corresponding to the maximal current phase. For the reverse bias phase of the PASEP when $q > 1$, the lattice path model was seen to be in an inflated phase, and the leading contribution to the lattice path partition function was identified with a half-filled state in the PASEP.

Interestingly, bicoloured Motzkin paths play a major role in various combinatorial bijections, and the ASEP and PASEP have recently been investigated for their combinatorial interest [27–31]. One common theme has been the appearance of various well-known combinatorial weights as special cases for the (P)ASEP normalization which can be related to the known results in the enumeration of permutations and other combinatorial objects. The approach in [29], in particular, is rather similar to that espoused here, namely going directly to a matrix representation for D and E and seeking a combinatorial interpretation (although the principal focus in [29] was on a different representation and permutation tableaux).

It would be an interesting exercise to relate the phase transitions discussed here for the Motzkin paths and the related phase structure of the PASEP to conformational transitions in other (weighted) combinatorial objects such as parallelogram polyominoes, binary trees and the permutation tableau of [29]. It would also be worth relating the lattice path picture to the discussions of the large deviation functional of the (P)ASEP, such as that in [32], where the appearance of a combination of a Brownian excursion and a Brownian walk in the discussion is strongly suggestive of a polyomino bijection from the Motzkin paths.

Acknowledgments

This work was partially supported by the EU RTN-Network 'ENRAGE': *Random Geometry and Random Matrices: From Quantum Gravity to Econophysics* under grant no. MRTN-CT-2004-005616. RAB thanks the RCUK for the support of an Academic Fellowship.

Appendix. Parameters appearing in the tridiagonal matrix representation for $\gamma, \delta \neq 0$

When all four boundary rates, α, β, γ and δ are nonzero, the tridiagonal D and E matrices (5.1) contain a number of parameters. These are

$$d_n^{\natural} = \frac{q^{n-1}}{(1-q^{2n-2}abcd)(1-q^{2n}abcd)} \times [bd(a+c)+(b+d)q - abcd(b+d)q^{n-1} - \{bd(a+c)+abcd(b+d)\}q^n - bd(a+c)q^{n+1} + ab^2cd^2(a+c)q^{2n-1} + abcd(b+d)q^{2n}],$$

$$e_n^{\natural} = \frac{q^{n-1}}{(1-q^{2n-2}abcd)(1-q^{2n}abcd)} \times [ac(b+d)+(a+c)q - abcd(a+c)q^{n-1} - \{ac(b+d)+abcd(a+c)\}q^n - ac(b+d)q^{n+1} + a^2bc^2d(b+d)q^{2n-1} + abcd(a+c)q^{2n}],$$

$$d_n^{\sharp} = \frac{1}{1-q^n ac}\mathcal{A}_n, \qquad e_n^{\sharp} = -\frac{q^n ac}{1-q^n ac}\mathcal{A}_n,$$

$$d_n^{\flat} = -\frac{q^n bd}{1-q^n bd}\mathcal{A}_n, \qquad e_n^{\flat} = \frac{1}{1-q^n bd}\mathcal{A}_n,$$

which involve the further parameters

$$a = \frac{1}{2\alpha}[(1-q-\alpha+\gamma)+\sqrt{(1-q-\alpha+\gamma)^2+4\alpha\gamma}],$$

$$b = \frac{1}{2\beta}[(1-q-\beta+\delta)+\sqrt{(1-q-\beta+\delta)^2+4\beta\delta}],$$

$$c = \frac{1}{2\alpha}[(1-q-\alpha+\gamma)-\sqrt{(1-q-\alpha+\gamma)^2+4\alpha\gamma}],$$

$$d = \frac{1}{2\beta}[(1-q-\beta+\delta)-\sqrt{(1-q-\beta+\delta)^2+4\beta\delta}],$$

$$\mathcal{A}_n = \left[\frac{(1-q^{n-1}abcd)(1-q^{n+1})(1-q^n ab)(1-q^n ac)(1-q^n ad)(1-q^n bc)(1-q^n bd)(1-q^n cd)}{(1-q^{2n-1}abcd)(1-q^{2n}abcd)^2(1-q^{2n+1}abcd)}\right]^{1/2}.$$

Finally, the constant appearing in (5.2) is

$$h_0 = \frac{(abcd;q)_\infty}{(q,ab,ac,ad,bc,bd,cd;q)_\infty}.$$

References

[1] Derrida B, Evans M R, Hakim V and Pasquier V 1993 *J. Phys. A: Math. Gen.* **26** 1493
[2] Blythe R A and Evans M R 2004 *J. Phys. A: Math. Gen.* **40** R333
[3] Sasamoto T 1999 *J. Phys. A: Math. Gen.* **32** 7109
[4] Blythe R A, Evans M R, Colaiori F and Essler F H L 2000 *J. Phys. A: Math. Gen.* **33** 2313
[5] Uchiyama M, Sasamoto T and Wadati M 2004 *J. Phys. A: Math. Gen.* **37** 4985
[6] Blythe R A, Janke W, Johnston D A and Kenna R 2004 *J. Stat. Mech.* P06001
[7] Blythe R A, Janke W, Johnston D A and Kenna R 2004 *J. Stat. Mech.* P10007
[8] Mallick K and Sandow S 1997 *J. Phys. A: Math. Gen.* **30** 4513
[9] Jafarpour F H 2000 *J. Phys. A: Math. Gen.* **33** 1797
[10] Wilf H S 2006 *Generatingfunctionology* 3rd edn (Wellesley, MA: Peters)
[11] Blythe R A and Evans M R 2002 *Phys. Rev. Lett.* **89** 080601
[12] Brak R and Essam J 2004 *J. Phys. A: Math. Gen.* **37** 4183
[13] Brak R, de Gier J and Rittenberg V 2004 *J. Phys. A: Math. Gen.* **37** 4303
[14] Flajolet P 1982 *Discrete Math.* **32** 125
[15] Wall H S 1948 *Analytic Theory of Continued Fractions* (Princeton, NJ: van Nostrand-Reinhold)
[16] Janse van Rensburg E J 2000 *The Statistical Mechanics of Interacting Walks, Polygons, Animals and Vesicles* (Oxford: Oxford University Press)
[17] Andews G E, Askey R and Roy R 2000 *Special Functions* (Cambridge: Cambridge University Press)
[18] Al-Salam W A and Chihara T S 1976 *SIAM J. Math. Anal.* **7** 16
[19] Askey R A and Wilson J A 1985 *Mem. Am. Math. Soc.* **54** 319
[20] Sasamoto T 2000 *J. Phys. Soc. Japan* **69** 1055
[21] Uchiyama M and Wadati M 2005 *J. Nonlinear Math. Phys.* **12** 676
[22] Prellberg T and Owczarek A L 1999 *Commun. Math. Phys.* **201** 493
[23] Murgan R, Nepomechie R I and Shi C 2006 *J. Stat. Mech.* P08006
[24] de Gier J and Essler F H L 2005 *Phys. Rev. Lett.* **95** 240601
[25] de Gier J and Essler F H L 2008 *J. Phys. A: Math. Theor.* **41** 485002
[26] Spridinov V and Zhedanov A 1997 *Duke Math. J.* **89** 283
[27] Brak R, Corteel S, Essam J, Parviainen R and Rechnitzer A 2006 *Elect. J. Combinatorics* **13** R108
[28] Corteel S and Williams L K 2007 *Int. Math. Res. Notes* RNM055
[29] Corteel S and Williams L K 2007 *Adv. App. Maths.* **39** 293
[30] Duchi E and Schaeffer G 2005 *J. Comb. Theory Ser.* A **110** 1
[31] Duchi E and Schaeffer G 2008 *Random Struct. Alg.* **33** 434
[32] Derrida B, Enaud C and Lebowitz J L 2004 *J. Stat. Phys.* **115** 365