Shift Equivalence of Measures and the Intrinsic Structure of Shocks in the Asymmetric Simple Exclusion Process

B. Derrida[1], S. Goldstein[2], J. L. Lebowitz[2,3], and E. R. Speer[2]

Abstract

We investigate properties of non-translation-invariant measures, describing particle systems on $\mathbb{Z}$, which are asymptotic to different translation invariant measures on the left and on the right. Often the structure of the transition region can only be observed from a point of view which is random—in particular, configuration dependent. Two such measures will be called *shift equivalent* if they differ only by the choice of such a viewpoint. We introduce certain quantities, called *translation sums*, which, under some auxiliary conditions, characterize the equivalence classes. Our prime example is the asymmetric simple exclusion process, for which the measures in question describe the microscopic structure of shocks. In this case we compute explicitly the translation sums and find that shocks generated in different ways—in particular, via initial conditions in an infinite system or by boundary conditions in a finite system—are described by shift equivalent measures. We show also that when the shock in the infinite system is observed from the location of a second class particle, treating this particle either as a first class particle or as an empty site leads to shift equivalent shock measures.

Date: February 12, 1998

Submitted to: Issue of *Journal of Statistical Physics* dedicated to Leo Kadanoff

Key Words: non-translation-invariant measures, shock structure, asymmetric simple exclusion process, second class particles, shift-coupling, orbit coupling

[1] Laboratoire de Physique Statistique, Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France; email derrida@lps.ens.fr

[2] Department of Mathematics, Rutgers University, New Brunswick, NJ 08903; email oldstein@math.rutgers.edu, lebowitz@math.rutgers.edu, speer@math.rutgers.edu

[3] Department of Physics, Rutgers University, New Brunswick, NJ 08903

1. Introduction

A major component of statistical mechanics, especially its mathematical aspect, is the study of measures or probability distributions for infinite particle systems. Such infinite systems represent idealizations of macroscopic physical systems whose spatial extension, although finite, is very large on the microscopic scale of interparticle distances or interactions. The advantage of this idealization is that many phenomena which are clearly manifested in real macroscopic systems, such as phase transitions, have precise counterparts in the behavior of the infinite volume measures. The inevitable boundary and finite size effects present in real systems, which are frequently irrelevant to the phenomena of interest, are eliminated in the thermodynamic (infinite volume) limit [1].

Our mathematical characterization of these measures for infinite particle systems is very good for situations in which the measures are translation invariant (TI) [1,2]. The situation becomes less transparent when dealing with spatially nonuniform measures. These can arise in various ways. A rather trivial case occurs when the interaction Hamiltonian or the dynamics specifying the evolution is position dependent. More interesting cases arise when the translation symmetry is broken "spontaneously" by the measure. We first illustrate these by a well known example from equilibrium statistical mechanics, then discuss a nonequilibrium example which is the main focus of this paper.

Consider the Gibbs measures for the nearest neighbor ferromagnetic Ising model on $\mathbb{Z}^d$ at low temperatures. In $d \geq 3$ there exist, in addition to the TI extremal measures μ_+ and μ_-, in which the spontaneous magnetizations are $\pm m^*$ with $m^* \neq 0$, many non-TI measures called *Dobrushin states*; a family of these can be obtained as the infinite volume limit of systems with $\pm$ boundary conditions in the $\hat{e}_1$-direction [3]. Specifically, let the domain Ω containing the system consist of sites $j = (j_1, \ldots, j_d)$ such that $j_1 \in \mathbb{Z}$ and $j_2, \ldots, j_d \in [-N, N]^{d-1}$, with all spins outside Ω being equal to $+1$ for $j_1 \geq 0$ and -1 for

$j_1 < 0$. In each system configuration we consider the set of $(d-1)$-dimensional surfaces separating $+1$ and -1 spins, formed of $(d-1)$-dimensional faces of cubes in the dual lattice $\mathbb{Z}^d + (1/2, \ldots, 1/2)$. The *Dobrushin interface* is the maximal connected component of this set which contains all such $(d-1)$-faces outside Ω. If $d \geq 3$ then at low enough temperatures (below the roughening transition) this interface remains localized near the $j_1 = 0$ plane as N increases. Consequently, in the resulting infinite volume states $\mu^{(\pm)}$ the expectation values (or correlations) depend on the $\hat{e}_1$ coordinate, e.g., if $\sigma_j = \pm 1$ is the spin at site $j \in \mathbb{Z}^d$ then $\langle \sigma_j \rangle_{\mu^{(\pm)}}$ is positive for $j_1 \geq 0$ and negative for $j_1 < 0$ [3]. This non-translation-invariant infinite volume Gibbs state is one of an infinite family obtained via translations in the $\hat{e}_1$ direction.

If $d = 2$, the same boundary conditions produce a translation invariant state in the infinite volume limit. This is because the interface, while remaining locally sharp, fluctuates in position with the variance of its displacement from the plane $j_1 = 0$ growing like N [4]. Consequently, the limiting measure (defined by the $N \to \infty$ limit of local correlation functions) is a superposition, with equal weights, of the extremal translation invariant measures μ_+ and μ_- [4,5]. Suppose, however, that we view the system from some point attached to the Dobrushin interface; for example, we might choose the point $(j^*, 0)$, where j^* is as large as possible so that $(j^* - 1/2, 0)$ is on the interface. Note that the value of j^* and hence the viewpoint will depend on the configuration under view. It seems clear that when $N \to \infty$ a limiting measure will exist which will not be translation invariant, but will instead approach the state μ_+ (respectively μ_-) as one goes to infinity in the positive (respectively negative) $\hat{e}_1$ direction; this has not been explicitly established but for results in this direction see [6]. Other viewpoints are of course possible, and the resulting measure will depend in a complicated way on the choice made. One might choose, for example, $(j^* + j_1, j_2)$ for some fixed (j_1, j_2), or $(j_*, 0)$ with j_* as small as possible so that $(j_* + 1/2, 0)$ is on the interface. One could even add an additional randomness by choosing either $(j^*, 0)$

or $(j_*, 0)$ with equal probability; this seems artificial in the current context but this sort of additional randomness is natural and necessary in the one-dimensional system to be studied shortly.

This example illustrates two ways in which non-TI measures arise. The non-TI measures for $d \geq 3$ arise in viewing the system from nonrandom, fixed, frames. Choice of a different frame simply effects a translation of the measure. In contrast, the non-TI measures for $d = 2$ can be seen only if one views the system from a random position—random in the sense that it depends on the configuration. Moreover, since the choice of a viewpoint is rather arbitrary and since the effect on the measure of a change in viewpoint is hardly transparent, one must now consider a large family of distinct measures arising from different viewpoints. Of course, even in $d \geq 3$ we could consider the measure as seen from a point attached to the interface. At low temperature there seems to be little to be gained from such an approach, but it might be of interest between the roughening and critical temperatures in $d = 3$, where the situation is expected to be similar to that in $d = 2$.

Several questions arise when the transition region must be described by non-TI measures obtained from configuration-dependent viewpoints. Is there a natural choice of viewpoint which will give a best or simplest description of the local structure of the transition region? How can one extract intrinsic properties of this region from such a description—properties independent of the choice of viewpoint? And, given two non-TI measures, how may one decide if they in fact describe the same system seen from different points of view? The purpose of the present work is to address such questions. Our motivation, and the focus of our study, is in fact not the above example but a nonequilibrium system, the one dimensional asymmetric simple exclusion process (ASEP). The results are, for the moment, also specific to one-dimensional systems.

We now describe the non-TI measures arising in the ASEP. The latter [7,8] is a model of particles moving on the lattice $\mathbb{Z}$; a configuration η of the system has the form

$\eta = (\eta(i))_{i\in\mathbb{Z}}$, where $\eta(i)$ is 0 or 1 at an empty or occupied site, respectively. Dynamically, each particle attempts to jump to a neighboring site, at random times with rate 1, choosing its right or left neighbor with probabilities p and q, respectively, where $p > 1/2$ and $q = 1 - p$. The jump takes place if and only if the target site is empty. The extremal stationary TI states of this system are the product (Bernoulli) measures ν_ρ with constant density ρ satisfying $0 \le \rho \le 1$ [8]. There exist also non-TI stationary states, which are product measures with nonuniform density approaching 0 as $i \to -\infty$ and 1 as $i \to \infty$. The latter are in fact special examples of the general class of non-TI measures we will study here: those which describe the microscopic structure of shocks present in the macroscopic description of the ASEP.

The ASEP is described on the (Euler) macroscopic space-time scale by the inviscid Burgers equation for the particle density $n(x,t) \in [0,1]$, where $x, t \in \mathbb{R}$ ([9–11]):

$$\frac{\partial n}{\partial t} + (p-q)\frac{\partial}{\partial x} n(1-n) = 0. \tag{1.1}$$

Equation (1.1) has shock solutions, $n(x,t) = u(x - Vt)$, where

$$u(y) = \begin{cases} \rho_- & \text{for } y < 0, \\ \rho_+ & \text{for } y > 0; \end{cases} \tag{1.2}$$

here $\rho_+ > \rho_-$ and the velocity is $V = (p-q)(1 - \rho_+ - \rho_-)$. A natural question then is what behavior of the ASEP system on the microscopic level corresponds to this shock solution. For example, one may take the initial state μ_0 of the system to be a product measure with density at site j given by ρ_- for $j < 0$ and ρ_+ for $j \ge 0$, and ask about the $t \to \infty$ limiting behavior of the state μ_t at time t. It might seem that if one were to view the system from a frame moving with the shock velocity V then one would see in this limit a non-TI state describing the intrinsic microscopic structure of the shock. But this is not true: because fluctuations in the shock position become unbounded on the microscopic scale as $t \to \infty$, the resulting measure is an equal superposition of the product measures ν_{ρ_+} and ν_{ρ_-} [12,13]. It has been shown [14,15], however, that there exists a (nonunique) time-dependent random position X_t such that the $t \to \infty$ limit of the measure μ_t *seen from the viewpoint* X_t, which we shall denote by μ', exists and is spatially asymptotic to the product measures ν_{ρ_+} and ν_{ρ_-}:

$$\lim_{k \to \pm\infty} T^{-k} \mu' = \nu_{\rho_\pm}. \tag{1.3}$$

Here T is the translation operator, which acts on configurations by $(T\eta)(i) = \eta(i-1)$, on functions of configurations by $(Tf)(\eta) = f(T^{-1}\eta)$, and on measures on configurations space by $\langle f \rangle_{T\mu} = \langle T^{-1} f \rangle_\mu$. The situation is thus analogous to that of the two dimensional Ising model: in the $t \to \infty$ limit here, and in the $N \to \infty$ limit there, one must look from a configuration-dependent viewpoint to see the non-TI state.

The random position X_t discussed above is given by the location of a single *second class particle* inserted into the system, which is then treated as an empty site in obtaining the measure μ'. This viewpoint is doubly random, in that the random configuration η does not completely determine the viewpoint, but only its *distribution*. The measure μ' is invariant under the ASEP dynamics for the system seen from the second class particle (we will describe this dynamics below). In previous works [16,17] explicit formulas were obtained for a measure $\widehat{\mu}$ invariant for this same dynamics and with the same spatial asymptotics (1.3), and it is this measure that will be our main example here; presumably $\widehat{\mu} = \mu'$, although this has not been established. (It is μ' which has shown to be obtained by the long time asymptotics described above.)

In this paper we will focus on questions like those raised above in the context of the Ising model, which arise from the possibility of different choices of viewpoint. In Section 2 we describe the evolution of the ASEP with a second class particle and the resulting viewpoint on the shock, as well as several other possible choices of viewpoint. In Section 3 we formalize, in a general one-dimensional context, the relation of *shift equivalence* on non-TI measures under which equivalent measures differ by a random change of viewpoint. There we define also certain quantities, called *translation sums*, which characterize this equivalence: two measures (which must satisfy certain additional conditions) are shift equivalent if and only if all the translation sums for the two measures agree. This result will be established in a separate paper [18].

In the remainder of the paper we apply these general ideas to the ASEP. In Section 4 we utilize the results of [17] to compute the translation sums explicitly. From this computation (and using the verification, here omitted, that the ASEP shock measures satisfy the additional conditions mentioned above) we establish both negative and positive results about the ASEP shock. In Section 5 we show that for certain values of the parameters the shock measure is not shift equivalent to any product measure with a monotone density; we show also that when the shock is observed from the location of a second class particle, treating this particle either as a first class particle or as an empty site leads to shift equivalent shock measures. Finally, in Section 6 we show that certain shocks arising in versions of the ASEP with different boundary conditions are in fact shift equivalent.

2. Points of view for the ASEP shock

In this section we illustrate the nature of shock measures by considering various viewpoints on the shock for the ASEP, beginning with the viewpoint from a second class particle. We first describe briefly the properties of the ASEP when a single second class particle is introduced into the system [12]. The second class particle has its own dynamics: it attempts to jump exactly as does an ordinary (first class) particle, succeeding only if the target site is empty; on the other hand, when a first class particle attempts to jump onto the site occupied by the second class particle, the jump succeeds and the two particles exchange sites. A configuration of this system is $\tau = (\tau(i))_{i\in\mathbb{Z}}$, where $\tau(i)$ is 0 if site i is unoccupied, 1 if it is occupied by one of the original particles, now called *first class particles*, and 2 if it is occupied by the second class particle. Let us denote the location of the second class particle by X. If λ_t is a measure on this system evolving under the above dynamics and X_t the corresponding location of the second class particle at time t, we write $T^{-X_t}\lambda_t$ for the measure describing the configurations as seen from the second class particle.

There is an alternate, equivalent way to describe the system with a second class particle [12]. Consider two copies of the ASEP system having configurations η_0 and η_1 which agree except at one site X, at which $\eta_0(X) = 0$ and $\eta_1(X) = 1$, i.e., system 0 has a hole and system 1 a particle. Allow this pair of systems to evolve under a coupled dynamics, so that attempts to jump from a given site to an adjacent one occur simultaneously. Then each time a jump occurs in either system the same jump occurs in the other, if possible; this synchronization can fail only when the extra particle in system 1 jumps, or when a particle in system 0 jumps on the extra hole, and in these cases the mismatch position X will move. From a configuration (η_0, η_1) of this doubled system we may obtain a configuration τ of the single ASEP with second class particle by taking $\tau(X) = 2$ and $\tau(i) = \eta_0(i) = \eta_1(i)$

when $i \neq X$; the dynamics for the doubled ASEP system corresponds to that described above of the system with a second class particle. Conversely, from a configuration τ we may obtain two distinct ASEP configurations η_0 and η_1 by restricting attention to one or the other of the paired ASEP systems or equivalently by replacing the second class particle by respectively a hole or a first class particle; we will write $\eta_0 = \Psi_0(\tau)$ and $\eta_1 = \Psi_1(\tau)$. Similarly, from a measure λ for the ASEP with second class particle we obtain ASEP measures $\Psi_0(\lambda)$ and $\Psi_1(\lambda)$ giving the distribution of η_0 and η_1 under λ. Clearly if λ_t is evolving under the second class particle dynamics then both $\Psi_0(\lambda_t)$ and $\Psi_1(\lambda_t)$ evolve under the simple ASEP dynamics.

A measure $\widehat{\lambda}$ describing this system from the viewpoint of the second class particle, i.e., in a reference frame in which $X = 0$, was constructed explicitly in [17]; the construction is summarized in Section 4. This measure has spatial asymptotics corresponding to the shock,

$$\lim_{k \to \pm\infty} T^{-k}\widehat{\lambda} = \nu_{\rho_\pm}, \tag{2.1}$$

and is invariant under the natural dynamics for the system seen from the second class particle, under which the second class particle is always at the origin and a jump of this particle in the original dynamics becomes a jump of the rest of the system in the opposite direction.

To obtain a measure on the original ASEP configurations from the measure $\widehat{\lambda}$ we may consider either $\Psi_0(\widehat{\lambda})$ or $\Psi_1(\widehat{\lambda})$. To be definite, let us focus for the moment on the former, which gives the distributions of the configuration η_0, and denote it by $\widehat{\mu}$; $\widehat{\mu}$ is obtained from $\widehat{\lambda}$ by replacing the second class particle at the origin with a hole. This is the measure referred to in the introduction. We may allow it to evolve to $\widehat{\mu}_t$ under the ASEP dynamics, or equivalently write $\widehat{\mu}_t = \Psi_0(\widehat{\lambda}_t)$; then viewed from the position X_t it is time invariant:

$$T^{-X_t}\widehat{\mu}_t = \widehat{\mu}. \tag{2.2}$$

Thus $\widehat{\mu}$ furnishes an invariant description of the shock itself, ignoring its location. It follows from (2.1) that $\widehat{\mu}$ has the spatial asymptotics (1.3).

Other points of view are possible. Thus one may obtain new descriptions of the shock by rather trivial shifts of viewpoint; for example, a constant one, to the second site to the right of the second class particle, a configuration-dependent shift, to the third empty site to its right, or a shift with additional randomness, such as a choice, with equal weights, between the two previous possibilities.

An alternate measure for the shock is implicit in the description of the system with second class particle as a coupled pair of ASEP systems: the measure $\widetilde{\mu}_t = \Psi_1(\widehat{\lambda}_t)$, obtained from $\widehat{\lambda}_t$ by replacing the second class particle by a first class particle, evolves with the ASEP dynamics, and is invariant in the sense of (2.2): $T^{-X_t}\widetilde{\mu}_t = \widetilde{\mu}_0$. We write $\widetilde{\mu} = \widetilde{\mu}_0 = \Psi_1(\widehat{\lambda})$; $\widetilde{\mu}$ is obtained from $\widehat{\mu}$ simply by replacing the empty site at the origin by a particle. Despite the fact that $\widetilde{\mu}$ and $\widehat{\mu}$ are equally valid as candidates for the description of the ASEP shock (in the terminology introduced in the next section, they are both invariant shock measures for the ASEP), it is not at all clear that they are shift equivalent, that is, differ by the sort of random change of viewpoint that we have been considering here. In Section 5 we will use the ideas of the next section to show that this is the case.

Other choices may for some purposes be more tractable. The existence of a shock measure was first proved by Ferrari, Kipnis, and Saada [14] (FKS) using a random viewpoint Z_t which Ferrari [15] later showed was related to X_t by a random translation of finite mean. To construct Z_t, consider again two copies of the ASEP system with configurations ζ_0 and ζ_1 satisfying $\zeta_0(k) \le \zeta_1(k)$ for all k, so that when there is a particle at site k in configuration ζ_0 there is also a particle at that site in configuration ζ_1. Allow the system to evolve under the coupled dynamics described above, so that again sites k at which $\zeta_0(k) = \zeta_1(k) = 1$ and those at which $\zeta_0(k) = 0$ and $\zeta_1(k) = 1$ obey the dynamics of first and second class particles, respectively. Let λ^* be a translation- and time-invariant measure for this system in which the densities are given by $\langle\zeta_0(0)\rangle_{\lambda^*} = \rho_-$ and $\langle\zeta_1(0)\rangle_{\lambda^*} = \rho_+$ (the existence of such a λ^* is established in [14]). Randomly select some second class particle, i.e., some discrepancy between ζ_0 and ζ_1 (more precisely, condition on the presence of such a particle at the origin at time 0), and let Z_t be its position at time t. At time $t = 0$, define an ASEP configuration η as follows: first, if $\zeta_0(k) = \zeta_1(k) = 1$ then $\eta(k) = 1$, and if $\zeta_0(k) = \zeta_1(k) = 0$ then $\eta(k) = 0$; second, if $\zeta_0(k) = 0$ and $\zeta_1(k) = 1$, and k is the j^{th} site at which such a discrepancy occurs, counting from $j = 0$ at Z_0, then $\eta(k) = 1$ with probability $1/(1 + (q/p)^j)$, and $\eta(k) = 0$ with the complementary probability, and all of these choices are independent. Allow the configuration η to evolve with ASEP dynamics coupled to that of ζ_0 and ζ_1, so that if a jump occurs in any of the three systems it also occurs in any others in which it is possible. Then the distribution of η at time t, viewed from the position Z_t, is time independent and has the shock asymptotics (1.3).

In all the examples considered so far the position of the viewpoint is not determined by knowledge of the ASEP configuration η: knowing η gives only the *distribution* of the random viewpoint. This is an additional randomness beyond the configuration dependence discussed for the $d = 2$ Ising model in the introduction. Our last example is a construction of a viewpoint ℓ on a non-TI measure μ which, as in the $d = 2$ equilibrium example, is intrinsic. This means, first, that the viewpoint $\ell(\eta)$ depends only on the configuration η, with no additional randomness, and second, that the viewpoint behaves covariantly under translations, so that

$$\ell(T\eta) = \ell(\eta) + 1. \tag{2.3}$$

The function $\ell(\eta)$ may be thought of as picking out a shock location in the configuration η. Its intrinsic nature means that if ν is any measure obtained from μ by a shift of viewpoint then ν and μ look the same from the configuration dependent viewpoint ℓ. In the construction of ℓ, μ need not be related to the ASEP dynamics; we require only that μ be a non-TI measure on ASEP configurations which has well-defined and configuration-independent asymptotic densities satisfying $\rho_+ > \rho_-$:

$$\rho_\pm = \lim_{N\to\pm\infty} \frac{1}{|N|+1} \sum_{k=0}^{N} \eta(k), \tag{2.4}$$

for μ-almost every η.

To define $\ell(\eta)$ we first define a function $h_\eta(j)$ to be the signed cumulative occupation from the origin to site j:

$$h_\eta(j) = \begin{cases} \sum_{i=1}^{j} \eta(i), & \text{if } j > 0, \\ 0, & \text{if } j = 0, \\ -\sum_{i=j+1}^{0} \eta(i), & \text{if } j < 0. \end{cases} \tag{2.5}$$

The graph of h_η has, for typical η, slope ρ_+ (on a large scale) far to the right of the origin and ρ_- far to the left. Now fix an irrational number ρ_* satisfying $\rho_- < \rho_* < \rho_+$, and for any η define $\ell(\eta)$ to be the integer j which minimizes $h_\eta(j) - \rho_* j$, whenever such a (necessarily unique) minimizing integer exists; for other η, $\ell(\eta)$ is undefined. The special role played by the origin in the definition (2.5) of h_η does not affect the value of $\ell(\eta)$. The construction is shown graphically in Figure 1. It is intuitively clear, and can be proved, that ℓ is well defined and finite with probability one relative to μ. The definition of ℓ depends strongly on the choice of ρ_*; slight changes in ρ_* will cause large changes in the viewpoint $\ell(\eta)$ for some configurations η.

Note that the sets $S_k = \{\eta \mid \ell(\eta) = k\}$ form a partition of the configuration space (up to a set of μ-measure zero) which is nicely mapped by translations: $T(S_k) = S_{k+1}$. Such partitions have been constructed for more general T in the context of discrete time dynamical systems by Gurevič and Oseledec [19].

The shock may look quite different from different viewpoints, in particular, from viewpoints ℓ defined with different values of ρ_*. In Figure 2 we show shock profiles (mean values $\langle\eta(k)\rangle$ at site k relative to the viewpoint adopted) for the shock in the totally asymmetric ($p=1$) model with densities $\rho_+=0.7$, $\rho_-=0.2$, seen from three different viewpoints: from the second class particle (that is, in the measure $\widehat{\mu}$) and from ℓ_1 and ℓ_2, the viewpoints constructed as above with $\rho_{*1}=\pi^{-1}\rho_+ + (1-\pi^{-1})\rho_-$ and $\rho_{*2}=(1-\pi^{-1})\rho_+ + \pi^{-1}\rho_-$.

3. Equivalence of measures under random shifts

In the previous section we have described implicitly an equivalence relation on probability measures on the ASEP configuration space $S=\{0,1\}^{\mathbb{Z}}$, under which two such measures μ_1 and μ_2 are equivalent if they differ by a configuration dependent random shift of viewpoint. Perhaps the simplest way to make this precise is in terms of a *coupling* for the two measures, that is, a measure on $S\times S$ with marginals μ_1 and μ_2 on the first and second components. We say that μ_1 and μ_2 are *shift equivalent* if there exists such a coupling μ^* and an integer-valued function Y such that for $(\eta_1,\eta_2)\in S\times S$, $\eta_2=T^{-Y(\eta_1,\eta_2)}\eta_1$ with μ^*-probability one; in this case we write

$$\mu_2 = T^{-Y}\mu_1. \tag{3.1}$$

Generalizations to measures on sets other than S on which a translation operator acts can easily be made. The coupling μ^* is sometimes referred to in the mathematical literature as a *shift coupling* [20,21] or an *orbit coupling* [22].

It is sometimes convenient to work with alternate formulations of this equivalence relation. One such is based on a variation of the well known Vasershtein distance [23] between two measures. Define $d(\eta_1,\eta_2)$ on $S\times S$ to be the minimum, over k, of the number of sites at which η_1 and $T^k\eta_2$ differ, and for measures μ_1 and μ_2 on S define $D(\mu_1, \mu_2)$ by

$$D(\mu_1, \mu_2) = \inf_{\nu^*} \langle d \rangle_{\nu^*}, \tag{3.2}$$

where the infimum is over couplings ν^* for μ_1 and μ_2. Then it can be shown [18] that μ_1 and μ_2 are shift equivalent if and only if $D(\mu_1, \mu_2) = 0$. The minimum in (3.2) is in fact achieved when $\nu^* = \mu^*$, with μ^* the coupling used to define (3.1).

A second reformulation of the relation of shift equivalence is obtained by noting that if (3.1) holds then clearly

$$\mu_1(A) = \mu_2(A) \qquad \text{for all } A \subset S \text{ and } T(A) = A, \tag{3.3}$$

that is, if A is translation invariant. Conversely, it can be shown [21,22] that if (3.3) holds then there exists a random position Y so that (3.1) is satisfied. For measures describing shocks, "interesting" TI sets A with nontrivial probability describe intrinsic properties of the shock; for example, we might take A to be the set of configurations η with a particle at the site three sites ahead of the position $\ell(\eta)$ defined in Section 2, so that $\mu(A) = \langle \eta(\ell(\eta) + 3) \rangle_\mu$.

Finally, as remarked earlier, when an intrinsic viewpoint $\ell(\eta)$ can be defined, that is, a function $\ell(\eta)$ defined almost everywhere with respect to both μ_1 and μ_2 and satisfying (2.3), then μ_1 and μ_2 are shift equivalent if and only if $T^{-\ell}\mu_1 = T^{-\ell}\mu_2$.

The concept of shift equivalence allows us to give precise definitions of two natural concepts for shock measures in the ASEP (or similar systems). We say that a measure μ is an *invariant shock measure* if it has the spatial asymptotics (1.3) for some $\rho_\pm$ and if, when μ_t is the measure evolving under the ASEP dynamics which satisfies $\mu_0 = \mu$, μ_t is shift equivalent to μ for all t; for example, the measures $\widehat{\mu}$ and $\widetilde{\mu}$ of the previous section are invariant shock measures in this sense. We say that the ASEP has a *unique* shock measure (for given $\rho_\pm$) if any two such invariant measures are shift equivalent. It seems natural to conjecture that the ASEP has a unique shock measure in this sense for all $\rho_\pm$ satisfying $\rho_+ > \rho_-$, but this has not been established.

Let us now restrict attention to measures on S which, like the ASEP shock measures considered in Section 2, converge under spatial translation to distinct TI states. Suppose then that $\mu_\pm$ are translation invariant probability measures on S with $\mu_+ \neq \mu_-$, and define a *ramp measures* to be a probability measure μ on S which is asymptotic to μ_+ to the right of the origin and to μ_- to the left:

$$\lim_{k\to\pm\infty} T^{-k}\mu = \mu_\pm. \tag{3.4}$$

Associated to each ramp measures is a family of *translation sums*. These sums, under rather mild additional technical conditions on the measures involved [18], are invariant under a shift of viewpoint and furnish a complete characterization of shift equivalence. In a sense, the equivalence of (3.1) and (3.3) also provides a set of invariant quantities which determine the equivalence class of a measure μ: the values $\mu(A)$ for all TI sets A. The example of the translation invariant set given above, however, suggests correctly that these quantities are difficult to calculate. We will see in the next section that the translation sums are calculable for the ASEP shock measure.

Suppose that μ is a ramp measure and that f is a function on S which depends on only finitely many occupation numbers and which satisfies $\mu_+(f) = \mu_-(f) = 0$; for example, $f(\eta) = \eta(1) - \eta(0)$ is such a function and, if the asymptotic states are product measures, i.e., if $\mu_\pm = \nu_{\rho_\pm}$, then so is $f(\eta) = \eta(k)(\eta(1) - \eta(0))$ whenever $k \neq 0, 1$. Then we may define the *translation sum*

$$\Delta_\mu(f) = \sum_{n=-\infty}^{\infty} \langle T^n f\rangle_\mu; \tag{3.5}$$

$\Delta_\mu(f)$ will be finite for any ramp measure μ for which the asymptotic behavior (3.4) is achieved with sufficient rapidity to guarantee that this sum converges. For example, if

$f(\eta) = \eta(1) - \eta(0)$ then $\Delta_\mu(f) = \rho_+ - \rho_-$ since (3.5) telescopes, and similarly $\Delta_\mu(f) = \langle g\rangle_{\mu_+} - \langle g\rangle_{\mu_-}$ if $f = Tg - g$ for some $g(\eta)$. However, we see no easy way to compute $\Delta_\mu(f)$ for general μ and f.

The values of the translation sums characterize the (shift) equivalence classes of ramp measures in the following sense: under additional conditions describing the convergence at $\pm\infty$, two ramp measures μ_1 and μ_2 are shift equivalent if and only if

$$\Delta_{\mu_1}(f) = \Delta_{\mu_2}(f) \tag{3.6}$$

for all f satisfying $\mu_+(f) = \mu_-(f) = 0$ [18].

As a simple application of this result, consider a product measure ν_1 on S with density $\rho_1(k) = \langle\eta(k)\rangle_{\nu_1}$ satisfying $\lim_{k\to\pm\infty}\rho_1(k) = \rho_\pm$, with $\rho_+ \neq \rho_-$; ν_1 is a ramp measure (asymptotic to ν_{ρ_+} and ν_{ρ_-}) and the additional technical considerations needed for the above result are satisfied if the asymptotic limit is achieved sufficiently rapidly. Let ν_2 be another such measure obtained by altering the density at the origin only: $\rho_2(0) \neq \rho_1(0)$, $\rho_2(k) = \rho_1(k)$ if $k \neq 0$. Then ν_1 and ν_2 will typically have different translation sums and hence not be shift equivalent; for example, consideration of the translations sums for the functions $f_k(\eta) = \eta(k)(\eta(1)-\eta(0))$, $k > 1$, shows that ν_1 and ν_2 will not be shift equivalent unless $\rho_1(k) + \rho_1(-k)$ is independent of k.

4. Translation sums for the ASEP shock measure

In this section we show how to compute the translation sums $\Delta_{\widehat{\mu}}(f)$ for the ASEP shock measure $\widehat{\mu}$ which, as discussed in Section 2, is obtained from the invariant shock measure $\widehat{\lambda}$ for the system with second class particle at the origin by replacing that particle by a hole. In this case the asymptotic measures $\mu_{\pm}$ are the product measures $\nu_{\rho_{\pm}}$. We will continue to denote a typical ASEP configuration by η ($\eta(i) = 0$ or 1) and a configuration in which second class particles may occur by τ ($\tau(i) = 0$, 1, or 2); in this section such a τ will always contain a single second class particle located at the origin.

In [17] (which was an extension to the general ASEP of the results of [16] for the totally asymmetric model, in which $p = 1$) it was shown that the measure $\widehat{\lambda}$ can be written in terms of two vectors $|v\rangle$ and $\langle w|$ and three operators A, D and E satisfying the following algebraic rules:

$$pDE - qED = (p - q)[(1 - \rho_-)(1 - \rho_+)D + \rho_-\rho_+E], \tag{4.1}$$

$$pAE - qEA = (p - q)(1 - \rho_-)(1 - \rho_+)A, \tag{4.2}$$

$$pDA - qAD = (p - q)\rho_+\rho_-A, \tag{4.3}$$

$$(D + E)|v\rangle = |v\rangle, \tag{4.4}$$

$$\langle w|(D + E) = \langle w|, \tag{4.5}$$

$$\langle w|A|v\rangle = 1. \tag{4.6}$$

Specifically, the probability of the set of configurations specified by the occupation numbers $\zeta(i)$ ($\zeta(i) = 0, 1$) of m consecutive sites to the left of the second class particle (which is located at the origin) and n consecutive sites to its right,

$$\widehat{\lambda}\big(\{\,\tau \mid \tau(i) = \zeta(i),\ i = -m, \ldots, -1 \text{ and } i = 1, \ldots, n\,\}\big) \equiv P_{m,n}(\zeta), \tag{4.7}$$

can be written as the matrix element in which a first class particle is represented by a matrix D, an empty site by a matrix E, and the second class particle by a matrix A:

$$P_{m,n}(\zeta) = \langle w| \left\{ \prod_{i=-m}^{-1} [\zeta(i)D + (1-\zeta(i))E] \right\} A \left\{ \prod_{j=1}^{n} [\zeta(j)D + (1-\zeta(j))E] \right\} |v\rangle. \quad (4.8)$$

For example, the probability of finding occupation numbers 1 0 1 immediately to the left and 0 1 1 0 0 immediately to the right of the second class particle, that is, of the local configuration 1 0 1 2 0 1 1 0 0, is given by $\langle w|DEDAED^2E^2|v\rangle$.

It follows from the above that the microscopic shock profile, defined as the average occupation $\langle \tau(n) \rangle_{\widehat{\lambda}} = \langle \eta(n) \rangle_{\widehat{\mu}}$ at site $n \neq 0$, is given for $n > 0$ by

$$\langle \tau(n) \rangle_{\widehat{\lambda}} = \langle w|A(D+E)^{n-1}D|v\rangle. \quad (4.9)$$

The exact expression of this profile was given in equations (4.1–4.5) of [17], where it was also shown that the profile has the symmetry property

$$\langle \tau(n) \rangle_{\widehat{\lambda}} + \langle \tau(-n) \rangle_{\widehat{\lambda}} = \rho_+ + \rho_-. \quad (4.10)$$

Let us call a finite product of D's and E's a *word*. To every function of k consecutive site occupation numbers, say $f(\eta(1), \eta(2), ..., \eta(k))$, there is naturally associated a linear combination of words of length k; for example, if $f = \eta(1)(1-\eta(3)) + 3\eta(5)\eta(6)$ then

$$W = D(D+E)E(D+E)^3 + 3(D+E)^4D^2. \quad (4.11)$$

We denote the word of length zero by 1 and associate it with the function $f = 1$. For every linear combination W of words we will define below a number $\Gamma(W)$ and will then develop, from the algebraic rules (4.1)–(4.6), new rules which enable us to calculate $\Gamma(W)$. This will determine the translation sums $\Delta_{\widehat{\mu}}(f)$ defined in (3.5), since we will show below that when f and W are related as above and in addition $\langle f \rangle_{\mu_\pm} = 0$, $\Delta_{\widehat{\mu}}(f) = \Gamma(W)$.

Suppose then that W is a linear combination of words and define

$$\Gamma(W) = \lim_{L,M\to\infty} \{\Gamma_{L,M}(W) - L\,r_+(W) - M\,r_-(W)\}\,. \tag{4.12}$$

Here

$$\Gamma_{L,M}(W) = \frac{d}{d\theta}\langle w|(\tilde D+\tilde E)^L \tilde W (\tilde D+\tilde E)^M|v\rangle\Big|_{\theta=0}\,, \tag{4.13}$$

with

$$\tilde D = D, \qquad\qquad \tilde E = E + \theta A, \tag{4.14}$$

and $\tilde W$ the operator obtained from W by replacing each D and E in W by $\tilde D$ and $\tilde E$; $r_\pm(W)$ denotes the number obtained by replacing each D and E in W by $\rho_\pm$ and $1-\rho_\pm$ respectively. It is clear from (4.13) and the definition (4.14) of $\tilde D$ and $\tilde E$ that the right hand side of (4.13) is a sum of matrix elements of products of the operators D, E, and A, with each product containing a single operator A; each such matrix element is computable from (4.1)–(4.6) and represents a probability calculated in the measure $\widehat\lambda$. Moreover, the limit in (4.12) exists because $\widehat\lambda$ converges exponentially fast to $\nu_{\rho_\pm}$ at $\pm\infty$ [17]. Thus $\Gamma(W)$ is well defined.

It follows directly from the definition of $\Gamma(W)$ that

$$\Gamma(a_1W_1 + a_2W_2) = a_1\Gamma(W_1) + a_2\Gamma(W_2), \qquad \text{for any } a_1 \text{ and } a_2; \tag{4.15}$$

$$\Gamma(W[D+E]) = \Gamma(W) + r_-(W); \tag{4.16}$$

$$\Gamma([D+E]W) = \Gamma(W) + r_+(W); \tag{4.17}$$

$$\Gamma(1) = 0. \tag{4.18}$$

Also, one can easily show that $\tilde D$ and $\tilde E$ satisfy (4.1), and hence for any W_1 and W_2,

$$\Gamma(W_1(pDE - qED)W_2) = (p-q)\Gamma(W_1[(1-\rho_-)(1-\rho_+)D + \rho_-\rho_+E]W_2). \tag{4.19}$$

The five rules (4.15)–(4.19), together with the value of $\Gamma(D)$, which we will derive below, allow one to calculate $\Gamma(W)$ for any linear combination of words W. The argument, based on a recursion on the length of a word W, is almost identical to that given in [24,17] to show that (4.1)–(4.6) are sufficient to calculate any matrix element of the form (4.8). Consider, for example, $\Gamma(DW_n)$, where W_n is a word of length $n \geq 1$ with k factors E:

$$\begin{aligned} p^n\Gamma(DW_n) &= q^k p^{n-k}\Gamma(W_n D) + \text{l.o.t.} \\ &= -q^k p^{n-k}\Gamma(W_n E) + q^k p^{n-k} r_-(W_n) + \text{l.o.t.} \\ &= -q^n\Gamma(EW_n) + q^k p^{n-k} r_-(W_n) + \text{l.o.t.} \\ &= q^n\Gamma(DW_n) - q^n r_+(W_n) + q^k p^{n-k} r_-(W_n) + \text{l.o.t.} \end{aligned} \tag{4.20}$$

Here l.o.t. (lower order terms) denotes linear combinations of $\Gamma(W)$ for words W of length n or less. Since $q < p$, (4.20) can be solved for $\Gamma(DW_n)$. For example, since (4.16) and (4.18) imply that $\Gamma(D+E) = 1$,

$$\begin{aligned} p\Gamma(D^2) &= -p\Gamma(DE) + p\Gamma(D) + p\rho_- \\ &= -q\Gamma(ED) - (p-q)\Gamma((1-\rho_+)(1-\rho_-)D + \rho_-\rho_+ E) + p\Gamma(D) + p\rho_- \\ &= q\Gamma(D^2) + p\rho_- - q\rho_+ - (p-q)\rho_-\rho_+ + (p-q)(\rho_+ + \rho_-)\Gamma(D), \end{aligned} \tag{4.21}$$

and hence

$$\Gamma(D^2) = \frac{p\rho_- - q\rho_+}{p-q} - \rho_+\rho_- + (\rho_+ + \rho_-)\Gamma(D). \tag{4.22}$$

If W is a word of length k which contains j factors E and if for $i = 1, \ldots, j$, $W^{(i)}$ is the operator product obtained from W by replacing the i^{th} factor E by A, then from (4.4), (4.5), (4.13), and (4.14),

$$\Gamma_{L,M}(W) = \sum_{i=0}^{L-1} \langle w|A(D+E)^i W|v\rangle + \sum_{i=1}^{j} \langle w|W^{(i)}|v\rangle + \sum_{i=0}^{M-1} \langle w|W(D+E)^i A|v\rangle. \tag{4.23}$$

More generally, if $f(\eta(1),\eta(2),...,\eta(k))$ is a function of the occupation numbers of sites $1,\ldots,k$ and W the corresponding linear combination of words of length k (see (4.11)), then (4.23) implies that

$$\Gamma_{L,M}(W) = \sum_{i=-M-k}^{L-1} \langle T^i f\rangle_{\hat\mu} \ ; \tag{4.24}$$

Note that, as is clear in (4.23), the averages are taken in the measure $\widehat{\mu}$ obtained from $\widehat{\lambda}$ by replacing the second class particle at the origin by a hole, that is, taking $\eta(0)=0$. For example, if $f=\eta(1)(1-\eta(3))+3\eta(5)\eta(6)$ then from (4.11) and (4.24),

$$\Gamma_{L,M}(D(D+E)E(D+E)^3+3(D+E)^4D^2) = \sum_{i=-M-5}^{L} \langle \eta(i)(1-\eta(i+2))+3\eta(i+4)\eta(i+5)\rangle_{\widehat{\mu}}. \tag{4.25}$$

Equations (4.24) and (4.10) imply that

$$\Gamma(D) = \lim_{L\to\infty}\left\{\sum_{i=-L}^{L}\langle\eta(i)\rangle_{\widehat{\mu}} - L(\rho_+ + \rho_-)\right\} = \langle\eta(0)\rangle_{\widehat{\mu}} = 0. \tag{4.26}$$

Moreover, (4.24) leads immediately to the calculation of translation sums $\Delta_{\widehat{\mu}}(f)$, for if f is a function satisfying $\langle f\rangle_{\mu_\pm}=0$ and W is the corresponding linear combination of words, then (4.12), (4.24), and the relations $\langle f\rangle_{\nu_{\rho_\pm}}=r_\pm(W)$ imply that

$$\Delta_{\widehat{\mu}}(f) = \Gamma(W). \tag{4.27}$$

The translation sums are in some cases related to expectation values in the measure $\widehat{\mu}$ in a surprisingly simple way. For example, let $f_1(\eta)$ be a function of $\eta(-m),\ldots,\eta(-1)$ and $f_2(\eta)$ a function of $\eta(2),\ldots,\eta(n)$, and let $h(\eta)=\eta(1)-\eta(0)$. Then $\mu_\pm(f_1hf_2)=0$, so that $\Delta_{\widehat{\mu}}(f_1hf_2)$ is defined. We will show that

$$\Delta_{\widehat{\mu}}(f_1hf_2) = (\rho_+ - \rho_-)\langle f_1(T^{-1}f_2)\rangle_{\widehat{\mu}}. \tag{4.28}$$

The operator corresponding to h is $A' = (D+E)D - D(D+E) = ED - DE$; in view of (4.8) and (4.27), (4.28) will follow if we show that for any W_1, W_2,

$$\Gamma(W_1 A' W_2) = (\rho_+ - \rho_-)\langle w|W_1 A W_2|v\rangle. \tag{4.29}$$

To verify (4.29) we show that, up to a normalization, $\Gamma(W_1 A' W_2)$ is determined by the same algebraic rules (4.1)-(4.6) which determined $\langle w|W_1 A W_2|v\rangle$. First, D and E of course satisfy (4.1), and this in turn implies that A' satisfies the analogue of (4.2) and (4.3):

$$pA'E - qEA' = (p-q)(1-\rho_-)(1-\rho_+)A', \tag{4.30}$$

$$pDA' - qA'D = (p-q)\rho_+\rho_- A'. \tag{4.31}$$

Moreover, since $r_+(W_1 A' W_2) = r_-(W_1 A' W_2) = 0$ for any W_1, W_2, (4.16) and (4.17) give

$$\Gamma([D+E]W_1 A' W_2) = \Gamma(W_1 A' W_2 [D+E]) = \Gamma(W_1 A' W_2), \tag{4.32}$$

which is the analogue of (4.4)–(4.5). Finally, the analogue of (4.6) is

$$\Gamma(A') = \Gamma([D+E]D) - \Gamma(D[D+E]) = \rho_+ - \rho_-. \tag{4.33}$$

Since the rules (4.1)–(4.6) determine all matrix elements $\langle w|W_1 A W_2|v\rangle$, (4.1) and (4.30)–(4.33) imply that these will agree with the $\Gamma(W_1 A' W_2)$ up to a normalization determined by comparing (4.6) and (4.33), that is, (4.29) holds.

5. Consequences of the calculation of the translation sums

In this section we apply the calculation of the translation sums $\Delta_{\widehat{\mu}}(f)$ outlined in Section 4 to discuss the shift equivalence of various measures describing the ASEP shock and to identify certain intrinsic features of the ASEP shock.

In Section 2 we observed that the measure $\widetilde{\mu}$, obtained from $\widehat{\lambda}$ by replacing the second class particle at the origin by a first class particle (or equivalently from $\widehat{\mu}$ by replacing the hole at the origin by a particle), stood on an equal footing with $\widehat{\mu}$ in providing a description of the shock. We now show that

$$\Delta_{\widehat{\mu}}(f) = \Delta_{\widetilde{\mu}}(f) \tag{5.1}$$

for any f satisfying $\langle f \rangle_{\mu_\pm} = 0$, so that by the general results referred to in Section 3, $\widehat{\mu}$ and $\widetilde{\mu}$ differ only by a shift of viewpoint (which in general will be a random shift with distribution depending on the configuration).

For suppose that, instead of using (4.14), we had defined $\tilde{D}$ and $\tilde{E}$ by $\tilde{D} = D + \theta A$ and $\tilde{E} = E$, arriving at new quantities $\tilde{\Gamma}(W)$. Then all the considerations of Section 4 would have been essentially unchanged, except that averages in (4.24) would be computed in the measure $\widetilde{\mu}$; as a result, (4.26) would become $\tilde{\Gamma}(D) = 1$. This would lead to

$$\begin{aligned}\tilde{\Gamma}(W) &= \Gamma(W) + [\tilde{\Gamma}(W) - \Gamma(W)]\frac{r_+(W) - r_-(W)}{\rho_+ - \rho_-}\\ &= \Gamma(W) + \frac{r_+(W) - r_-(W)}{\rho_+ - \rho_-};\end{aligned} \tag{5.2}$$

this equation may be verified by noting that $\Gamma_1(W) \equiv \tilde{\Gamma}(W) - \Gamma(W)$ and $\Gamma_2(W) \equiv r_+(W) - r_-(W)$ both satisfy (4.15), (4.18), (4.19), and the homogeneous versions of (4.16) and (4.17), and that (5.2) holds for $W = D$, so that the reduction procedure (4.20) yields (5.2) for all W. As a consequence, from (4.27) and the corresponding equation $\Delta_{\widetilde{\mu}}(f) = \tilde{\Gamma}(W)$, the translation sums satisfy (5.1).

We next turn to the calculation of the translation sums associated with the particular family of functions $f_n(\eta) = (\eta(0) - \rho_-)(\eta(n) - \rho_+)$, $n \geq 1$. Let us write

$$\Phi_n \equiv \Delta_{\widehat{\mu}}(f_n) = \sum_{i=-\infty}^{\infty} \langle (\eta(i) - \rho_-)(\eta(i+n) - \rho_+) \rangle_{\widehat{\mu}}. \tag{5.3}$$

Since $f_n(\eta) - f_{n+1}(\eta) = (\eta(n+1) - \eta(n))(\rho_- - \eta(0)) = (T^n F)(\eta)$, where $F(\eta) = (\eta(1) - \eta(0))(\rho_- - \eta(-n))$, (4.28) implies that

$$\Phi_n - \Phi_{n+1} = (\rho_+ - \rho_-)[\rho_- - \langle \eta(-n) \rangle_{\widehat{\mu}}] = (\rho_+ - \rho_-)[\langle \eta(n) \rangle_{\widehat{\mu}} - \rho_+], \tag{5.4}$$

where we have used the symmetry (4.10) of the profile. Thus

$$\Phi_n - \lim_{k\to\infty} \Phi_k = (\rho_+ - \rho_-) \sum_{i=n}^{\infty} [\langle \eta(i) \rangle_{\widehat{\mu}} - \rho_+]. \tag{5.5}$$

But Φ_1 can be evaluated from (4.15)–(4.18), (4.22), and (4.26):

$$\Phi_1 = (\rho_+ - \rho_-) \sum_{i=n}^{\infty} [\langle \eta(i) \rangle_{\widehat{\mu}} - \rho_+]. = \frac{p\rho_- - q\rho_+}{p - q} - \rho_+\rho_-. \tag{5.6}$$

The right hand side of (5.5) in the case $n = 1$ can be evaluated from formulas (4.1)–(4.5) of [17] and shown also to be given by (5.6); thus $\lim_{k\to\infty} \Phi_k = 0$ (as one would expect) and hence

$$\Phi_n = (\rho_+ - \rho_-) \sum_{i=n}^{\infty} [\langle \eta(i) \rangle_{\widehat{\mu}} - \rho_+]. \tag{5.7}$$

To understand the consequences of equation (5.7) we recall [17] that in general the profile $\langle \eta(n) \rangle_{\widehat{\mu}}$ decays exponentially to its asymptotic value: for $n \gg 1$,

$$\langle \eta(n) \rangle_{\widehat{\mu}} - \rho_+ \simeq C n^{\gamma} e^{-\alpha n}, \tag{5.8}$$

where $C > 0$ if $q/p < x^*$, $C = 0$ (and in fact $\langle \eta(n) \rangle_{\widehat{\mu}} = \rho_+$ for all $n > 0$) if $q/p = x^*$, and $C < 0$ if $q/p > x^*$, with $x^* = \rho_-(1-\rho_+)/\rho_+(1-\rho_-)$ (explicit values of α and γ are given in [17]). Thus (5.7) implies that, except on the line $q/p = x^*$, Φ_n decreases exponentially to 0 as $n \to \infty$, with characteristic length α^{-1}. Since the Φ_n are intrinsic properties, this gives an intrinsic characteristic size for the shock.

A somewhat surprising aspect of the shock profile, found in [16] for $p = 1$ and in [17] for all p satisfying $q/p < x^*$, is the "overshoot": $\langle \eta(n) \rangle_{\widehat{\mu}} > \rho_+$ for $n \geq 1$ (corresponding to $C > 0$ in (5.8)). A natural question is whether this overshoot is a by-product of the choice to view the shock from the second class particle, and might be eliminated by the adoption of another viewpoint. While we cannot answer this completely, (5.7) implies that for $q/p < x^*$ there is no viewpoint from which the shock measure is described by a product measure with density increasing monotonically from ρ_- to ρ_+, since such a measure would lead to a negative Φ_n. Similarly, for $q/p \neq x^*$ there is no point of view from which the shock measure would be a product measure with density ρ_+ to the right of the origin and density ρ_- to the left of the origin, since for such a measure, all the Φ_n would vanish.

6. Shocks in other ASEP models

Several different models, with ASEP dynamics but with specific initial conditions, boundary conditions or minor modifications, have been shown to give rise to shocks:

1. An infinite system with, as initial condition, a product measure with density ρ_+ to the right and ρ_- to the left of the origin. This is the case discussed in the introduction and in [14,15,17].
2. A system with ring geometry and a blockage bond, at which the jump rate is less than that at other bonds in the system [25].
3. A system with ring geometry and a slow second class particle [26,27].

4. A system with open boundary conditions on its first order transition line, discussed below and in [28].

Note that in examples 2 and 3 the shock in the ring geometry is caused by the blocking bond or particle but is located far from it, in a region where the usual ASEP dynamics holds. We believe that, when the size of the system goes to infinity, all of these different models lead in fact to shocks which are shift equivalent, i.e., that one may choose for each shock a viewpoint such that, seen from this viewpoint, all the shocks are described by the same ramp measure. (The relation of shocks observed in similar but more complicated systems [29] to those of the ASEP remains to be investigated.) This equivalence is not obvious *a priori* but can be verified, at least in cases where exact expressions permit exact computations, by showing that all the translation sums $\Delta(f)$ are identical to those found in Section 4.

We illustrate this by computing the translation sums for the ASEP with open boundary conditions, in the totally asymmetric ($p = 1$) case. This is a system with N sites, in which particles enter the system at site 1 with rate α and escape at site N with rate β. In [28], it was shown that the steady state of this system can be fully described by an algebra rather similar to the one discussed in Section 4 (the weight of a configuration can be written as the matrix element of a matrix product where a matrix D is used when a site is occupied and a matrix E when the site is empty). Using this algebra, in principle, any expectation can be calculated for systems of any size and for any values of the parameters α and β. In particular, it was shown that if $\alpha < \min\{\beta, 1/2\}$ the system is in a low density phase, described in the bulk by a product measure with density α, while if $\beta < \min\{\alpha, 1/2\}$ the system is in a similar high density phase, with bulk density $1 - \beta$.

On the line $\alpha = \beta < 1/2$ there is a first order phase transition between these two phases. When the system is on this phase coexistence line one can show [28] that for large

N and for i far from the boundaries ($i \gg 1$ and $N-i \gg 1$) the profile is linear in i/N,

$$\langle \eta(i) \rangle = \frac{i(1-\alpha) + (N-i)\alpha}{N} + O\left(\frac{1}{N}\right), \tag{6.1}$$

and by a similar computation, not given in [28], that nearest neighbor two point correlations behave in the same way:

$$\langle \eta(i)\eta(i+1) \rangle = \frac{i(1-\alpha)^2 + (N-i)\alpha^2}{N} + O\left(\frac{1}{N}\right). \tag{6.2}$$

The interpretation of this behavior is that there is a shock in the system, separating a region of low density α to its left and of high density $1-\alpha$ to its right, and that the linear dependence on i/N arises because the shock is equally likely to be anywhere in the system (apart from some corrections of order $1/N$ due to boundaries). It seems natural to suppose as well that the structure of the shock is independent of its location, so that if we determine the shock position in the configuration η by some function $\ell(\eta)$ as described in Section 2, and write μ_m for the system measure conditioned on $\ell(\eta) = m$, then the measures μ_m are just translates of one another, $\mu_m = T^{m-m'}\mu_{m'}$, for N large and m far from the boundaries.

If such an interpretation is correct, then the expectation of any function of the occupation numbers near site i should for large N show linear behavior in i/N similar to that of (6.1) and (6.2). The leading terms give no information about the shock structure, but it follows also from the picture described above that there is a probability of order $1/N$ that the shock is located near site i, so that to describe the shock properties one must compute the terms of order $1/N$ in expressions like (6.1) and (6.2). However, these order $1/N$ terms may contain not only contributions coming from the events where the shock is in the neighborhood of i, but also contributions from boundary effects, in particular from the precise definition of the coordinate i.

We now show that we may calculate these order $1/N$ terms and thus obtain additional support for this picture. If $f = f(\eta(i+1), \eta(i+2), ...\eta(i+k))$ is a function depending on the occupation numbers of sites between $i+1$ and $i+k$, with $k \ll N$, then it can be shown that for i far enough from the boundaries the expectation of f takes the form

$$\langle f \rangle = \frac{i\langle f\rangle_{\nu_{1-\alpha}} + (N+b-i-k)\langle f\rangle_{\nu_\alpha}}{N+b} + \frac{\gamma(W)}{N+b} + o\left(\frac{1}{N}\right) . \tag{6.3}$$

Here b is a constant of order 1, independent of f and i, and W is the linear combination of words of length k associated to f, as in Section 4 (see (4.11)). Equation (6.3) defines $\gamma(W)$; the correct choice of b, which may be regarded as representing the boundary effects, guarantees that $\gamma(W)$ is independent of i. Equation (6.3) is trivial for $f = 1$ (corresponding to $k = 0$) and leads to

$$\gamma(1) = 0. \tag{6.4}$$

For $f = \eta(i)$, (6.3) may be verified from the asymptotics derived in [28]; this calculation leads to explicit values for b and of $\gamma(D)$, but the values of these parameters are not needed for the calculation of the translation sums. For other words W one can use the algebra of [28] to verify (6.3) by induction on the length of W, and also show that the $\gamma(W)$ have the following properties:

$$\gamma(W(D+E)) = \gamma(W) + r_-(W) \tag{6.5}$$

$$\gamma((D+E)W) = \gamma(W) + r_+(W) \tag{6.6}$$

$$\gamma(a_1W_1 + a_2W_2) = a_1\gamma(W_1) + a_2\gamma(W_2) \tag{6.7}$$

$$\gamma(W_1DEW_2) = \alpha(1-\alpha)[\gamma(W_1DW_2) + \gamma(W_1EW_2)] \tag{6.8}$$

Here r_+ and r_- are defined with the densities $\rho_+ = 1-\alpha$ and $\rho_- = \alpha$. Thus the $\gamma(W)$ satisfy the same rules as $\Gamma(W)$ of Section 4, and one can conclude (see the argument following (5.2)) that the general expression of these $\gamma(W)$ is given by

$$\gamma(W) = \Gamma(W) + \gamma(D)\frac{\mu_+(W) - \mu_-(W)}{\rho_+ - \rho_-} \tag{6.9}$$

Now suppose that the function f satisfies $\langle f\rangle_{\nu_\alpha} = \langle f\rangle_{\nu_{1-\alpha}} = 0$, so that (6.3) becomes

$$\langle f\rangle = \frac{\gamma(W)}{N} + o(1/N). \tag{6.10}$$

This result has a simple interpretation in terms of the heuristic picture, described above, that the steady state of the open system contains a shock equally likely to be located at any site and with structure independent of its location. Thus in computing $\langle f\rangle$ we are averaging the expected value in the measure conditioned on the shock position being m, $\langle f\rangle_{\mu_m}$, over all N possible shock positions. Since the measures μ_m differ only by translation this is, up to a factor $1/N$, just the calculation of the translation sum $\Delta_{\mu_m}(f)$ for any fixed m (N must be large enough so that μ_m is effectively an infinite volume measure). Thus (6.10) implies that $\Delta_{\mu_m}(f) = \gamma(W)$. But (6.9) implies that for such a function f, $\gamma(W) = \Gamma(W)$, so that from (4.27), $\Delta_{\mu_m}(f) = \Delta_{\widehat{\mu}}(f)$: the translation sums for the shock in the open system are the same as those calculated in Section 4. Thus these two shocks are shift equivalent.

Acknowledgments

We thank P. Ferrari, Y. Sinai, and G. Steiff for helpful discussions. This work was supported in part by NSF grant DMS 95–04556, NSF grant DMR 95–23266, and AFSOR grant 4–26435. BD thanks the Institute for Advanced Study and JLL and ERS the Institut des Hautes Etudes Scientifiques for hospitality while portions of this work were carried out; BD and JLL would also like to thank DIMACS and its supporting agencies, the NSF under contract STC-91-19999 and the N. J. Commission on Science and Technology.

References

[1] D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, New York, 1969).

[2] H.-O. Georgii, *Gibbs Measures and Phase Transitions* (Walter de Gruyter, Berlin, 1988).

[3] R. Dobrushin, Gibbs states describing coexistence of phases for a three-dimensional Ising model, *Th. Prob. Appl.* **17**:582–600 (1972).

[4] G. Gallavotti, The phase separation line in the two-dimensional Ising model, *Commun. Math. Phys.* **27**:103–136 (1972).

[5] A. Messager and S. Miracle-Sole, Correlation functions and boundary conditions in the Ising ferromagnet, J. Stat. Phys **17**:245–262 (1977).

[6] J. Bricmont, J. L. Lebowitz, and C. E. Pfister, On the local structure of the phase separation line in the two-dimensional Ising system, *J. Stat. Phys.* **26**:313–332 (1981).

[7] F. Spitzer, Interaction of Markov processes, *Advances in Math.* **5**:246-290 (1970).

[8] T. M. Liggett, *Interacting Particle Systems* (Springer-Verlag, New York, 1985). See also references therein.

[9] H. Rost, Nonequilibrium behavior of many particle process: density profiles and local equilibria, *Z. Wahrsch. Verw. Gebiete* **58**:41–53, (1981).

[10] E. D. Andjel and M. E. Vares, Hydrodynamical equations for attractive particle systems on $\mathbb{Z}$, *J. Stat. Phys.* **47**:265–288 (1987).

[11] F. Razakhanlou, Hydrodynamic limit for attractive particle systems on $\mathbb{Z}^d$, *Commun. Math. Phys.* **140**:417–448 (1991).

[12] E. D. Andjel, M. Bramson, and T. M. Liggett, Shocks in the asymmetric exclusion process, *Probab. Theory Relat. Fields* **78**:231–247 (1988).

[13] P. Ferrari and L. R. G. Fontes, Shock fluctuations in the asymmetric simple exclusion process, *Probab. Theory Relat. Fields* **99**:305–319 (1994).

[14] P. Ferrari, C. Kipnis, and E. Saada, Microscopic structure of traveling waves in the asymmetric simple exclusion, *Ann. Probab.* **19**:226–244 (1991).

[15] P. Ferrari, Shock fluctuations in asymmetric simple exclusion, *Probab. Theory Relat. Fields* **91**:81–101 (1992).

[16] B. Derrida, S. A. Janowsky, J. L. Lebowitz, and E. R. Speer, Exact solution of the totally asymmetric simple exclusion process: shock profiles, *J. Stat. Phys.* **73**:813–842 (1993).

[17] B. Derrida, J. L. Lebowitz, and E. R. Speer, Shock profiles in the asymmetric simple exclusion process in one dimension, *J. Stat Phys.* **89**:135–167 (1997).

[18] Work in preparation.

[19] B. M. Gurevič and V. I. Oseledec, Gibbs distributions and dissipativeness of U-diffeomorphisms, *Soviet Math. Dokl* **14**:570–573 (1973).

[20] D. J. Aldous and H. Thorisson, Shift-coupling, *Stoch. Proc. Appl.* **44**:1–14 (1993).

[21] H. Thorisson, On time- and cycle-stationarity, *Stoch. Proc. Appl.* **55**:183–209 (1995).

[22] H.-O. Georgii, Orbit coupling, *Ann. Inst. Henri Poincaré* **33**:253–268 (1997).

[23] L. N. Vasershtein, Markov processes over denumerable products of spaces, describing large systems of automata, *Problems Inform. Transmission,* **5**:47–52 (1969).

[24] S. Sandow, Partially asymmetric exclusion process with open boundaries, *Phys. Rev. E* **50**:2660–2667 (1994).

[25] S. Janowsky and J. L. Lebowitz, Finite size effects and shock fluctuations in the asymmetric simple exclusion process, *Phys. Rev. A* **45**, 618–625 (1992), and Exact results for the asymmetric simple exclusion process with a blockage, *J. Stat Phys.* **77**:35–51 (1994).

[26] B. Derrida, Systems out of equilibrium: some exactly solved models, in *STATPHYS:*

the 19th IUPAP International Conference on Statistical Physics, ed. H. Bailin (World Scientific, Singapore, 1996).

[27] K. Mallick, Shocks in the asymmetric simple exclusion model with an impurity, *J. Phys. A* **29**:5375–5386 (1996).

[28] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, An exact solution of a 1D asymmetric exclusion model using a matrix formulation, *J. Phys. A* **26**:1493–1517 (1993).

[29] P. F. Arndt, T. Heinzel, and V. Rittenberg, First-order phase transitions in one-dimensional steady states, preprint SISSA Ref. 39/37/EP, cond-mat/9706114.

Figure captions

Figure 1. Construction of the intrinsic shock location function $\ell(\eta)$.

Figure 2. Shock profiles for $p = 1$, $\rho_+ = 0.7$, $\rho_- = 0.2$, from three different viewpoints: (i) the second class particle (diamonds, solid line), (ii) ℓ defined with $\rho_* = \pi^{-1}\rho_+ + (1 - \pi^{-1})\rho_-$ (plusses, dashed line), (iii) ℓ defined with $\rho_* = (1 - \pi^{-1})\rho_+ + \pi^{-1}\rho_-$ (squares, dotted line).

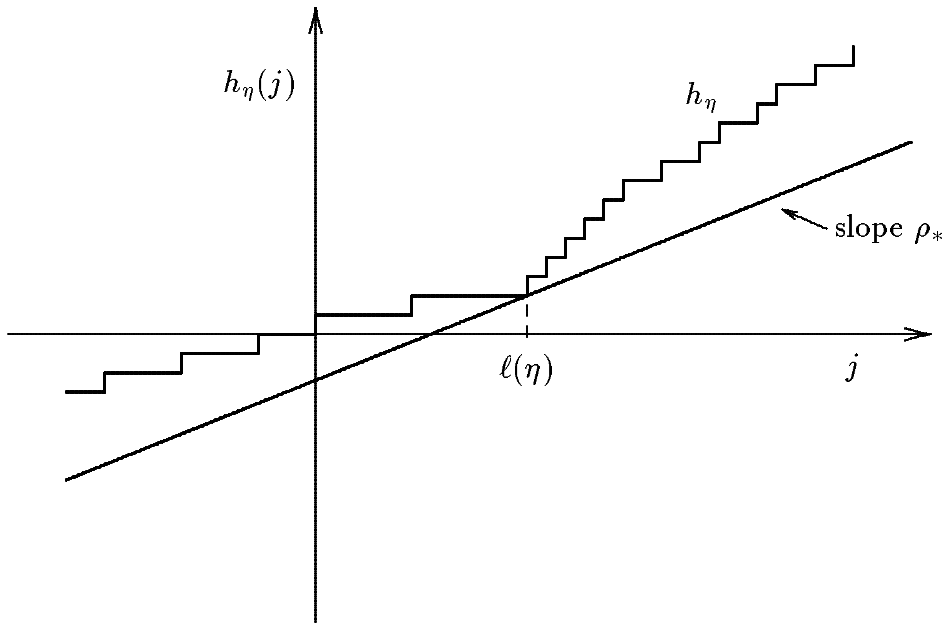

Figure 1

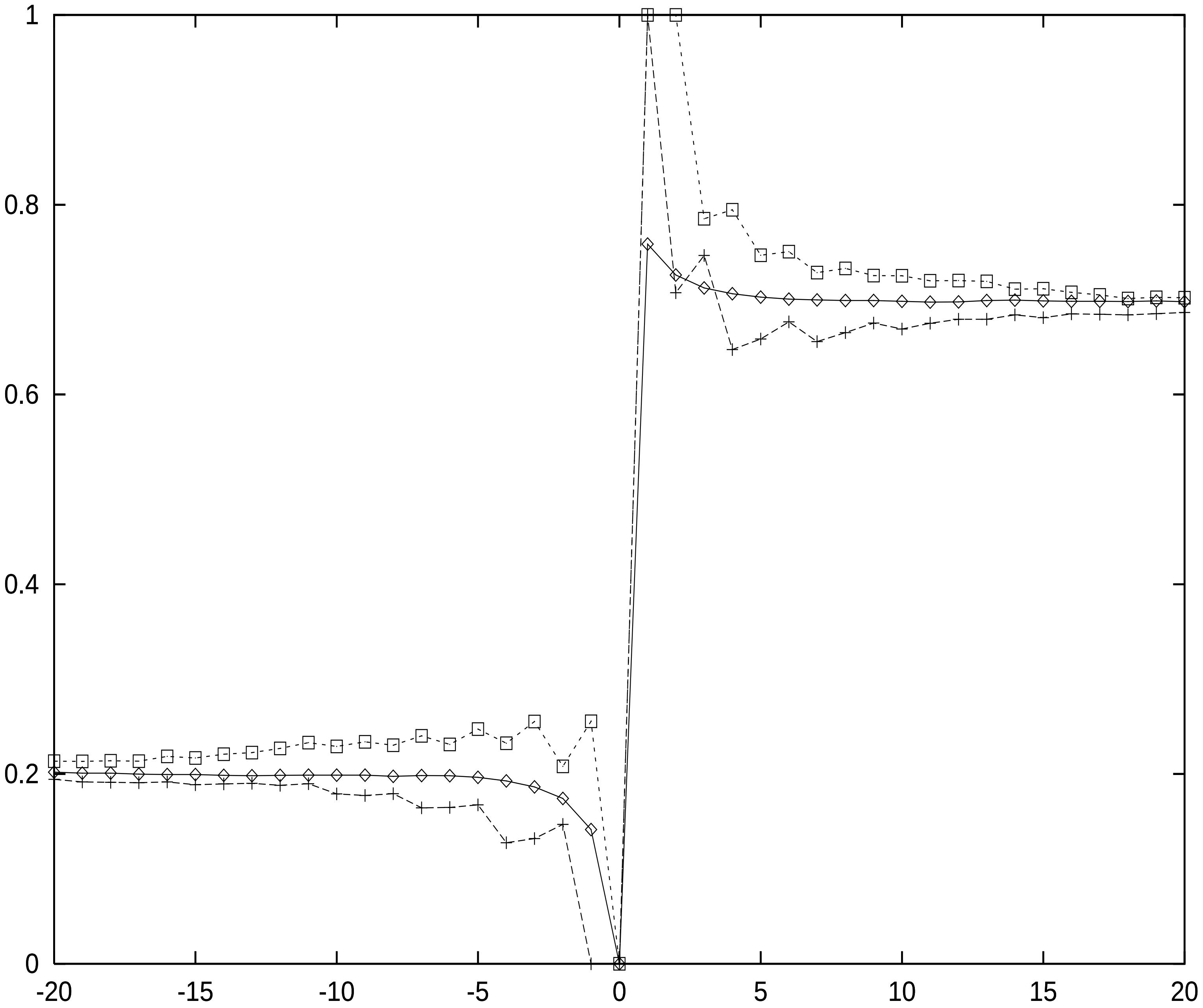

Figure 2