Exact Large Deviation Function in the Asymmetric Exclusion Process

Bernard Derrida† and Joel L. Lebowitz‡

† *Laboratoire de Physique Statistique, 24 rue Lhomond, 75231 Paris Cedex 05, France;* email derrida@physique.ens.fr

‡ *Department of Mathematics, Rutgers University, New Brunswick, NJ 08903;* email lebowitz@math.rutgers.edu

By an extension of the Bethe ansatz method used by Gwa and Spohn, we obtain an exact expression for the large deviation function of the time averaged current for the fully asymmetric exclusion process in a ring containing N sites and p particles. Using this expression we easily recover the exact diffusion constant obtained earlier and calculate as well some higher cumulants. The distribution of the deviation y of the average current is, in the limit $N \to \infty$, skew and decays like $\exp -(Ay^{5/2})$ for $y \to +\infty$ and $\exp -(A'|y|^{3/2})$ for $y \to -\infty$. Surprisingly, the large deviation function has an expression very similar to the pressure (as a function of the density) of an ideal Bose or Fermi gas in $3d$.

PACS:02.50.+s, 05.40.+j, 82.20-w

In the last decade, it has become clear that questions as different as $1d$ turbulence, growth of interfaces or directed polymers in a random medium in dimension $1+1$ dimension were different versions of the same problem described by the noisy Burgers or equivalently the Kardar-Parisi-Zhang (KPZ) equation [1,2]. One of the simplest lattice versions of this problem is the asymmetric exclusion process (ASEP) which also has a long history both in the mathematical and in the physical literature [3–5]. It describes a driven lattice gas with hard core exclusion and under suitable scalings the evolution of its macroscopic density is governed by a noisy Burgers equation in $1d$ which is equivalent to the KPZ equation.

Many properties of the ASEP have been calculated (steady states, mass gaps, diffusion constants) by various methods including exact solutions [6–13] and compared [1,14] with the predictions of the replica approach [15–17] or numerical simulations [18–20] for directed polymers or growth problems.

Here, extending the Bethe ansatz approach of Gwa and Spohn [8], we obtain the expression for a large deviation function which gives the exact probability for the time averaged current to take any, typical or untypical, value. Similar distributions have been studied numerically in the past [18–20], in particular to better understand the $n \to 0$ limit in the replica approach of directed polymers in a random medium. Other recent results include the numerical calculation of the distribution of the density for a randomly driven Burgers equation [21]. The results presented below are, to our knowledge, the first exact determination of a non-trivial probability distribution for a fluctuating quantity described by the noisy Burgers equation or equivalently the KPZ equation. Through the known equivalences between the different problems mentioned above, our results could easily be used to obtain the whole distribution of the ground state energy of a directed polymer in a random medium in dimension 1+1 for a finite lateral geometry.

We consider here a totally ASEP consisting of p particles on a ring of N sites. During any time interval dt, each particle jumps with probability dt to its right if the target site is empty. The unique stationary state of this system is one in which all $\binom{N}{p}$ different configurations C have equal weight [6]. We call Y_t the *total* distance covered by *all the particles* between time 0 and time t (thus $Y_0 = 0$). The probability $P_t(C)$ of finding the system in configuration

C at time t satisfies

$$\frac{d}{dt}P_t(C) = \sum_{C'}[M_0(C,C') + M_1(C,C')]P_t(C')$$

where $M_1(C,C')dt$ is the probability of going from C' to C during a time interval dt, increasing Y by 1, while M_0 is a diagonal matrix,

$$M_0(C,C) = -\sum_{C' \neq C} M_1(C',C)$$

Let $P_t(C,Y)$ be the joint probability of being, at time t, in the configuration C and having $Y_t = Y$, some non-negative integer. Define also $F_t(C)$ as $\sum_{Y=0}^{\infty} \exp(\alpha Y) P_t(C,Y)$. Then

$$\frac{\partial F_t(C)}{\partial t} = \sum [M_0(C,C') + e^{\alpha} M_1(C,C')]F_t(C')$$

Summing $F_t(C)$ over C we obtain $\langle e^{\alpha Y_t} \rangle$. It is easy to see that for large t

$$\langle e^{\alpha Y_t} \rangle \sim e^{\lambda(\alpha) t} \tag{1}$$

where $\lambda(\alpha)$ is the largest eigenvalue of the matrix $M_0 + e^{\alpha} M_1$ (we have omitted the dependence of λ on p and N). By $\sim$ we mean: take the logarithm of both sides of (1), divide by t and let $t \to \infty$ to obtain an equality. While (1) holds for every initial configuration, it is sometimes convenient to think of the initial state $P_0(C)$ as the stationary one.

In the present paper we show that the exact expression of $\lambda(\alpha)$, for arbitrary p and N, is given by the power series in α obtained by eliminating B between the following two series

$$\lambda(\alpha) = -p \sum_{q=1}^{\infty} B^q \frac{(Nq-2)!}{(pq)!\ (Nq-pq-1)!} \tag{2}$$

$$\alpha = -\sum_{q=1}^{\infty} B^q \frac{(Nq-1)!}{(pq)!\ (Nq-pq)!} \tag{3}$$

These series have a non-zero radius of convergence in B so $\lambda(\alpha)$ is analytic in α at least near $\alpha = 0$. Moreover, by the Perron-Frobenius theorem there is no crossing of the largest eigenvalue of the matrix $M_0 + e^{\alpha} M_1$ for real α and fixed finite N. Hence (2) and (3) determine $\lambda(\alpha)$ on the whole real axis as the analytic continuation of the expression in the vicinity of $\alpha = 0$. In the rest of this work, we first show how (2, 3) can be obtained using the Bethe ansatz. Then we investigate some consequences of this exact expression.

The Bethe ansatz we use is very similar to that used by Gwa and Spohn [8], except that here there is an extra parameter α. Specifying a configuration C by the positions $1 \leq x_1 < x_2 < ... < x_p \leq N$ of the p particles we write an eigenvector of the matrix $M_0 + e^{\alpha} M_1$ in the form

$$\sum_Q A_Q \prod_{j=1}^{p} [z_{Q(j)}]^{x_j} \tag{4}$$

where the sum is over all the permutations Q of $1, 2, ..., p$ and the p parameters z_k satisfy the following equations

$$\frac{(e^{\alpha} - z_k)^p}{z_k^N} = (-)^{p-1} \prod_{j=1}^{p} (e^{\alpha} - z_j) \tag{5}$$

$$z_1 z_2 ... z_p = 1 \tag{6}$$

The derivation of (5,6) is almost identical to that given by what Gwa and Spohn in [8] and we refer to their work.

For any solution $\{z_k\}$ of (5,6), expression (4) gives an eigenvector of the matrix $M_0 + e^\alpha M_1$ with eigenvalue $\lambda(\alpha)$ given by

$$\lambda(\alpha) = e^\alpha \left[\frac{1}{z_1} + \frac{1}{z_2} + ... + \frac{1}{z_p} \right] - p \tag{7}$$

The difficulty is that there are in general many solutions of (5,6) corresponding to different eigenvalues of the matrix $M_0 + e^\alpha M_1$. The one we need to choose is the largest eigenvalue $\lambda(\alpha)$ which, as already noted, can be followed when α varies, from the unique eigenvalue which vanishes when $\alpha \to 0$. Thus $\lambda(\alpha)$ corresponds to the solution such that all the $z_k \to 1$ as $\alpha \to 0$.

To see why (2,3) are equivalent to (5,6,7) when one chooses the above solution, we consider a function $h(z)$ analytic in the circle $|z - e^\alpha| < 2\epsilon$. It is then an easy consequence of the Cauchy theorem that

$$\sum_{1 \le k \le p} h(Z_k) = \oint \frac{dz}{2\pi i z} h(z) \frac{(z - e^\alpha)^{p-1}[pz - N(z - e^\alpha)]}{(z - e^\alpha)^p - Az^N} \tag{8}$$

where the contour is the circle $|z - e^\alpha| = \epsilon$, the parameter A is such that $|A| \ll \epsilon^p e^{-N\alpha}$ and the Z_k are the p solutions of

$$(z - e^\alpha)^p = Az^N \tag{9}$$

which tend to e^α as $A \to 0$. As (9) has a form identical to (5), one can use (8) to calculate $\lambda(\alpha)$ (by choosing $h(z) = -1 + e^\alpha / z$) and this gives (3) in terms of $B = Ae^{(N-p)\alpha}$. The constant B can then be determined by requiring that (6) is also satisfied. To do that one can use again the identity (8) for the function $h(z) = \log z$ and this gives (3).

From the exact expressions (2,3), one can expand $\lambda(\alpha)$ in a power series in α. In particular, one can check that

$$\begin{aligned} \lambda(\alpha) = & \frac{p(N-p)}{N-1}\alpha + \\ & \frac{p\ (p!)^2\ (N-p)!^2\ (2N-3)!}{(2p!)\ (N-1)!^2\ (2N-2p-1)!}\alpha^2 + O(\alpha^3) \end{aligned} \tag{10}$$

and using (1), this gives

$$\lim_{t \to \infty} \frac{\langle Y_t \rangle}{t} = \left. \frac{d\lambda(\alpha)}{d\alpha} \right|_{\alpha=0} = \frac{p(N-p)}{N-1} \tag{11}$$

which is known to be a simple consequence of the fact that in the steady state all configurations have equal weight [6,7]. (Note that if $P_0(C)$ was the stationary measure, (11) would in fact be valid for all t.) The expansion (10) also gives the diffusion constant

$$\lim_{t \to \infty} \frac{\langle Y_t^2 \rangle - \langle Y_t \rangle^2}{t} = \left. \frac{d^2\lambda(\alpha)}{d\alpha^2} \right|_{\alpha=0} \tag{12}$$

which agrees (up to a trivial factor p^2 due to a slight change in the definition of Y_t) with the expression obtained by an extension [10] of the matrix method [9]. Exact expressions of higher cumulants of Y_t may be obtained in a similar way from (1,2,3) by expanding $\lambda(\alpha)$ to higher orders in α. For example, the third derivative of λ at $\alpha = 0$ gives

$$\lim_{t\to\infty} \frac{\langle Y^3\rangle - 3\langle Y\rangle^2\langle Y\rangle + 2\langle Y\rangle^3}{t} = \tag{13}$$
$$\frac{3p(N-p)^3}{(2N-1)(N-1)} \binom{2N-1}{2p}^2 \binom{N-1}{p}^{-4}$$
$$-\frac{4p(N-p)^3}{(3N-1)(N-1)} \binom{3N-1}{3p} \binom{N-1}{p}^{-3}$$

For large positive α (i.e. $\alpha \gg 1$), the solutions z_k of (5,6), for $1 \le k \le p$ are given by

$$z_k = e^{(2k-p-1)\pi i/N} + O(e^{-\alpha})$$

and this leads to

$$\lambda(\alpha) = e^{\alpha}\, \frac{\sin(p\pi/N)}{\sin(\pi/N)} + O(1) \tag{14}$$

For large negative α (i.e. $-\alpha \gg 1$), one of the z_k becomes large (say $z_1 \simeq e^{(1-p)\alpha}$) and the $p-1$ other z_k become small ($z_k \simeq e^{\alpha}$ for $2 \le k \le p$). This leads to

$$\lambda(\alpha) \simeq \left\{ (Font) LaTeX Font Warning : Command invalid in math mode 1012 ptc \begin{array}{ll} -1 + e^{p\alpha} & \text{for} \quad p < N-p \\ -1 + 2e^{p\alpha} & \text{for} \quad p = N-p \\ -1 + e^{(N-p)\alpha} & \text{for} \quad p > N-p \end{array} \right. \tag{15}$$

(We can understand the limit $\alpha \to -\infty$ in (15) by noting that the probability of not making any jump during the time interval t is, for a configuration in which all particles form one cluster, just $\exp(-t)$.)

In the limit of large N and p, keeping the ratio $p/N = \rho$ fixed, one finds for the first four cumulants of Y (by calculating the first four derivatives of $\lambda(\alpha)$ at $\alpha = 0$)

$$\lim_{t\to\infty} \frac{\langle Y\rangle}{t} \simeq N\rho(1-\rho) \tag{16}$$

$$\lim_{t\to\infty} \frac{\langle Y^2\rangle - \langle Y\rangle^2}{t} \simeq N^{3/2}[\rho(1-\rho)]^{3/2}\frac{\sqrt{\pi}}{2} \tag{17}$$

$$\lim_{t\to\infty} \frac{\langle Y^3\rangle - 3\langle Y^2\rangle\langle Y\rangle + 2\langle Y\rangle^3}{t}$$
$$\simeq N^3[\rho(1-\rho)]^2 \left(\frac{3}{2} - \frac{8}{3^{3/2}}\right)\pi \tag{18}$$

$$\lim_{t\to\infty} \frac{\langle Y^4\rangle - 3\langle Y^2\rangle^2 - 4\langle Y^3\rangle\langle Y\rangle + 12\langle Y^2\rangle\langle Y\rangle^2 - 6\langle Y\rangle^4}{t}$$
$$\simeq N^{9/2}[\rho(1-\rho)]^{5/2} \left(\frac{15}{2} + \frac{9}{2^{1/2}} - \frac{24}{3^{1/2}}\right)\pi^{3/2} \tag{19}$$

In fact, in this limit $N \to \infty$ with $\rho = p/N$ fixed, the expressions (2) and (3) can be rewritten as

$$\sqrt{\frac{2\pi N^3}{\rho(1-\rho)}} \left[\lambda(\alpha) - \alpha\, \frac{p(N-p)}{N}\right] \simeq$$
$$G(\beta) \;\equiv\; \sum_{q\ge 1} \frac{(-)^{q+1}C^q}{q^{5/2}} \tag{20}$$

where

$$\alpha \sqrt{2\pi\rho(1-\rho)N^3} \ \simeq \ \beta \ \equiv \ \sum_{q\geq 1} \frac{(-)^{q+1}C^q}{q^{3/2}} \tag{21}$$

and $C = -BN^N p^{-p}(N-p)^{-N+p}$. This shows that for N large and β fixed, i.e. for $\alpha \sim N^{-3/2}$, the difference $\lambda(\alpha) - \alpha p(N-p)/N$ takes a scaling form

$$\lambda(\alpha) - \alpha \ \frac{p(N-p)}{N} \simeq \sqrt{\frac{\rho(1-\rho)}{2\pi N^3}} G\left(\alpha\sqrt{2\pi\rho(1-\rho)N^3}\right) \tag{22}$$

The two sums which define $G(\beta)$ in the right hand side of (20,21) are very reminiscent of the pressure and of the density of an Ideal Bose or Fermi gas (depending on the sign of C) in 3 dimension [22]: positive C and α correspond to the Fermi gas and negative C and α correspond to the Bose gas.

The range $N^{-3/2} \ll \alpha \ll 1$ (i.e. large positive C) correponds to the high pressure or low temperature limit of a Fermi gas and this analogy leads to

$$\lambda(\alpha) - \alpha \ \frac{p(N-p)}{N} \simeq \frac{(3\pi)^{2/3}}{5}[\rho(1-\rho)]^{4/3} N \alpha^{5/3} \tag{23}$$

Negative α can also be analysed: when C varies from 0 to -1 in (20,21), α decreases (this corresponds to a Bose gas in its high temperature phase) and α has a finite limit α_0 as $C \to -1$. To decrease further α, it turns out that one has to increase C from -1 to 0 and to replace the expressions (20) and (21) of $G(\beta)$ and β by their analytic continuations [23]

$$G(\beta) = \frac{4}{3}\sqrt{\pi}[-\log(-C)]^{3/2} + \sum_{q\geq 1} \frac{(-)^{q+1}C^q}{q^{5/2}} \tag{24}$$

$$\beta = -4\sqrt{\pi}[-\log(-C)]^{1/2} + \sum_{q\geq 1} \frac{(-)^{q+1}C^q}{q^{5/2}} \tag{25}$$

This analytic continuation is a major difference with the usual Bose-Einstein condensation where C remains fixed at the value -1 in the whole low temperature phase. Here, for $\alpha < \alpha_0$, one of the z_k becomes much larger than the others but, in contrast to the Bose condensate, the chemical potential C in (24) keeps varying with α. For the range $N^{-3/2} \ll -\alpha \ll 1$, this leads to

$$\lambda(\alpha) - \alpha \ \frac{p(N-p)}{N} \simeq \frac{-1}{24}[\rho(1-\rho)]^2 N^3 \alpha^3 \tag{26}$$

The knowlegde of the function $\lambda(\alpha)$ determines (through $y = d\lambda/d\alpha - p(N-p)/N$ and $f(y) = \lambda - \alpha[y + p(N-p)/N]$) the large deviation function $f(y)$ defined here as

$$f(y) = \lim_{t\to\infty} \frac{1}{t} \log\left\{ \text{Prob}\left(\frac{Y_t}{t} - N\rho(1-\rho) = y\right)\right\} \tag{27}$$

From (16,17,22,23,26) one can show that for N large, $f(y)$ takes the scaling form for $|y| \ll N$

$$f(y) \simeq \sqrt{\frac{\rho(1-\rho)}{\pi N^3}} H\left(\frac{y}{\rho(1-\rho)}\right) \tag{28}$$

where the function H has the following properties

$$H(y) \simeq -(y-1)^2 \quad \text{for} \quad |y-1| \ll 1 \tag{29}$$

$$H(y) \simeq -(2\sqrt{3}/(5\sqrt{\pi}))\ y^{5/2} \quad \text{for} \quad y \to +\infty \tag{30}$$

$$H(y) \simeq -(4\sqrt{2\pi}/3)\ |y|^{3/2} \quad \text{for} \quad y \to -\infty \tag{31}$$

Note that the difference between $\langle Y_t \rangle / t = p(N-p)/(N-1)$ and $N\rho(1-\rho)$ used in (27) is responsible for the fact that the maximum of $f(y)$ is at $y=1$ as in (29). The scaling form (28) was obtained (20-21) for $\alpha \sim N^{-3/2}$ and this means (27-31) that it is valid for y of order 1.

For y of order N, one expects from (14,15) the following two scaling forms for $f(y)$:

$$f(y) \simeq NH_+\left(yN^{-1}\right) \quad \text{for} \quad y > 0 \tag{32}$$

$$f(y) \simeq H_-\left(yN^{-1}\right) \quad \text{for} \quad y < 0 \tag{33}$$

As $y/(N\rho)$ is the deviation (from its most probable value $1-\rho$) in the velocity of the particles, the extra factor N in (32) has a simple interpretation: to speed up the traffic, one needs to speed up all the particles whereas to slow down the traffic, it is sufficient to slow down a single particle. Note that if one considers that the scaling forms (28,30,31) remain valid for $1 \ll |y| \ll N$, it is possible to obtain from (30,31) the small z behavior of $H_+(z)$ and $H_-(z)$.

The distribution of y is skew (13,18) and for large y (but still $|y| \ll N$), it decays (30,31) as the exponential of a power law with an exponent $5/2$ for $y \to \infty$ and $3/2$ for $y \to -\infty$. Skew distributions with similar tails have been observed in numerical simulations [1,18–20] and partly understood by the replica approach [15–17]. A precise comparison with existing numerical data cannot however be done because in the present work, we consider a finite geometry (a ring of N sites, taking the long time limit first, and then the limit of an infinite lattice ($N \to \infty$)) whereas the simulations had been usually done for directed polymers on an infinite lattice [1] (thus taking the limits in reverse order). It would certainly be interesting to try to extend the above Bethe ansatz calculation to this case. Other extensions could be the calculation of the diffusion constant [24], the higher cumulants of Y or the whole large deviation function in more general cases such as the partially asymmetric exclusion process [24–27] or the case of several species [28].

The similarity of the expression (20,21) of $G(\beta)$ as a function of β with that of the pressure of a Fermi or a Bose gas in $3d$ was a surprise to us and it would be nice to have a more direct explanation for it. A Bose-Einstein condensation for the ASEP has already been found in the presence of quenched disorder [29]. Here, the mechanism seems to be different as instead of observing the Bose-Einstein condensation, one finds (24,25) that $G(\beta)$ is analytically continued beyond the critical β_0.

Acknowledgments: We thank E. Speer for many very useful discussions. The research of JLL was supported by NSF Grant 95-23266 and AFOSR Grant 95-0159.

---

[1] T. Halpin-Healy and Y.C. Zhang, *Phys. Rep.* **254**, 215 (1995)

[2] J.P. Bouchaud, M. Mézard and G. Parisi, *Phys. Rev.* E **52**, 3656 (1995)

[3] T.M. Liggett, *Interacting Particle Systems* NY: Springer Verlag (1985)

[4] H. Spohn, *Large Scale Dynamics of Interacting Particles* NY: Springer Verlag (1991)

[5] J. Krug, *Phys. Rev. Lett.* **67**, 1882 (1991)

[6] P. Meakin, P. Ramanlal, L.M. Sander and R.C. Ball, *Phys. Rev.* A **34**, 5091 (1986)

[7] D. Dhar, *Phase Transitions* **9**, 51 (1987)

[8] L.H. Gwa and H. Spohn, *Phys. Rev. Lett.* **68**, 725 (1992); *Phys. Rev.* A **46**, 844 (1992)

[9] B. Derrida and M.R. Evans, V. Hakim and V. Pasquier, *J. Phys.* A **26**, 1493 (1993); B. Derrida, M.R. Evans, p 277 in *Non-equilibrium statistical mechanics in one dimension*, ed V. Privman (C.U.P.) 1997

[10] B. Derrida, M.R. Evans and D. Mukamel, *J. Phys.* A **26**, 4911 (1993)

[11] R.B. Stinchcombe and G.M. Schütz, *Europhys. Lett.* **29**, 663 (1995) and *Phys. Rev. Lett.* **75**, 140 (1995)

[12] G.M. Schütz, *J. Stat. Phys.* **88**, 427 (1997)

[13] S.A. Janowsky and J.L. Lebowitz, *Phys. Rev.* A **45**, 618 (1992)

[14] J. Krug and L.H. Tang, *Phys. Rev.* E **50**, 104 (1994)

[15] M. Kardar, *Nucl. Phys.* B **290** [FS20], 582 (1987)

[16] Y.C. Zhang, *J. Stat. Phys.* **57**, 1123 (1989)

[17] J.P. Bouchaud and H. Orland, *J. Stat. Phys.* **61**, 877 (1990)

[18] T. Halpin-Healy, *Phys. Rev.* A **44**, 3415 (1991)

[19] J.M. Kim, A.J. Bray and M.A. Moore, *Phys. Rev.* A **44**, 2345 (1991)

[20] J. Krug, P. Meakin and T. Halpin-Healy, *Phys. Rev.* A **45**, 638 (1992)

[21] V. Yakhot and A. Chekhlov, *Phys. Rev. Lett.* **77**, 3118 (1996)

[22] K. Huang, *Statistical Mechanics*, John Wiley & Sons 1987

[23] B. Derrida and R. Zeitak, *Phys. Rev. E* **54**, 2513 (1996)

[24] B. Derrida and K. Mallick, *J. Phys.* A **30**, 1031 (1997)

[25] S. Sandow, *Phys. Rev.* E **50**, 2660 (1994)

[26] D. Kim, *Phys. Rev.* E **52**, 3512 (1995); *J. Phys.* A **30**, 3817 (1997)

[27] F.H.L. Essler and V. Rittenberg, *J. Phys.* A **29**, 3375 (1996); K. Mallick and S. Sandow, *J. Phys.* A **30**, 4513 (1997)

[28] B. Derrida, S.A. Janowsky, J.L. Lebowitz and E.R. Speer, *J. Stat. Phys.* **73**, 813 (1993)

[29] M.R. Evans, *Europhys. Lett.* **36**, 13 (1996); *J. Phys.* A **30**, 5669 (1997)