Geometrical interpretation of fluctuating hydrodynamics in diffusive systems

Robert L. Jack

Department of Physics, University of Bath, Bath BA2 7AY, United Kingdom

Johannes Zimmer

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom

Abstract. We discuss geometric formulations of hydrodynamic limits in diffusive systems. Specifically, we describe a geometrical construction in the space of density profiles — the Wasserstein geometry — which allows the deterministic hydrodynamic evolution of the systems to be related to steepest descent of the free energy, and show how this formulation can be related to most probable paths of mesoscopic dissipative systems. The geometric viewpoint is also linked to fluctuating hydrodynamics of these systems via a saddle point argument.

1. Introduction

In many physical situations, one seeks to describe a complex system of many components by a simpler theory that operates on large length (and time) scales. Familiar examples include the description of molecular liquids by the continuum equations of fluid dynamics, or the use of a diffusion equation to describe the spreading of particles through a system. These are examples of hydrodynamic limits, where systems of discrete particles can be modelled by the evolution of continuous fields, such as local density and velocity. Here, we concentrate on a family of microscopic models with overdamped (stochastic) dynamics, in which hydrodynamic limits have been studied over many years [1, 2, 3]: these models have also attracted considerable recent interest [4, 5, 6, 7, 8, 9, 10]. In these *diffusive systems*, the theory of "fluctuating hydrodynamics" [1] captures the behaviour in the hydrodynamic limit, including both deterministic and stochastic effects. The models have been studied both in equilibrium and non-equilibrium settings. For example, they may be coupled to particle reservoirs with different chemical potentials, so that currents flow through the system. Several elegant results have been derived, including the exact analysis of long-ranged correlations that appear in the non-equilibrium states [4, 5, 6, 7, 8], and calculations of large deviations of currents and dynamical activities at equilibrium [9, 10].

The aim of this article is to show how recent results from mathematics [11, 12, 13] (see also [14, 15]) provide a geometrical interpretation of the theory of fluctuating hydrodynamics for these systems. In particular, we describe distance measures (metric structures) in the space of density profiles, such that the hydrodynamic limit equations for several diffusive systems correspond to "steepest descent" processes, or "gradient flows". That is, the systems evolve downhill in free energy, in the direction of the gradient, within the relevant metric. Further, the probability of large deviations from the most likely hydrodynamic behaviour [4, 5, 19] can be related to a geometrical "action functional" which also depends only on the free energy of the system and the relevant metric structure. The conclusion is that the theory of fluctuating hydrodynamics arises from a combination of a thermodynamic free energy functional, and a geometric structure that determines its dynamical evolution. The relevant geometric structures are based on the Wasserstein distance [20], and have a physical interpretation in terms of a cost that is required to transport density through the system.

Several aspects of the situation that we present have been noted in existing work. Formulae for large deviations in diffusive systems have been discussed in detail by Bertini *et al.* [4, 5, 6], who interpreted the most likely path in the system as the solution to a Hamilton-Jacobi equation. The path integral analyses that we will use to make the connection between geometrical structure and large deviations are standard, as summarised (for example) in [8]. In mathematics, connections between the Wasserstein geometry and hydrodynamic limit equations have been investigated [12], and generalisation of these results to large deviations (and fluctuating hydrodynamics) have also been discussed [14, 15]. Our main aim here is to interpret these mathematical

results in a physical setting, which we accomplish by means of path integral methods. The result is that the action functionals that appear in path integrals for diffusive systems can be given a geometrical interpretation in terms of steepest descent paths in the space of density profiles, providing an intuitive interpretation of these action functionals and (some of) their symmetry properties. One possible application of this geometrical formalism might that it faciliates the preservation of connections between the free energy of a system and its dynamics, in approximation schemes. (Preserving symmetries of the dynamics when making approximations can be difficult with existing methods [16, 17, 18].)

The form of the paper is as follows. In Section 2, we discuss the diffusive systems to which our results apply, and we summarize their most relevant properties. In Section 3, we describe the Wasserstein geometry, and we show how it provides a geometrical interpretation of several results for the hydrodynamic limit of a system of non-interacting particles. Then, in Section 4, we discuss how the Wasserstein geometry can be modified such that it applies to broad class of diffusive systems. We draw our conclusions and summarize outstanding issues in Section 5.

2. Model systems

2.1. Free particle diffusion

The results that we describe here are relevant for a range of model systems that include (for example) free-particle diffusion and the symmetric exclusion process. In all cases, the behaviour on hydrodynamic scales is captured by a locally-conserved density $\rho(\boldsymbol{r},t)$. The models also exhibit microscopic reversibility, which typically arises from a detailed balance relation at the level of the microscopic dynamics. The simplest model system contains N non-interacting particles diffusing in d-dimensional space. Starting from the master equation for an appropriate lattice model, standard methods lead to a path integral representation of the dynamics, valid on length scales much larger than the lattice spacing. The construction of this path integral is outlined in Appendix A. It is convenient to define $\varphi(\boldsymbol{r},t) = \frac{1}{N}\rho(\boldsymbol{r},t)$: since the system contains N particles, we have $\int \mathrm{d}\boldsymbol{r}\varphi(\boldsymbol{r}) = 1$. (Here and in the following, $\boldsymbol{r}$-integrals run over the entire space of interest.) The result (for large N) is that the expectation value of any density-dependent observable may be written as

$$\langle O\rangle = \frac{1}{Z}\int \mathcal{D}\varphi\mathcal{D}\hat{\rho}\, O[\varphi]\exp\left\{-N\int \mathrm{d}t\int \mathrm{d}\boldsymbol{r}\left[\hat{\rho}\partial_t\varphi - \mathcal{H}(\varphi,\hat{\rho},\nabla\hat{\rho})\right]\right\} \tag{1}$$

with the Hamiltonian (or Hamiltonian density)

$$\mathcal{H}(\varphi,\hat{\rho},\nabla\hat{\rho}) = -D\nabla\varphi\cdot\nabla\hat{\rho} + D\varphi|\nabla\hat{\rho}|^2 \tag{2}$$

and $Z = \int\mathcal{D}\varphi\mathcal{D}\hat{\rho}\,\exp[-N\int\mathrm{d}t\mathrm{d}\boldsymbol{r}\,(\hat{\rho}\partial_t\varphi - \mathcal{H})]$ for normalisation. Here and throughout, we use $\langle\cdot\rangle$ to represent an average (or expectation value), and $\int\mathcal{D}\varphi\mathcal{D}\hat{\rho}$ to indicate a functional integral over paths $\varphi(\boldsymbol{r},t)$ and 'response functions' $\hat{\rho}(\boldsymbol{r},t)$, subject to the boundary conditions that are relevant for the physical situation of interest.

The equilibrium state of this model has particle positions distributed independently and uniformly throughout the system: the microscopic dynamics obey detailed balance with respect to this distribution. In terms of the density φ, the free energy that corresponds to this distribution is

$$\beta F[\rho] = \int \mathrm{d}\boldsymbol{r}\, \varphi[\log(\varphi\lambda^d) - 1] \tag{3}$$

where β is an inverse temperature, and λ a constant with units of length. As expected, $F[\varphi]$ is the configurational part of the free energy of an ideal gas (which is purely entropic in this case).

At the level of the Hamiltonian (2), the time-reversal symmetry (detailed balance) of the model is not immediately apparent. To reveal this symmetry, one should notice that $\nabla\varphi = \varphi\nabla\frac{\delta}{\delta\varphi}(\beta F)$, and then make the change of variables $\varphi_{\mathrm{TR}}(\boldsymbol{r},t) = \varphi(\boldsymbol{r},-t)$, $\hat{\rho}_{\mathrm{TR}}(\boldsymbol{r},t) = -\hat{\rho}(\boldsymbol{r},-t) + \frac{\delta(\beta F)}{\delta\varphi}(\boldsymbol{r},-t)$ [8]. It follows that

$$\begin{aligned}\int_{-\tau}^{\tau}\mathrm{d}t\int\mathrm{d}\boldsymbol{r}[\hat{\rho}_{\mathrm{TR}}\partial_t\varphi_{\mathrm{TR}} - \mathcal{H}(\varphi_{\mathrm{TR}},\hat{\rho}_{\mathrm{TR}},\nabla\hat{\rho}_{\mathrm{TR}})] &= \beta F[\varphi(-\tau)] - \beta F[\varphi(\tau)] \\ &+ \int_{-\tau}^{\tau}\mathrm{d}t\int\mathrm{d}\boldsymbol{r}\,[\hat{\rho}\partial_t\varphi - \mathcal{H}(\varphi,\hat{\rho},\nabla\hat{\rho})]\,.\end{aligned} \tag{4}$$

This symmetry property of the path integral corresponds to the time-reversal symmetry of the underlying particle model. One of the outcomes of the analysis presented here is that while this construction may seem both complicated and rather arbitrary, it has a straightforward geometrical interpretation, within the Wasserstein metric: see Section 3.4 below.

2.2. Other diffusive systems

As well as this simple system of non-interacting particles, the results we will discuss also apply in a more general setting. We focus on hydrodynamic limits: Imagine observing a particle system on a large length scale, so that the fundamental particles are no longer visible, and it is convenient to think in terms of a smooth density profile $\rho(\boldsymbol{r})$. To take the limit of a large observation scale, it is mathematically convenient to rescale coordinates and instead consider a fixed region of space in which the number of particles $N \to \infty$. To enable this, we again define a rescaled density $\varphi(\boldsymbol{r},t) = \rho(\boldsymbol{r},t)/N$, so that the hydrodynamic limit is $N \to \infty$ at fixed φ. Then the class of systems that we consider have path integral representations of the form

$$\langle O\rangle = \frac{1}{Z}\int\mathcal{D}\varphi\mathcal{D}\hat{\rho}\,O[\varphi]\exp\left\{-N\int\mathrm{d}t\int\mathrm{d}\boldsymbol{r}\,[\hat{\rho}\partial_t\varphi - \mathcal{H}_m(\varphi,\hat{\rho},\nabla\hat{\rho})]\right\} \tag{5}$$

with the general Hamiltonian

$$\mathcal{H}_m(\varphi,\hat{\rho},\nabla\hat{\rho}) = -m(\varphi)D\nabla\hat{\rho}\cdot\nabla\frac{\delta}{\delta\varphi}\beta F[\varphi] + m(\varphi)D|\nabla\hat{\rho}|^2. \tag{6}$$

Here, $F[\varphi]$ is the free energy (per particle) of the diffusive system, and $m(\varphi)$ is a density-dependent local mobility. If we take F as in (3) and $m = \varphi$, we recover the

free-particle model. However, the class of models described by (5) also includes non-interacting particles diffusing in a potential, the symmetric simple exclusion process (SSEP), and the Kipnis-Marchioro-Presutti (KMP) model: we discuss these specific cases in Section 4.2 below. Our results may be further generalised to cases where the mobility m depends explicitly on $\boldsymbol{r}$, or to cases of anisotropic mobility (in which case m becomes a $d \times d$ matrix). It may easily be verified that systems described by (5) all have the time-reversal symmetry (4): they have time-reversal symmetric (equilibrium) steady states with free energy F.

Note that we have taken D to be a simple constant that sets the relative scaling of space and time. Alternatively we could incorporate the φ-dependence of m into a φ-dependent diffusion constant $m(\varphi)D \to D(\varphi)$ as in [6]: here, we anticipate that $m(\varphi)$ will have a geometrical interpretation (independent of time), so we separate it from the dynamical parameter D.

2.3. Saddle point analysis and large deviation functional

We now concentrate on the path integral (5), in the hydrodynamic (large-N) limit. In this limit, the integrand in (5) becomes sharply peaked about its maximal value. Assuming that O does not depend too strongly on N, one has

$$\langle O \rangle = O[\varphi^*], \tag{7}$$

where $\varphi^*(\boldsymbol{r}, t)$ is the path that maximises the exponential in (5), subject to any imposed boundary conditions. We first maximise the exponential over the response field $\hat{\rho}$, for which the Euler-Lagrange equation is

$$\partial_t \varphi = D\nabla \cdot m \left[\nabla \frac{\delta}{\delta \varphi} \beta F - 2\nabla \hat{\rho} \right]. \tag{8}$$

If we denote the (φ-dependent) solution of this equation by $\hat{\rho}^*$ and substitute into (5), the argument of the exponential reduces to $-N\mathcal{S}_m[\varphi]/2$ with

$$\mathcal{S}_m[\varphi] = 2 \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\, mD \left| \nabla \hat{\rho}^* \right|^2. \tag{9}$$

(The factor of 2 is incorporated into $\mathcal{S}_m$ for later convenience.) Minimising $\mathcal{S}_m$ then gives the most likely paths φ^* (clearly, paths with $\nabla\hat{\rho}^* = 0$ are minimisers of $\mathcal{S}_m$, but the existence of such paths may depend on the boundary conditions in the path integral). Hence, one may calculate expectation values of the form of (7).

To arrive at a stronger result, we restrict the integral in the numerator of (5) to trajectories φ that are close to some reference path $\varrho(\boldsymbol{r}, t)$. Integrating over the $\hat{\rho}$ field, one finds that the probability of observing such a path satisfies

$$\lim_{N\to\infty} \frac{1}{N} \log \mathrm{Prob}[\varphi(\boldsymbol{r}, t) \approx \varrho(\boldsymbol{r}, t)] = -\frac{1}{2}\mathcal{S}_m[\varrho]. \tag{10}$$

This is a large deviation principle [19], which determines the probability of observing a non-typical trajectory in this system, as $N \to \infty$. Such large deviation principles play a central role in several theories of hydrodynamic behaviour [2, 3, 6].

Equations (7-10) allow analysis of hydrodynamic limits in a range of diffusive systems. However, the physical interpretation of some of the quantities that appear in these equations is rather opaque. The central point of this paper is that the large deviation function $\mathcal{S}_m$ has a geometrical interpretation, as do the paths that minimise this function. Hence, this geometrical structure determines both the most likely paths in the hydrodynamic limit of these systems, as well as the probabilities of large deviations from these paths. As we shall see, it also clarifies the connection between the functional $\mathcal{S}$, the free energy F, and the time-reversal symmetry (detailed balance) of the microscopic dynamics in these systems.

3. Wasserstein geometry

The geometrical setting that we consider is called the Wasserstein geometry, which defines a metric structure in the space of density profiles $\rho(\boldsymbol{r})$. The Wasserstein distance is also called an 'earth-movers' distance: it arises in the theory of optimal transport, where we assume that some mass is initially distributed according to a density profile $\rho(\boldsymbol{r})$, and we consider how it can be transported with minimal cost, in order to arrive at a given final profile. For a detailed discussion, we refer the reader to [20] (see also [12, 14, 15]). The 'standard' Wasserstein geometry corresponds to the case $m = \varphi$ in the notation of (6): others cases correspond to modified geometrical structures, as discussed in Section 4. We first present this geometrical structure in an abstract setting based on the transport of density at minimal cost. We then show that the resulting cost function is directly connected to the probabilities of rare events in diffusive systems.

3.1. Wasserstein distance

We begin by defining the Wasserstein distance between two density profiles $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$, assuming that these profiles represent the same number N of particles, $\int \mathrm{d}\boldsymbol{r}\rho_1(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}\rho_2(\boldsymbol{r}) = N$. We motivate the definition by a construction that comes from the theory of optimal transport [20]. Starting with the profile $\rho_1(\boldsymbol{r})$, suppose that we redistribute the particle density over the whole space, with the fraction of the density from $\boldsymbol{r}$ that moves to $\boldsymbol{r}'$ being $q(\boldsymbol{r}'|\boldsymbol{r})$. Clearly $\int \mathrm{d}\boldsymbol{r}' q(\boldsymbol{r}'|\boldsymbol{r}) = 1$ for all $\boldsymbol{r}$. Also, for this process to generate the profile $\rho_2(\boldsymbol{r})$, one requires that $\int \mathrm{d}\boldsymbol{r} q(\boldsymbol{r}'|\boldsymbol{r})\rho_1(\boldsymbol{r}) = \rho_2(\boldsymbol{r}')$. If transporting density through the system requires some kind of 'cost' that is proportional to the square of the distance moved, then the total cost of this redistribution is

$$C[q] = \int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}'|\boldsymbol{r} - \boldsymbol{r}'|^2 q(\boldsymbol{r}', \boldsymbol{r}) \tag{11}$$

where $q(\boldsymbol{r}', \boldsymbol{r}) = q(\boldsymbol{r}'|\boldsymbol{r})\rho_1(\boldsymbol{r})$. At this stage, the cost C is a purely abstract quantity – its relation to the physics of diffusive systems will be discussed in Section 3.4 below. The (squared) Wasserstein distance between $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$ is then defined by

$$d[\rho_1, \rho_2]^2 = \inf_{q(\boldsymbol{r}', \boldsymbol{r})} C[q] \tag{12}$$

where the infimum (minimum) is taken over all distributions q which satisfy the constraints $\int \mathrm{d}\boldsymbol{r}' q(\boldsymbol{r}', \boldsymbol{r}) = \rho_1(\boldsymbol{r})$ and $\int \mathrm{d}\boldsymbol{r} q(\boldsymbol{r}', \boldsymbol{r}) = \rho_2(\boldsymbol{r}')$. That is, q is a distribution, whose marginals are ρ_1 and ρ_2. The interpretation of $d[\rho_1, \rho_2]^2$ as an 'earth-movers' distance reflects its connection to the distance by which density must be transported. In this setting, $q(\boldsymbol{r}'|\boldsymbol{r})$ is called a "transport plan".

3.2. Path lengths and path energies

It is useful to consider continuous paths in the space of density profiles, represented as $\rho(\boldsymbol{r}, s)$, where s is a progress co-ordinate along the path. We suppose that $\rho(\boldsymbol{r}, s)$ represents advection by an s-dependent velocity field $\boldsymbol{v}(\boldsymbol{r}, s)$, so that $\partial_s \rho = -\nabla \cdot (\rho \boldsymbol{v})$. This advective process transports the density from an initial profile $\rho(\boldsymbol{r}, a)$ to a final one $\rho(\boldsymbol{r}, b)$. Discretising the path into segments of length δs, we calculate the cost δC of each segment by using (11) with $q(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}' - \boldsymbol{v}(\boldsymbol{r})\delta s)\rho(\boldsymbol{r})$, where $\rho(\boldsymbol{r})$ is the profile at the start of the segment. Hence $\delta C = \int \mathrm{d}\boldsymbol{r} \rho |\boldsymbol{v}|^2 \delta s^2$. We then sum these costs over all segments within the path to obtain $\mathcal{E}_v[\rho(\boldsymbol{r}, s), \boldsymbol{v}(\boldsymbol{r}, s)] = \lim_{\delta s \to 0} \frac{1}{\delta s} \sum_k \delta C_k = \int_a^b \mathrm{d}s \int \mathrm{d}\boldsymbol{r}\, \rho |\boldsymbol{v}|^2$, which depends both on the path $\rho(\boldsymbol{r}, s)$ and the velocity field $\boldsymbol{v}(\boldsymbol{r}, s)$. For any path, $\mathcal{E}_v$ is minimised by a potential flow $\boldsymbol{v} = -\nabla\Phi$ [21]. We therefore define

$$\mathcal{E}[\rho(\boldsymbol{r}, s)] = \int_a^b \mathrm{d}s \int \mathrm{d}\boldsymbol{r}\, \rho |\nabla\Phi|^2 \tag{13}$$

where $\nabla\Phi$ satisfies

$$\partial_s \rho = \nabla \cdot (\rho \nabla \Phi). \tag{14}$$

The functional $\mathcal{E}$ depends only on the path $\rho(\boldsymbol{r}, s)$, and is obtained by summing the costs of the small segments that comprise this path.

It then follows [21] that the Wasserstein distance between two profiles $\rho_1(\boldsymbol{r})$ and $\rho_2(\boldsymbol{r})$ can be obtained by finding the path between these profiles that minimises $\mathcal{E}$. The result is that

$$d[\rho_1, \rho_2]^2 = (b - a) \inf_{\rho(\boldsymbol{r}, s)} \mathcal{E}[\rho] \tag{15}$$

where the minimisation is subject to $\rho(\boldsymbol{r}, a) = \rho_1$ and $\rho(\boldsymbol{r}, b) = \rho_2$. Note that $\mathcal{E}$ was defined by summing the costs δC of each segment along the path: from (12), the cost δC is a squared distance so $\mathcal{E}$ corresponds to a sum of squared segment lengths. However, the distance $d[\rho_1, \rho_2]$ should be equal to the sum of the segment lengths themselves: the content of (15) is that the infimum (minimum) is obtained when all segments are of equal length, in which case the distance d can be inferred from the (rescaled) sum of the squared segment lengths. The factor of $(b - a)$ in (15) and the factor of δs in the definition of $\mathcal{E}_v$ implement the required rescaling. Mathematically, this relation follows from a Cauchy-Schwarz inequality between path actions and path lengths: see Appendix B.

Hence, the path $\rho(\boldsymbol{r}, s)$ that realises the minimum in (15) is the geodesic (shortest path) connecting ρ_1 and ρ_2, in the Wasserstein geometry. The metric structure that

underlies these results is discussed in more detail in Appendix B. We have presented a heuristic argument for (15), but this result has been shown rigorously [21, Proposition 1.1].

3.3. Steepest descent with respect to a free energy

A central conclusion of this study is that diffusive processes are governed by two factors: the geometrical structure of the space of density profiles, and the free energies of these profiles. We therefore associate to each profile $\rho(\boldsymbol{r})$ a free energy $F[\rho]$. In this section, we focus on the case where F is given by (3), which corresponds to a microscopic model of non-interacting particles.

Given a free energy and a distance measure (metric), it is natural to define a steepest descent process. Paths of steepest descent may be constructed by a discrete-time process [11]: on each time step, a system with profile $\rho_t(\boldsymbol{r})$ evolves to the profile $\rho_{t+\delta t}(\boldsymbol{r})$ which minimises

$$S[\rho_{t+\delta t}] := \frac{1}{2D\delta t} d[\rho_t, \rho_{t+\delta t}]^2 + \beta F[\rho_{t+\delta t}]. \tag{16}$$

Here δt is the (small) time increment associated with each time step and D a constant that sets the units of time. (We note that d has units of length and Dt has units of length^2, while S and βF are dimensionless.) To make contact with the previous subsection, we identify Dt with the progress variable s.

For a given starting profile ρ_0, minimisation of (16) leads to paths of decreasing free energy F. Our aim in this section is to define a path action functional $\mathcal{S}_W[\rho(\boldsymbol{r},t)]$, which is minimal for these steepest descent paths. We define $\mathcal{S}_W[\rho(\boldsymbol{r},t)]$ by discretising the path $\rho(\boldsymbol{r},t)$ in time: for each time step, we measure the difference between $S[\rho_{t+\delta t}]$ and $\min_{\rho_{t+\delta t}} S[\rho_{t+\delta t}]$. This difference provides a measure of the deviation of the path $\rho(\boldsymbol{r},t)$ from a steepest descent path.

To make progress, we assume that δt is small, and use (13) and (15) to approximate

$$S[\rho_{t+\delta t}] \approx \frac{D\delta t}{2} \int \rho_t |\nabla\Phi|^2 \,\mathrm{d}\boldsymbol{r} + \beta F[\rho_{t+\delta t}] \tag{17}$$

where the dimensionless field $\Phi(\boldsymbol{r},t)$ solves

$$\partial_t \rho = \nabla\cdot(\rho D\nabla\Phi)\,, \tag{18}$$

subject to the boundary conditions $\rho(t)=\rho_t$ and $\rho(t+\delta t)=\rho_{t+\delta t}$.

To obtain steepest descent paths, we minimise (17) over $\rho_{t+\delta t}$, using a Lagrange multiplier $\lambda=\lambda(\boldsymbol{r})$ to enforce the constraint (18). The quantity to be extremised (over $\lambda, \Phi, \rho_{t+\delta t}$) is

$$\frac{D\delta t}{2} \int \rho_t |\nabla\Phi|^2 \mathrm{d}\boldsymbol{r} + \beta F[\rho_{t+\delta t}] - \int \lambda[\rho_{t+\delta t} - \rho_t - D\delta t \nabla\cdot(\rho\nabla\Phi)]\mathrm{d}\boldsymbol{r}. \tag{19}$$

Extremising over $\rho_{t+\delta t}$ yields $\lambda = \beta\delta F/\delta\rho$ and extremising over Φ gives $(D\delta t)\nabla\cdot(\rho\nabla\Phi) = (D\delta t)\nabla\cdot(\rho\nabla\lambda)$. We solve these equations by taking $\Phi=\lambda=\beta\delta F/\delta\rho$. It remains to compute the minimal value of S: we write $\beta F[\rho_{t+\delta t}] \approx \beta F[\rho_t] + (\rho_{t+\delta t}-\rho_t)\beta\delta F/\delta\rho$

and the constraint (18) implies that $\rho_{t+\delta t} - \rho_t = D\delta t \nabla \cdot (\rho \nabla \Phi)$. Substituting for $\beta F[\rho_{t+\delta t}]$ in (16) and using the values of Φ and $(\delta F/\delta\rho)$ that minimise (17) gives $\min_{\rho_{t+\delta t}} S[\rho_{t+\delta t}] = \beta F[\rho_t] - \frac{D\delta t}{2}\int_{\boldsymbol{r}} \rho_t |\nabla(\delta\beta F[\rho_t]/\delta\rho)|^2$.

Let ρ^* be the value of $\rho_{t+\delta t}$ that minimises (17). Clearly, minimising $S[\rho_{t+\delta t}]$ is equivalent to minimising $\tilde{S}[\rho_{t+\delta t}] := S[\rho_{t+\delta t}] - S[\rho^*]$. With this definition, we obtain

$$\tilde{S}[\rho_{t+\delta t}] = \frac{D\delta t}{2}\int \rho_t |\nabla\Phi|^2 \mathrm{d}\boldsymbol{r} + \beta F[\rho_{t+\delta t}] - \beta F[\rho_t] + \frac{D\delta t}{2}\int \rho_t \left|\nabla \frac{\delta}{\delta\rho}\beta F[\rho_t]\right|^2 \mathrm{d}\boldsymbol{r}, \qquad (20)$$

which is a non-negative functional that measures the difference between $S[\rho_{t+\delta t}]$ and its minimal value. (Recall that this minimal value is $S[\rho^*]$, where ρ^* is the value of $\rho_{t+\delta t}$ that would be obtained by a steepest descent process starting from ρ_t.)

Given a path $\rho(\boldsymbol{r}, t)$, we construct a discrete-time path action by discretising in time and calculating the value of $\tilde{S}$ for each segment, leading to the action $\delta t \sum_j \tilde{S}[\rho_{t_0+j\delta t}]$. We then take the limit of continuous time ($\delta t \to 0$) and move from this discrete-time construction to a functional for continuous paths. Using $\mathcal{E}[\rho] = \int \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\rho|\nabla\Phi|^2$ from Section 3.2 [with Φ being the solution of (18)], the sum over short path segments converges to

$$\mathcal{S}_{\mathrm{W}}[\rho(\boldsymbol{r}, s)] := \frac{D}{2}\mathcal{E}[\rho] + \beta F[\rho(\boldsymbol{r}, b)] - \beta F[\rho(\boldsymbol{r}, a)] + \frac{D}{2}\int_a^b \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\, \rho \left|\nabla \frac{\delta}{\delta\rho}(\beta F)\right|^2. \qquad (21)$$

Finally, noting that $F[\rho(\boldsymbol{r}, b)] - F[\rho(\boldsymbol{r}, a)] = \int_a^b \mathrm{d}t(\delta F/\delta\rho)\partial_t\rho$, one arrives at the path functional for steepest descent in the Wasserstein metric:

$$\mathcal{S}_{\mathrm{W}}[\rho(\boldsymbol{r}, t)] = \frac{1}{2}\int_a^b \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\, \rho D \left|\nabla\left(\Phi - \beta\frac{\delta F}{\delta\rho}\right)\right|^2 \qquad (22)$$

Summarising once again, this functional is minimal and equal to zero for steepest descent paths, which are those which minimise (16) for every step along the path. The value of $\mathcal{S}_{\mathrm{W}}[\rho(\boldsymbol{r}, t)]$ for a general path indicates the extent to which this path fails to minimise (16).

The physical interpretation of the paths that minimise (22) is that if $-D\nabla\Phi$ is the velocity field that advects the underlying particle density, the effect of the free energy F is to bias this advection. These paths (which have $\mathcal{S} = 0$) therefore satisfy

$$\partial_t\rho = \nabla \cdot \left[\rho D \nabla \frac{\delta}{\delta\rho}(\beta F)\right] =: -D\,\mathrm{grad}^{\mathrm{W}}(\beta F). \qquad (23)$$

Here, the 'gradient operator' $\mathrm{grad}^{\mathrm{W}}$ for the Wasserstein metric is defined by its action on a functional $\Psi[\rho(\boldsymbol{r})]$, as $\mathrm{grad}^{\mathrm{W}}(\Psi) = -\nabla \cdot \left(\rho\nabla\frac{\partial}{\partial\rho}\Psi\right)$. So (23) describes steepest descent with respect to the free energy F in the Wasserstein geometry. This formulation is analogous to steepest descent in Euclidean space, where $\gamma\partial_t\boldsymbol{x} = -\nabla_{\boldsymbol{x}}E$, in which γ is a friction constant and E a potential energy function. While this analogy looks at the moment formal and possibly arbitrary, since it is based on the definition of $\mathrm{grad}^{\mathrm{W}}$, it can be shown that the Wasserstein geometry indeed defines an infinite-dimensional geometric structure where $\mathrm{grad}^{\mathrm{W}}$ plays the part of a gradient. Elements of this theory are sketched in Appendix B.

3.4. Relation to diffusive systems: saddle point trajectories and large deviations

So far, our discussion of the Wasserstein geometry has been purely abstract, via the cost (11), the distance (12), and the steepest descent condition (16). We are now in a position to connect this geometrical construction to the results of Section 2. The results obtained so far correspond to the case $m = \varphi$ in that Section. The first thing to note is that the most likely paths φ^* in (7) are equal (up to a factor of N) to the solutions of the steepest descent equation (23). This follows because the most likely paths are those with $\nabla\hat{\rho}^* = 0$ [see (9)], in which case setting $m = \varphi$ in (8) yields (23). It is also instructive to substitute for F using (3), in which case (23) reduces to the diffusion equation

$$\partial_t \rho = -D\,\mathrm{grad}^{\mathrm{W}}(\beta F) = \nabla\cdot\left[\rho D\nabla\frac{\delta}{\delta\rho}(\beta F)\right] = D\nabla^2\rho. \tag{24}$$

Thus, the trajectories that dominate the hydrodynamic limit of this system correspond to steepest descent paths of the free energy, within the Wasserstein metric [11].

In addition, the large deviation function $\mathcal{S}_m$ in (9) corresponds (for $m = \varphi$) to the Wasserstein path action $\mathcal{S}_{\mathrm{W}}$ in (22). To see this, note that since $\hat{\rho}^*$ in (9) solves Eq. (8) then (for $m = \varphi$) we can identify $\hat{\rho} = \frac{1}{2}[\delta(\beta F)/\delta\varphi - \Phi]$ where Φ solves (18). Thus the action $\mathcal{S}_m$ that determines path probabilities in the large deviation principle (10) is the same as the Wasserstein path action $\mathcal{S}_{\mathrm{W}}$. That is, the Wasserstein metric and the free energy together specify not just the dominant hydrodynamic path but also the fluctuations about this path.

Finally, we note from (21) that the cost of a path depends on the direction with which the path is traversed: the free energy difference gives a contribution that is odd under time reversal while the first and last terms on the right-hand side of (21) are both even under time-reversal. The fact that the odd part of the cost is simply the free energy difference between start and end points is equivalent to the detailed balance symmetry of the microscopic model: for large N, if φ_{TR} is the time-reversed counterpart of a trajectory φ that runs from $t = a$ to $t = b$ then one has from (10) and (21) that $\mathrm{e}^{-\beta F(a)}\mathrm{Prob}[\varphi] = \mathrm{e}^{-\beta F(b)}\mathrm{Prob}[\varphi_{\mathrm{TR}}]$. The appearance of the free energy in (21) may appear coincidental from the derivation given here but this is a general property of steepest-descent processes (gradient flows), which follows from the the definition of the path cost (22): see also Appendix B. Hence, if a system obeys a large deviation function of the form of (10), where the action $\mathcal{S}$ corresponds to the path action for a steepest descent process, then one arrives at a detailed balance-like relation which relates the probability of trajectories to their time-reversed counterparts. The generalisation of this result to systems without detailed balance would presumably result in a fluctuation theorem similar to that of Crooks [26]: this would be an interesting direction for future study.

3.5. Fluctuating hydrodynamics

Since the Wasserstein metric provides a connection both to the most likely hydrodynamic path and to fluctuations about this path, it can also be used to interpret the theory if 'fluctuating hydrodynamics' for diffusive systems [1, 8]. Within this framework, one describes systems on macroscopic scales by Langevin equations, or stochastic partial differential equations. For free particle diffusion, the relevant equation is

$$\partial_t \rho = D[\nabla^2 \rho + \nabla \cdot (\boldsymbol{\eta}\sqrt{\rho})] \tag{25}$$

where $\boldsymbol{\eta}$ is a space-time white noise. [That is, a Gaussian-distributed random function with $\langle \boldsymbol{\eta} \rangle = 0$ and $\langle \eta^\mu(\boldsymbol{r},t)\eta^\nu(\boldsymbol{r}',t')\rangle = \delta^{\mu\nu}\delta(t-t')\delta(\boldsymbol{r}-\boldsymbol{r}')$, where μ,ν label Cartesian components of the vector $\boldsymbol{\eta}$.] In the hydrodynamic limit, it may be more convenient to write (25) as $\partial_t \varphi = D\left[\nabla^2\varphi + \frac{1}{\sqrt{N}}\nabla\cdot(\boldsymbol{\eta}\sqrt{\varphi})\right]$ to emphasise that the effect of the noise becomes increasingly weak as the large-N limit is approached. Following the procedure of Martin-Siggia-Rose-DeDominicis-Janssen [23, 24, 25], one may show that equation (25) is equivalent to the path-integral description (5): see Section 3 of Ref. [8] for a detailed discussion.

If the density profile at time t is $\rho_1(\boldsymbol{r}) = \rho(\boldsymbol{r},t)$, the Langevin equation (25) specifies a probability distribution for the density a short time δt later, $\rho_2(\boldsymbol{r}) = \rho(\boldsymbol{r},t+\delta t)$. In the steepest descent case (where the noise $\boldsymbol{\eta}$ is absent), we recall from (16) that ρ_2 may be obtained by minimising $\frac{1}{2}d[\rho_1,\rho_2]^2 + D\delta t\beta F[\rho_2]$. To incorporate the effects of noise, we start from (10) and time-discretise the action $\mathcal{S}$, from which we find that the probability distribution of ρ_2 is

$$p[\rho_2|\rho_1,\delta t] \propto \exp\left(-\frac{1}{4D\delta t}d[\rho_1,\rho_2]^2 - \frac{1}{2}\beta F[\rho_2]\right). \tag{26}$$

To be precise, $p[\rho_2]$ is a probability density: probabilities are obtained by functional integrals of the form $\int \mathcal{D}\rho_2\, p[\rho_2]$ where the functional integral runs over functions $\rho_2(\boldsymbol{r})$ with $\int \mathrm{d}\boldsymbol{r}\rho_2 = N$, so that the total density is conserved. To reiterate, the Langevin equation (25) is equivalent to the path integral (5), and for large-N the probabilities of trajectories in this system satisfy both (10) and (26). Thus, (26) is a large-N result for the stochastic evolution defined by (25). It means that the probability that a density profile ρ_1 evolves into ρ_2 over a short time period δt has a Gaussian dependence on the distance $d[\rho_1,\rho_2]$ and a simple exponential dependence on the free energy of the final state $\beta F[\rho_2]$.

It is useful to compare this result with the overdamped Langevin of a single particle in a potential $V(\boldsymbol{r})$. If the position of the particle is $\boldsymbol{x}(t)$, one writes

$$\partial_t \boldsymbol{x} = D\beta\nabla V + \boldsymbol{\eta}_0 \tag{27}$$

where γ is a friction constant, the noise $\boldsymbol{\eta}_0$ has covariance $\langle \eta_0^\mu(t)\eta_0^\nu(t')\rangle = 2D\delta^{\mu\nu}\delta(t-t')$, and β is the inverse temperature. This equation implies that if $\boldsymbol{x}(t) = \boldsymbol{x}_1$, then after a small time interval δt, the position of the particle $\boldsymbol{x}_2 = \boldsymbol{x}(t+\delta t)$ is distributed as

$$p(\boldsymbol{x}_2|\boldsymbol{x}_1,\delta t) \propto \exp\left[-\frac{1}{4D\delta t}|\boldsymbol{x}_2-\boldsymbol{x}_1|^2 - \frac{1}{2}\beta V(\boldsymbol{x}_2)\right]. \tag{28}$$

We again recover Gaussian dependence on the distance $|\boldsymbol{x}_2 - \boldsymbol{x}_1|$ and exponential dependence on the free energy difference (here the free energy is given simply by the energy). Based on the similarity between (26) and (28), we argue that (25) is the natural formulation of overdamped dynamics in the Wasserstein geometry – this is a geometrical interpretation of the theory of fluctuating hydrodynamics for this system. The particular noise in (25) has been connected to the Wasserstein geometry before on mathematical grounds [27]; here we arrive at the same result with a very different argument, starting from particles and a path integral formulation.

4. Generalisation to other systems

4.1. Modified Wasserstein metric

We have illustrated a connection between the Wasserstein distance as defined by (12) and the hydrodynamic limit for diffusion of free particles. [In the notation of Section 2, the results of Section 3 apply only in the case where $m = \varphi$ and F is given by (3).] To generalise this connection to the broader class of systems anticipated in (5), we define a generalised Wasserstein distance. To this end, define a modified path cost by analogy with (13): $\mathcal{E}_m[\rho(\boldsymbol{r}, s)] = \int_a^b \mathrm{d}s \int \mathrm{d}\boldsymbol{r}\, m|\nabla\Phi_m|^2$ where Φ solves $\partial_s\rho = \nabla\cdot(m\nabla\Phi_m)$. Then a construction of steepest descent trajectories as in Section 3.3 yields a generalised action functional

$$\mathcal{S}_m[\rho(\boldsymbol{r}, t)] = \frac{1}{2}\int_a^b \mathrm{d}t \int \mathrm{d}\boldsymbol{r}\, mD \left|\nabla\Phi_m - \nabla\frac{\delta(\beta F)}{\delta\rho}\right|^2, \tag{29}$$

where Φ_m solves $\partial_t\rho = \nabla\cdot(Dm\nabla\Phi_m)$. Repeating the analysis of Sec. 3.4, it is easily checked that this path action is equal to the large deviation function in (10): it follows that the dominant hydrodynamic trajectories are therefore steepest descent processes of the free energy within the relevant metric, that a detailed balance symmetry holds at the macroscopic level, and that one may apply the theory of fluctuating hydrodynamics in this more general case too. To make contact with previous work [6], it is useful to note that the operator $K(\rho)$ in [6] is the metric tensor for this generalised Wasserstein geometry: see Appendix B.

The key results are therefore that the action $\mathcal{S}_m$ appearing in the large deviation principle (10) is the relevant modified Wasserstein path action (29), and that the corresponding fluctuating hydrodynamic equation is

$$\partial_t\varphi = D\nabla\cdot\left[m\nabla\frac{\delta}{\delta\varphi}(\beta F) + \frac{1}{\sqrt{N}}(\boldsymbol{\eta}\sqrt{m})\right]. \tag{30}$$

The generalisation of (26) that is equivalent to this process is obtained by replacing the distance d in (26) by the modified distance d_m. We therefore interpret (30) as the natural generalisation of Langevin dynamics to the modified Wasserstein geometry. The dominant hydrodynamic path for this system is obtained by dropping the noise term from (30), leading to $\partial_t\varphi = D\,\mathrm{grad}_m^{\mathrm{W}}(\beta F)$ where $\mathrm{grad}_m^{\mathrm{W}}(\cdot) = \nabla\cdot[mD\nabla\frac{\delta}{\delta\varphi}(\cdot)]$ is the definition of the gradient operator within the modified Wasserstein geometry.

4.2. Diffusive systems described by modified Wasserstein geometries

To illustrate the range of physical systems to which this analysis applies, we now describe three physical systems that are described by path integrals of the form of (5), and we discuss the metric structures associated with the hydrodynamic limits of this models. The large deviation functions and the fluctuating hydrodynamic equations for all these models are related to modified Wasserstein distances, as we have outlined here. (We also note in passing that the case of additive noise, $m = 1$, leads to model-B dynamics [28].)

4.2.1. Diffusion of free particles in a potential The simplest generalisation of free particle diffusion is to introduce a potential $V(\boldsymbol{r})$ that is smooth on the hydrodynamic scale. In this case the free energy is simply

$$\beta F[\rho] = \beta F_{\mathrm{id}}[\rho] + \int \mathrm{d}\boldsymbol{r}\, \rho\beta V, \tag{31}$$

where $F_{\mathrm{id}}[\rho]$ is the non-interacting free energy given in (3). Constructing the path integral as in the free-particle case, one obtains a Hamiltonian of the form (6):

$$\mathcal{H} = -\nabla\hat{\rho}\cdot D(\nabla\varphi + \varphi\beta\nabla V) + D\varphi|\nabla\hat{\rho}|^2. \tag{32}$$

Comparing with (6), one identifies $m = \varphi$ as in the free particle case, and $m\nabla\frac{\delta}{\delta\varphi}(\beta F) = \nabla\varphi + \varphi\beta\nabla V$, as required. Hence, modifying free-particle diffusion by including an external potential preserves the connection to the Wasserstein geometry that was already identified in Sec. 3.4.

4.2.2. Symmetric exclusion process In the simple symmetric exclusion process (SSEP), particles hop between the sites of a lattice, subject to the constraint that at most one particle may occupy any site. This model is simple to define but interactions between particles are strong, and the behaviour of the system is richer as a result. The free energy for the SEP (on the hydrodynamic scale) is

$$\beta F = \int \mathrm{d}\boldsymbol{r}\, [\varphi\log\varphi + (1-\varphi)\log(1-\varphi)], \tag{33}$$

where φ is the local density, rescaled to lie between zero and unity, with $\varphi(\boldsymbol{r}) = 1$ corresponding to almost all sites being occupied in the vicinity of the point $\boldsymbol{r}$.

In the hydrodynamic limit, the system can be described [8] by a path integral of the form of (5), with

$$\mathcal{H} = -\nabla\hat{\rho}\cdot(D\nabla\varphi) + D\varphi(1-\varphi)|\nabla\hat{\rho}|^2. \tag{34}$$

Comparing with (6), we identify $m = \varphi(1-\varphi)$, so consistency between (6) and (33) requires $\varphi(1-\varphi)\nabla\frac{\delta}{\delta\varphi}(\beta F) = \nabla\varphi$. This may be verified from (33). Hence, a large deviation principle of the form of (10) applies to the SSEP as well as to non-interacting systems, where the large deviation function is the action functional for the relevatn modified Wasserstein metric. (Properties of this large deviation function have also been discussed extensively by Bertini *et al.* [4, 5]). For a fixed initial condition, the most likely trajectories in the SSEP are given by $\partial_t\varphi = D\nabla\cdot(m\nabla\frac{\delta}{\delta\varphi}\beta F[\varphi]) = D\nabla^2\varphi$, corresponding

to steepest descent in the relevant metric. We emphasise that while we recover the same diffusion equation as in the free-particle case, both the free energy F and the function m that determines the geometry are different, so the fluctuations about the most likely path differ strongly between the SEP and the free particle model. Specifically, the equation of fluctuating hydrodynamics for the symmetric exclusion process is

$$\partial_t \varphi = D[\nabla^2 \varphi + \nabla \cdot (\boldsymbol{\eta} \sqrt{\varphi(1-\varphi)/N})], \tag{35}$$

which differs from (25). Here, $\boldsymbol{\eta}$ is a space-time white noise, as above.

4.2.3. Kipnis-Marchioro-Presutti (KMP) model The KMP model [29] was developed as a model for heat conduction and has been studied extensively as a model diffusive system [30, 8, 10]. In this model, φ denotes a local energy density. The model is defined on a d-dimensional lattice, and at each time step, the energies of two neighbouring sites are redistributed between those sites. In the equilibrium state, the energy of each site is exponentially distributed, leading to a free energy on the hydrodynamic scale given by [30]

$$\beta F = \int \mathrm{d}\boldsymbol{r}\, [\beta\varphi - 1 - \log(\beta\varphi)]. \tag{36}$$

Constructing the path integral for the dynamical evolution of the model, the Hamiltonian is [8, 10]

$$\mathcal{H} = -\nabla\hat{\rho} \cdot (D\nabla\varphi) + D\varphi^2 |\nabla\hat{\rho}|^2, \tag{37}$$

so we identify $m = \varphi^2$. It is easily verified from (36) that $\varphi^2 \nabla \frac{\delta}{\delta\varphi}(\beta F) = \nabla\varphi$, as required for consistency with (6). Thus the deterministic (hydrodynamic) equation is again $\partial_t \varphi = D\nabla^2\varphi$, as for free particles and for the SEP, but in the KMP model this limit equation arises from a different combination of a free energy and a (generalised Wasserstein) geometrical structure.

5. Conclusion

We have shown how large deviations in the hydrodynamic theory of several diffusive systems can be interpreted geometrically, in terms of the Wasserstein geometry and its generalisations. In particular, the most likely trajectories for these systems in the hydrodynamic limit are given by steepest descent of the free energy, in the appropriate metric. We also argued that the equations of fluctuating hydrodynamics are the natural generalisations of Langevin dynamics, within this geometrical structure. The relation between the large deviations of the time-dependent density and steepest descent processes clarifies the decomposition of the large deviation function into parts that are even and odd under time reversal, which can be traced back to detailed balance properties of the original stochastic processes. It would be interesting to understand whether other properties of these systems such as responses to boundary driving [5, 8] or large deviations of time-integrated quantities [9, 10] can be related to properties of the

Wasserstein geometry (see [31] for nonlinear diffusion with inhomogeneous boundary data). In particular, one might expect that the mapping from boundary-driven to equilibrium systems considered in [8] should have an interpretation in terms of the Wasserstein geometry: this remains a topic for future study.

More generally, for models where the general structure of Section 4 applies, all hydrodynamic properties are determined by the (thermodynamic) free energy functional F, and the geometrical function m, which determines the 'cost' of density redistribution in the system. The idea that the dynamical evolution of stochastic models arises from a combination of a free energy and a metric structure has been discussed in other contexts too [32]: the freedom to choose different metrics while preserving the same free energy functional means that systems with the same thermodynamic properties can have very different dynamical behaviour [33, 34].

It would be interesting to find other examples of stochastic processes whose behaviour can be analysed using metric structures in infinite-dimensional spaces, such as the space of density profiles considered here. One particular aspect here is the derivation of systems driven by energy and entropy. In particular, there is a rich class of equations of fluctuating hydrodynamics with the dissipative part being of the form (30) [3, 35]. This might provide a route to generalise the construction shown here to models without microscopic reversibility, such as active-matter systems [37]. Some recent results for systems without microscopic reversibility [31, 38] indicate that progress in this direction may indeed be possible.

Acknowledgments

We would like to thank Alan McKane, Vivien Lecomte, Mark A. Peletier, Max von Renesse, Tim Rogers, Peter Sollich, and Fred van Wijland for helpful discussions and comments on the manuscript. RLJ was supported by the EPSRC through grant EP/I003797/1. JZ gratefully acknowledges financial support from the EPSRC (EP/K027743/1) and the Leverhulme Trust (RPG-2013-261).

Appendix A. Path integral construction

In this section, we give a very brief review of the construction of the path integral expression (5), in order that the presentation of this paper be as self-contained as possible, and to emphasise that this path integral comes from a microscopic description of a specific particle system. The microscopic model involves particles hopping on a (hyper)cubic lattice in d-dimensions. The lattice spacing is l_0 and particles hop independently, with a rate D/l_0^2 for hopping along each available bond on the lattice. The number of particles on site i is n_i, and a configuration $\mathcal{C}$ is specified by the values of all the n_i.

Let $P(\mathcal{C}, t)$ be the probability of finding the system in configuration $\mathcal{C}$ at time t. This quantity evolves in time by a master equation. To obtain a representation of this

equation, it is useful [22] to write the probability distribution as a high-dimensional vector $|P\rangle = \sum_{\mathcal{C}} P(\mathcal{C},t)|\mathcal{C}\rangle$. Within this vector space, the operator $a_i^\dagger$ acts on $|\mathcal{C}\rangle$ by adding a particle to site i, while a_i multiplies by n_i before removing a particle from site i (the occupancies of all sites $j \neq i$ are unchanged). Hence $a_i^\dagger a_i|\mathcal{C}\rangle = n_i|\mathcal{C}\rangle$ and one has also the commutation relation $[a_i, a_j^\dagger] \equiv a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij}$. It is also useful to denote the state with no particles at all by $|0\rangle$. Then the master equation may be written as [22]

$$\partial_t|P\rangle = \mathbb{W}|P\rangle \tag{A.1}$$

with

$$\mathbb{W} = Dl_0^{-2}\sum_{\langle ij\rangle}(a_i^\dagger - a_j^\dagger)(a_j - a_i) \tag{A.2}$$

where the sum runs over pairs of nearest neighbours on the lattice.

The formal solution of the master equation is $|P(t)\rangle = \mathrm{e}^{\mathbb{W}t}|P(0)\rangle$. If we assume for convenience that the operator O in (5) depends only on the density at a single time t_1 then it may be shown that $\langle O\rangle = \langle 0|\mathrm{e}^{\sum_i a_i}\hat{O}\mathrm{e}^{\mathbb{W}t_1}|P(0)\rangle$ where $\hat{O}$ is an operator that depends on the density $\hat{\rho} = \sum_i a_i^\dagger a_i$ in the same way that O depends on $\rho(t_1)$. Generalisation to other observables O is straightforward but we omit it, for brevity.

To make further progress, one makes a time discretisation, writing $\mathrm{e}^{\mathbb{W}t_1} = \prod_{k=1}^M \mathrm{e}^{\mathbb{W}\delta t}$ with $\delta t = t_1/M$. One also requires a formula for the identity operator [22]:

$$\hat{1} = \int\left[\prod_i \mathrm{d}(z_i, z_i^*)\right]\ \mathrm{e}^{-\sum_i z_i z_i^*}\mathrm{e}^{\sum_i z_i a_i^\dagger}|0\rangle\langle 0|\mathrm{e}^{\sum_i z_i^* a_i} \tag{A.3}$$

where z_i^* is the complex conjugate of z_i, and the notation $[\prod_i \mathrm{d}(z_i, z_i^*)]$ means that each z_i is integrated over the entire complex plane, with one integration variable z_i for each site on the lattice. Inserting this resolution of the identity between each factor of $\mathrm{e}^{\mathbb{W}\delta t}$ in the representation of $\mathrm{e}^{\mathbb{W}t_1}$, one arrives at

$$\langle O\rangle = \int\left[\prod_{ik}\mathrm{d}(z_{ik}, z_{ik}^*)\right]\langle 0|\mathrm{e}^{\sum_i a_i}\hat{O}\mathrm{e}^{\sum_i z_{iM}a_i^\dagger}|0\rangle\mathrm{e}^{-S_0}\mathrm{e}^{-\sum_i z_{i0}^* z_{i0}}\langle 0|\mathrm{e}^{\sum_i (z_{i0})^* a_i}|P(0)\rangle \tag{A.4}$$

where

$$\mathrm{e}^{-S_0} = \prod_{k=1}^M \mathrm{e}^{-\sum_i z_{ik}^* z_{ik}}\langle 0|\mathrm{e}^{\sum_i z_{ik}^* a_i}\mathrm{e}^{\mathbb{W}\delta t}\mathrm{e}^{\sum_i z_{i,k-1}a_i^\dagger}|0\rangle \tag{A.5}$$

Then, for small enough δt, one may use the explicit form of $\mathbb{W}$ to approximate the product as

$$\mathrm{e}^{-S_0} = \prod_{k=1}^M \exp\left[-\sum_i z_{ik}^*(z_{ik} - z_{i,k-1}) - Dl_0^{-2}\delta t\sum_{\langle ij\rangle}(z_{ik}^* - z_{jk}^*)(z_{i,k-1} - z_{j,k-1}) + O(\delta t)^2\right] \tag{A.6}$$

The presence of finite differences and lattice derivatives makes this expression very unwieldy. It is therefore conventional to combine (A.4) and (A.6) into a formal expression, based on the assumption that the time step δt and the lattice space l_0 are small enough that all quantities of interest vary little between adjacent points in the

space-time discretisation. Returning to a general case where O may be any functional of the density, the result is

$$\langle O\rangle = \int \mathcal{D}(z,z^*)O[zz^*]\exp\left[-\int \mathrm{d}t\int \mathrm{d}\boldsymbol{r}\,(z^*\partial_t z) - D(\nabla z^*)\cdot(\nabla z)\right] \tag{A.7}$$

where the integral over values of z at a large number of space-time points has been replaced by an integral over a (rescaled) function $z(\boldsymbol{r},t)$, and differences between adjacent points have been approximated by considering derivatives of this function. the dependence of $\hat{O}$ on $\hat{\rho}$ is reduced to a simple functional dependence on zz^*. There are boundary conditions on the functional integral that encode the initial condition $P(0)$ and the behaviour at the final time t.

Finally, to arrive at (5), one makes the change of variables

$$zz^* = \rho, \qquad z^* = \mathrm{e}^{\hat{\rho}}. \tag{A.8}$$

The transformation has unit Jacobian so the integration measure at a single spacetime point $\mathrm{d}(z_{ik}, z^*_{ik})$ becomes simply $\mathrm{d}\rho_{ik}\mathrm{d}\hat{\rho}_{ik}$, with the integration contour for $\hat{\rho}$ lying along the 'imaginary direction' in the complex plane (see also [36]). Within the path integral (5), the measure $\mathcal{D}\rho\mathcal{D}\hat{\rho}$ indicates $\prod_{ik}\mathrm{d}\rho_{ik}\mathrm{d}\hat{\rho}_{ik}$, similarly to the measure in (A.7). As long as the functions $\rho,\hat{\rho}$ can be safely assumed to be smooth, the transformation (A.8) maps $z^*\partial_t z$ to $\hat{\rho}\partial_t\rho$ and $(\nabla z^*)\cdot(\nabla z)$ to $(\nabla\hat{\rho})\cdot(\nabla\rho) - \rho|\nabla\hat{\rho}|^2$. In the hydrodynamic limit, one expects that that $\rho,\hat{\rho}$ are indeed smooth enough that higher-order derivatives can be neglected. Thus one makes the passage from (A.7) to (5), which should be valid for large N.

Appendix B. Metric structure

In this section, we sketch the metric structure associated with the Wasserstein geometry, to provide extra context for our main discussion. On a formal level, the relationships that we quote are straightforward generalisations of results from differential geometry to the space of density profiles, included here for completeness and to illustrate how the geometry of the space of density profiles can be related to the geometry of more familiar finite-dimensional curved spaces. A mathematically rigorous formulation exists but is far beyond the scope of this review [13].

The main object of interest is the metric tensor. In physics, metric tensors are familiar from differential geometry: we consider two paths $x(s)$ and $y(s)$ in a d-dimensional curved space, with $y(0) = x(0) = a$. Let the "tangent vectors" to these paths at the point a be $u = dx/ds$ and $v = dy/ds$. Then, the inner product between these two tangent vectors is $u^\mu g_{\mu\nu} v^\nu$, where $g_{\mu\nu}$ is the metric tensor (evaluated at a); we use the Einstein convention of implicit summation of repeated upper and lower indices. One can think of g as a symmetric matrix with strictly positive eigenvalues, which ensures that $u^\mu g_{\mu\nu} u^\nu \geq 0$, with equality only when $u = 0$, as expected for an inner product. In this curved d-dimensional space, the length of a path $x(s)$ is $\ell_d[x(s)] = \int \mathrm{d}s \sqrt{\frac{\mathrm{d}x^\mu}{\mathrm{d}s} g_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}s}}$ and one may also define a path action $\mathcal{E}_d[x(s)] = \int \mathrm{d}s\, \frac{\mathrm{d}x^\mu}{\mathrm{d}s} g_{\mu\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}s}$. If one minimises, for

fixed initial and final position, the functionals $\mathcal{E}_d$ and ℓ_d, then one finds that minimisers (critical points) of $\mathcal{E}_d$ are minimisers (critical points) of ℓ_d, but the converse is only true is the curve is reparametrised proportional to arc-length (the functional ℓ_d is invariant under reparametrisations, $\mathcal{E}_d$ is not). For a path from $s = a$ to $s = b$ one also has

$$\ell_d^2[x(s)] \leq (b-a)\mathcal{E}_d[x(s)]. \tag{B.1}$$

To generalise these results to the present context, the d-dimensional vector x is replaced by a density profile $\rho = \rho(\boldsymbol{r})$. One may consider the metric tensor $g_{\mu\nu}$ as a bilinear function $g(u,v) = u^\mu g_{\mu\nu} v^\nu$. The analogue of this function for the Wasserstein geometry in the space of density profiles is

$$G[u_1,u_2] = \int \mathrm{d}\boldsymbol{r}\rho\nabla\Phi_1\cdot\nabla\Phi_2 \tag{B.2}$$

where the functions $\Phi_{1,2}(\boldsymbol{r})$ solve $u_{1,2} = \nabla\cdot(\rho\nabla\Phi_{1,2})$. Just like $g(u,v)$, the functional G is bilinear, $G(au_1, bu_2) = ab\cdot G(u_1,u_2)$, as well as symmetric under interchange of u_1 and u_2. From (13) and the associated discussion, one identifies the path cost functional, $\mathcal{E}[\rho(s)]$ with $\int \mathrm{d}s\, G[\partial_s\rho, \partial_s\rho]$, the natural generalisation of the finite dimensional path action to the Wasserstein geometry. The result (15) is therefore analogous to (B.1). Also, steepest descent processes in finite-dimensional curved spaces obey $g_{\mu\nu}(\mathrm{d}/\mathrm{d}t)x^\nu = -\partial_\mu E$ where E is the energy and $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$ as usual. If $g^{\mu\nu}$ is the inverse metric tensor then one may write instead $(\mathrm{d}/\mathrm{d}t)x^\mu = -g^{\mu\nu}\partial_\nu E$. Comparing with the discussion of Sec. 3.3, we identify the operator $\mathrm{grad}^W(\cdot) = -\nabla\cdot\left[\rho\nabla\frac{\delta}{\delta\rho}(\cdot)\right]$ as the analogue of $g^{\mu\nu}\partial_\nu(\cdot)$.

So far we have shown how various quantities from the main text have analogies in finite-dimensional curved spaces. Finally, we show how these analogies provide an explanation for the apparently coincidental simplification of (22) into the form given in (21). The path cost (22) associated with steepest descent may be identified as

$$\mathcal{S}[\rho] = \frac{1}{2D}\int_a^b \mathrm{d}t\, G[\partial_t\rho + D\mathrm{grad}^W(\beta F), \partial_t\rho + D\mathrm{grad}^W(\beta F)]. \tag{B.3}$$

However, since the grad^W operator is analogous to $g^{\mu\nu}\partial_\nu$, the finite dimensional analogue of this cost is $\mathcal{S}_d = \frac{1}{2}\int \mathrm{d}t\,(\partial_t x^\mu + g^{\mu\nu}\partial_\nu E)g_{\mu\lambda}(\partial_t x^\lambda + g^{\lambda\xi}\partial_\xi E)$, where we use an energy function E as the analogue of βF, as above. If we then note that $g_{\mu\lambda}g^{\lambda\xi} = \delta_\lambda^\xi$ since $g^{\mu\nu}$ is (by definition) the inverse metric tensor, then one arrives at $\mathcal{S}_d = \frac{1}{2}\int \mathrm{d}t[(\partial_t x^\mu)g_{\mu\lambda}(\partial_t x^\lambda) + 2(\partial_t x^\mu)(\partial_\mu E) + (\partial_\mu E)g^{\mu\nu}(\partial_\nu E)]$. Noting that $(\partial_t x^\mu)(\partial_\mu E) = \partial_t E$ by the chain rule, we may integrate the second term may to obtain $E(b) - E(a)$, so that

$$\mathcal{S}_d = E(b) - E(a) + \frac{1}{2}\int \mathrm{d}t\,[(\partial_t x^\mu)g_{\mu\lambda}(\partial_t x^\lambda) + (\partial_\mu E)g^{\mu\nu}(\partial_\nu E)]. \tag{B.4}$$

Since the last two terms in this equation are time-reversal symmetric while the combination of the first two terms is odd under time-reversal, it follows that if the probability of a path $x(t)$ is given by $\mathrm{e}^{-\mathcal{S}_d}$ then there is a detailed balance symmetry between $x(t)$ and its time-reversed counterpart $x_{\mathrm{TR}}(t)$, which is $\mathrm{Prob}[x(t)]\mathrm{e}^{-E(a)} = \mathrm{Prob}[x_{\mathrm{TR}}(t)]\mathrm{e}^{-E(b)}$. To re-cast (B.4) in the Wasserstein setting, it is useful to define

a bilinear function $\tilde{G}[x,y] = \int \mathrm{d}t\, \rho(\nabla x)\cdot(\nabla y)$ which is the analogue of $x_\mu g^{\mu\nu} y_\nu$. Then, using the analogy with the finite-dimensional case, in the Wasserstein setting,

$$\mathcal{S}[\rho] = \beta F(b) - \beta F(a) + \frac{1}{2D}\int \mathrm{d}t\, G[\partial_t\rho, \partial_t\rho] + \frac{D}{2}\int \mathrm{d}t\, \tilde{G}\left[\beta\frac{\delta F}{\delta\rho}, \beta\frac{\delta F}{\delta\rho}\right]. \tag{B.5}$$

This result is equivalent to (21) which confirms its validity (it was derived independently in Section 3.3). However, writing this result as in (B.5) emphasises that it is connected to the metric tensor and its inverse, and clarifies its relation to (B.3) and hence to (22).

References

[1] H. Spohn, J. Phys. A **16**, 4275 (1983)
[2] C. Kipnis, S. Olla and S. R. S. Varadhan, Commun. Pure Appl. Math. **42**, 115 (1989).
[3] G. L. Eyink, J. Stat. Phys. **61**, 533 (1990).
[4] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Phys. Rev. Lett. **87**, 040501 (2001); J. Stat. Phys. **107**, 625 (2002).
[5] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, J. Stat. Phys. **135**, 857 (2009).
[6] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio and C. Landim, arXiv:1404.6466.
[7] J. Tailleur, J. Kurchan and V. Lecomte, Phys. Rev. Lett. **99**, 150602 (2007).
[8] J. Tailleur, J. Kurchan and V. Lecomte, J. Phys. A **41**, 505001 (2008)
[9] A. Imparato, V. Lecomte and F. van Wijland, Phys. Rev. E **80**, 011131 (2009)
[10] V. Lecomte, A. Imparato, and F. van Wijland, Prog. Theor. Phys. Supp. **184**, 276 (2010).
[11] R. Jordan, D. Kinderlehrer, F. Otto, SIAM J. Math. Anal. **29**, 1 (1998)
[12] F. Otto, Commun. Partial Differ. Equations **26**, 101 (2001).
[13] L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures (Lectures in Mathematics ETH Zürich)*, (Birkhäuser, 2005)
[14] S. Adams, N. Dirr, M. A. Peletier and J. Zimmer, Commun. Math. Phys. **307**, 791 (2011).
[15] S. Adams, N. Dirr, M. A. Peletier and J. Zimmer, Phil. Trans. Roy. Soc. A **371**, 20120341 (2013).
[16] H. Jacquin and F. van Wijland, Phys. Rev. Lett. **106**, 210602 (2011)
[17] A. Andreanov, G. Biroli and A. Lefevre, J. Stat. Mech. **2006**, P07008
[18] B. Kim and K. Kawasaki, J. Phys. A **40**, F33 (2007).
[19] H. Touchette, Phys. Rep. **478**, 1 (2009).
[20] C. Villani, *Topics in Optimal Transportation* (Am. Math. Soc., 2003).
[21] J.-D. Benamou, Y. Brenier, Numerische Mathematik **84**, 375 (2000).
[22] M. Doi, J. Phys. A **9**, 1465 (1976); L. Peliti, J. Physique **46**, 1469 (1985).
[23] P. C. Martin, E. D. Siggia and H. A. Rose, Phys. Rev. A **8**, 423 (1973).
[24] C. De Dominicis, Lett. Nuovo Cimento, **12**, 567 (1975).
[25] H.-K. Janssen, Z. Phys. B **23**, 377 (1976).
[26] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[27] M. v. Renesse, K.-T. Sturm, Ann. Probab. **37**, 1114 (2009).
[28] P. C. Hohenberg, B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[29] C. Kipnis, C. Marchioro, E. Presutti, J. Stat. Phys. **27**, 65 (1982).
[30] L. Bertini, D. Gabrielli J. L. Lebowitz, J. Stat. Phys. **121**, 843 (2005).
[31] T. Bodineau, J. L. Lebowitz, C. Mouhot, C. Villani, arXiv:1305.7405.
[32] S. Whitelam, J. P. Garrahan, J. Phys. Chem. B **108**, 6611 (2004).
[33] F. Ritort, P. Sollich, Adv. Phys. **52**, 219 (2003).
[34] H. Jacquin and F. van Wijland, arXiv:1206.1586.
[35] G. L. Eyink, J. L. Lebowitz, H. Spohn, J. Stat. Phys. **83**, 385 (1996).
[36] K. Itakura, J. Ohkubo and S. Sasa, J. Phys. A **43**, 125001 (2010)

[37] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao and R. A. Simha, Rev. Mod. Phys. **85**, 1143 (2013).
[38] M. H. Duong, M. A. Peletier and J. Zimmer, Nonlinearity **26**, 2951 (2013).