ERGODIC THEORY OF THE BURGERS EQUATION.

YURI BAKHTIN

1. Introduction

The ultimate goal of these lectures is to present the ergodic theory of the Burgers equation with random force in noncompact setting. Developing such a theory has been considered a hard problem since the first publication on the randomly forced inviscid Burgers [EKMS00]. It was solved in a recent work [BCK] for the forcing of Poissonian type.

The Burgers equation is a basic fluid dynamic model, and the main motivation for the study of ergodicity for Burgers equation probably comes from statistical hydrodynamics where one is interested in description of statistically steady regimes of fluid flows. It can also be interpreted as a growth model and the main idea of [BCK] is to look at the Burgers equation as a model of last passage percolation type. This allowed to use various tools from the theory of first and last passage percolation.

First several sections play an introductory role. We begin with an introduction to stochastic stability in Section 2. In Section 3 we briefly discuss the progress in the ergodic theory of another important hydrodynamics model, the Navier–Stokes system with random force. Section 4 is an introduction to the Burgers equation. Section 5 is a discussion of the ergodic theory of Burgers equation with random force in compact setting. In Section 6 we introduce the Poissonian forcing for the Burgers equation. In Section 7 we state the ergodic results from [Bak12] on quasi-compact setting. In Section 8 we state the main results and the proof is given in Sections 9 through 13. In Section 14 we give some concluding remarks.

Although many of the proofs given here are detailed and rigorous, often we give only the ideas behind the proof referring the reader to the details in [BCK].

2. Stability in stochastic dynamics

We begin with a very simple example, a deterministic linear dynamical system with one stable fixed point.

A discrete dynamical system is given by a transformation f of a phase space $\mathbb{X}$. For our example, we take the phase space $\mathbb{X}$ to be the real line $\mathbb{R}$ and define the transformation f by $f(x) = ax$, $x \in \mathbb{R}$, where a is a real number between 0 and 1. To any point $x \in \mathbb{X}$ one can associate its forward orbit $(x_n)_{n=0}^{\infty}$, a sequence of points obtained from $x_0 = x$ by iterations of the map f, i.e., $x_n = f(x_{n-1})$ for all

Lecture notes for the Summer School on Probability and Statistical Physics in St.Petersburg, June 2012.

The author was partially supported by NSF CAREER Award DMS-0742424.

$n \in \mathbb{N}$:

$$\begin{aligned}
x_0 &= x = f^0(x) \\
x_1 &= f(x_0) = f^1(x), \\
x_2 &= f(x_1) = f \circ f(x_0) = f^2(x), \\
x_3 &= f(x_2) = f \circ f \circ f(x) = f^3(x), \\
&\ldots.
\end{aligned}$$

A natural question in the theory of dynamical systems is the behavior of the forward orbit $(x_n)_{n=0}^{\infty}$ as $n \to \infty$, where n plays the role of time. In other words, we may be interested in what happens to the initial condition x in the long run under evolution defined by the map f. In our simple example, the analysis is straightforward.

Namely, zero is a unique fixed point of the transformation: $f(0) = 0$, and since $x_n = a^n x$, $n \in \mathbb{N}$ and $a \in (0,1)$ we conclude that as $n \to \infty$, x_n converges to that fixed point exponentially fast. Therefore, 0 is a stable fixed point, or a one-point global attractor for the dynamical system $(\mathbb{R}, f)$, i.e., its domain of attraction coincides with $\mathbb{R}$. So, due to the contraction that is present in the map f, there is a fast loss of memory in the system, and no matter what the initial condition is, it gets forgotten in the long run and the points $x_n = f^n(x)$ approach the stable fixed point 0 as $n \to \infty$.

Notice that so far we have always assumed that the evolution begins at time 0, but the picture would not change if we assume that the evolution begins at any other starting time $n_0 \in \mathbb{Z}$. In fact, since the map f is invertible in our example, the full (two-sided) orbit $(x_n)_{n\in\mathbb{Z}} = (f^n x)_{n\in\mathbb{Z}}$ indexed by $\mathbb{Z}$ is well-defined for any $x \in \mathbb{R}$.

Let us now modify the dynamical system of our first example a little by adding noise, i.e., a random perturbation that will kick the system out of equilibrium. Let us consider some probability space $(\Omega, \mathcal{F}, \mathsf{P})$ rich enough to support a sequence $(\xi_n)_{n\in\mathbb{Z}}$ of independent Gaussian random variables with mean 0 and variance σ^2. For every $n \in \mathbb{Z}$ we will now define a random map $f_{n,\omega} : \mathbb{R} \to \mathbb{R}$ by

$$f_{n,\omega}(x) = ax + \xi_n(\omega).$$

This model is known as an autoregressive-moving-average (ARMA) model of order 1.

A natural analogue of a forward orbit from our first example would be a stochastic process $(X_n)_{n\geq n_0}$ emitted at time n_0 from point x, i.e., satisfying $X_{n_0} = x$ and, for all $n \geq n_0 + 1$,

$$X_{n+1} = aX_n + \xi_n. \tag{1}$$

We want to describe the long term behavior of the resulting random dynamical system and study its stability properties. However, it is not as straightforward as in the deterministic case. It is clear that there is no fixed point that serves all maps $f_{n,\omega}$ at the same time. The solution of the equation $f_{n,\omega}(x) = x$ for some n may be irrelevant for all other values of n. Still, the system exhibits a pull from infinity towards the origin and contraction. In fact,

$$|f_{n,\omega}(x) - f_{n,\omega}(y)| = |(ax + \xi_n) - (ay + \xi_n)| = a|x - y|, \quad x, y \in \mathbb{R},$$

and for the random map $\Phi_\omega^{m,n}$ corresponding to the random dynamics between times m and n and defined by

$$\Phi_\omega^{m,n}(x) = f_{n,\omega} \circ f_{n-1,\omega} \circ \ldots \circ f_{m+2,\omega} \circ f_{m+1,\omega}(x).$$

we obtain

$$|\Phi_\omega^{m,n}(x) - \Phi_\omega^{m,n}(y)| = a^{n-m}|x-y|. \tag{2}$$

The contraction established above should imply some kind of stability. In the random dynamics setting there are several nonequivalent notions of fixed points and their stability.

One way to deal with stability is at the level of distributions. We notice that due to the i.i.d. property of the sequence (ξ_n), the process X_n defined above is a homogeneous Markov process with one-step transition probability

$$P(x,A) = \frac{1}{\sqrt{2\pi}\sigma}\int_{x\in A} e^{-\frac{(y-ax)^2}{2\sigma^2}}dy.$$

If instead of a deterministic initial condition x, say, at time 0 we have a random initial condition X_0 independent of $(\xi_n)_{n\ge 1}$ and distributed according to a distribution μ_0, then the distribution of X_1 is given by

$$\mu_1(A) = \int_{x\in\mathbb{R}} \mu_0(dx)P(x,A),$$

and a probability measure μ on $\mathbb{R}$ is called invariant or stationary if

$$\mu(A) = \int_{x\in\mathbb{R}} \mu(dx)P(x,A),$$

for all Borel sets A. If the dynamics is initiated with initial value distributed according to an invariant distribution, then the resulting process (X_n) is stationary, i.e., its distribution is invariant under time shifts.

Studying the stability of a Markov dynamical system at the level of distributions involves identification of invariant distributions and establishing convergence of the distribution of X_n to one of the stationary ones.

In our example, if X_0 is independent of ξ_1 and normally distributed with zero mean and variance D, then $X_1 = aX_0 + \xi_1$ is also centered Gaussian with variance $a^2D + \sigma^2$. So the distributions at time 0 and time 1 coincide if $a^2D + \sigma^2 = D$, i.e., $D = \sigma^2/(1-a^2)$. Therefore, the centered Gaussian distribution with variance $\sigma^2/(1-a^2)$ is invariant and gives rise to a stationary process. There are several ways to establish uniqueness of this invariant distribution, e.g., the celebrated coupling method introduced by Doeblin in 1930's which also allows to prove that for any deterministic initial data, the distribution of X_n converges exponentially fast to the unique invariant distribution in total variation as $n \to \infty$.

Another way to approach stability is studying random attractors. Let us convince ourselves that in our example, the random attractor contains only one point and that point is a global solution $(X_n)_{n\in\mathbb{Z}}$ defined by

$$X_n = X(\xi_n, \xi_{n-1}, \xi_{n-2}, \ldots) = \xi_n + a\xi_{n-1} + a^2\xi_{n-2} + \ldots.$$

Clearly, this series converges with probability 1. Moreover, for any $n \in \mathbb{Z}$,

$$\begin{aligned} aX_{n-1} + \xi_n &= \xi_n + a(\xi_{n-1} + a\xi_{n-2} + a^2\xi_{n-3} + \ldots) \\ &= \xi_n + a\xi_{n-1} + a^2\xi_{n-2} + a^3\xi_{n-3} + \ldots \\ &= X_n, \end{aligned}$$

and $(X_n)_{n\in\mathbb{Z}}$ is indeed a two-sided orbit, i.e., a global (in time) solution of equation (1) defined on the entire $\mathbb{Z}$. Notice that X_n is a functional of the history of the process ξ up to time n. Since the process ξ is stationary, thus constructed $(X_n)_{n\in\mathbb{Z}}$ is also a stationary process. In fact, X_n is centered Gaussian with variance

$$\sigma^2 + a^2\sigma^2 + a^4\sigma^2 + \ldots = \sigma^2/(1-a^2),$$

which confirms our previous computation of the invariant distribution. Let us now interpret the global solution $(X_n)_{n\in\mathbb{Z}}$ as a global attractor. We know from the contraction estimate (2) that for any $x \in \mathbb{R}$,

$$|\Phi^{m,n}(x) - X_n| = |\Phi^{m,n}(x) - \Phi^{m,n}(X_m)| = a^{n-m}|x - X_m|. \tag{3}$$

Using the stationarity of X_m and applying integration by parts, we see that

$$\sum_{m<0} \mathsf{P}\{|X_m| > |m|\} \le \sum_{m<0} \mathsf{P}\{|X_0| > |m|\} \le \mathsf{E}|X_0| < \infty.$$

The Borel–Cantelli lemma implies now that X_m grows at most linearly in m (in fact, one can prove much better estimates on the growth rate of X_m since it is a Gaussian stationary process), and (3) implies that for any $x \in \mathbb{R}$,

$$\lim_{m\to-\infty} |\Phi^{m,n}(x) - X_n| = 0.$$

In words, if we fix an initial condition x and run the random dynamics from time m to time n, then the result of this evolution converges to the special global solution X_n as we pull the starting time m back to $-\infty$. This allows us to call $(X_n)_{n\in\mathbb{Z}}$ a one-point pullback attractor for our random dynamical system.

We can adapt the reasoning above to show that there are no other global stationary solutions of (1). In fact, if $(Y_n)_{n\in\mathbb{Z}}$ is such a stationary solution then there is a number R such that for almost all ω there is a sequence $m_k(\omega) \to -\infty$ such that $|Y_{m_k}| < R$. Using the contraction estimate (2), we see that for any n and k

$$|X_n - Y_n| = |\Phi^{m_k,n}X_{m_k} - \Phi^{m_k,n}Y_{m_k}| = a^{n-m_k}|X_{m_k} - Y_{m_k}|.$$

Using that $|Y_{m_k}| < R$ and that $|X_{m_k}|$ grows not faster than linearly in m_k, we take m_k to $-\infty$ and conclude that $X_n = Y_n$ thus proving our uniqueness claim.

We can rephrase the uniqueness and convergence statements as the following One Force — One Solution Principle (1F1S): with probability 1, at any given time there is a unique value of X_n compatible with the history of the "forcing" $(\xi_k)_{k\le n}$. One can also say that X_n is a unique value worked out by the dynamics in the past up to n.

In our example where 1F1S is valid, the global solution plays the role of a one-point random attractor, and the unique invariant distribution can be recovered as the distribution of this random point at, say, time 0. In general, the picture can be more complicated. If a random dynamical system admits an invariant distribution, then at a given time there can be more than one point compatible with the history of random maps. One can consider the union of these points as a random attractor and, moreover, introduce a natural distribution on these points called the sample measure associated to the history of forcing, see [Cra08],[LY88], and [CSG11].

Notice that the invariant distribution in our example results from a balance between two factors. One factor is the decay or dissipation due to the contractive dynamics. In the absence of randomness the system would simply equilibrate to the stable fixed point. The second one is the "random forcing" that keeps the system from the rest at equilibrium, and the stronger that influence is the more the resulting stationary distribution is spread out. This mechanism is typical for many physical systems.

One can loosely define ergodic theory as the study of statistical patterns in dynamical systems in stationary regimes. One of the basic questions of ergodic theory of a dynamical system is the description of the stationary regimes. In fact, taking measurements of a system and averaging them over time makes sense only in a stationary regime. Moreover, different stationary regimes may produce different limiting values for these averages, thus making it an important task to characterize all stationary regimes for a system.

The main content of this paper is the ergodic theory (including a form of 1F1S) for the randomly forced Burgers equation on the real line. The Burgers equation is a basic fluid dynamics model, and before studying it we provide a brief view into the development of ergodic theory of stochastic hydrodynamics in the last two decades.

3. Ergodic theory for the Navier–Stokes system with random forcing

In this section we give a brief view into the ergodic theory of the Navier–Stokes system with random force. Since our goal is mainly to draw a parallel with the study of the Burgers equation in the forthcoming sections, we will avoid precise and technical formulations in our sketchy exposition.

The Navier–Stokes system describing incompressible flows of Newtonian fluids is one of the most important models in fluid dynamics. Incompressibility means that the density of the fluid is constant. Assuming that the units are chosen so that the density is identically equal to 1, the system writes as

$$\partial_t u + (u \cdot \nabla) u = \nu \Delta u - \nabla p + f, \tag{4}$$

$$(\nabla \cdot u) = 0. \tag{5}$$

Here u is the velocity profile, i.e., $u(x,t)$ denotes the vector of velocity of the particle that at a time t is located at a spatial location x. The equation is valid for describing 2-dimensional or 3-dimensional flows, so the dimension of the velocity vectors matches the dimension of the space. The left-hand side of the first equation computed at a space-time point (x,t) is the acceleration of the particle at point x at time t, and the right-hand side represents the forces exerted on the particle. The first term on the right-hand side is due to stress. Here $\nu > 0$ is the viscosity constant, and Δ means the Laplace operator. The second term ∇p is the gradient of the unknown pressure field p. The third contribution is the external volume force f.

The incompressibility of the flow is expressed by the second equation saying that the divergence of the velocity field is identical 0.

The theory of this system studies the evolution of the velocity field u in various 2- and 3-dimensional domains under a variety of boundary and initial conditions. The methods of the theory are too heavy to address in this section. Its results can be briefly summarized as follows: (i) the evolution in reasonably nice 2-dimensional domains with reasonable boundary conditions is well defined for all positive times for tame initial conditions and the solutions are smooth; (ii) much less regularity is known in 3 dimensions. For this reason we mostly restrict ourselves to 2 dimensions in this section.

If the external forcing f is random then under appropriate conditions the system (4)–(5) becomes a random dynamical system in an appropriate functional space.

The existence of an invariant distribution for the Navier–Stokes system in a bounded 2-dimensional domain D with zero boundary conditions (i.e., the fluid does not move at the boundary of the domain) and random forcing was first obtained in [Fla94] with the help of a compactness argument. Essentially, the existence of invariant distributions is based on the balance between the injection of energy into the system by the random forcing and the dissipation of energy due to the friction represented by the viscosity term involving the Laplace operator. In this respect the situation is very similar to our elementary example from Section 2.

Uniqueness turned out to be a much more intricate matter. First ([FM95]) it was established for the case where the noise excites sufficiently strongly all eigendirections of the Laplacian in D, but then this unnatural assumption was removed in [EMS01] where the case of the Navier–Stokes system on the two-dimensional torus $\mathbb{T}^2$ was considered, and uniqueness was established for the situation where sufficiently many (but finitely many) Fourier modes were excited by the noise. Eventually, with the help of Malliavin calculus, the uniqueness of invariant distributions for the stochastic Navier–Stokes system on $\mathbb{T}^2$ and similar systems was established in [HM06],[HM08], and [HM11] even in the highly hypoelliptic situation where the forcing is allowed to be highly degenerate. Other important early contributions to the problem of uniqueness of invariant distributions of the stochastic Navier–Stokes system are [BKL01] and[KS00].

All these and related results (also for other nonlinear stochastic PDEs, e.g., the Boussinesq system and reaction-diffusion systems) concern the compact case where the domain is bounded. However, as was noticed and studied in [Kuk04],[Kuk07], and [Kuk08], the invariant distributions obtained in the work cited above behave in a way that contradicts the established physics knowledge. Namely, the properties of these distributions become different from those predicted by the Kraichnan theory as viscosity tends to zero. This discrepancy can be explained by finite size effects since the inverse cascade that Kraichnan's theory of 2-dimensional turbulence is based upon is impossible in a bounded domain. This naturally brings us to the problem of confirming the existing physics theories by rigorous ergodic theory of the Navier–Stokes system on the entire $\mathbb{R}^2$ with space-time stationary noise and no assumption of compactness or periodicity. To our best knowledge, no significant progress have been made in this direction. The only ergodic result for Navier–Stokes system in the entire space seems to be [Bak06], where the Navier–Stokes dynamics in $\mathbb{R}^3$ is considered and under certain conditions on the decay of the noise at infinity a unique invariant distribution on Le Jan–Sznitman uniqueness class is constructed. The Le Jan–Sznitman setting automatically means that the viscosity is large and the solutions decay fast in space, and the situation is very far from the desired space-time homogeneous case.

Why does the noncompactness pose a serious obstacle to proving ergodic results? One can imagine the noncompact space to be split into countably many compact cells interacting in a nonlinear way. In each cell the dynamics might be nice, so it tends to bring the cell to statistical equilibrium. However, the actual closeness to the equilibrium may differ from one cell to another due to random fluctuations, and the cells that are far from equilibrium can, due to nonlinear interactions, destroy the near-equilibrium states of other cells. So, to prove ergodic results one typically has to exclude such situations.

The rest of these notes are devoted to the Burgers equation. The ergodic theory of the Burgers equation was initially created in [Sin91] and [EKMS00] for the case of dynamics on the circle. Later it was extended to the Burgers dynamics on a multi-dimensional torus and some quasi-compact situations. However, developing the ergodic theory of Burgers dynamics in the fully noncompact setting on the entire real line with space-time stationary noise has been an open problem for more then a decade. We will explain a solution of this problem that is based on the tools from the theory of last passage percolation and was obtained in [BCK].

4. Basics on the Burgers equation

The Burgers equation is a basic model of fluid dynamics. It was introduced by J.M.Burgers in late 1930's ([Bur39]) as a simplified model for turbulence:

$$\partial_t u(x,t) + u(x,t)\cdot\partial_x u(x,t) = \nu\partial_{xx}u(x,t). \tag{6}$$

Here $t\in\mathbb{R}$ is the time variable, $x\in\mathbb{R}$ is the space variable, $u(x,t)$ represents the value of the velocity of the particle located at point x at time t, and $\nu\geq 0$ is the viscosity parameter. The quadratic nonlinearity and the diffusive term of this equation are indeed similar to the Navier–Stokes system, and so are invariances and conservation laws. However, this equation turned out to be a poor turbulence model mainly due to lack of sensitivity of solutions to small perturbations of the initial data and thus lack of chaos typical for turbulent flows. Despite this fact, the Burgers equation and its generalization often appear in various contexts, from traffic modeling to studying the large scale structure of the Universe. When supplied with a random forcing term, it is also related to the KPZ universality class. We refer to [BK07] for a relatively recent survey of research on Burgulence.

In these notes we will study the inviscid Burgers equation ($\nu = 0$). What does the equation mean? We start with noticing that the left-hand side represents the acceleration of the particle at point x at time t: if the trajectory of the particle is given by a function $x(t)$, then $\dot{x}(t) = u(x(t),t)$ and, according to the chain rule,

$$\ddot{x}(t) = \partial_x u(x(t),t)\dot{x}(t) + \partial_t u(x(t),t) = u(x(t),t)\cdot\partial_x u(x(t),t) + \partial_t u(x(t),t).$$

Therefore, the inviscid Burgers equation describes a flow of particles that move with zero acceleration, i.e., with constant velocity. The characteristics of the equation, i.e., space-time curves along which the information propagates, coincide with particle trajectories, i.e., straight lines: in our case it is the information about the velocity that is carried by particles.

This description works perfectly well only for a short time interval until particles start bumping into each other. Clearly, if there are fast particles behind slow ones and each particle moves with constant velocity, then sooner or later the faster ones

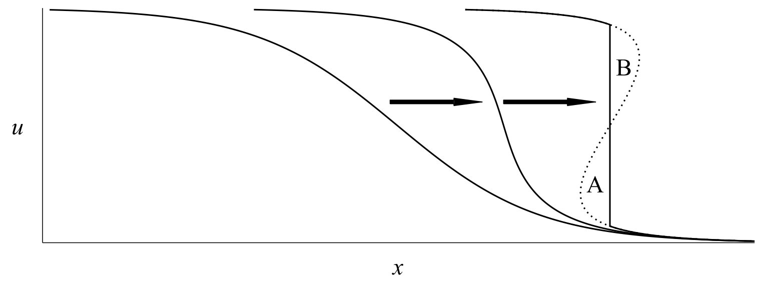

FIGURE 1. Shock formation in Burgers turbulence. The leftmost velocity profile is the initial condition at time 0. The middle one represents the result of evolution at time 1, no shock has been formed, although the profile has developed a strong negative slope — fast particles on the left are catching up with the slow particles on the right. The rightmost velocity profile is the result of the evolution at time 2. The fast particles have started bumping into the slow particles in front of them and the shock has formed. The areas A and B are equal to ensure the conservation of momentum.

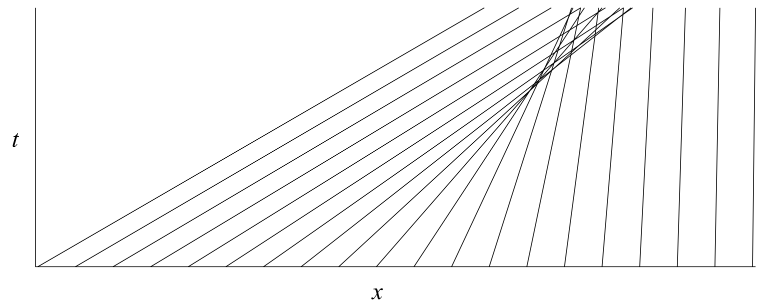

FIGURE 2. Characteristics of the Burgers equation, i.e., space-time trajectories of dust particles moving with constant velocities. At some point these lines begin to cross. This means that particles start bumping into each other.

will catch up with the slow ones, see Figures 1 and 2. If one allows particles to go through each other, then this results in multivalued velocity profiles since there will be spatial locations with multiple particles present simultaneously. However, if we insist that the solution has to be univalued, and the particles are not allowed to pass through one another, at each spatial location we have to choose one of the branches of the multivalued function thus introducing a jump discontinuity. The integral $\int u(x,t)dx$ is a conserved quantity while solutions stay smooth, and it is natural to require that this conservation law is still true after the emergence of discontinuities as well. This implies that the downward jump has to be chosen so that the areas A and B on Figure 1 are equal.

The solution with a downward jump or shock obtained from the conservation law coincides with the point-wise limit of smooth solutions to the viscous Burgers

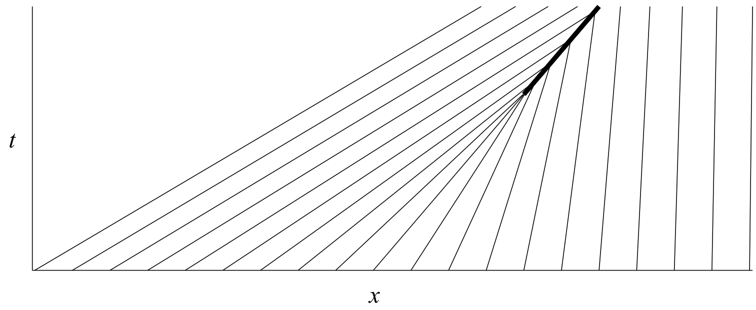

FIGURE 3. At some point the characteristics start crossing each other. This means that particles start bumping into each other.

equation as the viscosity tends to zero (the viscous Burgers equation can be explicitly solved by the Hopf–Cole logarithmic substitution reducing it to the linear heat equation). These solutions are called viscosity solutions. They can be understood as generalized solutions of the Burgers equation. In fact the class of generalized solutions of the Burgers equation is vast and includes a lot of unphysical solutions with upward jumps, but the viscosity solution (also called the entropy solution) admitting only downward jumps is unique and meaningful from the physics point of view, so we will study only entropy solutions. We postpone a precise variational characterization for entropy solutions of the initial value problem to the end of this section where a more general case of the Burgers equation with external forcing will be addressed.

After shocks are formed they keep moving in space absorbing incoming particles from left and right, see Figure 3. The dynamics of the shock location satisfies the Rankine–Hugoniot condition:

$$\dot{x} = \frac{u(x+0) + u(x-0)}{2},$$

i.e., the shock moves with velocity that is the average between the velocities of incoming particles being absorbed by the shock on the left and on the right. One can view the shock as a clump of particles stuck together and thus the Burgers equation can be said to describe the pressureless dynamics of sticky dust, i.e., particles that do not interact until they collide and stick together. In fact the shocks can never disappear, but two shocks can coalesce. The result is a hierarchical tree-like structure of shocks in space-time, see Figure 4.

Another way to look at the discontinuities of the Burgers equation is to interpret the particles absorbed by the shocks as disappearing from the system. Thus the dissipation of energy in the system happens at the shocks: for smooth solutions the integral $\int u^2(x,t)dx$ is a conserved property, but if discontinuities are present, this integral will decay.

Due to this dissipation, the solution to the unforced Burgers equation will (under fairly general conditions) relax to one of equilibrium steady states, i.e., profiles with constant velocity. More interesting dynamics will arise if we start injecting energy into the system by external perturbations thus keeping the system away from the equilibrium. If one adds a forcing term into the right-hand side of the inviscid

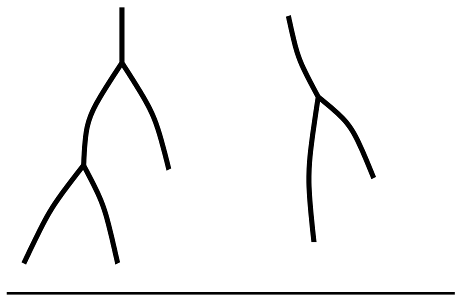

FIGURE 4. Hierarchical structure of shocks in space-time.

Burgers equation, the resulting equation will be

$$\partial_t u(x,t) + u(x,t) \cdot \partial_x u(x,t) = f(x,t), \tag{7}$$

and the particles that in the unforced case moved with zero acceleration are now supposed to move in the field of prescribed accelerations given by the term $f(x,t)$. If $x(t)$ is the trajectory of a particle, then

$$\ddot{x}(t) = f(x(t),t). \tag{8}$$

The characteristics of the system, i.e., the particle trajectories, are not straight lines any more, they are curved, but aside from that the picture stays the same: one has to introduce entropy solutions with downward shocks to deal with particles bumping into each other.

A detailed description of these solutions is given by the Lax–Oleinik variational principle which allows for an efficient analysis of the system via studying the minimizers of the corresponding Lagrangian system. Namely, the velocity field can be represented as $u(x,t) = \partial_x U(x,t)$, where the potential $U(x,t)$ is a solution of the Hamilton–Jacobi equation

$$\partial_t U(x,t) + \frac{(\partial_x U(x,t))^2}{2} - F(x,t) = 0, \tag{9}$$

F being the forcing potential: $\partial_x F(x,t) = f(x,t)$. The entropy solution of the Cauchy problem for this equation with initial data $U(\cdot,t_0) = U_0(\cdot)$ can be written as

$$U(x,t) = \inf_{\gamma:[t_0,t]\to\mathbb{R}} \left\{ U_0(\gamma(t_0)) + \frac{1}{2}\int_{t_0}^t \gamma^2(s)ds + \int_{t_0}^t F(\gamma(s),s)ds \right\}, \tag{10}$$

where the infimum is taken over all absolutely continuous curves γ satisfying $\gamma(t) = x$. The functional of paths in the right-hand side of this equation is called the action.

Also, optimal paths providing the minimum value in the minimization problem can be identified with particle trajectories. In particular, the velocity value $u(x,t) = \partial_x U(x,t)$ can also be found as the velocity $\dot{\gamma}(t)$ of the optimal path γ at the terminal point. For most of the space-time points (x,t) the minimizer is unique, but at there are points where there is more than one minimizer, and these nonuniqueness points correspond to shocks where particles coming from different initial locations meet.

We will denote by $\Psi_f^{t_0,t} w$ the solution at time t of the initial-value problem with forcing f and initial velocity w given at time t_0.

In $F \equiv 0$, then action minimizers are straight lines which is consistent with our picture of unforced Burgers equation:

Lemma 4.1. *Suppose $F(x,t) \equiv 0$. Then for any two points $x_0, x_1 \in \mathbb{R}$ and any two times $t_0 < t_1$ the minimum of kinetic action*

$$S^{t_0,t_1}(\gamma) = \frac{1}{2}\int_{t_0}^{t} \dot{\gamma}^2(s)ds$$

over paths satisfying $\gamma(t_0) = x_0$ and $\gamma(t_1) = x_1$ is attained on the straight line

$$\gamma^*(t) = x_0 + (t - t_0)\frac{x_1 - x_0}{t_1 - t_0}, \quad t \in [t_0, t_1],$$

corresponding to motion with constant velocity $v = (x_1 - x_0)/(t_1 - t_0)$. The minimal action is

$$S^{t_0,t_1}(\gamma^*) = \frac{v^2(t_1 - t_0)}{2} = \frac{(x_1 - x_0)^2}{2(t_1 - t_0)} = \frac{(x_1 - x_0)v}{2}. \tag{11}$$

PROOF: From variational calculus we know that minimizers must satisfy the Euler–Lagrange equation, which is, in our case, equation (8) with zero right-hand side. So, all minimizers have constant velocity, i.e., they are straight lines. Relation (11) is a result of a direct computation. □

If the forcing potential is random, then, due to the variational principle, to understand the long-term behavior of the Burgers dynamics one needs to study the behavior of random action minimizers over long time intervals. This task can be viewed as an asymptotic problem in random media.

The optimal paths in the variational principle benefit from large in magnitude negative values of the potential term F so they tend to visit the space-time spots where F is large negative. However, the kinetic action term containing $\dot{\gamma}^2/2$ penalizes large velocities which makes it impossible for the optimal path to freely move between best spots. The interaction of these two terms is the main subject of the analysis.

5. The Burgers equation with random forcing in compact setting

The first ergodic results for the inviscid Burgers equation concerned the case of dynamics on the unit circle $\mathbb{T}^1 = \mathbb{R}^1/\mathbb{Z}^1$ that can be viewed as the segment $[0,1]$ with identified endpoints. An equivalent view is considering the equation with forcing and velocity profiles that are 1-periodic in space (this is often called periodic boundary conditions). In [EKMS00] the forcing $f(x,t) = -\partial_x F(x,t)$ was assumed to be smooth in space and white noise type in time:

$$F(x,t) = \sum_{j=1}^{n} V_j(x)\dot{W}_j(t),$$

where $n \in \mathbb{N}$, $V_j, j = 1, \ldots, n$ are smooth functions on $\mathbb{T}^1$ (or, equivalently, periodic functions on $\mathbb{R}^1$), and $W_j, j = 1, \ldots, n$ are independent Wiener processes.

The Burgers dynamics defined by (10) preserves the velocity integral $\int u = \int_{\mathbb{T}^1} u(x)dx$, so it is sufficient to study sets $\mathbb{X}_c = \{u : \int u = c\}$, $c \in \mathbb{R}$ separately. Moreover, the dynamics commutes with Galilean shear transformations $(x,t) \to (x + vt, t)$ that correspond to switching to a new coordinate system moving with constant velocity v with respect to the old one. Due to this shear invariance, it is sufficient to study only the set $\mathbb{X}_0$.

The main statement of [EKMS00] is a 1F1S for the randomly forced Burgers equation on $\mathbb{X}_0$ (and thus on $\mathbb{X}_c$ for every c). There is a functional Φ such that

$$u(\cdot, t) = \Phi(\pi_t f)$$

is a unique in $\mathbb{X}_0$ stationary global solution of the Burgers equation with forcing f, where $\pi_t f$ is the history of forcing f up to time t. The distribution of $\Phi(\pi_0 f)$ is a unique invariant distribution on $\mathbb{X}_0$ for the associated Markov semigroup. The uniqueness of an invariant distribution is often called "unique ergodicity".

For any initial profile $w \in \mathbb{X}_0$ and every $t \in \mathbb{R}$, $\Psi_f^{t_0,t} w \to u(\cdot, t)$ as $t_0 \to -\infty$, so the global solution $u(\cdot, t), t \in \mathbb{R}$, plays the role of a one point pullback attractor.

The method of constructing the global solution and proving its attractor property is the following. Let us fix an initial condition $w \in \mathbb{X}_0$. To find $\Psi_f^{t_0,t} w$ one has to solve the variational problem (10) for all $x \in \mathbb{T}^1$. For most points x, the optimal action is attained by a unique minimizer $\gamma_{t_0,t,x}$. The key claim is that as $t_0 \to -\infty$, these minimizers converge to a limiting one-sided infinite trajectory $\gamma_{-\infty,t,x}$, uniformly on bounded intervals. Moreover, the limiting paths do not depend on $w \in \mathbb{X}_0$. They are actually infinite one-sided minimizers of the action, i.e., if γ is another absolutely continuous trajectory defined on $(-\infty, t]$ such that $\gamma(t) = x$ and $\gamma(s) = \gamma_{-\infty,t,x}(s)$ for some $t_0 < t$ and all $s \leq t_0$, then

$$A^{t_0,t}(\gamma) \geq A^{t_0,t}(\gamma_{-\infty,t,x}),$$

where

$$A^{t_0,t}(\gamma) = \frac{1}{2} \int_{t_0}^{t} \dot{\gamma}^2(s) ds + \int_{t_0}^{t} F(s, \gamma(s)) ds.$$

Given the field of one-sided minimizers, the global solution can be defined as $u(x, t) = \dot{\gamma}_{-\infty,x,t}(t)$.

The program developed in [EKMS00] includes proving existence and uniqueness of one-sided minimizers, convergence of finite minimizers to infinite ones, and their hyperbolicity property (i.e., every two one-sided minimizers converge to each other exponentially fast in reversed time).

Let us explain one step in this program. Let us consider one point (x, t). We claim that there is a number $C > 0$ depending on the random realization of the forcing in the past such that for any initial condition $w \in \mathbb{X}_0$ the minimizer γ for the corresponding variational problem on any time interval $[t_0, t]$ satisfies $|\dot{\gamma}(t)| \leq C$.

An optimal trajectory must solve the Euler–Lagrange equation (8) (which is consistent with our interpretation of minimizing paths as particle trajectories). Since $\gamma(t) = x$ is fixed, the entire solution of (8) is uniquely determined by the terminal condition $\dot{\gamma}(t) = a$. If a is sufficiently large, then the velocity $\dot{\gamma}$ stays large for values of time sufficiently close to t. Therefore, γ will quickly go around the circle and come back to the initial position x. Such a path will accumulate large kinetic action, and it cannot be optimal since the path staying at x accumulates zero kinetic action. Of course, for both paths there will be a contribution from the external potential, but these contributions are bounded on a bounded time interval, and the discrepancy between the kinetic actions of both paths can be made arbitrarily large by choosing $|a|$ to be large. This reasoning shows that the terminal velocity of minimizers over a finite interval $[t_0, t]$ is bounded by a number that does not depend on t_0.

It follows that there is a sequence of initial times such that the corresponding sequence of terminal velocities converges to a limiting value b. The corresponding paths converge (uniformly on any finite interval) to the solution of the Euler–Lagrange equation with terminal position x and velocity b. It is easily checked that the limit of a sequence of minimizers is a minimizer. Therefore, the limiting trajectory is a minimizer on any bounded interval and thus a one-sided minimizer.

The uniqueness of the limiting one-sided infinite trajectory and independence of the result of this procedure of the initial condition $w \in \mathbb{X}_0$ can be explained in the following way: in the variational problem, the contribution from the initial condition is bounded, so in the long run it is dominated by the contribution from the external forcing. This is a useful argument that also explains the loss of memory and long-term contraction in the system, but the rigorous proof is much more involved than this description and we refer to [EKMS00] for details.

The ergodic theory of the Burgers equation has also been developed on multidimensional tori $\mathbb{T}^d$ and stochastic forcing in [GIKP05], [IK03] and for the one-dimensional case with random boundary conditions in [Bak07]. In these situations a version of a variational principle holds and the theory is based on the analysis of minimizers in the respective compact domain.

It is important to stress that compactness plays a very important role. For example, the above argument for boundedness of the velocity of minimizers depends crucially on the compactness of the circle. In the noncompact situation on the entire real line without periodicity this argument fails, and the question about the behavior of minimizers over growing time intervals is more intricate. Even the fact that the variational problem has a well-defined solution on a finite interval needs an explanation because in principle if paths are allowed to travel arbitrarily far, they can visits spots with arbitrarily large in magnitude negative F. Further, as t_0 is pulled to $-\infty$, new spots with low values of F will be uncovered arbitrarily far in the space, and potentially this can lead to unboundedness of the velocity of minimizers at the terminal point.

So, due to these considerations, the long-term behavior of the Burgers equation with random force in noncompact setting has been considered a hard problem. The ergodic theory of the Burgers equation was constructed in [Bak12] for quasi-compact case and in [BCK] for fully noncompact case of space-time stationary noise. These papers use a special kind of forcing introduced in [Bak12] — the Poissonian forcing concentrated at the points of a Poisson point field. We proceed to describe this model.

6. The Burgers equation with Poissonian forcing

From now on we adopt the picture where the space axis of $\mathbb{R}^2$ is horizontal and the time axis is vertical and directed upward as on Figures 2–4.

In the variational problem (10) the paths accumulate random contributions from the external forcing potential. The goal of this section is to introduce a special random potential that is somewhat generic and simultaneously is easy to visualize, analyze, and perform computations with. We will assume that the forcing is concentrated on a random discrete set of points called the Poisson point field, each point (x_i, s_i) coming with a weight ξ_i :

$$F(x,t) = \sum_i \xi_i \delta_{(x_i,s_i)}(x,t).$$

So we are going to replace the term $\int_{t_0}^{t} F(\gamma(s), s)ds$ in (10) by $\sum_{(x_i,s_i)} \xi_i$ where the summation is taken over the Poissonian points that the path γ passes through. In fact, points (x_i, s_i) with $\xi_i \geq 0$ will be bypassed by the optimal paths, so the situation will not change if we assume that $\xi_i < 0$. Moreover, for simplicity we are going to set $\xi_i \equiv -1$, although most of our results will be valid if the distribution of ξ_i has exponential tails. If we set $\xi_i \equiv -1$ then the contribution from the external forcing is simply minus number of Poissonian points visited by γ. Now let us be more precise.

Formally, we will consider a Poissonian point field ω on space-time $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ with intensity measure $\mu(dt \times dx)$. Since we want the forcing to be stationary in time, we shall always assume that the intensity is a product measure:

$$\mu(dx \times dt) = m(dx) \times dt.$$

We will always consider one of the following two cases: (i) the quasi-compact case where the measure m is finite and most of the points are concentrated in a compact spatial domain; and (ii) the fully noncompact case of space-time homogeneous Poisson point field where m is the Lebesgue measure on $\mathbb{R}$ and μ is the Lebesgue measure on $\mathbb{R}^2$.

The configuration space Ω is the space of all locally finite point sets in space-time. For a Borel set $A \subset \mathbb{R}^2$, we use $\omega(A)$ to denote the number of Poissonian points in A, and introduce the σ-algebra $\mathcal{F}$ generated by maps $\omega \mapsto \omega(A)$ with A running through all bounded Borel sets. The probability measure P is such that for any bounded Borel set A, $\omega(A)$ is a Poisson random variable with mean equal to $\mu(A)$, and for disjoint bounded Borel sets $A_1, \ldots, A_m$, the random variables $\omega(A_1), \ldots, \omega(A_m)$ are independent.

Often, we treat the point configuration ω as a locally bounded Borel measure with a unit atom at each point of the configuration. For background on Poisson point fields, also called Poisson processes, we refer to [DVJ03].

There is a natural family of time-shift operators $(\theta^t)_{t\in\mathbb{R}}$ on the Poisson configurations: the configuration $\theta^t\omega$ is obtained from ω by shifting each point $(x, s) \in \omega$ to $(x, s - t)$.

The space of velocity potentials that we will consider will be $\mathbb{H}$, the space of all locally Lipschitz functions $W : \mathbb{R} \to \mathbb{R}$ satisfying

$$\liminf_{x\to+\infty} \frac{W(x)}{x} > -\infty,$$
$$\limsup_{x\to-\infty} \frac{W(x)}{x} < +\infty.$$

Although it is possible to work with weaker conditions, some restrictions on the growth rate of $W(x)$ as $x \to \pm\infty$ are necessary to control velocities of particles coming from $\pm\infty$.

Let us define random Hamilton–Jacobi–Hopf–Lax–Oleinik (HJHLO) dynamics on $\mathbb{H}$ (the terminology is borrowed from [Vil09, Definition 7.33]). For a function $W \in \mathbb{H}$, a Poissonian configuration ω, and an absolutely continuous trajectory (path) γ defined on $[t_0, t]$, we introduce the action

$$A_{\omega}^{t_0,t}(W, \gamma) = W(\gamma(s)) + S^{t_0,t}(\gamma) - \omega^{t_0,t}(\gamma), \tag{12}$$

where S is the kinetic action, and $\omega^{t_0,t}(\gamma) = \omega(\{(\gamma(s), s) : \ r \in [t_0, t)\})$ denotes the number of configuration points that γ passes through. The last term in (12) is responsible for the interaction with the external forcing potential corresponding to the realization of the Poissonian field.

We now consider the following minimization problem:

$$\begin{aligned} A^{t_0,t}_{\omega}(W,\gamma) &\to \inf, \\ \gamma(t) &= x. \end{aligned} \tag{13}$$

Notice that the optimal trajectories are given by straight lines for any time interval on which the trajectory stays away from the configuration points. Since Poissonian configurations are locally finite, it is sufficient to take the minimum over broken lines with vertices at configuration points.

Inspecting (12), we see that the optimal paths solving (13) try to visit as many Poissonian points as possible, but the kinetic action term penalizes large velocities, thus preventing too wild oscillations of the minimizers in their chase after Poissonian points.

The variational problem (13) has a well-defined solution:

Lemma 6.1. *With probability 1, for any $W \in \mathbb{H}$, any $x \in \mathbb{R}$ and any t_0, t with $t_0 < t$ there is a path γ^* that realizes the minimum in* (13)*. The path γ^* is a broken line with finitely many segments, all its vertices belong to ω.*

The main potential problem in proving this lemma is that paths can wander arbitrarily far in space. However the problem can easily be compactified: the kinetic action is at least quadratic in the total distance traveled by the path and the number of Poissonian points in both quasi-compact and space-homogeneous case grows at most linearly with the area, so the problem reduces to finding an optimal path within a large but compact rectangle. For details, see Lemma 2.1 in [BCK].

Let us denote the infimum (minimum) value in the variational problem (13) by ${}^{t_0,t}_{\omega}W(x)$. Our main goal is to understand the asymptotics of random nonlinear operators $\Phi^{t_0,t}_{\omega}$ as $t - t_0 \to \infty$.

The following statement can be called the cocycle property for the operator family $(\Phi^{s,t}_{\omega})$ (for general background on the cocycle property, we refer to [Arn98]). It is a direct consequence of Bellman's principle of dynamic programming.

Lemma 6.2. *For almost all If ω, the following is true simultaneously for all $W \in \mathbb{H}$ and any s, r, t satisfying $s < r < t$: $\Phi^{r,t}_{\omega}\Phi^{s,r}_{\omega}W$ is well-defined and equals $\Phi^{s,t}_{\omega}W$. If γ is an optimal path realizing $\Phi^{s,t}_{\omega}W(x)$, then the restrictions of γ on $[s, r]$ and $[r, t]$ are optimal paths realizing $\Phi^{s,r}_{\omega}W(\gamma(r))$ and $\Phi^{r,t}_{\omega}(\Phi^{s,r}_{\omega}W)(x)$.*

Introducing $\Phi^{t}_{\omega} = \Phi^{0,t}_{\omega}$ we can rewrite the cocycle property as

$$\Phi^{0,t+s}_{\omega}W = \Phi^{s}_{\theta^t\omega}\Phi^{t}_{\omega}W, \quad s, t > 0.$$

Since potentials are naturally defined up to an additive constant, it is convenient to work with $\hat{\mathbb{H}}$, the space of equivalence classes of potentials from $\mathbb{H}$. The cocycle Φ can be projected on $\hat{\mathbb{H}}$ in a natural way. We denote the resulting cocycle on $\hat{\mathbb{H}}$ by $\hat{\Phi}$.

Let us now explain how the dynamics that we consider is connected to the classical inviscid Burgers equation. One way to describe this connection is to introduce a mollification of the Poisson integer-valued measure. Let us take smooth kernels $\phi, \psi : \mathbb{R} \to [0, \infty)$ with bounded support, satisfying $\int_{\mathbb{R}} \phi(t)dt = 1$ and $\max_{x\in\mathbb{R}} \psi(x) = 1$, and for each $\varepsilon > 0$ consider the potential of shot-noise type:

$$F_\varepsilon(x,t) = -\frac{1}{\varepsilon} \sum_{(y,s)\in\omega} \phi\left(\frac{t-s}{\varepsilon}\right) \psi\left(\frac{x-y}{\alpha(\varepsilon)}\right),$$

where α is any function satisfying $\lim_{\varepsilon\downarrow 0} \alpha(\varepsilon) = 0$.

Lemma 6.3. *With probability 1, for all $s, t, x \in \mathbb{R}$ and $W \in \mathbb{H}$, the entropy solution $U_\varepsilon(x,t)$ of the Cauchy problem for the Hamilton–Jacobi equation with smooth forcing potential $F_\varepsilon(\cdot,\cdot)$ converges, as $\varepsilon \to 0$, to $U(x,t) = \Phi^{s,t}_\omega W(x)$.*

The next statement shows that away from the Poissonian points the system we consider behaves as unforced Burgers dynamics.

Lemma 6.4. *For all $\omega \in \Omega_1$, $s \in \mathbb{R}$, $W \in \mathbb{H}$, the function $U(x,t) = \Phi^{s,t}_\omega W(x)$ is an entropy solution of the Hamilton–Jacobi equation*

$$\partial_t U + \frac{(\partial_x U)^2}{2} = 0. \tag{14}$$

in $((s,\infty)\times\mathbb{R}) \setminus \omega$.

Equivalently, $u(x,t) = \partial_x U(x,t)$ is an entropy solution of the Burgers equation (7) with $f \equiv 0$ in $((s,\infty)\times\mathbb{R}) \setminus \omega$. Of course, for each t, $u(x,t)$ is a piecewise continuous function of x, and at each of the countably many discontinuity points it makes a negative jump. Since a velocity field determines its potential uniquely up to an additive constant we can also introduce dynamics on velocity fields. We can introduce the space $\mathbb{H}'$ of functions w (actually, classes of equivalence of functions since we do not distinguish two functions coinciding almost everywhere) such that for some function $W \in \mathbb{H}$ and almost every x, $w(x) = W'(x)$. This allows us to introduce the Burgers dynamics. We will say that $w_2 = \Psi^{s,t}_\omega w_1$ if $w_1 = W_1'$, $w_2 = W_2'$, and $W_2 = \Phi^{s,t}_\omega W_1$ for some $W_1, W_2 \in \mathbb{H}$. The main results of these notes describe the ergodic properties of the random dynamical system (cocycle) Ψ.

The solutions of the Burgers equation belonging to $\mathbb{H}'$ often have negative jump discontinuities (shocks). Although it is not essential, we can require the functions in $\mathbb{H}'$ to be right-continuous.

We already know that away from the Poissonian points the system that we consider behaves exactly as the solution of the classical Burgers equation (Lemma 6.4). So what is the behavior of the system at or near a Poissonian point? Let us first consider a model situation where a smooth beam of Burgers particles encounters a forcing point at the origin at time 0. Let us assume that at time 0, the velocity vector field near 0 is $u_0(y) = a + by$, where $b > 0$.

It is clear that for every (t, x) with $t > 0$ and x close to the origin there are two minimizer candidates. The minimizer either passes through the origin or it does not. If it does then (assuming there are no other point sources of forcing) it has to be a straight line connecting the origin to (t, x), and the accumulated action is $A_1(t,x) = x^2/(2t) - 1$, where -1 is the contribution of the forcing point at the origin, and $x^2/(2t)$ is the action accumulated while moving with constant velocity x/t between 0 and t. If the minimizer does not pass through the origin, then it is a straight line connecting some point $(0, x_0)$ to (t, x). On the one hand, the velocity of the particle associated with the minimizer is $(x-x_0)/t$. On the other hand, it has to coincide with $u_0(x_0) = a + bx_0$. Therefore, we can find $x_0 = (x - at)/(1 + bt)$.

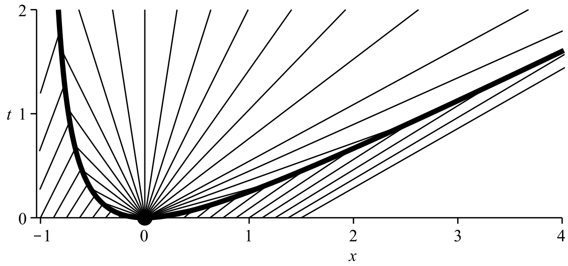

FIGURE 5. Minimizers around a Poissonian point

Taking into account that $U_0(x_0) = ax_0 + bx_0^2/2$, we can compute that the total action of that path is $A_2(t,x) = (bx^2 + 2ax - a^2t)/(2(1+bt))$.

To see which of the two cases is realized for (t,x) we must compare $A_1(t,x)$ and $A_2(t,x)$. If $A_1(t,x) < A_2(t,x)$, then the particle arriving to x at time t is at the origin at time 0. If $A_1(t,x) > A_2(t,x)$, then the particle arriving to x at time t is one of the particles that moved with constant velocity and was a part of the incoming beam. If $A_1(t,x) = A_2(t,x)$, then both of these paths are minimizers, and at time t there is a shock at point x. The relation $A_1(t,x) = A_2(t,x)$ can be rewritten as

$$(x - at)^2 = 2t(1 + bt).$$

For small values of t, the set of points satisfying this relation looks like a parabola $(x - at)^2 = 2t$, see Figure 6 where an example with $a = 1$ and $b = 1/2$ is shown.

We see that when a Poissonian point appears, it emits a continuum of particles each moving with constant velocity, creating two shock fronts moving (at least for a short time) to the left and right.

It is important to notice that in our model case with a forcing point at the origin, $u(t,x) = x/t$ for all points connected to the origin by a minimizing segment. It means that for each time t, the velocity is linear in the domain of influence of the forcing point, and the velocity gradient decays with time as $1/t$.

In general, the behavior of this kind occurs near each forcing point, and in the long run more and more points of the space-time plane get assigned to forcing points. Grouping together points assigned to the same forcing point, we obtain a tessellation of space-time into domains of influence of forcing points. Inside each domain or cell the velocity field is linear in x if the time t is fixed. The following lemma serves as a rigorous description of this picture.

For a Borel set $B \subset \mathbb{R}^2$, we denote the restriction of ω to B by $\omega\big|_B$.

Lemma 6.5. *For almost every $\omega \in \Omega$, for all $W \in \mathbb{H}$, $s, t \in \mathbb{R}$ with $s < t$, the following holds true:*

(1) *For any $p \in \omega|_{\mathbb{R}\times[s,t)}$, the set O_p of points $x \in \mathbb{R}$ such that p is the last configuration point visited by a unique minimizer for problem* (13)*, is open. Also, the set of points with a unique minimizer that does not pass through any configuration points is open. The union of these open sets is dense in $\mathbb{R}$.*

(2) *If x_0 belongs to one of these open sets, $\gamma(t) = x_0$, and $A^{s,t}_\omega(W,\gamma) = \Phi^{s,t}_\omega W(x_0)$, then $\Phi^{s,t}_\omega W(x)$ is differentiable at x_0 w.r.t. x, and*

$$\frac{d}{dx}\Phi^{s,t}_\omega W(x)\big|_{x=x_0} = \dot\gamma(t).$$

At a boundary point x_0 of any of the open sets introduced above, the right and left derivatives of $\Phi^{s,t}_\omega W(x)$ w.r.t. x are well defined. They are equal to the slope of, respectively, the leftmost and rightmost minimizers realizing $\Phi^{s,t}_\omega W(x_0)$.

(3) *For any $p \in \omega|_{\mathbb{R}\times[s,t)}$, the function $x \mapsto \frac{d}{dx}\Phi^{s,t}_\omega W(x)$ is linear in O_p.*

(4) *The function $x \mapsto \Phi^{s,t}_\omega W(x)$ is locally Lipschitz.*

Proof: Part 1 follows from the observation that the action depends on the path continuously if the family of configuration points visited by the path is fixed.

In the situation where γ does not pass through any configuration points, Part 2 follows from the theory for unforced Burgers equation, see e.g., [Lio82, Section 1.3]. If the minimizer γ for x has a straight line segment connecting a configuration point $p = (y, r)$ to (x, t), then

$$\frac{d}{dx}(\Phi^{s,t}_\omega W)(x) = \frac{d}{dx}\frac{(x-y)^2}{2(t-r)} = \frac{x-y}{t-r} = \dot\gamma(t), \tag{15}$$

and Part 2 is proven for points with unique minimizers. The case of the boundary points where the minimizers are not unique is considered similarly.

Part 3 is a direct consequence of Part 2. Part 4 follows from local boundedness of the slope of minimizers. For minimizers not passing through any configuration points, this is a consequence of the classical theory of unforced Burgers equation, and for minimizers passing through some configuration points it follows from (15). □

Let us recall that the energy is dissipated at the shocks. By seeding new particles at each Poissonian point, the forcing pumps energy into the system and, therefore, we can hope that there is a dynamical or statistical energy balance in the system necessary for existence of invariant distributions studied in our ergodic results.

Let us explain why we introduce a new kind of forcing. The continuous white-noise forcing presents considerable technical difficulties which is probably the main reason why even the existence of one-sided minimizers in fully noncompact setting was not known before [BCK]. The Poisson forcing is a model that, compared to white noise, is relatively easy to visualize, argue about, perform computations with. On the other hand, at large scales the differences between the behavior of the two models should not be essential.

Another reason for introducing Poissonian forcing is that the resulting model is similar to the Hammersley process which in turn is related to last passage percolation and longest increasing subsequences, the models that have attracted a lot of attention in last two decades with considerable progress achieved and tools developed.

Let us recall what the Hammersley process is. Let us consider a Poisson point field with Lebesgue intensity measure on $\mathbb{R}^2$. For any two points $x, y \in \mathbb{R}^2$ we consider the following maximization problem: among up-right paths (i.e., paths $(x_1(t), x_2(t))$ with $x_1(\cdot)$ and $x_2(\cdot)$ nondecreasing functions of time) starting at x

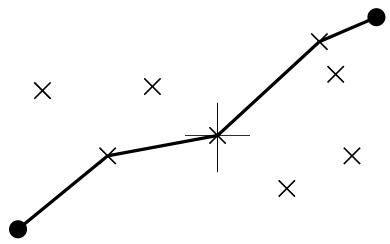

FIGURE 6. Hammersley's process.

and terminating at y, find a path that maximizes the number of Poissonian points it visits, see Figure 6.

This in indeed very similar to our optimization problem (12)–(13). The paths in Hammersley process, however, are subject to hard-core constraints. Namely, the up-right property means that the paths are identified with graphs of 1-Lipschitz functions if looked at the coordinate system where the line $x_1 = x_2$ plays the role of time. In other words, the velocity of these paths in an appropriate coordinate frame is bounded by 1. In (12)–(13) there are no hard core constraints, they are replaced by soft penalization: the kinetic action term only penalizes large velocities but does not prohibit them. This comparison gives hope that methods developed for Hammersley process, its analogues and generalizations in [AD95],[Wüt02],[CP11],[CP] can be used to study the Burgers equation with Poisson forcing. It is also clear that potential unboundedness of path velocity is going to be a major source of difficulties in implementing these methods.

Many of the tools that we are using were, in fact, invented and developed in [New95], [HN01], [HN99], [HN97], [Kes93] for first-passage percolation models.

Let us finish this section with a simulation showing what the optimal paths look like in practice. Figure 7 shows minimizing paths for variational problem (13) with zero initial condition and a simulated Poissonian environment with Lebesgue intensity measure restricted to the square $[0, 300] \times [0, 300]$. Only paths with Poissonian endpoints are shown.

7. Quasi-compact case

In this section based on the results from [Bak12] we briefly discuss the quasi-compact case where $m(\mathbb{R}) < \infty$. In this case, with probability 1, in each strip $\mathbb{R} \times [s, t]$ there are finitely many Poissonian points. Also, most Poissonian points appear in a large but compact interval. Therefore, most of the minimizers do not spend much time far away from the origin. Using this along with sub-additivity arguments and large deviation estimates, one can prove the following lemma.

Lemma 7.1. *There are random variables r and $(\tau_R)_{R>0}$ such that for any $R > 0$, with probability 1, the following implication holds true. If points x, y satisfy $|x|, |y| < R$, and times t_-, t_+ satisfy $t_-, t_+ > \tau_R$, then any action minimizer γ connecting point x at time $-t_-$ to point y at time t_+ satisfies $|\gamma(0)| < r$.*

The meaning of this main localization lemma is that there is a random radius r such that minimizers over long time intervals containing time zero are necessarily

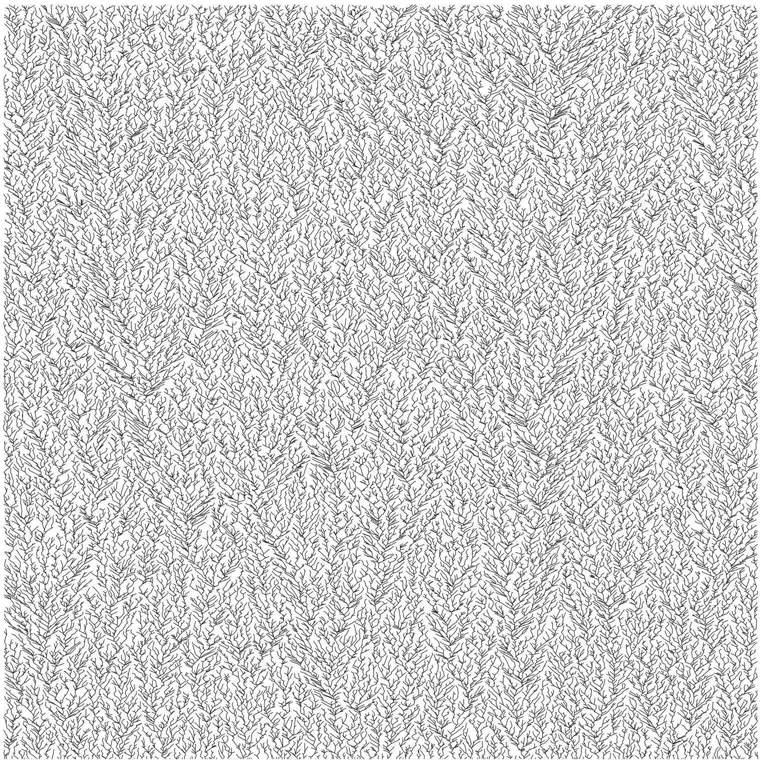

FIGURE 7. Optimal paths in a square 300 by 300 with zero initial conditions. The space axis is horizontal, and time is directed upward. Simulation and graphics by Gautam Goel.

within distance r from the origin. It is natural to call r the radius of localization at time 0.

Applying Lemma 7.1 to time shifts $\theta^t\omega$ of the Poisson points, we conclude that there is a stationary stochastic process $r_t(\omega) = r(\theta^t\omega)$ such that long minimizers have to be within distance r_t of the origin at time t, see Figure 8, so that the problem becomes almost compact. However, the localization radius can significantly depend on time.

This automatically solves the existence of one-sided minimizers problem since it is easy to take partial limits of long minimizers that have to stay within random localization sausage. What about uniqueness? One can also prove the following result:

Lemma 7.2. *There is time T and radius R such that with positive probability, $r(\omega) < R, r(\theta^T\omega) < R, \tau_R(\omega) < T, \tau_R(\theta^T\omega) < T$ and there is a Poissonian point (x^*, t^*) with $t^* \in [0, T]$ such that every minimizer connecting a point $x \in [-R, R]$ at time 0 to point $y \in [-R, R]$ at time T passes through (x^*, t^*).*

Since the probability the event described by this lemma is positive and the Poisson point field is a time stationary ergodic process, we conclude that with probability 1 infinitely many time translates of this event will happen. Therefore,

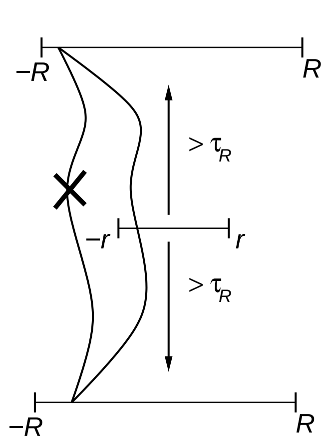

FIGURE 8. Main localization lemma.

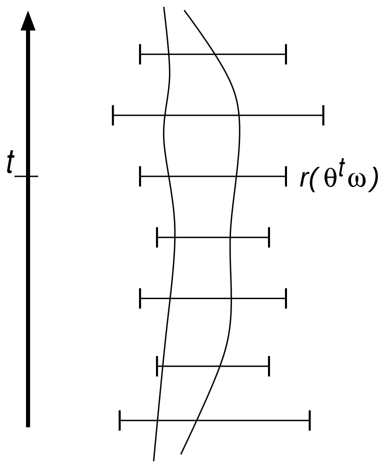

FIGURE 9. Localization sausage (the process of localization radii) — effective compactification of the problem.

any two infinite one-sided minimizers will pass through the same Poissonian point in the past and, moreover, will coincide at all times before that Poissonian point.

The random solution u_ω of the Burgers equation obtained through the velocities of the minimizers constructed above has a specific structure. It turns out that there is a nonrandom number q such that with probability 1,

$$\lim_{x\to\infty} \frac{U_\omega(x)}{x} = q,$$

and, if one makes an additional assumption

$$\int_{\mathbb{R}} (1+|x|)m(dx) < \infty,$$

then, moreover,

$$\lim_{x\to\pm\infty} (u_\omega(x)\cdot \operatorname{sgn} x) = q.$$

The idea is that if we want to consider, an optimal path terminating at a space-time point point (x, T) for large values of x and T and zero initial condition at time 0, then the path naturally decomposes into two parts. Most Poissonian points are scattered over a compact domain, so in a certain time interval $[0, t]$ the path mostly stays in a compact domain around the origin collecting action at approximately linear rate $S < 0$ (it is negative since each Poisson point contributes -1), and then it leaps from the compact domain straight to x roughly with constant speed between t and T, hardly meeting any Poissonian points in this regime and collecting approximately $x^2/(2(T-t))$ action. Finding the minimum of

$$St + \frac{x^2}{2(T-t)}, \quad t \in [0, T],$$

we obtain that the optimal t satisfies

$$\frac{x}{t} = \sqrt{-2S}.$$

Therefore $q = \sqrt{-2S}$. This reasoning can be made precise, see Section 6 of [Bak12].

The 1F1S principle for this system can be formulated in the following way. If one considers an initial condition $v = V'$ satisfying

$$\liminf_{x \to \infty} \frac{V(x)}{x} > -q, \tag{16}$$

then

$$\Psi_\omega^{t_0,t} v \to u_\omega, \quad t_0 \to -\infty. \tag{17}$$

In particular, $u_\omega = U_\omega'$ is a unique stationary global solution satisfying

$$\liminf_{x \to \infty} \frac{U_\omega(x)}{x} > -q.$$

So, if one starts with an initial condition that sends particles from infinity towards zero with speed that is less than q, see condition (16), then this inbound flow is not strong enough to compete with the outbound flow of particles developed due to the noise, and in the long run it is dominated by the latter. If the condition (16) is violated, then the long term properties of solutions are sensitive to the details of the behavior of the initial condition at infinity because the inbound flow of particles may be stronger than the outbound one.

The results of this Section crucially depend on the quasi-compact properties of the driving Poisson process. The external forcing mostly acts on a compact part of the real line, which leads to localization of long minimizers. Of course we do not expect anything like this for the case where the Poissonian forcing is also stationary in space.

8. Main results for space-time stationary Poissonian forcing

In this section we state the main results on the Burgers equation driven by Poissonian noise with Lebesgue intensity measure.

To formulate our results, we need to fix a value $v \in \mathbb{R}$ and work with the set of velocity profiles with average velocity v. It turns out that we can treat each of these sets as an ergodic component. On each of these sets a 1F1S principle holds, i.e., there is a unique stationary solution with average velocity v, at any time t it depends only on the history of the forcing up to time t. Moreover, this global solution uniquely determined by the past of the forcing is a random one-point pullback attractor, and we will describe its basin of attraction. This picture is very similar to that of the random Burgers dynamics on the circle. It is true though that there are initial conditions irregular enough (in particular, the average velocity has to be undefined for these erratic velocity profiles) that do not belong to the domain of attraction of any of the global solutions of the Burgers equation described above. However, if one requires the initial velocity profile to be, say, a typical realization of an integrable stationary process, then, due to the ergodic theorem, its average is well-defined, and so it belongs to one of the domains of attraction.

Let us now introduce some notation and state the main results.

We say that $u(t,x) = u_\omega(t,x)$ is a global solution for the cocycle Ψ if there is a set Ω' with $\mathsf{P}(\Omega') = 1$ such that for all $\omega \in \Omega'$, all s and t with $s < t$, we have $\Psi^{s,t}_\omega u_\omega(s,\cdot) = u_\omega(t,\cdot)$. We can also introduce the global solution as a skew-invariant function: $u_\omega(x)$ is called skew-invariant if there is a set Ω' with $\mathsf{P}(\Omega') = 1$ such that for any $t \in \mathbb{R}$, $\theta^t\Omega' = \Omega'$, and for any $t > 0$ and $\omega \in \Omega'$, $\Psi^t_\omega u_\omega = u_{\theta^t\omega}$. If $u_\omega(x)$ is a skew-invariant function, then $u_\omega(x,t) = u_{\theta^t\omega}(x)$ is a global solution. One can naturally view the potentials of $u_\omega(x)$ and $u_\omega(s,x)$ as a skew-invariant function and global solution for the cocycle $\hat{\Phi}$.

Let us denote $\mathbb{H}(v_-,v_+) = \{W \in \mathbb{H} : \lim_{x\to\pm\infty}(W(x)/x) = v_\pm\}$. The spaces $\hat{\mathbb{H}}(v_-,v_+)$ are defined as classes of potentials in $\mathbb{H}(v_-,v_+)$ coinciding up to an additive constant. It can be shown that spaces $\mathbb{H}(v_-,v_+)$ are invariant under Φ, i.e., if $W \in \mathbb{H}(v_-,v_+)$ for some v_-,v_+, then $\Phi^{s,t}_\omega W \in \mathbb{H}(v_-,v_+)$ for all $s < t$.

Our first result is the description of global solutions.

Theorem 8.1. *For every $v \in \mathbb{R}$ there is a unique (up to zero-measure modifications) skew-invariant function $u_v : \Omega \to \mathbb{H}'$ such that for almost every $\omega \in \Omega$, the potential $U_{v,\omega}$ defined by $U_{v,\omega}(x) = \int^x u_{v,\omega}(y)dy$ belongs to $\hat{\mathbb{H}}(v,v)$.*

The potential $U_{v,\omega}$ is a unique skew-invariant potential in $\hat{\mathbb{H}}(v,v)$. The skew-invariant functions $U_{v,\omega}$ and $u_{v,\omega}$ are measurable w.r.t. $\mathcal{F}|_{\mathbb{R}\times(-\infty,0]}$, i.e., they depend only on the history of the forcing. With probability 1, the realizations of $(u_{v,\omega}(y))_{y\in\mathbb{R}}$ are piecewise linear with negative jumps between linear pieces. The spatial random process $(u_{v,\omega}(y))_{y\in\mathbb{R}}$ is stationary and mixing.

Remark 8.2. *This theorem can be interpreted as a 1F1S Principle: for any velocity value v, the solution at time 0 with mean velocity v is uniquely determined by the history of the forcing: $u_{v,\omega} \overset{\text{a.s.}}{=} \chi_v(\omega|_{\mathbb{R}\times(-\infty,0]})$ for some deterministic functional χ_v of the point configurations on half-plane $\mathbb{R}\times(-\infty,0]$ of the past (we actually construct χ_v in the proof). Since the forcing is stationary in time, we obtain that $u_{v,\theta^t\omega}$ is a stationary process in t, and the distribution of $u_{v,\omega}$ is an invariant distribution for the corresponding Markov semi-group, concentrated on $\mathbb{H}'(v,v)$.*

The next result shows that each of the global solutions constructed in Theorem 8.1 plays the role of a one-point pullback attractor. To describe the domains of attraction we will make assumptions on initial potentials $W \in \mathbb{H}$. Namely, we will assume that there is $v \in \mathbb{R}$ such that W and v satisfy one of the following sets of conditions:

$$\begin{aligned} &v=0,\\ &\liminf_{x\to+\infty}\frac{W(x)}{x}\geq 0,\\ &\limsup_{x\to-\infty}\frac{W(x)}{x}\leq 0, \end{aligned} \tag{18}$$

or

$$\begin{aligned} &v>0,\\ &\lim_{x\to-\infty}\frac{W(x)}{x}=v,\\ &\liminf_{x\to+\infty}\frac{W(x)}{x}>-v, \end{aligned} \tag{19}$$

or

$$\begin{aligned} &v<0,\\ &\lim_{x\to+\infty}\frac{W(x)}{x}=v,\\ &\limsup_{x\to-\infty}\frac{W(x)}{x}<-v. \end{aligned} \tag{20}$$

Condition (18) means that there is no macroscopic flux of particles from infinity toward the origin for the initial velocity profile W'. In particular, any $W\in\mathbb{H}(0,0)$ or any $W\in\mathbb{H}(v_-,v_+)$ with $v_-\leq 0$ and $v_+\geq 0$ satisfies (18). It is natural to call the arising phenomenon a rarefaction fan. We will see that in this case the long-term behavior is described by the global solution u_0 with mean velocity $v=0$.

Condition (19) means that the initial velocity profile W' creates the influx of particles from $-\infty$ with effective velocity $v\geq 0$, and the influence of the particles at $+\infty$ is not as strong. In particular, any $W\in\mathbb{H}(v,v_+)$ with $v\geq 0$ and $v_+>-v$ (e.g., $v_+=v$) satisfies (19). We will see that in this case the long-term behavior is described by the global solution u_v.

Condition (20) describes a situation symmetric to (19), where in the long run the system is dominated by the flux of particles from $+\infty$.

The following precise statement supplements Theorem 8.1 and describes the basins of attraction of the global solutions u_v in terms of conditions (18)–(20).

Theorem 8.3. *There is a set $\Omega''\in\mathcal{F}$ with $\mathsf{P}(\Omega'')=1$ such that if $\omega\in\Omega''$, $W\in\mathbb{H}$, and one of conditions* (18),(19),(20) *holds: then $w=W'$ belongs to the domain of pullback attraction of u_v in the following sense: for any $t\in\mathbb{R}$ and any $R>0$ there is $s_0=s_0(\omega)<t$ such that for all $s<s_0$*

$$\Psi^{s,t}_\omega w(x)=u_{v,\omega}(x,t),\quad x\in[-R,R].$$

In particular,

$$\mathsf{P}\left\{\Psi^{s,t}_\omega w\big|_{[-R,R]}=u_{v,\omega}(\cdot,t)\big|_{[-R,R]}\right\}\to 1,\quad s\to-\infty.$$

Remark 8.4. The last statement of the theorem implies that for every $v\in\mathbb{R}$, the invariant measure on $\mathbb{H}'(v,v)$ described in Remark 8.2 is unique and for any initial condition $w=W'\in\mathbb{H}'$ satisfying one of conditions (18),(19), and (20), the distribution of the random velocity profile at time t converges to the unique stationary distribution on $\mathbb{H}'(v,v)$ as $t \to \infty$. However, our approach does not produce any convergence rate estimates.

Remark 8.5. Using space-time Galilean transformations, it is easy to obtain a version of Theorem 8.3 for attraction in a coordinate frame moving with constant velocity, but we omit it for brevity.

The proofs of Theorems 8.1 and 8.3 are given in Sections 12 and 13, but most of the preparatory work is carried out in Sections 9, 10, and 11.

The long-term behavior of the cocycles Φ and Ψ defined through the optimization problem (13) depends on the asymptotic behavior of the action minimizers over long time intervals. The natural notion that plays a crucial role in this paper is the notion of backward one-sided infinite minimizers or geodesics. A curve $\gamma : (-\infty, t] \to \mathbb{R}$ with $\gamma(t) = x$ is called a backward minimizer if its restriction onto any time interval $[s,t]$ provides the minimum to the action $A^{s,t}_\omega(W,\cdot)$ defined in (12) among paths connecting $\gamma(s)$ to x.

It can be shown (see Lemma 11.4) that any backward minimizer γ has an asymptotic slope $v = \lim_{t\to-\infty}(\gamma(t)/t)$. On the other hand, for every space-time point (x,t) and every $v \in \mathbb{R}$ there is a backward minimizer with asymptotic slope v and endpoint (x,t). The following theorem describes the most important properties of backward minimizers associated with the Poisson point field.

Theorem 8.6. *For every $v \in \mathbb{R}$ there is a set of full measure Ω' such that for all $\omega \in \Omega'$ and any $(x,t) \in \omega$ there is a unique backward minimizer with asymptotic slope v. For any $(x_1,t_1),(x_2,t_2) \in \omega$ there is a time $s \le \min\{t_1,t_2\}$ such that both minimizers coincide before s, i.e., $\gamma_1(r) = \gamma_2(r)$ for $r \le s$.*

The proof of this core statement of this paper is spread over Sections 9 through 11. In Section 9 we apply the sub-additive ergodic theorem to derive the linear growth of action. In Section 10 we prove quantitative estimates on deviations from the linear growth. We use these results in Section 11 to analyze deviations of optimal paths from straight lines and deduce the existence of infinite one-sided optimal paths and their properties.

9. Optimal action asymptotics and the shape function

In this section begin the study the asymptotic behavior of the optimal action between space-time points (x,s) and (y,t) denoted by

$$
\begin{aligned}
A^{s,t}(x,y) = A^{s,t}_\omega(x,y) &= \min_{\gamma:\gamma(s)=x,\gamma(t)=y} (A^{s,t}_\omega(0,\gamma)) \\
&= \min_{\gamma:\gamma(s)=x,\gamma(t)=y} \left(S^{s,t}(\gamma) - \omega^{s,t}(\gamma)\right)
\end{aligned}
\tag{21}
$$

(the minimum is taken over all absolutely continuous paths γ or, equivalently, over all piecewise linear paths with vertices at configuration points). Although to construct stationary solutions for the Burgers equation, we will need the asymptotic behavior as $s \to -\infty$, it is more convenient and equally useful (due to the obvious symmetry in the variational problem) to formulate most results for the limiting behavior as $t \to \infty$, and so we will here and in the next two sections.

Our distant goal is to show that long optimal paths do not deviate a lot from straight lines. In this section we make some first steps, and our main goal here is to use sub-additive ergodic theorem (see, e.g., [Lig85]) to prove that with probability one, $A^{0,t}(0,vt)=\alpha(v)t+o(t)$ as $t\to\infty$, where the nonrandom shape function $\alpha(v)$ satisfies $\alpha(v)=\alpha_0-v^2/2$.

We begin with some simple observations on Galilean shear transformations of the point field.

Lemma 9.1. *Let $a,v\in\mathbb{R}$ and let L be a transformation of space-time defined by $L(x,s)=(x+a+vs,s)$.*

(1) *Suppose that γ is a path defined on a time interval $[t_0,t_1]$ and let $\bar\gamma$ be defined by $(\bar\gamma(s),s)=L(\gamma(s),s)$. Then*
$$S^{t_0,t_1}(\bar\gamma)=S^{t_0,t_1}(\gamma)+(\gamma(t_1)-\gamma(t_0))v+\frac{(t_1-t_0)v^2}{2}.$$
(2) *Let $L(\omega)$ be the point configuration obtained from $\omega\in\Omega$ by applying L point-wise. Then $L(\omega)$ is also a Poisson process with Lebesgue intensity measure.*
(3) *Let $\omega\in\Omega$. For any time interval $[t_0,t_1]$ and any points $x_0,x_1,\bar x_0,\bar x_1$ satisfying $L(x_0,t_0)=(\bar x_0,t_0)$ and $L(x_1,t_1)=(\bar x_1,t_1)$,*
$$A^{t_0,t_1}_{L(\omega)}(\bar x_0,\bar x_1)=A^{t_0,t_1}_{\omega}(x_0,x_1)+(x_1-x_0)v+\frac{(t_1-t_0)v^2}{2},$$
and L maps minimizers realizing $A^{t_0,t_1}_{\omega}(x_0,x_1)$ onto minimizers realizing $A^{t_0,t_1}_{L(\omega)}(\bar x_0,\bar x_1)$.
(4) *For any points $x_0,x_1,\bar x_0,\bar x_1$ and any time interval $[t_0,t_1]$,*
$$A^{t_0,t_1}(\bar x_0,\bar x_1)\stackrel{distr}{=}A^{t_0,t_1}(x_0,x_1)+(x_1-x_0)v+\frac{(t_1-t_0)v^2}{2},$$
where
$$v=\frac{(\bar x_1-x_1)-(\bar x_0-x_0)}{t_1-t_0}.$$

Proof: The first part of the Lemma is a simple computation:
$$\begin{aligned}S^{t_0,t_1}(\bar\gamma)&=\frac12\int_{t_0}^{t_1}(\dot\gamma(s)+v)^2ds\\&=\frac12\int_{t_0}^{t_1}\dot\gamma^2(s)ds+\int_{t_0}^{t_1}\dot\gamma(s)vds+\frac12\int_{t_0}^{t_1}v^2ds.\end{aligned}$$
The second part holds since L preserves the Lebesgue measure. The third part follows from the first one since the images of paths transformed by L are also paths passing through the L-images of configuration points. The last part is a consequence of the previous two parts, since the appropriate Galilean transformation sending (x_0,t_0) to $(\bar x_0,t_0)$ and (x_1,t_1) to $(\bar x_1,t_1)$ preserves the Lebesgue measure and the distribution of the Poisson process. □

The next useful property is the sub-additivity of action along any direction: for any velocity $v\in\mathbb{R}$, and any $t,s\geq 0$, we have
$$A^{0,t+s}(0,v(t+s))\leq A^{0,t}(0,vt)+A^{t,t+s}(vt,v(t+s)).$$
This means that we can apply Kingman's sub-additive ergodic theorem to the function $t\mapsto A^{0,t}(0,vt)$ if we can show that $-\mathsf{E}A^{0,t}(0,vt)$ grows at most linearly in t. We claim this linear bound in the following proposition:

Lemma 9.2. *Let $v \in \mathbb{R}$. There exists a constant $C = C(v) > 0$ such that for all $t \geq 0$*

$$\mathsf{E}|A^{0,t}(0, vt)| \leq Ct.$$

PROOF: Lemma 9.1 implies that it is enough to prove this for $v = 0$. So in this proof we work with $A^t = A^t(0,0)$.

Let $\gamma : [0,t] \to \mathbb{R}$ be a path realizing A^t. We have $\gamma(0) = \gamma(t) = 0$. Let us split up $\mathbb{R}^2$ into unit blocks $B_{i,j} = [i, i+1) \times [j, j+1)$, for $i, j \in \mathbb{Z}$. We define $\mathcal{A}$ as the union of all indices (i,j) such that γ passes through $B_{i,j}$. The set $\mathcal{A}$ is a lattice animal, i.e., it is a connected set that contains the origin $(0,0) \in \mathbb{Z}^2$ (see, e.g.,[GK94]). Let us introduce the event $E_{n,t} = \{\#\mathcal{A} = n\}$.

Lemma 9.3. *There are constants $C_1, C_2, R, t_0 > 0$ such that if $t \geq t_0$ and $n \geq Rt$, then*

$$\mathsf{P}(E_{n,t}) \leq C_1 \exp(-C_2 n^2/t).$$

PROOF: We define $X_{i,j} = \omega(B_{i,j})$, the number of Poisson points in $B_{i,j}$. Define the weight of the animal $\mathcal{A}$ as

$$N_{\mathcal{A}} = \sum_{\nu \in \mathcal{A}} X_\nu.$$

Clearly, the number of Poisson points picked up by γ between 0 and t, is upper bounded by $N_{\mathcal{A}}$. Define $k_j = \#\{i \in \mathbb{Z} \ : \ (i,j) \in \mathcal{A}\}$, the number of blocks hit on the j^{th} row. These blocks will form a connected row of length k_j, and the kinetic action accumulated between j and $(j+1) \wedge t$ can therefore be bounded by

$$\frac{1}{2} \int_j^{(j+1)\wedge t} \dot{\gamma}^2(s)\, ds \geq \frac{1}{2} (k_j - 2)_+^2 .$$

Here, $a_+ = \max(0, a)$. This leads to the following bound on the action:

$$A^t \geq \frac{1}{2} \sum_{0 \leq j < t} (k_j - 2)_+^2 - N_{\mathcal{A}}.$$

On $E_{n,t}$ we have $\sum_{0 \leq j < t} k_j = n$. Since $a \mapsto (a-2)_+^2$ is convex, we can use Jensen's inequality to see that

$$\frac{1}{2} \sum_{0 \leq j < t} (k_j - 2)_+^2 \geq \frac{1}{2} \lceil t \rceil \left(\frac{n}{\lceil t \rceil} - 2 \right)_+^2 .$$

Therefore,

$$A^t \geq \frac{1}{2} \lceil t \rceil \left(\frac{n}{\lceil t \rceil} - 2 \right)_+^2 - N_{\mathcal{A}}. \tag{22}$$

We also know that $A^t \leq 0$ since we can use the identical zero path on $[0,t]$. Hence, on $E_{n,t}$ we have

$$N_{\mathcal{A}} \geq \frac{1}{2} \lceil t \rceil \left(\frac{n}{\lceil t \rceil} - 2 \right)_+^2 .$$

Furthermore, if N_n is the weight of the greedy animal of size n (i.e., the animal of size n with greatest weight), then $N_n \geq N_{\mathcal{A}}$, and

$$E_{n,t} \subset \left\{ N_n \geq \frac{1}{2} \lceil t \rceil \left(\frac{n}{\lceil t \rceil} - 2 \right)_+^2 \right\}.$$

Let us recall that the reasoning in [CGGK93] after equation (2.12) implies that, due to standard large deviation estimates and the exponential growth of the number of lattice animals as a function of size n, there are constants $K_1, K_2, y_0 > 0$, such that if

$$y \geq y_0, \tag{23}$$

then

$$\mathsf{P}\{N_n \geq yn\} \leq K_1 \exp(-K_2 ny). \tag{24}$$

We now need to make sure that (23) holds for $y = \frac{1}{2n}\lceil t\rceil \left(\frac{n}{\lceil t\rceil} - 2\right)_+^2$. If we require $n \geq \max(4, 8y_0)\lceil t\rceil$, then

$$\frac{1}{2n}\lceil t\rceil \left(\frac{n}{\lceil t\rceil} - 2\right)_+^2 = \frac{1}{2n}\lceil t\rceil \left(\frac{n - 2\lceil t\rceil}{\lceil t\rceil}\right)_+^2 \geq \frac{1}{2n}\lceil t\rceil \left(\frac{n - \frac{n}{2}}{\lceil t\rceil}\right)^2 \geq \frac{1}{8}\frac{n}{\lceil t\rceil} \geq y_0,$$

and the lemma follows from (24). □

Remark 9.4. We will choose the constant R to be an integer, making it larger if needed.

From (22) we already know that on $E_{n,t}$ we have $0 \geq A^t \geq -N_n$. We wish to use this to estimate $\mathsf{E}|A^t|$, but we need an extension of (24).

Lemma 9.5. *For any $k \geq 1$, there is $c_k > 0$ such that for all $n \geq 1$,*

$$\mathsf{E}N_n^k \leq c_k n^k.$$

Proof: Clearly,

$$\mathsf{E}N_n^k = \sum_{i=0}^{\lfloor y_0 n\rfloor} i^k \mathsf{P}\{N_n = i\} + \sum_{i=\lfloor y_0 n\rfloor+1}^{\infty} i^k \mathsf{P}\{N_n = i\}.$$

We can bound the first term simply by

$$\sum_{i=0}^{\lfloor y_0 n\rfloor} i^k \mathsf{P}\{N_n = i\} \leq (y_0 n)^k.$$

For the second term we can use (24):

$$\sum_{i=\lfloor y_0 n\rfloor+1}^{\infty} i^k \mathsf{P}\{N_n = i\} \leq \sum_{i=\lfloor y_0 n\rfloor+1}^{\infty} K_1 i^k \exp(-K_2 i).$$

The right-hand side is bounded in n and the proof is complete. □

Lemma 9.2 now follows from Lemmas 9.3 and 9.5:

$$\begin{aligned}
\mathsf{E}|A^t| &= \sum_{n\leq Rt} \mathsf{E}|A^t|\mathbf{1}_{E_{n,t}} + \sum_{n>Rt} \mathsf{E}|A^t|\mathbf{1}_{E_{n,t}} \\
&\leq \mathsf{E}N_{[Rt]} + \sum_{n>Rt} \mathsf{E}N_n \mathbf{1}_{E_{n,t}} \\
&\leq Rc_1 t + \sum_{n>Rt} \sqrt{\mathsf{E}N_n^2}\sqrt{\mathsf{P}(E_{n,t})} \\
&\leq Rc_1 t + \sqrt{c_2}\sum_{n>Rt} \sqrt{C_1} n \exp(-C_2 n^2/(2t)) \\
&\leq Ct,
\end{aligned}$$

for C big enough. □

In fact, we can use the last calculation to obtain the following generalization of Lemma 9.2 for higher moments of A^t:

Lemma 9.6. *Let $k \in \mathbb{N}$. Then there are constants $C(k), t_0(k) > 0$ such that*

$$\mathsf{E}(|A^t|^k) \le C(k)t^k, \quad t \ge t_0(k).$$

Now a standard application of the sub-additive ergodic theorem shows that there exists a *shape function* $\alpha(v)$ such that

$$\frac{A^{0,t}(0,vt)}{t} \to \alpha(v), \quad \text{a.s. and in } L^1, \quad t \to \infty. \tag{25}$$

Furthermore, $\alpha(0) < 0$, since $A^t \le 0$ and $\alpha(0) \le \mathsf{E}(A^t) < 0$. It turns out that the shape function $\alpha(v)$ is quadratic in v:

Lemma 9.7. *The shape function satisfies*

$$\alpha(v) = \alpha(0) + \frac{v^2}{2}, \quad v \in \mathbb{R}.$$

Proof: The Galilean shear map $(x,t) \mapsto (x+vt,t)$ transforms the paths connecting $(0,0)$ to $(0,t)$ into paths connecting $(0,0)$ to (vt,t). Lemma 9.1 implies that under this map the optimal action over these paths is altered by a deterministic correction $v^2t/2$. Since α is a constant almost surely we obtain the statement of the lemma. □

Let us recall that our goal is to prove that optimal paths are almost straight. We can already derive that it is unlikely that a path following some direction v_1 for a large time t_1 and then some other direction v_2 for a large time t_2 can be optimal. In fact, we know from (25) that $A^{0,t}(0,vt) \sim \alpha(v)t$ for large t, and the strong convexity of function α implies that for $\bar{v} = (v_1t_1 + v_2t_2)/(t_1+t_2)$,

$$\alpha(\bar{v})(t_1+t_2) < \alpha(v_1)t_1 + \alpha(v_2)t_2, \tag{26}$$

and our path that switches the slopes from v_1 to v_2 accumulates much more action the the optimal path. Notice that this estimate would be almost useless if the graph of shape function α had flat pieces — in that case inequality (26) would not be strict.

However, arguing at the level of convexity and inequality (26) is not enough for our purposes since on top of the deterministic linear growth there are noisy fluctuations. So, we need quantitative estimates on deviations of $A^{0,t}(0,vt)$ from $\alpha(v)t$. This is the material of the next section.

10. Concentration inequality for optimal action

The goal of this section is to prove a concentration inequality for $A^t(vt) = A^t(0,vt) = A^t_\omega(0,vt) = A^{0,t}_\omega(0,vt)$:

Theorem 10.1. *There are positive constants c_0, c_1, c_2, c_3, c_4 such that for any $v \in \mathbb{R}$, all $t > c_0$, and all $u \in (c_3 t^{1/2}\ln^2 t, c_4 t^{3/2}\ln t]$,*

$$\mathsf{P}\{|A^t(0,vt) - \alpha(v)t| > u\} \le c_1 \exp\left\{-c_2 \frac{u}{t^{1/2}\ln t}\right\}.$$

We give a complete proof since the tools we are invoking are useful in similar contexts. The central role of the proof is played by Kesten's martingale concentration inequality that estimates the deviations of a martingale from its mean, and we use it in a very similar way as Kesten did in his original paper [Kes93] on first passage percolation. To use Kesten's lemma we need to introduce an appropriate martingale and to make two truncations. One of them is made to make sure that the optimal paths realizing the minimal action does not deviate too far. The other one is made to ensure that the concentration of Poisson points is not abnormally high. Then we can check the conditions of Kesten's lemma and take care of the additional errors introduced by truncations. Notice, however, that the statement of Theorem 10.1 involves deviations of $A^t(0, vt)$ not from its mean but from an approximating value $\alpha(v)t$. To tackle this additional discrepancy we use a method of Howard and Newman based on their interpolation lemma on approximately additive functions. This lemma estimates how far a function is from a linear one if it is known how it behaves under doubling of its argument.

Due to the invariance under shear transformations (Lemmas 9.1 and 9.7), it is sufficient to prove Theorem 10.1 for $v = 0$. We will first derive a similar inequality with $\alpha(0)t$ replaced by $\mathsf{E}A^t$, and then we will have to estimate the corresponding approximation error. We recall that $A^t \le 0$.

Lemma 10.2. *There are positive constants b_0, b_1, b_2, b_3 such that for all $t > b_0$ and all $u \in (0, b_3 t^{3/2} \ln t]$,*

$$\mathsf{P}\{|A^t - \mathsf{E}A^t| > u\} \le b_1 \exp\left\{-b_2 \frac{u}{t^{1/2}\ln t}\right\}.$$

The method of proof is derived from that for the generalized Hammersley's process in [CP11], but we have to take into account that the optimal paths are allowed to travel arbitrarily far within any bounded time interval in search for areas rich with configuration points. However, the situation where they decline too far from the kinetically most efficient path is not typical. In the remaining part of this section we will often use the following lemma showing that with high probability the minimizer γ connecting $(0,0)$ to $(0,t)$ stays within distance Rt from the origin, where R was introduced in Lemma 9.3.

Lemma 10.3. *There is a constant C_3 such that if $t \ge t_0$ and $u \ge Rt$ then*

$$\mathsf{P}\left\{\max_{s\in[0,t]} |\gamma(s)| > u\right\} \le C_3 \exp(-C_2 u^2/t),$$

where constants C_2, R, t_0 were introduced in Lemma 9.3.

PROOF: If $\max\{|\gamma(s)| : s \in [0,t]\} > u$, then the size of the lattice animal $\mathcal{A}$ traced by γ is at least u. Lemma 9.3 implies

$$\mathsf{P}\left\{\max_{s\in[0,t]} |\gamma(s)| > u\right\} \le \sum_{n \ge u} C_1 \exp(-C_2 n^2/t) \le C_3 \exp(-C_2 u^2/t)$$

for a constant C_3, since the first term of the series is bounded by $C_1 \exp(-C_2 u^2/t)$ and the ratio of two consecutive terms is bounded by $\exp(-C_2 R)$. □

Having Lemma 10.3 in mind, we define $\tilde{A}^t$ to be the optimal action over all paths connecting $(0,0)$ to $(0,t)$ and staying within $[-Rt, Rt]$.

Lemma 10.4. *Let constants t_0, R, C_2, C_3 be defined in Lemmas 9.3 and 10.3. For any $t > t_0$,*

$$\mathsf{P}\{A^t \neq \tilde{A}^t\} \leq C_3 \exp(-R^2 C_2 t).$$

PROOF: It is sufficient to notice that

$$\mathsf{P}\{A^t \neq \tilde{A}^t\} \leq \mathsf{P}\left\{\max_{s\in[0,t]} |\gamma(s)| > Rt\right\}$$

and apply Lemma 10.3. □

Lemma 10.5. *There is a constant D_1 such that for all $t > t_0$,*

$$0 \leq \mathsf{E}\tilde{A}^t - \mathsf{E}A^t \leq -\mathsf{E}A^t \mathbf{1}_{\left\{\sup_{s\in[0,t]} |\gamma(s)|>Rt\right\}} \leq D_1.$$

PROOF: The first two inequalities are obvious, since we have that $0 \geq \tilde{A}^t \geq A^t$. For the last one, we have

$$\begin{aligned} -\mathsf{E}A^t \mathbf{1}_{\left\{\sup_{s\in[0,t]} |\gamma(s)|>Rt\right\}} &\leq \sum_{n>Rt} \mathsf{E}(N_n \mathbf{1}_{E_{n,t}}) \\ &\leq \sum_{n>Rt} \sqrt{\mathsf{E}N_n^2}\sqrt{\mathsf{P}(E_{n,t})} \\ &\leq \sum_{n>Rt} \sqrt{c_2}n\sqrt{C_1}\exp(-C_2 n^2/(2t)), \end{aligned}$$

where we used Lemmas 9.3 and 9.5. The statement follows since the last series is uniformly convergent for $t > t_0$. □

To obtain a concentration inequality for $\tilde{A}$, we will apply the following lemma by Kesten [Kes93]:

Lemma 10.6. *Let $(\mathcal{F}_k)_{0\leq k\leq N}$ be a filtration and let $(U_k)_{0\leq k\leq N}$ be a family of nonnegative random variables measurable with respect to $\mathcal{F}_N$. Let $(M_k)_{0\leq k\leq N}$ be a martingale with respect to $(\mathcal{F}_k)_{0\leq k\leq N}$. Assume that for some constant $c > 0$ the increments $\Delta_k = M_k - M_{k-1}$ satisfy*

$$|\Delta_k| < c, \quad k = 1, \ldots, N,$$

and

$$\mathsf{E}(\Delta_k^2 | \mathcal{F}_{k-1}) \leq \mathsf{E}(U_k | \mathcal{F}_{k-1}).$$

Assume further that for some positive constants c_1, c_2 and some $x_0 \geq e^2 c^2$ we have

$$\mathsf{P}\left\{\sum_{k=1}^N U_k > x\right\} \leq c_1 \exp(-c_2 x), \quad x \geq x_0.$$

Then

$$\mathsf{P}\{M_N - M_0 \geq x\} \leq c_3 \left(1 + c_1 + \frac{c_1}{c_2 x_0}\right) \exp\left(-c_4 \frac{x}{x_0^{1/2} + c_2^{-1/3} x^{1/3}}\right), \quad x > 0,$$

where c_3, c_4 are universal positive constants that do not depend on N, c, c_1, c_2, x_0, nor on the distribution of $(M_k)_{0\leq k\leq N}$ and $(U_k)_{0\leq k\leq N}$. In particular,

$$\mathsf{P}\{M_N - M_0 \geq x\} \leq c_3 \left(1 + c_1 + \frac{c_1}{c_2 x_0}\right) \exp\left(-c_4 \frac{x}{2\sqrt{x_0}}\right), \quad x \leq c_2 x_0^{3/2}.$$

To use this lemma in our framework, we must introduce an appropriate martingale. For a given t we consider the rectangle $Q(t) = [-Rt, Rt] \times [0, t]$ and partition it into $N = 2Rt \cdot t = 2Rt^2$ disjoint unit squares: $Q(t) = \bigcup_{k=1}^{N} B_k$. The order of enumeration is not important. Here we assume that $t \in \mathbb{N}$, but it is easy to adapt the reasoning to the case of non-integer t.

We introduce a filtration $(\mathcal{F}_k)_{0\le k\le N}$ in the following way. We set $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and

$$\mathcal{F}_k = \sigma\left(\omega\big|_{\bigcup_{j=1}^{k} B_j}\right), \quad k = 1, \dots, N.$$

We introduce a martingale $(M_k, \mathcal{F}_k)_{0\le k\le N}$ by

$$M_k = \mathsf{E}(\tilde{A}^t|\mathcal{F}_k), \quad 0 \le k \le N.$$

We denote by P_k the distribution of $\omega\big|_{B_k}$ on the sample space Ω_k of finite point configurations in B_k. For $\omega, \sigma \in \prod_{k=1}^{N} \Omega_k$ we write

$$[\omega, \sigma]_k = (\omega_1, \dots, \omega_k, \sigma_{k+1}, \dots, \sigma_N) \in \prod_{k=1}^{N} \Omega_k.$$

Then

$$\begin{aligned}\Delta_k(\omega_1, \dots, \omega_k) &:= M_k - M_{k-1}\\ &= \int \tilde{A}^t_{[\omega,\sigma]_k} \prod_{j=k+1}^{N} dP_j(\sigma_j) - \int \tilde{A}^t_{[\omega,\sigma]_{k-1}} \prod_{j=k}^{N} dP_j(\sigma_j)\\ &= \int \left(\tilde{A}^t_{[\omega,\sigma]_k} - \tilde{A}^t_{[\omega,\sigma]_{k-1}}\right) \prod_{j=k}^{N} dP_j(\sigma_j).\end{aligned}$$

Lemma 10.7. *Let I_k denote the indicator that the minimizer connecting $(0,0)$ to $(0,t)$ and staying in $[-Rt, Rt]$ passes through a Poissonian point in B_k. Then*

$$|\tilde{A}^t_{[\omega,\sigma]_k} - \tilde{A}^t_{[\omega,\sigma]_{k-1}}| \le \max\{I_k([\omega,\sigma]_k), I_k([\omega,\sigma]_{k-1})\} \max\{\omega(B_k), \sigma(B_k)\}.$$

PROOF: Suppose we delete the points of ω in B_k. When we then consider the minimizer for $[\omega,\sigma]_k$, we decrease the number of Poissonian points contributing to the action by at most $\omega(B_k)$, and only decrease the kinetic action. Comparing the resulting path with the minimizer for $[\omega,\sigma]_{k-1}$, we obtain

$$\tilde{A}^t_{[\omega,\sigma]_{k-1}} \le \tilde{A}^t_{[\omega,\sigma]_k} + \omega(B_k).$$

Similarly, we get

$$\tilde{A}^t_{[\omega,\sigma]_k} \le \tilde{A}^t_{[\omega,\sigma]_{k-1}} + \sigma(B_k).$$

This shows that

$$|\tilde{A}^t_{[\omega,\sigma]_k} - \tilde{A}^t_{[\omega,\sigma]_{k-1}}| \le \max\{\omega(B_k), \sigma(B_k)\}.$$

Now remark that if none of the two minimizers (for $[\omega,\sigma]_k$ and $[\omega,\sigma]_{k-1}$) passes through a Poissonian point inside B_k, then $\tilde{A}^t_{[\omega,\sigma]_k}$ and $\tilde{A}^t_{[\omega,\sigma]_{k-1}}$ coincide. This completes the proof. □

The next step is to define a truncated Poissonian configuration $\bar{\omega}$ by erasing all Poissonian points of ω in each block B_j with $\omega(B_j) > b \ln t$, where $b > 0$ will be chosen later. The restrictions of $\bar\omega$ to B_j, $j = 1,\dots,N$ are jointly independent. Lemma 10.7 applies to truncated configurations as well and we obtain

$$|\tilde A^t_{[\bar\omega,\bar\sigma]_k} - \tilde A^t_{[\bar\omega,\bar\sigma]_{k-1}}| \le b\ln t \max\{I_k([\bar\omega,\bar\sigma]_k), I_k([\bar\omega,\bar\sigma]_{k-1})\},$$

where $\bar\sigma$ is obtained from σ in the same way as $\bar\omega$ from ω. Therefore,

$$|\Delta_k(\bar\omega_1,\dots,\bar\omega_k)| \le b\ln t \int \max\{I_k([\bar\omega,\bar\sigma]_k), I_k([\bar\omega,\bar\sigma]_{k-1})\} \prod_{j=k}^N dP_j(\sigma_j) \le b\ln t.$$

We must now estimate the increments of the martingale predictable characteristic. This estimate is a straightforward analogue of Lemma 4.3 of [CP11].

Lemma 10.8. *Let $U_k = 2(b\ln t)^2 I_k$. Then, with probability 1, $|U_k(\bar\omega)| \le 2(b\ln t)^2$ and*

$$\mathsf{E}(\Delta_k^2(\bar\omega_1,\dots,\bar\omega_k)|\mathcal{F}_{k-1}) \le \mathsf{E}(U_k(\bar\omega)|\mathcal{F}_{k-1}).$$

PROOF:

$$\begin{aligned}
\mathsf{E}(\Delta_k^2(\bar\omega_1,\dots,\bar\omega_k)|\mathcal{F}_{k-1}) &= \int\left(\int\left(\tilde A^t_{[\bar\omega,\bar\sigma]_k} - \tilde A^t_{[\bar\omega,\bar\sigma]_{k-1}}\right)\prod_{j=k}^N dP_j(\sigma_j)\right)^2 dP_k(\omega_k)\\
&\le \int\left(\int \max\{I_k([\bar\omega,\bar\sigma]_k), I_k([\bar\omega,\bar\sigma]_{k-1})\}\cdot b\ln t\prod_{j=k}^N dP_j(\sigma_j)\right)^2 dP_k(\omega_k)\\
&\le \int\int \max\{I_k([\bar\omega,\bar\sigma]_k), I_k([\bar\omega,\bar\sigma]_{k-1})\}\cdot (b\ln t)^2\prod_{j=k}^N dP_j(\sigma_j)dP_k(\omega_k)\\
&\le \int\int (I_k([\bar\omega,\bar\sigma]_k) + I_k([\bar\omega,\bar\sigma]_{k-1}))\cdot (b\ln t)^2\prod_{j=k}^N dP_j(\sigma_j)dP_k(\omega_k)\\
&= \mathsf{E}(U_k(\bar\omega)|\mathcal{F}_{k-1}).
\end{aligned}$$

□

We have

$$\sum_{k=1}^N U_k(\bar\omega) = 2(b\ln t)^2\sum_{k=1}^N I_k(\bar\omega). \tag{27}$$

Since

$$\sum_{k=1}^N I_k(\bar\omega) \le \#\mathcal{A}(\bar\omega),$$

we can write

$$\mathsf{P}\left\{\sum_{k=1}^N U_k(\bar\omega) > x\right\} \le \mathsf{P}\left\{\#\mathcal{A}(\bar\omega) > \frac{x}{2(b\ln t)^2}\right\} \le \sum_{n > x/(2(b\ln t)^2)} \mathsf{P}\{\bar\omega \in E_{n,t}\}.$$

It is easy to see that Lemma 9.3 applies to $\bar\omega$ as well as to ω, since its proof depends only on the tail estimate for the number of configuration points in each block. We can conclude that

$$\mathsf{P}\{\bar\omega \in E_{n,t}\} \le C_1\exp(-C_2 n^2/t), \quad n \ge Rt,\ t \ge t_0,$$

where C_1, C_2, R, t_0 were introduced in Lemma 9.3.

Combining the last two inequalities and choosing $x_0 = 2Rt(b\ln t)^2$, we can write for $x > x_0$

$$\begin{aligned}\mathsf{P}\left\{\sum_{k=1}^{N} U_k(\bar\omega) > x\right\} &\le C_1 \sum_{n > x/(2(b\ln t)^2)} \exp(-C_2 n^2/t)\\ &\le C_4 \exp(-C_2 x^2/(4t(b\ln t)^4))\\ &\le C_4 \exp(-C_2 x x_0/(4t(b\ln t)^4))\\ &\le C_4 \exp(-C_5 (b\ln t)^{-2} x).\end{aligned}$$

The above estimates on $\Delta_k(\bar\omega)$ and $U_k(\bar\omega)$ allow to apply Kesten's lemma with $c = 2b\ln t$, $c_1 = C_4$, $c_2 = C_5(b\ln t)^{-2}$, $x_0 = 2Rt(b\ln t)^2$ and obtain the following statement:

Lemma 10.9. *There are constants $C_6, C_7, C_8, t_0 > 0$ such that for $t > t_0$ and $x \le C_8 b t^{3/2} \ln t$,*

$$\mathsf{P}\{|\tilde A^t(\bar\omega) - \mathsf{E}\tilde A_t(\bar\omega)| > x\} \le C_6 \exp\left(-C_7 \frac{x}{bt^{1/2}\ln t}\right).$$

Lemma 10.10. *With probability 1,*

$$\tilde A^t(\omega) \le \tilde A^t(\bar\omega).$$

Also, we can choose b and t_0 such that for all $t > t_0$ and $x > 0$,

$$\mathsf{P}\{\tilde A^t(\bar\omega) - \tilde A^t(\omega) > x\} \le 2e^{-x}.$$

PROOF: The first statement of the lemma is obvious, and we have

$$0 \le \tilde A^t(\bar\omega) - \tilde A^t(\omega) \le \sum_{k=1}^{N} \omega(B_k) \mathbf{1}_{\{\omega(B_k) > b\ln t\}}.$$

By Markov's inequality and mutual independence of $\omega|_{B_k}$, $k = 1, \ldots, N$,

$$\mathsf{P}\left\{\sum_{k=1}^{N} \omega(B_k)\mathbf{1}_{\{\omega(B_k) > b\ln t\}} > x\right\} \le e^{-x}\left[\mathsf{E}e^{\omega(B_1)\mathbf{1}_{\{\omega(B_1)>b\ln t\}}}\right]^N.$$

The lemma will follow from

$$\lim_{t\to\infty}\left[\mathsf{E}e^{\omega(B_1)\mathbf{1}_{\{\omega(B_1)>b\ln t\}}}\right]^{2Rt^2} = 1, \tag{28}$$

which is implied by

$$\mathsf{E}e^{\omega(B_1)\mathbf{1}_{\{\omega(B_1)>b\ln t\}}} \le 1 + \frac{\mathsf{E}e^{2\omega(B_1)}}{e^{b\ln t}} \le 1 + \frac{\mathsf{E}e^{2\omega(B_1)}}{t^b},$$

if we choose $b > 2$. □

The only missing part in the proof of Lemma 10.2 is the following corollary of Lemma 10.10:

Lemma 10.11. *There is a constant D_2 such that for all $t > t_0$,*

$$0 \le \mathsf{E}\tilde A^t(\bar\omega) - \mathsf{E}\tilde A^t(\omega) < D_2.$$

PROOF OF LEMMA 10.2: Lemmas 10.5 and 10.11 imply that for $u > D_1 + D_2$

$$\begin{aligned}\mathsf{P}\{|A^t(\omega) - \mathsf{E}A^t(\omega)| > u\} \le &\mathsf{P}\{|A^t(\omega) - \tilde{A}^t(\omega)| > (u - D_1 - D_2)/3\}\\ &+ \mathsf{P}\{|\tilde{A}^t(\omega) - \tilde{A}^t(\bar\omega)| > (u - D_1 - D_2)/3\}\\ &+ \mathsf{P}\{|\tilde{A}^t(\bar\omega) - \mathsf{E}\tilde{A}^t(\bar\omega)| > (u - D_1 - D_2)/3\}.\end{aligned}$$

The Lemma follows from the estimates of the three terms provided by Lemmas 10.4, 10.9, and 10.10. □

The following lemma quantifies how the growth of $-\mathsf{E}A^t$ deviates from the linear one under argument doubling. We will use this lemma to find an estimate on $\mathsf{E}A^t - \alpha(0)t$ which makes it possible to fill the gap between Lemma 10.2 and Theorem 10.1.

Lemma 10.12. *There is a number $b_0 > 0$ such that for any $t > t_0$,*

$$0 \le 2\mathsf{E}A^t - \mathsf{E}A^{2t} \le b_0 t^{1/2} \ln^2 t.$$

PROOF: The first inequality follows from $A^{0,2t}(0,0) \le A^{0,t}(0,0) + A^{t,2t}(0,0)$. Let us prove the second one.

Let γ be the minimizer from $(0,0)$ to $(0,2t)$. Then

$$A^{2t} \ge \min_{|x|\le 2Rt} A^{0,t}(0,x) + \min_{|x|\le 2Rt} A^{t,2t}(x,0) + A^{2t}\mathbf{1}_{\{\max_{s\in[0,2t]}|\gamma(s)|>2Rt\}}.$$

Therefore, by symmetry with respect to t and Lemma 10.5,

$$\mathsf{E}A^{2t} \ge 2\mathsf{E}\min_{|x|\le 2Rt} A^t(0,x) - D_1. \tag{29}$$

For $k \in I_t = \{-2Rt, \dots, 2Rt-2, 2Rt-1\}$, we define a unit square $B_k = [k, k+1] \times [t-1, t]$.

Let now γ be the minimizer from $(0,0)$ to (x,t), with $x \in [k, k+1]$ for some $k \in I_t$. Denote $t' = \sup\{s \le t : \gamma(s) \notin B_k\}$ and $x' = \gamma(t')$.

If $x' < k+1$, then by reconnecting (x', t') to (k, t) we obtain

$$A^t(k) \le A^{t'}(x') + 1/2 \le A^t(x) + \omega(B_k) + 1/2.$$

If $x' = k+1$, then by reconnecting (x', t') to $(k+1, t)$ we obtain

$$A^t(k+1) \le A^{t'}(x') \le A^t(x) + \omega(B_k).$$

Therefore,

$$A^t(x) \ge \min\{A^t(k), A^t(k+1)\} - \omega(B_k) - 1/2,$$

and (29) implies

$$\mathsf{E}A^{2t} \ge 2\mathsf{E}\min_{k\in I_t} A^t(k) - \mathsf{E}\max_{k\in I_t}\omega(B_k) - 1/2 - D_1.$$

The second term grows logarithmically in t. Hence, for some constant $c > 0$,

$$\mathsf{E}A^{2t} \ge 2\mathsf{E}\min_{k\in I_t} A^t(k) - c(\ln t + 1).$$

Lemma 9.1 implies

$$\min_x \mathsf{E}A^t(x) = \mathsf{E}A^t(0).$$

Therefore,

$$\begin{aligned}\mathsf{E}A^{2t} &\ge 2\mathsf{E}\min_{k\in I_t} A^t(k) - c(\ln t + 1)\\ &\ge 2\min_{k\in I_t}\mathsf{E}A^t(k) - 2\mathsf{E}X_t - c(\ln t + 1)\\ &\ge 2\mathsf{E}A^t - 2\mathsf{E}X_t - c(\ln t + 1),\end{aligned} \tag{30}$$

where

$$X_t = \max_{k\in I_t}\{(\mathsf{E}A^t(k) - A^t(k))_+\}.$$

For a constant r to be determined later, we introduce the event

$$E = \{X_t \le r(\ln^2 t)\sqrt{t}\}.$$

Then

$$X_t \le r(\ln^2 t)\sqrt{t}\mathbf{1}_E + X_t\mathbf{1}_{E^c}.$$

Therefore,

$$\mathsf{E}X_t \le r(\ln^2 t)\sqrt{t} + \sqrt{\mathsf{E}(X_t)^2\mathsf{P}(E^c)}. \tag{31}$$

Let us estimate the second term. According to Lemma 9.1, the random variables $A^t(k) - \mathsf{E}A^t(k)$, $k \in I_t$ have the same distribution, so replacing the maximum in the definition of X_t^2 with summation we obtain

$$\mathsf{E}X_t^2 \le 4Rt\mathsf{E}(A^t - \mathsf{E}A^t)_+^2 \le 4Rt\mathsf{E}(A^t)^2 \le Ct^3, \tag{32}$$

for some $C > 0$ and all t exceeding some t_0, where we used Lemma 9.6 in the last inequality.

Also, Lemma 10.2 shows that

$$\begin{aligned}\mathsf{P}(E^c) &\le \sum_{k\in I_t}\mathsf{P}\left\{A^t(k) - \mathsf{E}A^t(k) > r(\ln^2 t)\sqrt{t}\right\}\\ &\le 4Rtb_1\exp\{-b_2 r\ln t\}.\end{aligned} \tag{33}$$

We can now finish the proof by choosing r to be large enough and combining estimates (30)–(33). □

With this lemma at hand we can now use the following statement ([HN01, Lemma 4.2]):

Lemma 10.13. *Suppose the functions $a : \mathbb{R}_+ \to \mathbb{R}$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy the following conditions: $a(t)/t \to \nu \in \mathbb{R}$ and $g(t)/t \to 0$ as $t \to \infty$, $a(2t) \ge 2a(t) - g(t)$ and $\psi \equiv \limsup_{t\to\infty} g(2t)/g(t) < 2$. Then, for any $c > 1/(2-\psi)$, and for all large t,*

$$a(t) \le \nu t + cg(t).$$

Taking $a(t) = \mathsf{E}A^t$, $\nu = \alpha(0)$, $g(t) = b_0 t^{1/2}\ln^2 t$, $\psi = \sqrt{2}$, $c = 2$, we conclude that for $b_0' = 2b_0$ and large t,

$$0 \le \mathsf{E}A^t - \alpha(0)t \le b_0' t^{1/2}\ln^2 t,$$

and Theorem 10.1 follows from this estimate, Lemma 10.2, and the shear invariance established in Lemma 9.1. □

11. One-sided minimizers: existence, uniqueness, and coalescence

In this section we are going to prove Theorem 8.6 that asserts the existence, uniqueness, and coalescence of one-sided infinite minimizers for any given asymptotic slope v. We begin though with approximate straightness of minimizers over finite time intervals. This is where Theorem 10.1 plays a crucial role, along with strong convexity of the shape function α.

11.1. **δ-Straightness.** First, we need some notation. Let us recall that our space-time is $\mathbb{R}^2$. For a point $p \in \mathbb{R}^2$ the numbers p_1 and p_2 are, respectively, space and time components of p.

For a speed $v > 0$ we introduce the following symmetric cone in space-time:

$$\mathrm{Co}(v) = \{p \in \mathbb{R} \times \mathbb{R}^+ : \ |p_1| \le p_2 v\}.$$

For $p \in \mathbb{R} \times \mathbb{R}^+$ and for $L > 0$

$$\mathcal{C}(p, L) := \left\{ q \in \mathbb{R}^2 : \ q_2 \in (p_2, 2p_2] \text{ and } \left| \frac{q_2}{p_2} p_1 - q_1 \right| \le L \right\}.$$

So $\mathcal{C}(p, L)$ is a parallelogram of width $2L$ along $[p, 2p]$ (for any two points p, q on the plane, $[p, q]$ denotes the straight line segment connecting these two points). We need to consider the side-edges of this parallelogram:

$$\partial_S \mathcal{C}(p, L) := \left\{ q \in \mathbb{R}^2 : \ q_2 \in (p_2, 2p_2] \text{ and } \left| \frac{q_2}{p_2} p_1 - q_1 \right| = L \right\}.$$

For $p \in \mathbb{R}^2$, we define its square neighborhood in space-time:

$$K(p, R) := \left\{ q \in \mathbb{R}^2 : \ |q_1 - p_1| \le R \text{ and } |q_2 - p_2| \le R \right\}. \tag{34}$$

For $p, q \in \mathbb{R}^2$ satisfying $p_2 < q_2$, we denote by $\gamma_{p,q}$ the optimal path from p to q. For $p, z \in \mathbb{R}^2$ with $0 < p_2 < z_2$, we define the event

$$G(p, z) = \left\{ \exists \tilde{0} \in K((0, 0), 1), \tilde{z} \in K(z, 1) : \ \gamma_{\tilde{0},\tilde{z}} \cap K(p, 1) \ne \emptyset \right\}.$$

This says that a geodesic starting close to $(0, 0)$ and ending near z, passes close to p.

Using the strong convexity of α and Theorem 10.1 one can prove the following inequality:

Lemma 11.1. *Fix $\delta \in (0, 1/4)$ and $v > 0$. There are constants $c_1, c_2, M > 0$ (independent of δ), such that for all $p \in \mathrm{Co}(v)$ with $p_2 > M$ and $z \in \partial_S \mathcal{C}(p, p_2^{1-\delta})$, we have*

$$\mathsf{P}(G(p, z)) \le c_1 \exp\left(-c_2 p_2^{1/2 - 2\delta} / \log(p_2) \right).$$

With this lemma at hand one can show that a minimal path starting close to the origin and passing close to p, with high probability will not exit the slanted cylinder $\mathcal{C}(p, p_2^{1-\delta})$ through the sides, see Figure 10. Define the event

$$G(p) = \left\{ \exists \ \tilde{0} \in K((0, 0), 1) \ \exists \ z \in \partial_S \mathcal{C}(p, p_2^{1-\delta}) : \ \gamma_{\tilde{0},z} \cap K(p, 1) \ne \emptyset \right\}.$$

Lemma 11.2. *Fix $\delta \in (0, 1/4)$ and $v > 0$. There are constants $c_1, c_2, \kappa, M > 0$, such that for all $p \in \mathrm{Co}(v)$ with $p_2 > M$ we have*

$$\mathsf{P}(G(p)) \le c_1 \exp(-c_2 p_2^{\kappa}).$$

We will need one more result on straightness.

For all $x \in \mathbb{R} \times \mathbb{R}^+$ and $\eta > 0$ we introduce

$$\mathrm{Co}(x, \eta) = \{z \in \mathbb{R} \times \mathbb{R}^+ \ : \ |z_1/z_2 - x_1/x_2| \le \eta\}, \tag{35}$$

which is the cone (rooted at the origin) of all points z that have the corresponding velocity closer than η to the speed of x. For a path γ and $t \in \mathbb{R}$, we define

$$\gamma^{\mathrm{out}}(t) = \{(s, \gamma(s)) \ : \ s \ge t\}.$$

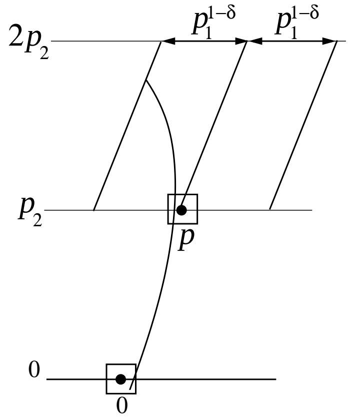

FIGURE 10. Probability of this violation of straightness is bounded by $c_1 \exp(-c_2 p_2^{\kappa})$

Lemma 11.3 (**δ-straightness**). *For $\delta \in (0, 1/4)$ and $v > 0$, there exist constants $M, R, \kappa, C_1, C_2 > 0$ such that*

(1) *with probability one, for all $\tilde{0} \in K((0,0),1)$, for all $z \in \mathbb{R} \times \mathbb{R}^+$ and for all $p \in \gamma(\tilde{0}, z) \cap \mathrm{Co}(v)$ with $p_2 > M$, we have*
$$\gamma^{\mathrm{out}}_{\tilde{0},z}(p_2) \subset \mathrm{Co}(p, Rp_2^{-\delta});$$
(2) *for all $n \geq M$*
$$\mathsf{P}(G_n) \leq C_1 e^{-C_2 n^{\kappa}},$$
where
$$G_n = \Big\{ \exists \tilde{0} \in K((0,0),1), z \in \mathbb{R} \times \mathbb{R}^+, p \in \gamma(\tilde{0}, z) \text{ with } p_2 > n \text{ and } p \in \mathrm{Co}(v) : \gamma^{\mathrm{out}}(p_2) \not\subset \mathrm{Co}(p, Rp_2^{-\delta}) \Big\}.$$

This lemma states that if a geodesic starting near $(0,0)$ passes through a remote point p, it has to stay in a narrow cone around the ray $\mathbb{R}^+ \cdot p$.

PROOF: Let us consider events $G(\bar{p})$ for all $\bar{p} \in \mathbb{Z} \times \mathbb{Z}^+ \cap \mathrm{Co}(v')$, with $v' > v$. Event G_n implies that for some $\bar{p}$ with $\bar{p}_2 > n$ event $G(\bar{p})$ happens. The lemma then follows from estimating the probabilities of $G(\bar{p})$ with Lemma 11.2 and using the Borel–Cantelli lemma. In fact, if for sufficiently large $\bar{p}_2$, events $G(\bar{p})$ do not happen, then each long optimal path never crosses side boundaries of a telescopic tower of parallelograms with sublinearly growing width, see Figure 11, and each such tower can be imbedded in a narrow cone. See Section 6 of [BCK] for more details. □

11.2. **Existence and uniqueness of one-sided minimizers.** With approximate straightness in hand, we can prove some important properties of minimizing paths. A semi-infinite minimizer starting at $(x,t) \in \mathbb{R}^2$ is a path $\gamma : [t, \infty) \to \mathbb{R}$ such that $\gamma(t) = x$ and the restriction of γ to any finite time interval is a minimizer. We call (x,t) the endpoint of γ.

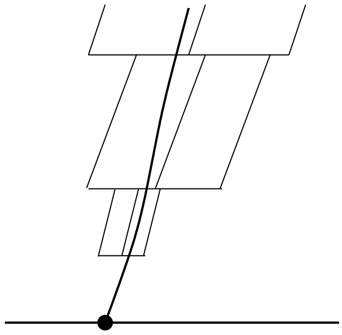

FIGURE 11. The tower of parallelograms in the straightness lemma

Lemma 11.4. *With probability one, all semi-infinite minimizers have an asymptotic slope (velocity, direction): for every minimizer γ there exists $v \in \mathbb{R} \cup \{\pm\infty\}$ depending on γ such that*

$$\lim_{t\to\infty} \frac{\gamma(t)}{t} = v.$$

PROOF: Let us fix a sequence $v_n \to \infty$. Using the translation invariance of the Poisson point field, with probability one, for any $q \in \mathbb{Z}^2$ we can choose a corresponding sequence of constants $M_n(q) > 0$ such that the statement in Lemma 11.3 holds for the entire sequence, for paths starting in $K(q, 1)$.

Let us take some one-sided minimizer γ. If $\gamma(t)/t \to +\infty$ or $-\infty$, then the desired statement is automatically true. In the opposite case we have

$$\liminf_{t\to\infty} \frac{|\gamma(t)|}{t} < \infty.$$

This implies that there exist $n \geq 1$ and a sequence $t_m \to \infty$ such that $|\gamma(t_m)|/t_m \leq v_n$. We define $y_m = (\gamma(t_m), t_m)$ and choose $q \in \mathbb{Z}^2$ such that $y_1 \in K(q, 1)$. For m large enough, we will have that $t_m > M_n(q)$ and, therefore,

$$\gamma^{\text{out}}(y_m) \subset q + \text{Co}(y_m - q, R|y_m - q|^{-\delta}),$$

for some constant $R > 0$ and m large enough. Clearly this implies that γ must have a finite asymptotic slope. □

Lemma 11.5. *With probability one, for every $v \in \mathbb{R}$ and for every sequence $(y_n, t_n) \in \mathbb{R}^2$ with $t_n \to \infty$ and*

$$\lim_{n\to\infty} \frac{y_n}{t_n} = v,$$

and for every $x \in \mathbb{R}^2$, there exists a subsequence (n_k) such that the minimizing paths $\gamma_{x,(y_{n_k}, t_{n_k})}$ are an increasing collection of paths that converge to a semi-infinite minimizer starting at x and with asymptotic slope equal to v.

PROOF: The proof of the lemma is based on a compactness argument. We give only a sketch of the proof and refer to Section 6 of [BCK] for details.

The straightness estimates and the Borel–Cantelli lemma imply that for sufficiently large t and sufficiently large n all minimizing paths to (x_n, t_n) have to pass through the rectangle with sides t and $t^{1/4}$ shown on Figure 12. One can also show that (if t is sufficiently large) an optimal path has to visit at least one Poissonian

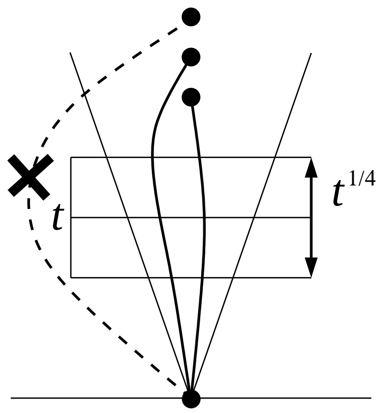

FIGURE 12. Due to straightness estimates, minimizers have to pass through the rectangle for large t_n.

point in that rectangle. Since there are only finitely many of those, there will be one such that an infinite number of minimizers from our sequence will visit it. Taking that subsequence, increasing t, repeating the same argument iteratively, and taking the diagonal subsequence from the sequence of thus constructed subsequences of minimizers, we obtain a limiting trajectory. Being a limit of minimizing paths, it is the desired one-sided minimizer . □

Lemma 11.6. *With probability 1, the following statement holds: if γ_1 and γ_2 are two (finite-time) geodesics, starting at the same Poisson point p, and for some $t > p_2$ we have $\gamma_1(t) < \gamma_2(t)$, then for all (relevant) $s > t$ we have $\gamma_1(s) < \gamma_2(s)$.*

PROOF: The probability that there are two Poisson point connected by two distinct geodesics is zero. So we only have to consider the situation where two paths with vertices $p, p_1, \ldots, p_n$ and $p, q_1, \ldots, q_m$ intersect transversally, i.e., for some k, j, $[p_j, p_{j+1}] \cap [q_k, q_{k+1}] = \{x\}$ for some point $x \notin \omega$. In this case, the total actions of the two paths can be improved by switching to paths with vertices $p, p_1, \ldots, p_j, q_{k+1}, \ldots, q_m$ and, respectively, $p, q_1, \ldots, q_k, p_{j+1}, \ldots, p_n$. Therefore, we obtain a contradiction with the optimality of the original paths. □

Lemma 11.7. *Let $v \in \mathbb{R}$. With probability one, every Poisson point belongs to at most one semi-infinite minimizer with asymptotic slope v.*

PROOF: A triple of distinct Poisson points (p, q_1, q_2) is called a bifurcation triple for v if there exist two distinct semi-infinite minimizers γ_1 and γ_2 with asymptotic slope v that both start at p, then one goes directly (at constant velocity) to q_1 and the other goes directly to q_2. We choose q_1 such that γ_1 lies to the left of γ_2.

Lemma 11.6 implies that a triple (p, q_1, q_2) can be a bifurcation triple no more than for one direction v. Therefore, each realization ω of Poissonian point field gives at most countably many bifurcation triples. For $v \in \mathbb{R}$ we define B_v as the event consisting of all point configurations with at least one bifurcation triple for v. Let $f > 0$ be a bounded probability density on $\mathbb{R}$. Then

$$\mathsf{E} \int_{\mathbb{R}} f(v) \mathbf{1}_{B_v}(\omega) dv = \mathsf{E}0 = 0,$$

since $\mathbf{1}_{B_v}(\omega) \neq 0$ for at most countably many values of v. Changing the order of integration, we see that the left-hand side equals

$$\int_{\mathbb{R}} f(v)\mathsf{P}(B_v)dv = \int_{\mathbb{R}} f(v)\mathsf{P}(B_0)dv = \mathsf{P}(B_0),$$

where we used shear invariance to conclude that $\mathsf{P}(B_v) = \mathsf{P}(B_0)$ for all v. Therefore, $\mathsf{P}(B_0) = 0$ and the lemma follows. □

With uniqueness in hand we can strengthen Lemma 11.5:

Lemma 11.8. *With probability one, for every $v \in \mathbb{R}$ and for every sequence $(y_n, t_n) \in \mathbb{R}^2$ with $t_n \to \infty$ and*

$$\lim_{n\to\infty} \frac{y_n}{t_n} = v,$$

and for every Poissonian point $p \in \mathbb{R}^2$, the minimizing paths $\gamma_{p,(y_n,t_n)}$ converge to a unique semi-infinite minimizer $\gamma_{p,v}$ starting at p and with asymptotic speed equal to v.

Proof: Let us assume that the convergence does not hold, i.e., there is a sequence (n') such that the restrictions of $\gamma_{p,(y_{n'},t_{n'})}$ and $\gamma_{p,v}$ on some finite time interval I do not coincide for all n'. Lemma 11.5 allows to choose a subsequence (n'') from (n') such that for sufficiently large n'' the restrictions of $\gamma_{p,(y_{n''},t_{n''})}$ on I coincide with the restrictions of some infinite one-sided geodesics γ'. The uniqueness established in Lemma 11.7 guarantees that γ' coincides with $\gamma_{p,v}$, and the resulting contradiction shows that our assumption was false, completing the proof. □

11.3. Coalescence of minimizers.

Here we prove that any two one-sided minimizers with the same asymptotic slope coalesce. In a sense, this is a strengthening of the hyperbolicity property that holds true for minimizers of smooth random action in compact setting. Hyperbolicity means that different minimizers approach each other exponentially fast in reverse time. In the Poissonian setting every two minimizers meet in finite time, and this property is called hyperhyperbolicity in [Bak12].

Lemma 11.9. *With probability one it holds that for every $v \in \mathbb{R}$ and for every pair of semi-infinite minimizers, starting at different Poisson points, with asymptotic speed v, these minimizers either do not touch, or they coalesce at some Poisson point.*

Proof: Suppose for some $v \in \mathbb{R}$, there do exist two semi-infinite minimizers with asymptotic speed v that touch, but do not coalesce. If the two minimizers γ_1 and γ_2 contain the same Poisson point p, then they must stay together for all times above p according to Lemma 11.7. Therefore, the only option is that γ_1 and γ_2 cross, i.e., they consecutively visit Poissonian points $p_1, p_2, \ldots$, and, respectively, $q_1, q_2, \ldots$, and $[p_1, p_2] \cap [q_1, q_2] = \{x\}$, for some $x \in \mathbb{R}^2$.

The sequence $\{q_m : m \geq 1\}$ satisfies the conditions of Lemma 11.8, which means that the minimizers γ_{p_1,q_m} converge to γ_1. However, we claim that with probability 1, none of the minimizers γ_{p_1,q_m} contain any of the p_n ($n \geq 2$). In fact, if this claim is violated for some m, n, then due to a.s.-uniqueness of a geodesic between any two Poisson points, we know that γ_{p_1,q_m} passes through p_2 and x. This implies that action picked up by γ_2 between x and q_m must be equal to the action picked up by γ_{p_1,q_m} between x and q_m. However, this contradicts the optimality of γ_2 as the comparison with the path connecting q_1 directly to p_2 and then following γ_{p_1,q_m} shows. The proof is complete. □

For every $v \in \mathbb{R}$, coalescence of one-sided geodesics with asymptotic slope v generates an equivalence relation on Poissonian points. We call each equivalence class a coalescence component.

Lemma 11.10. *Let $v \in \mathbb{R}$. With probability 1, every coalescence component is unbounded below in time.*

Proof: We give only a sketch of proof, see [BCK] for details. Due to shear invariance, it is sufficient to consider only $v = 0$. Suppose that the probability of existence of a coalescence component bounded from below is positive. Then the indicators $I_{ij}, (i,j) \in \mathbb{Z}^2$ of the fact that such a component begins within $[i, i+1] \times [j, j+1]$ form a stationary process in space-time. If a coalescence component begins between the one-sided optimal paths γ_+ and γ_- emitted from the origin 0 with asymptotic slopes 1 and -1, respectively, then it is trapped between them for all times. Due to the Tempelman multi-parametric ergodic theorem, (see, e.g., [Kre85, Chapter 6]), the number of these components emitted from squares $[i, i+1] \times [j, j+1]$ with $j \le n$ between γ_+ and γ_- is approximately linear in the number of these squares, i.e., grows quadratically in n. However, the spread between γ_+ and γ_- grows only linearly, and thus it is too small to contain Poissonian points from all quadratically many disjoint components. □

Lemma 11.11. *Let $v \in \mathbb{R}$. With probability one, every two semi-infinite minimizers with asymptotic slope v coalesce.*

Proof: Let us only give the main ideas behind the proof, see [BCK] for details. The first statement is: if, with positive probability, there are two noncoalescent minimizers, then, probability that there are three noncoalescent minimizers is also positive. This follows from stationarity.

The main part of the reasoning is to use thus constructed event of positive probability to construct another event of positive probability on which there is a coalescence component bounded from below. This is possible to do by erasing Poissonian points in the rectangular domain shown along with the three noncoalescent minimizers on Figure 13: we erase all Poissonian points in the rectangle except those in a small neighborhood of the leftmost and rightmost minimizers. One can prove that (i) the resulting set of configurations still has positive probability, and (ii) the coalescence component of the minimizer in the middle does not extend below the zero time level and hence is bounded from below. The contradiction with Lemma 11.10 finishes the proof. □

Now we have the picture where for fixed v, with probability 1, each Poissonian point comes with a unique one-sided minimizer of asymptotic slope v, and every two of these minimizers coalesce. The reasoning above does not give any information about the statistics of the time it takes for two minimizers to coalesce. All we know is that this time is finite with probability 1. Figure 7 computed for a finite domain gives some idea of what goes on. Minimizers for terminal points that are close to each other tend to coalesce fast, but some Poissonian points that are close to each other are separated by a shock (shocks roughly correspond to white space on the picture) with a long history, and it takes a long time for their respective minimizers to meet. In fact, it would be interesting and useful to understand the statistics of coalescence times or shock ages.

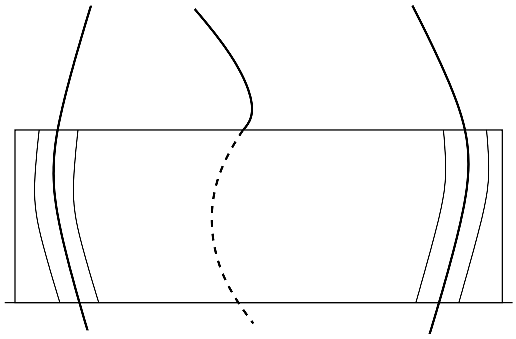

FIGURE 13. By erasing the Poissonian points in the rectangle we obtain a situation where the coalescence component of the minimizer in the middle cannot be extended below 0 level.

What if we do not restrict ourselves to Poissonian points as endpoints of minimizers? All the minimizers for other endpoints have the same structure: they first go to one of the Poissonian points and then follow the unique minimizer attached to that point. So the entire space-time is tessellated into domains such that all points in one domain have minimizers visiting the same Poissonian point first. The uniqueness is violated at the boundaries of these domains and the points on these boundaries have more than one minimizer and correspond to shocks.

12. Busemann functions and stationary solutions of the Burgers equation

In this Section we use the one-sided minimizers to construct global solutions of the Burgers equation, thus proving the existence part of Theorem 8.1.

In several last sections we worked with forward minimizers that were obtained from a limiting procedure as the terminal time approached $+\infty$. All the same conclusions are valid for one-sided backward minimizers that can be obtained from a limiting procedure in the reverse time. This amounts only to switching the direction of time due to the symmetry of the action with respect to this transformation.

Let us summarize some facts on one-sided backward minimizers. For any velocity $v \in \mathbb{R}$, the following holds with probability 1. For every point $p = (x, t)$ there is a non-empty set $\Gamma_{v,p}$ of one-sided action minimizers $\gamma : (-\infty, t] \to \mathbb{R}$ with asymptotic slope v

$$\lim_{s\to-\infty} \frac{\gamma(s)}{s} = v,$$

ending at p. They all coalesce, i.e., they coincide on $(-\infty, t_{v,p}]$ for some $t_{v,p} < t$. For most points $p \in \mathbb{R}^2$, $\Gamma_{v,p}$ consists of a unique minimizer $\gamma_{v,p}$, but even if the uniqueness does not hold, there is the right-most minimizer $\gamma_{v,p} \in \Gamma_{v,p}$ such that $\gamma_{v,p}(s) \geq \gamma(s)$ for $s \leq t$ and any other minimizer $\gamma \in \Gamma_{v,p}$.

For every two points $p_1 = (x_1, t_1)$ and $p_2 = (x_2, t_2)$, all their one-sided minimizers coalesce, i.e., there is a time $t_v = t_v(p_1, p_2)$ such that $\gamma_{v,p_1}(s) = \gamma_{v,p_2}(s)$ for all $s \leq t_v$.

This allows us to define Busemann functions for slope v:

$$B_v(p_1, p_2) = B_{v,\omega}(p_1, p_2) = A_\omega^{t_v(p_1,p_2),t_2}(\gamma_{v,p_2}) - A_\omega^{t_v(p_1,p_2),t_1}(\gamma_{v,p_1}),\ p_1, p_2 \in \mathbb{R}^2.$$

Although t_v is not defined uniquely, the definition clearly does not depend on a concrete choice of t_v or $\gamma_{v,p_1}, \gamma_{v,p_2}$. One can also choose t_v to be the maximal of all possible coalescence times.

Some properties of Busemann functions are summarized in the following lemma:

Lemma 12.1. *Let B_v be defined as above for $v \in \mathbb{R}$.*

(1) *The distribution of B_v is translation invariant: for any $\Delta \in \mathbb{R}^2$,*

$$B_v(\cdot + \Delta, \cdot + \Delta) \stackrel{distr}{=} B_v(\cdot, \cdot).$$

(2) *B_v is antisymmetric:*

$$B_v(p_1, p_2) = -B_v(p_2, p_1), \quad p_1, p_2 \in \mathbb{R}^2,$$

in particular $B_v(p, p) = 0$ for any $p \in \mathbb{R}^2$.

(3) *B_v is additive:*

$$B_v(p_1, p_3) = B_v(p_1, p_2) + B_v(p_2, p_3), \quad p_1, p_2, p_3 \in \mathbb{R}^2.$$

(4) *For any $p_1, p_2 \in \mathbb{R}^2$, $\mathsf{E}|B_v(p_1, p_2)| < \infty$.*

First three parts of the Lemma are straightforward. The proof of part 4 is not trivial, and we refer to Section 7 of [BCK] for details.

Having the Busemann function at hand, one can define

$$U_v(x, t) = B((0, 0), (x, t)), \quad (x, t) \in \mathbb{R}^2.$$

The main claim of this Section is that thus defined U_v is skew-invariant under cocycle Φ and its space derivative is the global solution of the Burgers equation.

Let us recall that Φ is given by

$$\Phi^{s,t} W(y) = \inf_{x \in \mathbb{R}} \{W(x) + A^{s,t}(x, y)\}, \quad s \le t, \quad y \in \mathbb{R}, \tag{36}$$

where $A^{s,t}(x, y)$ has been defined in (21).

Lemma 12.2. *Function U_v defined above is a global solution of the Hamilton–Jacobi equation. If $s \le t$, then*

$$\Phi^{s,t} U_v(\cdot, s)(x) = U_v(x, t).$$

Proof: Let γ_v be a minimizer through (x, t) with slope v. Then

$$\begin{aligned} U_v(x,t) =& U_v(\gamma_v(s), s) + (U_v(x,t) - U_v(\gamma_v(s), s)) \\ =& U_v(\gamma_v(s), s) + A^{s,t}(\gamma_v(s), x). \end{aligned}$$

We need to show that the right-hand side is the infimum of $U_v(y, s) + A^{s,t}(y, x)$ over all $y \in \mathbb{R}$. Suppose that for some $y \in \mathbb{R}$,

$$U_v(y, s) + A^{s,t}(y, x) < U_v(\gamma_v(s), s) + A^{s,t}(\gamma_v(s), x). \tag{37}$$

Let us take any minimizer $\bar\gamma_v$ originating at (y, s) and denote by $\tau < s$ the time of coalescence of $\bar\gamma_v$ and γ_v. We claim that

$$A^{\tau,s}(\bar\gamma_v) + A^{s,t}(y, x) < A^{\tau,t}(\gamma_v(\tau), x) = A^{\tau,t}(\gamma_v), \tag{38}$$

which contradicts the minimizing property of γ. In fact, (38) is a consequence of

$$\begin{aligned} A^{\tau,s}(\bar\gamma_v) - A^{\tau,s}(\gamma_v) &= U_v(y, s) - U_v(\gamma_v(s), s) \\ &< A^{s,t}(\gamma_v(s), x) - A^{s,t}(y, x). \end{aligned}$$

where the second inequality follows from (37). □

Another way to approach the Burgers equation is to consider, for $p=(x,t)$,

$$u_v(x,t) = \dot{\gamma}_{v,p}(t).$$

Then $U_v(x,t) - U_v(0,t) = \int_0^x u_v(y,t)dy$. We recall that $\Psi^{s,t}w$ denotes the solution at time t of the Burgers equation with initial condition w imposed at time s.

Lemma 12.3. *The function u_v defined above is a global solution of the Burgers equation. If $s \le t$, then*

$$\Psi^{s,t} u_v(\cdot, s) = u_v(\cdot, t), \quad s \le t.$$

Proof: This statement is a direct consequence of Lemmas 6.5, 12.2, and the definition of the Burgers cocycle Ψ. □

The function $u_v(\cdot,t)$ is piecewise linear with respect to the space coordinate, with downward jumps, each linear regime corresponding to the configuration point visited last by one-sided minimizers, see the last paragraph of Section 11.

To prove that $U_v(\cdot,t) \in \mathbb{H}(v,v)$ for all t, we will compute the expectation of its spatial increments (we already know that it is well defined due to part 4 of Lemma 12.1), and prove that $u_v(\cdot,t)$ is mixing with respect to the spatial variable.

Lemma 12.4. *For any $(x,t) \in \mathbb{R}^2$,*

$$\mathsf{E}(U_v(x+1,t) - U_v(x,t)) = \mathsf{E}B_v((x,t),(x+1,t)) = v.$$

Proof: First, we consider the case $v=0$. Due to the distributional invariance of Poisson process under reflections,

$$\mathsf{E}B_0((x,t),(x+1,t)) = \mathsf{E}B_0((x+1,t),(x,t)).$$

Combining this with the anti-symmetry of B_0, we obtain $\mathsf{E}B_0((x+1,t),(x,t)) = 0$, as required.

In the general case, we can apply the shear transformation $L : (y,s) \mapsto (y + (t-s)v, s) = (y + vt - vs, s)$. Due to Lemma 9.1, the one-sided minimizers of slope v will be mapped onto one-sided minimizers of slope 0 for the new Poissonian configuration $L(\omega)$. We already know that

$$\mathsf{E}B_{0,L(\omega)}((x+1,t),(x,t)) = 0,$$

and a direct computation based on Lemma 9.1 gives

$$B_{0,L(\omega)}((x,t),(x+1,t)) = B_{v,\omega}((x,t),(x+1,t)) + v,$$

and our statement follows since L preserves the distribution of Poisson process. □

Lemma 12.5. *Let $v \in \mathbb{R}$. For any t, the process $u_v(\cdot,t)$ is mixing.*

Roughly speaking, mixing means asymptotic independence of $u_v(x,t)$ and $u_v(y,t)$ as $x - y \to \infty$. The idea of the proof is that if the distance between x and y is large, then the corresponding one-sided minimizers explore disjoint domains in space-time for a long time, thus collecting input from independent patches of Poissonian points. See Section 7 of [BCK] for a formal proof of mixing.

Mixing implies ergodicity. Therefore, combining Lemmas 12.4 and 12.5 with the ergodic theorem, we can conclude that the Birkhoff space averages of u_v have a well-defined, deterministic limit v, so $U_v \in \mathbb{H}(v,v)$.

13. Stationary solutions: uniqueness and basins of attraction

In this section we prove Theorem 8.3 and the uniqueness part in Theorem 8.1. In both cases the idea is to show that minimizers for the corresponding variational problems over long time intervals tend to coincide with the one-sided minimizers constructed and studied in the previous sections.

The key step in the proof of Theorem 8.3 is the following observation.

Lemma 13.1. *Let $t \in \mathbb{R}$ and suppose that an initial condition W satisfies one of the conditions* (18),(19),(20)*. With probability one, the following holds true for every $y \in \mathbb{R}$. Let $y^*(s)$ be a solution of the optimization problem* (36)*. Then*

$$\lim_{s\to-\infty} \frac{y^*(s)}{s} = v.$$

Proof of Theorem 8.3: Let us take any rectangle $Q = [-R, R] \times [t_0, t_1]$ and set $t = t_1$. We can find points $a, b \in \mathbb{R}$ satisfying $a < -R < R < b$, not coinciding with any of the points $x_k, k \in \mathbb{Z}$ and such that one-sided backward minimizers $\gamma_{(a,t),v}$ and $\gamma_{(b,t),v}$ do not cross Q.

Applying Lemma 13.1 to $x = a, b$, we see that the corresponding points $a^*(s)$ and $b^*(s)$ satisfy $a^*(s)/s \to v$ and $b^*(s)/s \to v$ as $s \to -\infty$.

Let $p = (x_0, \tau_0)$ be the point of coalescence of the one-sided minimizers $\gamma_{(a,t),v}$ and $\gamma_{(b,t),v}$. We automatically have $\tau_0 < t_0$. There is $\tau_1 < \min\{\tau_0, 0\}$ such that for $s < \tau_1$, the restrictions of the finite minimizers connecting $(a^*(s), s)$ to (a, t) and $(b^*(s), s)$ to (b, t) on $[\tau_0, t]$ coincide with the restrictions of $\gamma_{(a,t),v}$ and $\gamma_{(b,t),v}$ (this also implies that we can choose $a^*(s) = b^*(s)$).

Since Q is trapped between $\gamma_{(a,t),v}$ and $\gamma_{(b,t),v}$, and minimizing paths cannot cross each other, we conclude that for any $s < \tau_1$, and any $(x, t) \in Q$, the minimizers connecting $(x^*(s), s)$ to (x, t) (where x^* is a solution of the optimization problem (36)) have to pass through p. In particular, the slopes of these minimizers determining the evolution of the Burgers velocity field in $[-R, R]$ throughout $[t_0, t_1]$ do not change (and coincide with the slopes of one-sided backward minimizers) as long as $s < \tau_1$, which completes the proof. □

Proof of uniqueness in Theorem 8.1: We will prove that any skew-invariant function u with average velocity v coincides with the global solution u_v at time 0.

Let us take an arbitrary interval $I = (a, b)$. Lemma 13.1 implies that for any W satisfying $\mathbb{H}(v, v)$, there is a time $T_0(a, b, W) \geq 0$ such that if $s < -T_0$, then there is a point $a^* \in \mathbb{R}$ that solves the optimization problem (36) for $t = 0$ and for all points $y \in I$ at once, and the respective finite minimizers on $[s, 0]$ have the same velocity at time 0 as the infinite one-sided minimizers of asymptotic slope v.

Suppose now that $U_\omega(x, t) = U_{\theta^t\omega}(x)$ is a global solution in $\hat{\mathbb{H}}(v, v)$. Then $T_0(a, b, U_{\theta^t\omega}(\cdot)) > 0$ is a stationary process. In particular, this means that with probability 1, there is $R > 0$ and a sequence of times $s_n \downarrow -\infty$ such that for all $n \in \mathbb{N}$, we have $T_0(a, b, U_{\theta^{s_n}\omega}(\cdot)) < R$. Therefore, there is n such that $s_n < -T_0(a, b, U_{\theta^{s_n}\omega}(\cdot))$. This and the fact that U at time 0 is the solution of problem (36) for $t = 0$, $s = s_n$ and initial condition $W = U_{\theta^{s_n}\omega}(\cdot)$, we conclude that U and the global solution U_v coincide on I at time 0, and the proof is complete. □

14. Conclusion

We finish these notes with a short discussion of future directions.

A natural step in developing a general theory of random Lagrangian systems is to eliminate the singular character of the forcing. There are several settings where such elimination looks plausible. For example, one can consider shot-noise type random forcing where a singular contribution to the potential from each Poissonian point is replaced by a smooth one. In this situation, large parts of the program described in our notes go through, but we cannon expect coalescence of one-sided minimizers. We expect that this property will be replaced by hyperbolicity analogously to the situation on the circle, but the study of hyperbolicity will require new tools.

Another interesting related set of problems concerns discrete-time systems or systems with kick forcing where the external potential is concentrated at integer times and is smooth in space.

To extend our program to higher dimensions is also an interesting problem. Existence of one-sided minimizers can be obtained in the same way as in one dimension, but the rest of the program in our notes critically depends on the geometry of the space-time plane — we often use the fact that the minimizers can be ordered from left to right, can be trapped between other minimizers, etc.

Another set of problems is connected to the case of positive viscosity ν. Our variational approach has to be replaced then by stochastic control. Using the latter has proved to be productive in compact setting. It would be very interesting to see if the invariant measures corresponding to $\nu > 0$ converge in some sense to the invariant measures for the inviscid case, as they do in compact setting.

Finally, the Burgers equation is believed to belong to the KPZ universality class, and it might be easier to obtain the KPZ scaling exponents 1/3 and 2/3 for the optimal paths in the Burgers setting than for other models, due to the fact that the shape function is precisely quadratic.

References

[AD95] D. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields*, 103(2):199–213, 1995.

[Arn98] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[Bak06] Yuri Bakhtin. Existence and uniqueness of stationary solutions for 3D Navier-Stokes system with small random forcing via stochastic cascades. *J. Stat. Phys.*, 122(2):351–360, 2006.

[Bak07] Yuri Bakhtin. Burgers equation with random boundary conditions. *Proc. Amer. Math. Soc.*, 135(7):2257–2262 (electronic), 2007.

[Bak12] Yuri Bakhtin. Burgers equation with Poisson random forcing. *Accepted at Ann. of Probab., available at http://arxiv.org/abs/1109.5668*, 2012.

[BCK] Yuri Bakhtin, Eric Cator, and Konstantin Khanin. Space-time stationary solutions for the Burgers equation. *Accepted at JAMS, available at http://arxiv.org/abs/1205.6721.*

[BK07] Jérémie Bec and Konstantin Khanin. Burgers turbulence. *Phys. Rep.*, 447(1-2):1–66, 2007.

[BKL01] J. Bricmont, A. Kupiainen, and R. Lefevere. Ergodicity of the 2D Navier-Stokes equations with random forcing. *Comm. Math. Phys.*, 224(1):65–81, 2001. Dedicated to Joel L. Lebowitz.

[Bur39] J. M. Burgers. Mathematical examples illustrating relations occurring in the theory of turbulent fluid motion. *Verh. Nederl. Akad. Wetensch. Afd. Natuurk. Sect. 1.*, 17(2):53, 1939.

[CGGK93] J. Theodore Cox, Alberto Gandolfi, Philip S. Griffin, and Harry Kesten. Greedy lattice animals. I. Upper bounds. *Ann. Appl. Probab.*, 3(4):1151–1169, 1993.

[CP] Eric Cator and Leandro Pimentel. Busemann functions and equilibrium measures in last passage percolation models. *Probability Theory and Related Fields*, pages 1–37. 10.1007/s00440-011-0363-6.

[CP11] Eric Cator and Leandro P.R. Pimentel. A shape theorem and semi-infinite geodesics for the Hammersley model with random weights. *ALEA*, 8:163–175, 2011.

[Cra08] H. Crauel. Measure attractors and Markov attractors. *Dyn. Syst.*, 23(1):75–107, 2008.

[CSG11] Mickaël D. Chekroun, Eric Simonnet, and Michael Ghil. Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Phys. D*, 240(21):1685–1700, 2011.

[DVJ03] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[EKMS00] Weinan E, K. Khanin, A. Mazel, and Ya. Sinai. Invariant measures for Burgers equation with stochastic forcing. *Ann. of Math. (2)*, 151(3):877–960, 2000.

[EMS01] Weinan E, J. C. Mattingly, and Ya. Sinai. Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation. *Comm. Math. Phys.*, 224(1):83–106, 2001. Dedicated to Joel L. Lebowitz.

[Fla94] Franco Flandoli. Dissipativity and invariant measures for stochastic Navier-Stokes equations. *NoDEA Nonlinear Differential Equations Appl.*, 1(4):403–423, 1994.

[FM95] Franco Flandoli and Bohdan Maslowski. Ergodicity of the 2-D Navier-Stokes equation under random perturbations. *Comm. Math. Phys.*, 172(1):119–141, 1995.

[GIKP05] Diogo Gomes, Renato Iturriaga, Konstantin Khanin, and Pablo Padilla. Viscosity limit of stationary distributions for the random forced Burgers equation. *Mosc. Math. J.*, 5(3):613–631, 743, 2005.

[GK94] Alberto Gandolfi and Harry Kesten. Greedy lattice animals. II. Linear growth. *Ann. Appl. Probab.*, 4(1):76–107, 1994.

[HM06] Martin Hairer and Jonathan C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)*, 164(3):993–1032, 2006.

[HM08] Martin Hairer and Jonathan C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[HM11] Martin Hairer and Jonathan C. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electron. J. Probab.*, 16:no. 23, 658–738, 2011.

[HN97] C. Douglas Howard and Charles M. Newman. Euclidean models of first-passage percolation. *Probability Theory and Related Fields*, 108:153–170, 1997. 10.1007/s004400050105.

[HN99] C. Douglas Howard and Charles M. Newman. From greedy lattice animals to Euclidean first-passage percolation. In *Perplexing problems in probability*, volume 44 of *Progr. Probab.*, pages 107–119. Birkhäuser Boston, Boston, MA, 1999.

[HN01] C. Douglas Howard and Charles M. Newman. Geodesics and spanning trees for Euclidean first-passage percolation. *Ann. Probab.*, 29(2):577–623, 2001.

[IK03] R. Iturriaga and K. Khanin. Burgers turbulence and random Lagrangian systems. *Comm. Math. Phys.*, 232(3):377–428, 2003.

[Kes93] Harry Kesten. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(2):296–338, 1993.

[Kre85] Ulrich Krengel. *Ergodic theorems*, volume 6 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1985. With a supplement by Antoine Brunel.

[KS00] Sergei Kuksin and Armen Shirikyan. Stochastic dissipative PDEs and Gibbs measures. *Comm. Math. Phys.*, 213(2):291–330, 2000.

[Kuk04] Sergei B. Kuksin. The Eulerian limit for 2D statistical hydrodynamics. *J. Statist. Phys.*, 115(1-2):469–492, 2004.

[Kuk07] S. B. Kuksin. Eulerian limit for 2D Navier-Stokes equation and damped/driven KdV equation as its model. *Tr. Mat. Inst. Steklova*, 259(Anal. i Osob. Ch. 2):134–142, 2007.

[Kuk08] Sergei B. Kuksin. On distribution of energy and vorticity for solutions of 2D Navier-Stokes equation with small viscosity. *Comm. Math. Phys.*, 284(2):407–424, 2008.

[Lig85] Thomas M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.*, 13(4):1279–1285, 1985.

[Lio82] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[LY88] F. Ledrappier and L.-S. Young. Entropy formula for random transformations. *Probab. Theory Related Fields*, 80(2):217–240, 1988.

[New95] Charles M. Newman. A surface view of first-passage percolation. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 1017–1023, Basel, 1995. Birkhäuser.

[Sin91] Ya. G. Sinaĭ. Two results concerning asymptotic behavior of solutions of the Burgers equation with force. *J. Statist. Phys.*, 64(1-2):1–12, 1991.

[Vil09] Cédric Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

[Wüt02] Mario V. Wüthrich. Asymptotic behaviour of semi-infinite geodesics for maximal increasing subsequences in the plane. In *In and out of equilibrium (Mambucaba, 2000)*, volume 51 of *Progr. Probab.*, pages 205–226. Birkhäuser Boston, Boston, MA, 2002.

School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA, 30332-0160, USA

E-mail address: bakhtin@math.gatech.edu