Kinetic theory and Lax equations for shock clustering and Burgers turbulence

Govind Menon\textsuperscript{1} and Ravi Srinivasan\textsuperscript{2}

September 21, 2009

Abstract

We study shock statistics in the scalar conservation law $\partial_t u + \partial_x f(u) = 0$, $x \in \mathbb{R}$, $t > 0$, with a convex flux $f$ and random initial data. We show that a large class of random initial data (Markov processes with downward jumps and derivatives of Lévy processes with downward jumps) is preserved by the entropy solution to the conservation law and we derive kinetic equations that describe the evolution of shock statistics. These kinetic equations are equivalent to a Lax pair. Moreover, they admit remarkable exact solutions for Burgers equation ($f(u) = u^2/2$) suggesting the complete integrability of Burgers turbulence.

MSC classification: 60J75, 35R60, 35L67, 82C99

1 Introduction

We consider the scalar conservation law

$$\partial_t u + \partial_x f(u) = 0, \quad x \in \mathbb{R}, t > 0, \quad u(x, 0) = u_0(x),$$

with a strictly convex, $C^1$ flux $f$ and random initial data $u_0$. Our approach to the study of (1) is rooted in the study of Burgers equation ($f(u) = u^2/2$) with white noise as initial data. This example was introduced by Burgers in his study of turbulence \cite{burgers1948, burgers1954, burgers1955}. Burgers turbulence also refers to the study of Burgers equation with random forcing, but we do not consider this here. While the model fails to describe turbulence in incompressible fluids, it still serves as a widely useful benchmark for theoretical methods and computations in turbulence. It also has fascinating links with combinatorics, mathematical physics and statistics. Some of these are indicated below.

\textsuperscript{1}Division of Applied Mathematics, Box F, Brown University, Providence, RI 02912. Email: menon@dam.brown.edu

\textsuperscript{2}Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, Email: rav@math.utexas.edu
The Hopf-Lax formula

Let us first recall the notion of the **entropy solution** to (1) [13, 27]. Characteristics for (1) are lines in space-time along which \( u \) is constant. If a unique characteristic connects \((a, 0)\) to \((x, t)\) then \( u(x, t) = u_0(a) \). However, there may be many characteristics that pass through \((x, t)\). Well-posedness is resolved by adding a small dissipative term \( \varepsilon u_{xx} \) to the right-hand side of (1) and passing to the limit \( \varepsilon \downarrow 0 \). This was first carried out by Hopf in his pioneering work on Burgers equation [21], and later generalized to convex \( f \) by Lax [26].

Let \( f^*(s) = \sup_{u \in \mathbb{R}} (su - f(u)) \) denote the Legendre transform of \( f \) and call \( U_0(s) = \int_0^s u_0(r) dr \) the initial potential. We define the **Hopf-Lax functional**

\[
I(s; x, t) = U_0(s) + tf^* \left( \frac{x - s}{t} \right). \tag{2}
\]

The ‘correct’ characteristic through \((x, t)\) is given by the variational principle

\[
a(x, t) = \sup \left\{ s \in \mathbb{R} : I(s; x, t) = \inf_{r \in \mathbb{R}} I(r; x, t) \right\}. \tag{3}
\]

We will always assume that \( U_0 \) has no upward jumps and \( I \) satisfies the growth condition

\[
\lim_{|s| \to \infty} I(s; x, t) = \infty. \tag{4}
\]

This ensures the infimum of \( I \) is a minimum, and that \( a(x, t) \) is finite. We write

\[
a(x, t) = \arg^+ \min_{s \in \mathbb{R}} I(s; x, t). \tag{5}
\]

This is the **Hopf-Lax formula** for the inverse Lagrangian function \( a(x, t) \). The \( ^+ \) in (5) denotes that we choose \( a(x, t) \) to be the largest location if \( I \) is minimized at more than one point. Of particular importance is Burgers equation with \( f(u) = u^2/2 \). In this case, (5) is called the Cole-Hopf formula, and takes the form

\[
a(x, t) = \arg^+ \min_{s \in \mathbb{R}} \left\{ U_0(s) + \frac{(x - s)^2}{2t} \right\}, \quad u(x, t) = \frac{x - a(x, t)}{t}. \tag{6}
\]

For fixed \( t \), \( a \) defined by (5) is non-decreasing in \( x \). If \( x \) is a point of continuity of \( a(\cdot, t) \), the velocity field is given implicitly by

\[
f'(u(x, t)) = \frac{x - a(x, t)}{t}. \tag{7}
\]

\( u(x, t) \) is well-defined because \( f' \) is continuous and strictly increasing. In particular, if \( a(x, t) \) is constant for \( x \) in an interval, we obtain a **rarefaction wave**. \( a \) may also have upward jumps. These arise if the minimum of \( I \) is attained at more than one point. Jumps in \( a \) give rise to **shocks** in \( u \). The left and
right limits $u_\pm = u(x_\pm, t)$ exist and the velocity of the shock is given by the Rankine-Hugoniot condition

$$u(x, t) = f(u_-) - f(u_+) =: [f]_{u_-.u_+}.$$  

(8)

We stress that the entropy solution $u(x, t)$ has only downward jumps in $x$. This will play a key role in our analysis.

1.2 Exact solutions for Burgers turbulence

Let us now consider random initial data. The solution immediately develops shocks that are separated by rarefaction waves (see for example, the computations in [32]). The shocks move with speeds given by the Rankine-Hugoniot condition, and cluster when they meet. Our goal is to understand this process. In order to explain the context and genesis of our results, we first focus on two exact solutions for Burgers equation.

Lévy process initial data on a half-line

Here we assume

$$u_0(x) = \begin{cases} 
0, & x \leq 0, \\
X_x, & x > 0,
\end{cases}$$  

(9)

where $X$ is a Lévy processes with only downward jumps (a spectrally negative Lévy process). A particularly interesting case is when $X$ is a standard Brownian motion (that is, a Brownian motion with $\mathbb{E}(X_x^2) = x$). This problem was solved formally by Carraro and Duchon [9, 10] and rigorously by Bertoin [5]. The key to their solution is a closure property of (6): if $u_0(x), x > 0$ is a spectrally negative Lévy processes , then so is $u(x, t) - u(0, t), x > 0$. These processes are characterized by their Laplace exponent $\psi(q,t)$ defined by

$$\mathbb{E} \left( e^{q(u(x,t)-u(0,t))} \right) = e^{x\psi(q,t)}, \quad x, q, t > 0.$$  

(10)

Thus, the problem is reduced to determining the evolution of $\psi$. Remarkably, $\psi(q,t)$ satisfies Burgers equation!

$$\partial_t \psi + \psi \partial_q \psi = 0, \quad \psi(q, 0) = \psi_0(q).$$  

(11)

If $X$ is a Brownian motion, $\psi_0(q) = q^2$, and we obtain the self-similar solution

$$\psi(q, t) = \frac{1}{t^2} \psi_*(qt), \quad \psi_*(q) = q + \frac{1}{2} - \sqrt{q + \frac{1}{4}}.$$  

(12)

Various explicit formulas are summarized in [30].

The interpretation of (11) in terms of shock clustering is as follows. $\psi$ satisfies the celebrated Lévy-Khintchine formula

$$\psi(q,t) = \int_0^\infty (e^{-qs} - 1 + qs) \Lambda(s) ds, \quad q \geq 0, t > 0,$$  

(13)

V
where $\Lambda_t(ds)$ is the jump measure at time $t$. The evolution of $\psi$ by (11) also induces evolution of the jump measure $\Lambda_t$. But the jumps in the process $u(\cdot, t)$ are the shocks, which evolve in a simple manner: shocks move at speed given by the Rankine-Hugoniot condition (8) and stick, conserving momentum, when they meet. Equation (11) captures this process. It is equivalent to the fact that $\Lambda_t$ satisfies a kinetic equation, Smoluchowski’s coagulation equation with additive kernel, which describes the evolution and coalescence of shocks. A derivation of the kinetic equation from this perspective may be found in [30]. An excellent survey of several links between stochastic coalescence and Burgers turbulence is [7].

White noise initial data

Here we must characterize the law of $u(x, t)$ when the initial velocity is white noise. Precisely, let us suppose that the initial potential $U_0(x) = \sigma B_x$ where $B$ is a standard two-sided Brownian motion pinned at the origin and $\sigma$ a fixed scale parameter. This problem also arises in statistics, and it was in this context that Groeneboom first characterized the law of the process $u(x, t)$, $x \in \mathbb{R}$ [20]. He showed that for every $t > 0$, $u(x, t)$, $x \in \mathbb{R}$ is a Markov process with only downward jumps (a spectrally negative Markov process) that is stationary in $x$. He then computed the generator of this Markov process explicitly in terms of Airy functions. Here is a brief summary of his solution.

By (6) and the scaling invariance of Brownian motion, we see that

$$a_\sigma(x, t) \overset{\mathcal{L}}{=} (\sigma t)^{\frac{2}{3}} a_1 \left(x(\sigma t)^{-\frac{2}{3}}, 1\right), \quad u_\sigma(x, t) \overset{\mathcal{L}}{=} \sigma \frac{t}{3} u_1 \left(x(\sigma t)^{-\frac{2}{3}}, 1\right)$$

(14)

where $\overset{\mathcal{L}}{\equiv}$ denotes equality in law and the subscript $\sigma$ refers to the variance of $U_0$. It is simplest to state the formulas under the assumption that $\sigma^2 = 1/2$. Let $u_x$ denote the process $u_{2^{-1/2}}(x, 1)$. The generator of $u_x$ is an operator defined by its action on a test function $\varphi$ in its domain as follows:

$$\mathcal{A}_\varphi(y) = \lim_{x \downarrow 0} \frac{\mathbb{E}_y(\varphi(u_x)) - \varphi(y)}{x},$$

(15)

where $\mathbb{E}_y$ denotes the law of the process with $u_0 = y$. In our case, $\mathcal{A}$ is an integro-differential operator of the form

$$\mathcal{A}_\varphi(y) = \varphi'(y) + \int_{-\infty}^{y} (\varphi(z) - \varphi(y)) n_*(y, z) \, dz.$$  

(16)

Groeneboom showed that the jump density $n_*$ of the integral operator is

$$n_*(y, z) = \frac{J(z)}{J(y)} K(y-z), \quad y > z.$$  

(17)

Here $J$ and $K$ are positive functions defined on the line and positive half-line respectively, whose Laplace transforms

$$j(q) = \int_{-\infty}^{\infty} e^{-qy} J(y) \, dy, \quad k(q) = \int_{0}^{\infty} e^{-qy} K(y) \, dy,$$

(18)
are meromorphic functions on \( \mathbb{C} \) given by

\[
    j(q) = \frac{1}{\operatorname{Ai}(q)}, \quad k(q) = -2 \frac{q^2}{dq^2} \log \operatorname{Ai}(q).
\]

(19)

\( \operatorname{Ai} \) denotes the first Airy function as defined in [2, 10.4]. Our normalization of \( J \) and \( K \) differs from [20] in two aspects. First, the choice \( \sigma^2 = 1/2 \) helps avoid several factors of \( 2^{1/3} \) while stating the main formulas. Second, the above definition of \( K \) is suggestive of Dyson’s formula in the theory of inverse scattering.

We shall return to this solution at several points in this article. The one-point and two-point distribution functions can be computed once the generator is known. For example, \( p(y)dy = P(u_x \in (y,y+dy)) = J(y)J(-y)dy \). The distribution of the shock sizes and the two-point distribution is given in [16]. Tauberian arguments yield precise asymptotics of these distributions.

1.3 Lax equations

Despite the elegance of the solution procedure for Lévy process data, it does not apply to Burgers equation with broader classes of random initial data (e.g. white noise), or to the general scalar conservation law (1). There are two obstructions. First, there is no reason to expect that a shock connecting states \( y \) and \( z \) appears at a rate that depends only on \( y - z \). Second, the pathwise structure of Lévy processes is rigid – if \( u(\cdot, t) \) is a Lévy process, then all rarefaction waves must be straight lines with the same slope. But rarefactions are nonlinear if the flux function \( f(u) \) is not quadratic (this follows from (7)), so the entropy solution cannot preserve Lévy processes. Our goal in this article is to develop a kinetic theory that addresses the general problem.

We build on two main sources: the exact solution for Burgers equation with white noise initial data [16, 20], and an analysis of closure properties for Burgers turbulence [11, 17]. Groeneboom’s computation of the generator is a tour de force of hard analysis and is the basis for a detailed understanding of the shock structure [19]. He works directly with the variational principle (6) at \( t = 1/2 \) to compute the probability of Brownian excursions above a parabola. Girsanov’s theorem is crucial in the analysis. Most past work in the turbulence community, is also based on a similar point of view. This begins with Burgers’ analysis which was skilfully completed by Frachebourg and Martin [8, 16]. However, the focus on white noise initial data and the use of Girsanov’s theorem obscures to some extent an understanding of the dynamic process of shock clustering for general initial data (see for example, the concluding remarks in [6]).

The idea of seeking kinetic equations for shock evolution originates with Kida [24]. More recently, Frachebourg et al showed that the hierarchy of \( n \)-point functions for ballistic aggregation can be closed at the level of the 2-point function [17]. In closely related work, Chabanol and Duchon showed formally that if a weak solution to Burgers equation preserves the Markov property, then one can derive evolution equations for the generator of the Markov process [11]. Moreover, they showed that Groeneboom’s solution yields a self-similar solution to this evolution equation.
This is the starting point for our work. What is truly new here is a simplifying viewpoint: Rather than seek a specific exact solution, we look for a class of (natural) random processes that is preserved by the entropy solution to (1). Here it is the class of spectrally negative Markov processes. Our main results are:

- **a closure theorem**: Suppose $f$ is strictly convex and $C^1$. If $u_0$ is a spectrally negative Markov process, so is the entropy solution to (1).

- **kinetic equations**: We derive evolution equations for the generator of the Markov process. These are state-dependent coagulation equations that describe the clustering of shocks. Remarkably, these have the structure of a Lax pair. Moreover, the Lax equations admit the above exact solutions when $f(u) = u^2/2$ in a manner strongly reminiscent of complete (Liouville) integrability.

In order to explain the Lax equation, let us assume that $u(x,t)$ is a stationary, spectrally negative Feller process whose sample paths have bounded variation. As in the Lévy-Khintchine formula, such a process is characterized by its generator $A(t)$, which acts on test functions $\varphi$ via

$$A\varphi(y) = b(y,t)\varphi'(y) + \int_{-\infty}^{y} (\varphi(z) - \varphi(y)) n(y,dz,t). \quad (20)$$

Here and in what follows, our treatment of generators is quite formal. The basic notions needed may be gleaned from the introduction to [33].

We now introduce an operator $B$ associated to $A$ and the flux function $f$. The operator $B$ is defined by its action on test functions as follows

$$B\varphi(y) = -f'(y)b(y,t)\varphi'(y) - \int_{-\infty}^{y} \frac{f(y) - f(z)}{y-z} (\varphi(z) - \varphi(y)) n(y,dz,t). \quad (21)$$

Then the evolution of $A$ is given by the Lax equation

$$\partial_t A = [A,B] = AB - BA. \quad (22)$$

The equations in [11] are written differently. We find the form above particularly transparent.

We present four different derivations of (22): (1) as a compatibility condition for martingales in $x$ and $t$; (2) from kinetic theory as in [30]; (3) using BV calculus as in [15]; (4) using Hopf’s functional calculus as in [11]. The shortest (and most heuristic) of these is the first, and goes as follows.

The main observation is that we have a two-parameter random process, and formally $B$ may be viewed as the ‘generator’ in $t$. To see this, fix $x \in \mathbb{R}$ and consider the random process $u(x,t)$, $t > 0$. Then the multipliers $-f'(y)$ and $-(f(y) - f(z))/(y-z)$ in (21) simply correspond to the evolution (in $t$) of the drift and jumps. Indeed, if the path $u(x,t)$ is differentiable at $x$ with $u(x,t) = y$, $\partial_x u = b$, then $\partial_t u = -f'(y)b$ by (1). Similarly, if we have a shock connecting left
and right states $y$ and $z$, then the shock speed is given by (8). Thus, formally, one may consider $\mathcal{B}$ as a generator of the process $u(x, t)$, $t > 0$. Of course, this is only formal, because this process typically does not have the Markov property, and we cannot expect these multipliers to be positive in general. Nevertheless, let us persist with this analogy. The evolution of the one-point distributions is then given by Kolmogorov’s forward equations

$$\partial_x p = A^1 p, \quad \partial_t p = \mathcal{B}^1 p. \quad (23)$$

We now seek martingales in $x$ and $t$. To this end, fix $\alpha > 0$, $(x_0, t_0)$, and consider the processes $\varphi(u(x_0 + s, t_0))$ and $\varphi(u(x_0, t_0 + s))$ with $s \in [0, \alpha]$. These processes are (formally!) martingales if $\varphi$ solves Kolmogorov’s backward equations

$$\partial_x \varphi + A \varphi = 0, \quad \partial_t \varphi + B \varphi = 0, \quad (24)$$

in the domain $(x, t) \in [x_0, x_0 + \alpha] \times [t_0, t_0 + \alpha]$, and $y, z \in \mathbb{R}$. If the compatibility condition $\varphi_{xt} = \varphi_{tx}$ holds for a sufficiently rich class of functions $\varphi$, we find

$$\partial_t A - \partial_x B = [A, \mathcal{B}]. \quad (25)$$

In this form, the Lax equation is akin to zero curvature conditions in integrable systems. If the process is stationary (in $x$), $\partial_x \mathcal{B}$ vanishes and we obtain (22).

### 1.4 Kinetic theory

When we expand the commutator in (22), and separate the evolution of the drift $b$ and the jump measure $n$ we obtain a kinetic equation that describes shock clustering. The drift satisfies the differential equation

$$\partial_t b(y, t) = -f''(y) b^2(y, t). \quad (26)$$

Note that the drift does not depend on the jump measure. The jump density $n(y, z, t) dz = n(y, dz, t)$ satisfies the kinetic equation

$$\partial_t n(y, z, t) + \partial_y (nV_y(y, z, t)) + \partial_z (nV_z(y, z, t)) = Q(n, n) + \quad \quad \quad (27)$$

Here the velocities $V_y$ and $V_z$ in (27) are given by

$$V_y(y, z, t) = ([f]_{y,z} - f'(y)) b(y, t), \quad V_z(y, z, t) = ([f]_{y,z} - f'(z)) b(z, t) \quad (28)$$

and the collision kernel $Q$ is

$$Q(n, n)(y, z, t) = \int_{-\infty}^{y} ([f]_{y,w} - [f]_{w,z}) n(y, w, t)n(w, z, t) dw \quad (29)$$

$$- \int_{-\infty}^{y} ([f]_{y,z} - [f]_{z,w}) n(y, z, t)n(z, w, t) dw$$

$$- \int_{-\infty}^{y} ([f]_{y,w} - [f]_{y,z}) n(y, z, t)n(y, w, t) dw.$$
We have assumed for convenience that the jump measure has a density, but the above equations extend naturally to general jump measures.

In the next section we derive these equations from the perspective of kinetic theory. We consider single shocks and rarefaction waves as building blocks, and use this to derive a natural Boltzmann-like equation. This will yield the equations above. We then show that these kinetic equations are equivalent to the Lax equation (22). That calculation reflects the fact that operators of the form (20) formally constitute a Lie algebra.

1.5 The broader context of our work

We conclude this introduction by connecting our work with some other problems in mathematical physics and statistics.

1.5.1 Complete integrability, random matrices and Burgers turbulence

The techniques of integrable systems have played an important role in the recent analysis of some exactly solvable models in statistical physics. We present some analogies between our approach and these methods.

The solution to Burgers turbulence with spectrally negative Lévy process data (see §1.2) is obtained from (22) as follows. Suppose $u(x,t) - u(0,t)$, $x > 0$ is a spectrally negative Lévy process with bounded variation and mean zero. If we denote the jump measure $\Lambda_t$, then $b(t) = \int_0^\infty s \Lambda_t(ds)$, and the generators $A$ and $B$ take the form

$$ A(t)\phi(y) = \int_0^\infty (\phi(y - s) - \phi(y) + s\phi'(y)) \Lambda_t(ds), $$

$$ B(t)\phi(y) = -y A(t)\phi(y) + \frac{1}{2} \int_0^\infty (\phi(y - s) - \phi(y)) s \Lambda_t(ds). $$

In particular, when $\phi(y) = e^{qy}$, Re$\{q\} > 0$, we have

$$ A(t)e^{qy} = \psi(q,t)e^{qy}, \quad B(t)e^{qy} = -\left(y\psi(q,t) + \frac{1}{2} \partial_q \psi\right)e^{qy}. $$

We substitute in (22) to obtain (11).

This calculation is akin to inverse scattering and suggests completely integrability of (22) for Lévy process initial data. The transformation to the spectral variable $\psi(q,t)$ can be considered a change of variables to action-angle coordinates. Indeed, if we write (11) along characteristics we have the Hamiltonian system

$$ \frac{d\psi}{dt} = 0, \quad \frac{dq}{dt} = \psi. $$

Thus, we may consider $q \in (0, \infty)$ as the angles and $\psi$ as the actions, and the flow has been linearized. A more complete description of complete integrability must elucidate the Hamiltonian structure of the evolution of generators.
We next note that the solution (12) can be mapped to Wigner’s semicircle law in the theory of random matrices. Dyson observed that the eigenvalues of a standard matrix valued Brownian motion $M_t$ in the group of $n \times n$ symmetric, hermitian or symplectic matrices satisfy the stochastic differential equation \[ \frac{d\lambda_k}{\lambda_k - \lambda_j} + \beta dB_k, \quad 1 \leq k \leq n, \] (33)

where $B_k$, $1 \leq k \leq n$ are independent Brownian motions, and $\beta = 1, 2$ or 4 for the ensembles above. That is, the eigenvalues behave like repulsive unit charges on the line perturbed by independent white noise. The law of large numbers for this ensemble is as follows. As $n \to \infty$ the spectral measure of $n^{-1/2}M_t$ converges to Wigner’s semicircle law:

$$
\mu_t(dx) = \frac{1}{2\pi t} \sqrt{4t - x^2} \, dx, \quad |x| < 2\sqrt{t}.
$$
(34)

Voiculescu noticed that the Cauchy transform of $\mu_t$,

$$
g(z, t) = \int \frac{1}{z - x} \mu_t(dx), \quad z \in \mathbb{C} \setminus [-2\sqrt{t}, 2\sqrt{t}],
$$
(35)

solves Burgers equation with a simple pole as initial data [35]. That is,

$$
\partial_t g + g \partial_z g = 0, \quad g(z, 0) = \frac{1}{z}.
$$
(36)

More precisely, $g$ solves (36) in the slit plane $\mathbb{C} \setminus [-2\sqrt{t}, 2\sqrt{t}]$, and has the form

$$
g(z, t) = \frac{1}{\sqrt{t}} g_s \left( \frac{z}{\sqrt{t}} \right), \quad g_s(z) = \frac{1}{2} \left( z - \sqrt{z^2 - 4} \right), \quad |z| \geq 2.
$$
(37)

It now transpires that the self-similar solution (12) can be transformed to (37) by a simple change of variables.

$$
\frac{\psi^*_s(q)}{q} = g_s(z), \quad z = 2 + \frac{1}{q}, \quad \text{or} \quad \frac{g(t^{1/2}z, t)}{t^{1/2}} = \frac{\psi(t^{-1}q, t)}{t^{-1}q}.
$$
(38)

In his beautiful thesis [23], Kerov found a deeper interpretation of (36) based on the representation theory of the symmetric group. He introduced a Markov process for the growth of Young diagrams (Plancherel growth), and derived (36) in a mean-field limit. More general initial conditions may also be included in (36) and he showed that the evolution of $g$ by Burgers equation is equivalent to a kinetic equation for $\mu_t$ [23, Ch. 4.5]. Thus, the transformation (38) links $\mu_t$ to $\Lambda_t$, and Plancherel growth with Smoluchowski’s coagulation equation (see (13)). We do not have a deeper (i.e. stochastic process) explanation of this relation yet.

Airy functions and the Painlevé transcendents arise in the scaling limit of fluctuations from Wigner’s law at the edge of the spectrum. The fluctuations
are given by the celebrated Tracy-Widom distributions involving a solution to Painlevé-II [34]. We now point out that the function \( l = j'/j \) (\( j \) as in (18)) solves the Riccati equation \( dl/dq = -q + l^2 \) and is therefore an Airy solution to Painlevé-II [1, Ch. 7]. This is used to verify that Groeneboom’s solution satisfies the kinetic equation (22) in §6.

These analogies and the fact that the dynamics are described by a Lax pair, suggest that (22) is completely integrable for \( f(u) = u^2/2 \) and spectrally negative Markov process data. Completely integrable systems in 2 + 1 dimensions are rare, and it is of considerable interest to investigate this further.

1.5.2 Shell models of turbulence and their continuum limits

It is also of interest to consider (1) on the half-line \( x > 0 \) with random forcing at \( x = 0 \). This problem arises as a continuum limit of shell models of turbulence [28, 29]. Shell models are lattice equations of the form \( \dot{c}_k = J_{k-1} - J_k, \quad k = 1, 2, \ldots \). Here \( c_k \geq 0 \) models the energy in the \( k \)-th Fourier mode (shell), and \( J_k \) the flux from shell \( k \) to \( k + 1 \). \( J_k \) is expressed in terms of \( c_k \) and its nearest neighbors by a suitable constitutive relation. The main question is to understand how randomness spreads through the system under the assumption that \( J_0 \) is a prescribed random forcing (a stationary Feller process, for example). In particular, it is of interest to understand whether the system has a unique invariant measure. In the continuum limit, these questions may be treated by our approach. We note that the derivation of (25) does not depend on assumptions of stationarity and is independent of boundary conditions. Therefore, (25) holds in the domain \( x, t > 0 \) and must be augmented with a boundary condition on the line \( x = 0 \). In particular, as \( t \to \infty \) invariant measures are solutions to

\[
-\partial_s B = [B, A],
\]

with a boundary condition at \( x = 0 \) matching \( B \) and the generator of the forcing.

1.5.3 Applications to statistics

In statistics, (6) with \( U_0 \) a two-sided Brownian motion first arose in the following estimation problem [12]. Suppose \( X_1, \ldots, X_n \) are independent, identically distributed (iid) samples from a distribution with a smooth unimodal density \( \rho \) with mode \( m \) and finite variance. Let us consider a naive ‘binning’ strategy to estimate the mode \( m \). We fix a bin width \( w \) and count the number of samples \( N_n(s) = \# \{ X_k \in (s-w, s+w) \} \) in the bin centered at \( s \). We estimate the mode by \( m_n = \arg\max_s N_n(s) \). Chernoff observed that the fluctuations \( m_n - m \) are \( O(n^{-1/3}) \). Precisely, for suitable \( c(\rho, w) > 0 \), the rescaled random variables \( cn^{1/3}(m_n - m) \) converge in law to Chernoff’s distribution

\[
Z = \arg\max_s \{ U_0(s) - s^2 \}.
\]

The quadratic term \(-s^2\) arises from the Taylor expansion of \( \rho \) around \( m \) – the first order term vanishes since \( m \) is the maximum of \( \rho \). By the symmetry of Brownian motion, it is clear that \( Z \) has the same law as \( a(0, 1/2) \).
This is not an isolated result: ‘cube-root’ fluctuations appear naturally in a wide class of estimation problems [3]. Kim and Pollard proved functional limit theorems for several such estimators, with the law of the limit characterized by
\[
\arg^+ \max_{s \in \mathbb{R}^d} \{ U_0(s) - |s|^2 \}
\]
where \( U_0(s) \) is a continuous Gaussian process in \( \mathbb{R}^d \) pinned at the origin [25]. These correspond to solutions to Burgers equation in \( \mathbb{R}^d \) with random initial data \( U_0 \), but this connection has not been explored in the literature.

1.6 Outlook

Let us conclude by pointing out some significant shortcomings in our work. Much of this article relies on formal calculations. But these calculations are often interesting, and it seems more fruitful to present them in a transparent and suggestive manner, rather than as rigorous statements burdened by technicalities. The most significant gap is that we do not prove (22). This is because our closure theorem is not strong enough. We only establish that the entropy condition preserves the Markov property. In order to rigorously establish (22) it is necessary to prove that the entropy condition preserves Feller processes. This is an assertion of regularity, whereas we only establish measurability. This is closely tied to establishing a satisfactory well-posedness theory for (22). A suitable well-posedness theorem would also yield a probabilistic proof of the existence of a two-parameter family of self-similar solutions to (22). These solutions are generated by considering the flux functions \( f(u) = |u|^p/p, 1 < p < \infty \) and an initial potential that is an \( \alpha \)-stable spectrally negative Lévy process. It seems challenging to prove this analytically starting with (22).

Also, while our work yields a deeper understanding of Groeneboom’s solution, it does not, as yet, constitute an independent proof of his results. We have only been able to verify that his solution satisfies (22). We have been unable to derive it using (22) alone.

The rest of this article is organized as follows. We first derive (26) and (27) from the standpoint of kinetic theory in the next section. This is followed by the rigorous closure theorems. We then derive the Lax equations by BV calculus and by Hopf’s method. Finally, we consider Groeneboom’s self-similar solution for Burgers equation in Section 6.

2 Kinetic equations

2.1 Introduction

In this section, we assume the velocity \( u \) is a stationary Feller process in \( x \), and derive the kinetic equations of Section 1.4. We use the evolution of a single shock and rarefaction wave to derive a Boltzmann-like equation for the evolution of the density of shocks. We conclude this section by showing that the Lax equation (22) is equivalent to the kinetic equations (27)–(29). The main observation is that the space of operators of the form (20) is formally a Lie algebra.
2.2 Conservation of total number density

Let \( p(y,t) \) denote the stationary 1-point density, i.e. \( p(y,t)dy = P(u(x,t) \in (y, y + dy)) \), and \( F(y, z, t) \) denote the total number density, i.e. the expected number of jumps per unit length from states \( y \) to \( z \). Then

\[
F(y, z, t) = p(y, t)n(y, z, t). \tag{41}
\]

The total number density changes because of a flux of shocks and shock collisions, and we have the general conservation law of Boltzmann-type

\[
\partial_t F + \partial_y (F V_y) + \partial_v (F V_z) = C(F, F). \tag{42}
\]

Here \( V_y \) and \( V_z \) denote ‘velocities’ in the \((y, z)\) ‘phase space’, and \( C \) is a binary collision kernel. This is the general structure of the equation. We now derive the evolution equation for \( b \), the velocities \( V_y, V_z \), and the collision kernel \( C \), based on elementary solutions to the scalar conservation law (1).

2.3 Decay of the drift

First consider how affine data evolves under (1). Let \( u(x, t) \) solve (1) with \( u_0(x) = \alpha_0 + \beta_0x \). For \( x, t \approx 0 \), we have to leading order \( u(x, t) \approx \alpha(t) + \beta(t)x \), so that

\[
\partial_t u \approx \dot{\alpha} + \dot{\beta}x, \quad \partial_x f(u) \approx f'(\alpha)\beta + f''(\alpha)\beta^2 x. \tag{43}
\]

We now balance terms in (1) to obtain

\[
\dot{\alpha} = -f'(\alpha)\beta, \quad \dot{\beta} = -f''(\alpha)\beta^2. \tag{44}
\]

The second equation expresses the decay of rarefaction waves when \( \beta_0 > 0 \). The connection between this elementary solution and the generator of the process is the following. The drift coefficient \( b(y, t) \) corresponds locally to an affine profile as above with \( \alpha = y \) and \( \beta = b \). Thus, the second equation above is simply (26).

2.4 Decay of shocks and \( V_y, V_z \)

The ‘velocities’ \( V_y \) and \( V_z \) in \((y, z)\) space arise because of the decay of shocks. In order to derive these velocities, we fix \( u_- > u_+ \) and consider piecewise affine initial data

\[
u_0(x) = \begin{cases} 
            u_- + b_-x, & x < 0, \\
            u_+ + b_+x, & x > 0.
          \end{cases} \tag{45}
\]

Let \( s(t) \) denote the path of the shock. Then by (8), for small \( t \)

\[
s(t) = [f]_{u_-, u_+}(t + o(t)). \tag{46}
\]

\( s \) is also given by the kinematic condition

\[
s(t) = a_\pm(t) + f'(u_\pm)t + o(t). \tag{47}
\]
where \( a_\pm \) denotes the left and right inverse Lagrangian points at time \( t \). Thus,

\[
a_\pm(t) = ([f]_{u_-,u_+} - f'(u_\pm)) t + o(t). \tag{48}
\]

Since \( u_- > u_+ \) and \( f \) is convex, we see that \([f]_{u_-,u_+} - f'(u_-) < 0\). Similarly, \([f]_{u_-,u_+} - f'(u_+) > 0\). As a consequence, a shock initially connecting states \( u_\pm \) decays to a shock connecting states

\[
u_\pm + ([f]_{u_-,u_+} - f'(u_\pm)) b_\pm t + o(t). \tag{49}
\]

This decay gives rise to a flux of \( F \). To first approximation, the flux is linear in \( F \), and of the form (42) with the drift velocities given by (28).

### 2.5 The collision kernel \( C(F,F) \)

Binary collisions of shocks occur at a rate determined by the Markov property (in \( x \)) and the relative velocity of shocks given by (8). In order to simplify notation, we suppress the \( t \)-dependence for the process \( u \) and write \( u_x \) for \( u(x,t) \). We also denote a shock connecting states \( u_- \) and \( u_+ \) by \( \{u_-,u_+\} \). Shock clustering involves the following events:

- **growth**: \( \{y,w\} + \{z,w\} = \{y,z\}, \quad z < w < y \) \( \tag{50} \)
- **decay**: \( \{y,z\} + \{z,w\} = \{y,w\}, \quad -\infty < w < z \) \( \tag{51} \)
- **decay**: \( \{w,y\} + \{y,z\} = \{w,z\}, \quad y < w < \infty \). \( \tag{52} \)

The computation of rates for these events is similar. To be concrete, let us first consider (50). Fix \( z < w < y \) and consider small \( \Delta x_1 > 0, \Delta x_2 > 0 \). Then formally, by the Markov property

\[
dP(u_0 = y, u_{\Delta x_1} = z) \approx p(y)n(y,z) dy dz \Delta x_1, \tag{53}
\]

and similarly,

\[
dP(u_0 = y, u_{\Delta x_1} = w, u_{\Delta x_1 + \Delta x_2} = z) \approx p(y)n(y,w)n(w,z) dy dz dw \Delta x_1 \Delta x_2. \tag{54}
\]

The relative velocity of these shocks is \([f]_{y,w} - [f]_{w,z}\) to leading order. We thus set \( \Delta x_2 = ([f]_{y,w} - [f]_{w,z}) \Delta t \) to compute the number of collisions in time \( \Delta t \). We now sum over \( w \) in the range \( z < w < y \) to obtain the growth term in \( C(F,F) \):

\[
C_1 := \int_{z}^{y} p(y)n(y,w)n(w,z) ([f]_{y,w} - [f]_{w,z}) dw. \tag{55}
\]

The computation for the events (51) is similar. We now find the decay terms

\[
C_2 := -\int_{-\infty}^{z} p(y)n(y,z)n(z,w) ([f]_{y,z} - [f]_{z,w}) dw, \tag{56}
\]

and

\[
C_3 := -\int_{y}^{\infty} p(w)n(w,y)n(y,z) ([f]_{w,y} - [f]_{y,z}) dw. \tag{57}
\]
2.6 Kinetic equations for \( n \)

We have now defined all the terms in (42). In order to obtain an equation in terms of \( n \) alone we use Kolmogorov’s forward equations to eliminate the 1-point distribution \( p(y, t) \). The first equation in (23) is now \( 0 = \mathcal{A}^tp \) since the process is stationary in \( x \). The second equation in (23) implies

\[
p(y)\partial_t n(y, z) = \partial_t F(y, z) - n(y, z)\mathcal{B}^t p(y),
\]

where

\[
\mathcal{B}^t p(y) = \partial_y \left( f'(y)b(y)p(y) \right) - \int_{\mathbb{R}} (p(w)n(w, y) - p(y)n(y, w)) [f]_{y, w} \, dw.
\]

(59)

(It is convenient to denote the domain of integration by \( \mathbb{R} \), noting that it is actually a half-line because \( n(y, z) = 0 \) for \( y < z \).)

To isolate the main cancellations on the right hand side of (58), we note that the integral term in \( C_3 - n(y, z)\mathcal{B}^t p(y) \) is

\[
n(y, z) \left( [f]_{y, z} \int_{\mathbb{R}} p(w)n(w, y) \, dw - p(y) \int_{\mathbb{R}} n(y, w)[f]_{y, w} \, dw \right).
\]

(60)

The first integral above can be simplified further. Since \( \mathcal{A}^t p = 0 \), we also have

\[
\int_{\mathbb{R}} p(w)n(w, y) \, dw = p(y) \int_{\mathbb{R}} n(y, w) \, dw + \partial_y \left( b(y)p(y) \right).
\]

(61)

Therefore, we may rewrite the expression in (60) as

\[
p(y)n(y, z) \int_{-\infty}^{y} ([f]_{y, z} - [f]_{y, w}) n(y, w) \, dw + n(y, z)[f]_{y, z} \partial_y \left( b(y)p(y) \right).
\]

(62)

We now collect all terms on the right hand side of (58) using (42), (28), (55), (56), and (62). We then have

\[
p(y)\partial_t n(y, z) = p(y)Q(n, n)(y, z)
\]

\[
+ n(y, z) \left( [f]_{y, z} \partial_y \left( b(y)p(y) \right) - \partial_y \left( f'(y)b(y)p(y) \right) \right)
\]

\[
- \partial_y \left( F(y, z)V_y(y, z) \right) - \partial_z \left( F(y, z)V_z(y, z) \right),
\]

(63)

\[
= p(y) \left( n \left( [f]_{y, z} - f'(y) \right) \partial_y b - bf''(y) \right) - \partial_y \left( nV_y(y, z) \right) - \partial_z \left( nV_z(y, z) \right).
\]

(64)

(65)

Under the assumption that the Markov process \( u \) has a strictly positive stationary density, we may cancel \( p(y) \) on both sides of (62). We then have the kinetic equations (27).
2.7 Equivalence of the Lax equations and kinetic equations

The equivalence of the Lax equations (22) and the kinetic equations (27)–(29) follows from the algebraic structure of operators of the form (20) and (21). The space of operators of the form

\[ A \varphi(y) = b(y)\varphi(y) + \int_{\mathbb{R}} n(y, z) (\varphi(z) - \varphi(y)) \, dz, \quad (67) \]

with smooth \( b, n \) and the bracket \([\cdot, \cdot]\) is formally a Lie algebra. That is, if \( A_i, i = 1, 2, 3 \) are operators as above, then \([A_1, A_2]\) is an operator of the same form, and the following Jacobi identity holds:

\[
[[A_1, A_2], A_3] + [[A_2, A_3], A_1] + [[A_3, A_1], A_2] = 0. \quad (68)
\]

We do not assume that \( b \) and \( n \) are positive or that their support is a half-line. Both assertions rely on tedious, but direct calculations. We will omit the proof of (68), and simply summarize the calculation for \([A_1, A_2]\). We have

\[
[A_1, A_2] \varphi(y) = (b_1 b_2' - b_1' b_2) \varphi(y) + \int_{\mathbb{R}} (\nu + \sigma) (\varphi(z) - \varphi(y)) \, dz, \quad (69)
\]

with

\[
\nu(y, z) = b_1(y) \partial_y n_2 - b_2(y) \partial_y n_1 + \partial_z (b_1(z) n_2 - b_2(z) n_1), \quad (70)
\]

and

\[
\sigma(y, z) = \int_{\mathbb{R}} (n_1(y, w) n_2(y, z) - n_2(y, w) n_1(y, z)) \, dw \quad (71)
\]

\[
+ \int_{\mathbb{R}} n_2(y, z) n_1(z, w) - n_1(y, z) n_2(z, w) \, dw
\]

\[
+ \int_{\mathbb{R}} n_1(y, z) n_2(y, w) - n_2(y, z) n_1(y, w) \, dw.
\]

We apply (69) to (22) as follows. Let \( A_1 = A \) and \( A_2 = B \), with \( A \) and \( B \) as in (20) and (21). We then find immediately from (70) that the drift coefficient of \([A, B]\) is

\[
-b(y) (f'(y) b(y))' + f'(y) b(y) b'(y) = -f''(y) b^2(y) \quad (72)
\]

as in (26). Similarly, we use (70) to obtain the second, third and last term in (27). Finally, we substitute the jump measures in (71) to obtain the collision kernel \( Q(n, n) \).

3 Closure theorems

3.1 Introduction

In this section we show that the entropy solution to (1) preserves the class of spectrally negative Markov processes. The Markov property of \( u \) was implicitly
used by Burgers, and made explicit in [4]. Our work is based on previous results by Bertoin [5] and Winkel [38]. We extend these results to general convex fluxes \( f \) and a large class of noise initial data following their methods.

### 3.2 Splitting times

The main technical tool we need is the decomposition of Markov processes at random times so that the past and future are conditionally independent. This certainly holds at a stopping time. However, it also holds for a broader class of **splitting times**. We use the following results. The first is a theorem of Getoor on the preservation of the Markov property at a last-passage time.

**Theorem 1** (Getoor [18]). Let \( M \subset \mathbb{R} \) be a fixed set and let \( L = \sup \{ x \in \mathbb{R} : X_x \in M \} \) be the end of \( M \). Then given a càdlàg strong Markov process \( X_x \), the post-\( L \) process \( \{ X_x \}_{x \geq L} \) is independent of \( \{ X_x \}_{x < L} \) given \( X_L \).

The second is a theorem of Millar on splitting at the ultimate minimum of a functional of the process. It is a special case of Theorem 1.

**Theorem 2** (Millar [31]). Let \( X_s \) be a càdlàg strong Markov process and \( g : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \) continuous. Let \( L \) be the last time the process \( g(X_{s-}, X_s) \) hits its ultimate minimum. Then \( \{ X_s \}_{s \geq L} \) is conditionally independent of \( \{ X_s \}_{s < L} \) given \( X_L \) and the auxiliary variable \( m = \inf_{s \leq L} g(X_{s-}, X_s) = g(X_{L-}, X_L) \).

In both theorems, the transition semigroup of the pre- and post-\( L \) processes can be explicitly determined from that of \( X \). It is an interesting and subtle fact that these transition semigroups are typically different. Both theorems hold in greater generality than presented here.

### 3.3 Closure for spectrally negative initial velocity

We work within the canonical framework. Let \( u_0 = (\Omega, \mathcal{F}, \mathcal{F}_s, u_0(s), \mu_0) \) be a spectrally negative Markov process on the space \( \Omega \) of càdlàg paths on \(( -\infty, \infty)\) endowed with the Skorohod topology. Here, \( \mu_0 \) is the law of the coordinate process \( u_0(\omega; s) = \omega(s) \) on \( \Omega \) and \( \mathcal{F}_s = \sigma(u_0(r), r \leq s) \) is the natural filtration of \( u_0 \) extended to be right-continuous and complete. Note that in our work \( x \) plays the role of “time” as traditionally used in the theory of stochastic processes. \( t \) acts as a parameter in the following discussion, and will be dropped from the notation when convenient.

We rewrite the Hopf-Lax functional and inverse Lagrangian function as

\[
I(s; x, t) = \int_0^t \left( u_0(r) - (f')^{-1} \left( \frac{x - t'}{t} \right) \right) dr \tag{73}
\]

\[
a(x, t) = \arg^+ \min_{s \in \mathbb{R}} I(s; x, t). \tag{74}
\]

To ensure that (74) is well-defined and finite, we assume (4) holds a.s. This is certainly true if \( u_0 \) is a stationary random process and \( f^* \) grows fast enough.
at infinity. Since \( u(x,t) \) and \( a(x,t) \) are related by (7), in order to obtain the closure property we only need to show that \( a(x,t) \) is a Markov process. We prove the following:

**Theorem 3.** Let \( u_0 \) be a spectrally negative strong Markov process such that (4) holds a.s. Under the law \( \mu_0 \) and for any fixed \( t > 0 \), the inverse Lagrangian \( a(x,t) \) is a Markov process. Thus, \( u(x,t) \) is a spectrally negative Markov process.

**Proof.** Fix \( x \in \mathbb{R} \). Without loss of generality it suffices to take \( t = 1 \). We also suppress the dependence on \( t \) in the notation for clarity. The proof consists of three steps, the first two of which are entirely deterministic.

1. **Dependence structure of \( (a(x))_{x \in \mathbb{R}} \):** Since \( a(x) \) is increasing, for \( h > 0 \)

   \[
   a(x + h) = \arg^+ \min_{s \in \mathbb{R}} I(s; x + h)
   = \arg^+ \min_{s \geq a(x)} \{ I(a(x); x + h) + (I(s; x + h) - I(a(x); x + h)) \}
   = a(x) + \arg^+ \min_{s \geq 0} \left\{ \int_a^{a(x) + h} \left( u_0(r) - \left( f' \right)^{-1}(x - r) \right) dr \right\}.
   \]

   \( (a(x + h) - a(x))_{h > 0} \) therefore depends on \( u_0 \) only through \( (u_0(s))_{s > a(x)} \). The same argument shows that \( (a(x + h) - a(x))_{h > 0} \) depends only on \( (u_0(s))_{s < a(x)} \).

2. **Downward jumps determine \( u_0(a(x)) \):** If \( u_0 \) has only downward jumps then \( \partial_s I(a(x); x) = 0 \). That is,

   \[
   u_0(a(x)) = f'^{-1}(x - a(x)).
   \]

   This may be seen easily by sketching a picture. Upward jumps in \( u_0 \) give rise to a potential with a corner that is convex, and downward jumps give a potential with a corner that is concave. In the second case, the minimum of \( I \) can never be achieved at the corner. Thus, it is always obtained at a point of continuity, implying (76).

3. **\( a(x) \) is a splitting time for \( u_0 \):** It is not obvious that \( a(x) \) is a Markov time for \( a_0 \). It is not a stopping time since the event \( \{ u(x) \leq s \} \) is not \( \mathcal{F}_s \)-measurable. Indeed, \( a(x) \) is the last time (73) is minimized and depends on \( (u_0(r))_{r > s} \) as well. \( a(x) \) is, however, a splitting time for a certain functional of \( u_0 \). To see this, first define

   \[
   m(s; x) = \min_{r \leq s} I(r; x), \quad D(s; x) = (I - m)(s; x).
   \]

   Dropping the explicit dependence on \( x \) from the notation, consider the bivariate process \( (u_0, D) \). Since \( D(s) \) is entirely dependent on \( (u_0(r))_{r \leq s} \), \( u_0, D(s) \) is \( \mathcal{F}_s \)-measurable. For any \( \mathcal{F} \)-stopping time \( \xi \)

   \[
   m(\xi + s) = m(\xi) \wedge \left( \min_{\xi < r \leq \xi + s} I(r) \right) = I(\xi) + \left\{ -D(\xi) \wedge \min_{0 < r \leq s} (I(\xi + r) - I(\xi)) \right\}
   \]
and

\[ D(\xi + s) = (I(\xi + s) - I(\xi)) - \left\{ -D(\xi) \wedge \min_{0 \leq r \leq s} (I(\xi + r) - I(\xi)) \right\}. \]

For \( s > 0 \) the increment \( I(\xi + s) - I(\xi) \) depends on \( \mathcal{F}_\xi \) only through \( u_0(\xi) \) by the strong Markov property of \( u_0 \). Therefore, \((u_0, D)\) is also a càdlàg, strong Markov process.

We now show that \( a(x) \) is a splitting time for \((u_0, D)\). By definition (74) and step 2,

\[ a(x) = \sup \left\{ s \in \mathbb{R} : (u_0, D)(s) = ((f')^{-1}(x - s), 0) \right\}. \]

Thus, \( a(x) \) is the last time \((u_0, D)\) hits the fixed set \( \{(f')^{-1}(x - s), 0) : s \in \mathbb{R}\} \) and we can use Theorem 1 to split \( u_0 \) at \( a(x) \)

\[ (u_0, D)_{s < a(x)} \text{ and } (u_0, D)_{s > a(x)}, \]

and consequently,

\[ (u_0(s))_{s < a(x)} \text{ and } (u_0(s))_{s > a(x)} \]

are conditionally independent given \((f')^{-1}(x - a(x))\) —that is, given \( a(x) \).

This shows that \( a \) is a Markov process (note that the law of the increments of the process may vary with \( x \)). Therefore, \( u(x) = (f')^{-1}(x - a(x)) \) is a spectrally negative Markov process.

\textbf{Remark 1.} Theorem 3 reduces to Bertoin’s closure theorem [5, Thm.2] if \( f(u) = u^2/2 \) (Burgers flux) and \( u_0 \) is a spectrally negative Lévy process. To see this, use (76) in (75):

\[ a(x + h) - a(x) = \arg^+ \min_{s \geq 0} \left\{ \int_{a(x)}^{a(x)+s} (u_0(r) - u_0(a(x)) + r - a(x)) \, dr \right\} \]

\[ = \arg^+ \min_{s \geq 0} \left\{ \int_{0}^{s} (u_0(a(x)) + r - u_0(a(x)) + r) \, dr \right\}. \]

Since \( u_0 \) has independent increments, the integrand above is independent of \((u_0(s))_{s < a(x)}\) and has the same law as \( u_0(r) - u_0(0) + r \). So, the law of \( a(x + h) - a(x) \) is independent of \( a(x - h) - a(x) \) and does not vary with \( x \). Note that it is necessary to use \( f(u) = u^2/2 \) to obtain this result, and to show that the increments of \( a(x) \) are identical in law to those of the first hitting process \( T(x) = \inf \{ s \geq 0 : u_0(s) + s \geq x \} \).

### 3.4 Closure for noise initial data

We now extend Theorem 3 to noise initial data. Here we prescribe the law of the potential as follows. Let

\[ U_0(x) = \begin{cases} X_x & x \geq 0 \\ -\tilde{X}_{-x} & x < 0 \end{cases}, \quad (77) \]
where $X, \tilde{X}$ are independent copies of a spectrally negative additive process starting at 0. If $X$ also has stationary increments it is a Lévy process, but it is not necessary to assume this. In addition, assume the growth condition (4) almost surely.

**Theorem 4.** Suppose $U_0 = (\Omega, \mathcal{G}, \mathcal{G}_s, U_0(s), M_0)$ is a two-sided spectrally negative process with independent increments and satisfies (4) a.s. Then under $M_0$ and for all $t > 0$, $u(x, t)$ is a spectrally negative Markov process.

**Proof.** Again we fix $x \in \mathbb{R}$, take $t = 1$, and drop the $t$-dependence in our notation.

1. **Dependence structure of $(a(x))_{x \in \mathbb{R}}$:** As before, for $h > 0$

$$a(x + h) = \arg^+ \min_{s \in \mathbb{R}} I(s; x + h)$$

$$= a(x) + \arg^+ \min_{s \geq 0} \{ I(a(x) + s; x + h) - I(a(x); x + h) \}$$

$$= a(x) + \arg^+ \min_{s \geq 0} \{ U_0(a(x) + s) - U_0(a(x)) + f^*(x + h - (a(x) + s)) - f^*(x + h - a(x)) \} \quad (78)$$

Given $a(x)$, the increment $a(x + h) - a(x)$ therefore depends on $U_0$ only through the increment $U_0(a(x) + s) - U_0(a(x))$.

2. **Spectral negativity of $U_0$:** If $U_0$ is spectrally negative, $U_0(s) \leq U_0(s^-)$. In particular

$$U_0(a(x)^- \wedge U_0(a(x)) = U_0(a(x)).$$

Without spectral negativity, $U_0(a(x)^-) \wedge U_0(a(x))$ cannot be reduced to any simpler form.

3. **Markov property of $U_0$ at $a(x)$:** We use Theorem 2. Consider the functional

$$g(s) = (U_0(s^-) + f^*(x - s)) \wedge (U_0(s) + f^*(x - s))$$

$$= (U_0(s^-) \wedge U_0(s)) + f^*(x - s).$$

Then the inverse Lagrangian satisfies

$$a(x) = \sup \left\{ s \in \mathbb{R} : g(s) = \min_{r \in \mathbb{R}} g(r) \right\}.$$ 

Since $U_0$ is strong Markov, Theorem 2 implies that $(U_0(s))_{s \geq a(x)}$ is independent of $(U_0(s))_{s < a(x)}$ given $(U_0(a(x)), g(a(x)))$ —that is, given

$$(U_0(a(x)), \{U_0(a(x)^-) \wedge U_0(a(x))\} + f^*(x - a(x))) \quad (79)$$

Now we use the spectral negativity of $U_0$. Since $U_0(a(x)^- \wedge U_0(a(x)) = U_0(a(x))$, the information in (79) is equivalent to that in $(U_0(a(x)), a(x))$. Note that if we did not have spectral negativity then we would have to be given $g(a(x))$. 

19
We combine steps 1 and 3, to see that the increment $a(x+h) - a(x)$ is independent of $(a(x-h) - a(x))_{h \geq 0}$ given $(U_0(a(x)), a(x))$. Since $a(x+h) - a(x)$ only depends on $U_0$ through the increment $U_0(a(x) + s) - U_0(a(x))$, it is independent of $U_0(a(x))$ as well. Therefore, $a(x)$ is a Markov process and $u(x)$ is a spectrally negative Markov process. Notice that if $u_0$ was not spectrally negative or did not have independent increments, then $a(x+h) - a(x)$ would depend on $g(a(x))$ or $U_0(a(x))$ in addition to $a(x)$, destroying the Markov property.

**Remark 2.** For symmetric fluxes $f(u) = f(-u)$, the condition of spectral negativity of $U_0$ is somewhat artificial since it is entirely dependent on the continuity properties of the potential. If we switch from càdlàg paths to those that are left-continuous with right limits, then the closure argument also holds true for $U_0$ spectrally positive. Nevertheless, given a fixed coordinate framework we are inevitably led to a sign condition on the jumps of the process $U_0$.

## 4 BV calculus and the Lax equations

In this section, we derive the Lax equation (22) using the conservation law (1) and the Vol’pert chain rule for BV functions. This calculation is similar in spirit to [15]. We assume that for every $t > 0$, $u(x,t)$ is a spectrally negative, stationary Feller process. The generator of $u$ is given by (20) and the operator $B$ is defined by (21). In addition, we define one and two-point operators as follows. We associate a linear functional $P$ to the stationary one-point distribution $p(y,t)$ via

$$P(\varphi) = E(\varphi(u(x,t))) = \int_{\mathbb{R}} \varphi(y) p(y) dy.$$  \hspace{1cm} (80)

(Here and in what follows, it is convenient to suppress $t$ in the notation.) Similarly, given $h > 0$ we denote the transition kernel for the process $u$ by $q_h$ and define the associated transition operator

$$(Q_h \varphi)(y) = \int_{\mathbb{R}} q_h(y,z) \varphi(z) dz.$$ \hspace{1cm} (81)

### 4.1 The 1-point function

We recall that the entropy solutions to (1) are in $BV_{loc}(\mathbb{R}_+ \times \mathbb{R})$. Vol’pert showed that one may extend the chain rule to such functions in a natural manner. Let $u_\pm = u(x_\pm, t)$ denote the right and left limits of $u(x,t)$ and for any test function $\varphi$, consider the composition $\varphi(u(x,t))$ and set

$$[\varphi] = \int_{0}^{1} \varphi(u_- + \beta(u_+ - u_-)) d\beta.$$ \hspace{1cm} (82)

Then for every smooth test function with compact support, the entropy solution to (1) satisfies [36, p.248]

$$\partial_t \varphi(u(x,t)) = -[\varphi][f] \partial_x u(x,t), \quad x \in \mathbb{R}, t > 0.$$ \hspace{1cm} (83)
The law of the 1-point function is determined by \( \mathbb{E}(\varphi(u(x,t))) \) for arbitrary \( \varphi \).

We use equation (83) to obtain

\[
\partial_t \mathbb{E}(\varphi(u(x,t))) = -\mathbb{E}(\langle \varphi[f] \partial_x u(x,t) \rangle). \tag{84}
\]

The left hand side of (92) is \((\partial_t P) \varphi\). We must determine the right hand side.

To this end, fix \( x \in \mathbb{R} \) and for \( h > 0 \) let us denote \( u_{-h} = u(x-h,t) \). Our calculation relies on the following unjustified interchange of limits:

\[
\mathbb{E}(\langle \varphi[f] \partial_x u(x,t) \rangle) = \lim_{h \to 0} \mathbb{E}\left( \frac{1}{h} (\varphi(u_+) - \varphi(u_{-h})) \frac{(f(u_+) - f(u_{-h}))}{u_+ - u_{-h}} \right). \tag{85}
\]

We assume (85), and show that

\[
\lim_{h \to 0} \mathbb{E}\left( \frac{1}{h} (\varphi(u_+) - \varphi(u_{-h})) \frac{(f(u_+) - f(u_{-h}))}{u_+ - u_{-h}} \right) = -\mathcal{P}(B \varphi), \tag{86}
\]

As a consequence,

\[
(\partial_t P) \varphi = \mathcal{P}B \varphi, \tag{87}
\]

which is the formal forward equation (23).

We now establish (86) for a monomial \( f(u) = u^n \). Since \( B \) depends linearly on \( f \), (86) then also holds for polynomials and by approximation for \( C^1 \) fluxes with polynomial growth. For \( f(u) = u^n \) we expand the jump term

\[
\frac{f(u_+) - f(u_{-h})}{u_+ - u_{-h}} = \sum_{j=0}^{m-1} u_{-h}^{m-1-j} u_+^j. \tag{88}
\]

Therefore, the limit in (86) is a sum of terms of the form

\[
(\varphi(u_+) - \varphi(u_{-h})) u_{-h}^{m-1-j} u_+^j = u_{-h}^{m-1-j} \left( \varphi(u_+ u_+^j - \varphi(u_{-h}) u_+^j - \varphi(u_{-h}) (u_+^j - u_{-h}^j) \right).
\]

The limit of each of these terms is

\[
\lim_{h \to 0} \frac{1}{h} \mathbb{E}(u_{-h}^{m-1-j} u_+^j) = \lim_{h \to 0} \frac{1}{h} \mathcal{P}(y^{m-1-j} (Q_h(y^j \varphi) - y^j \varphi - \varphi(Q_h(y^j) - y^j)))
\]

\[= \mathcal{P}(y^{m-1-j} (A(y^j \varphi) - \varphi A(y^j))). \tag{89}\]

The term \( A(y^j \varphi) - \varphi A(y^j) \) may be computed using (20). First the drift term is

\[y^{m-1-j}(b(y^j \varphi)' - \varphi b(y^j)'') = by^{m-1} \varphi'. \tag{90}\]

Similarly, the jump term simplifies to

\[y^{m-1-j} \int_{\mathbb{R}} n(y,z) z^j (\varphi(z) - \varphi(y)) \, dz. \tag{91}\]
We now sum over all terms in the expansion (88) to obtain

\[
\lim_{h \to 0} \mathbb{E} \left( \frac{1}{h} (\varphi(u_+) - \varphi(u_-)) \frac{(f(u_+) - f(u_-))}{u_+ - u_-} \right) = \sum_{j=0}^{m-1} \mathcal{P} \left( y^{m-1-j} (A(y^j \varphi) - \varphi A y^j) \right)
\]

\[
= \mathcal{P} \left( \sum_{j=0}^{m-1} b y^{m-1-j} \varphi' + \int_{\mathbb{R}} n(y,z) y^{m-1-j} z^j (\varphi(z) - \varphi(y)) \, dz \right)
\]

\[
= \mathcal{P} \left( b m y^{m-1} \varphi' + \int_{\mathbb{R}} n(y,z) z^m - y^m z - y (\varphi(z) - \varphi(y)) \, dz \right) = -\mathcal{P} \mathcal{B} \varphi.
\]

### 4.2 The 2-point function

Fix \( x \in \mathbb{R} \) and \( \alpha > 0 \). Let \( \varphi \) and \( \psi \) be two test functions. The law of the 2-point function is described completely by

\[
\mathbb{E} (\varphi(u(x,t)) \psi(u(x + \alpha,t))) = \mathcal{P} (\varphi \mathcal{Q}_\alpha \psi).
\]

The BV chain rule extends to a product rule as follows

\[
\partial_t (\varphi \mathcal{Q}_\alpha \psi) = \mathcal{Q}_\alpha \mathcal{B} \varphi + \varphi \mathcal{B} \mathcal{Q}_\alpha \psi,
\]

where

\[
\mathcal{B}(x,t) = \frac{1}{2} (\varphi(x,t_-) + \varphi(x,t_+)).
\]

We combine the BV chain rule and the conservation law to obtain

\[
\partial_t \varphi(u(x,t)) = -[\varphi][f] \partial_x u(x,t), \quad \partial_t \psi(u(x + \alpha,t)) = -[\psi][f] \partial_x u(x + \alpha,t).
\]

Therefore,

\[
\partial_t (\varphi(u(x,t)) \psi(x + \alpha,t)) = -\mathcal{B} [\varphi]' [f'] \partial_x u(x,t) - \mathcal{B} [\psi]' [f'] \partial_x u(x + \alpha,t).
\]

We compute the expected value of each of these terms in turn. First, as earlier, the main assumption is

\[
\mathbb{E} (\mathcal{B} [\varphi]' [f'] \partial_x u(x,t)) = \lim_{h \to 0} \frac{1}{h} \mathbb{E} \left( \frac{f(u_+) - f(u_-)}{u_+ - u_-} (\varphi(u_+) - \varphi(u_-)) \left( \frac{\psi_+ + \psi_-}{2} \right) \right),
\]

where \( \psi_+ = \psi(u(x + \alpha,t)) \) and \( \psi_- = \psi(u(x + \alpha - h,t)) \).

We can do away with the mean value. To be explicit, we write the above expectation as

\[
\int_{\mathbb{R}} \int_{\mathbb{R}} p(y) g_h(y,z) \left( \frac{f(z) - f(y)}{z - y} (\varphi(z) - \varphi(y)) \right) (\mathcal{Q}_{\alpha+h} + \mathcal{Q}_\alpha) \mathcal{B} \frac{\psi(z)}{2} \, dz \, dy.
\]
We see that the last term converges to \( Q_\alpha \psi(y) \) as \( h \to 0 \). Therefore,

\[
\mathbb{E}(\psi'[f'] \partial_x u(x, t)) = \lim_{h \to 0} \frac{1}{h} \mathbb{E} \left( \frac{f(u_+)}{u_+ - u_h} (\varphi(u_+) - \varphi(u_-)) Q_\alpha \psi(u_+) \right) \]

\[
= \lim_{h \to 0} \frac{1}{h} \mathbb{E} \left( \frac{f(u_+)}{u_+ - u_h} (\varphi Q_\alpha \psi(u_+) - \varphi Q_\alpha \psi(u_-)) - \varphi(u_-) \frac{f(u_+)}{u_+ - u_h} (Q_\alpha \psi(u_+) - Q_\alpha \psi(u_-)) \right),
\]

\[
= -\mathcal{P} (\mathcal{B} (\varphi Q_\alpha \psi) - \varphi Q_\alpha \mathcal{B} \psi),
\]  

(95)

where we used (86) to compute the limits in (95) and (96).

We now compute the second term on the right hand side of (92). As in the calculation above, we can do away with the mean value, and we have

\[
\mathbb{E}(\psi'[f'] \partial_x u(x + \alpha, t)) = \lim_{h \to 0} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} p(y) \varphi(y) q_{\alpha - h}(y, z) \frac{q_h(z, w)}{h} \left( \frac{f(w)}{w} - \frac{f(z)}{z} \right) (\psi(w) - \psi(z)) \, dw \, dz \, dy.
\]

As in the computation of (86) we find that the limit of the innermost integral is \( \mathcal{B} \psi(z) \). Therefore,

\[
\mathbb{E}(\psi'[f'] \partial_x u(x + \alpha, t)) = -\mathcal{P} (\varphi Q_\alpha \mathcal{B} \psi).
\]  

(98)

We combine (92), (97) and (98) to find

\[
\partial_t (\mathcal{P} (\varphi Q_\alpha \psi)) = \mathcal{P} (\mathcal{B} (\varphi Q_\alpha \psi) - \varphi Q_\alpha \mathcal{B} \psi + \varphi Q_\alpha \mathcal{B} \psi).
\]  

(99)

But the left hand side is simply

\[
(\partial_t \mathcal{P}) \varphi Q_\alpha \psi + \mathcal{P} (\varphi \partial_t Q_\alpha \psi),
\]

(100)

and by (87)

\[
(\partial_t \mathcal{P}) \varphi Q_\alpha \psi = \mathcal{P} \mathcal{B} (\varphi Q_\alpha \psi).
\]  

(101)

We now combine (99), (100) and (101) to obtain

\[
\mathcal{P} (\varphi \partial_t Q_\alpha \psi) = \mathcal{P} (\varphi (Q_\alpha \mathcal{B} - \mathcal{B} Q_\alpha) \psi).
\]  

(102)

This equation holds for all \( \varphi \) and \( \psi \) in the domain of \( \mathcal{A} \). Thus, we may write

\[
\partial_t Q_\alpha = (Q_\alpha \mathcal{B} - \mathcal{B} Q_\alpha) = [Q_\alpha, \mathcal{B}],
\]

(103)

and in the limit \( \alpha \downarrow 0 \) we have

\[
\partial_t \mathcal{A} = [\mathcal{A}, \mathcal{B}].
\]  

(104)
5 Hopf’s method

In this section we derive the Lax equation (22) following Hopf’s method [22]. This method was used by Chabanol and Duchon to derive their kinetic equations. We show that the calculation works for any convex flux $f$.

Though our calculations are mainly formal, we begin with the canonical framework of § 3. Recall that the initial data $u_0$ has law $\mu_0$ on $\Omega$. Let $u(t; u_0) : [0, \infty) \rightarrow \Omega$ be a weak solution of (1) with initial data $u_0$, which induces a sequence of probability measures $(\mu_t)_{t>0}$ on $\Omega$. To derive an equation for the flow of measures $(\mu_t)_{t>0}$, make the assumption that $u$ is differentiable in $t$ so that the conservation law can be written for all $\phi \in C_\infty^\infty(\mathbb{R})$ as

$$\partial_t \langle u(t; u_0), \phi \rangle - \langle f(u(t; u_0)), \phi' \rangle = 0. \quad (105)$$

Here, $' = \partial_\tau$ and $\langle \cdot, \cdot \rangle$ is the standard duality pairing. Define the Hopf characteristic functional $\hat{\mu}_t$ of the law $\mu_t$

$$\hat{\mu}_t(\phi) = \int_\Omega e^{i\langle u, \phi \rangle} \mu_t(du) \quad (106)$$

which evolves according to

$$\partial_t \hat{\mu}_t(\phi) = \partial_\tau \int_\Omega e^{i\langle u, \phi \rangle} \mu_t(du) = \partial_\tau \int_\Omega e^{i\langle u(t; u_0), \phi \rangle} \mu_0(du_0)$$

$$= \int_\Omega i \langle f(u(t; u_0), \phi'), e^{i\langle u(t; u_0), \phi \rangle} \mu_0(du_0)$$

$$= \int_\Omega i \langle f(u), \phi' \rangle e^{i\langle u, \phi \rangle} \mu_t(du). \quad (107)$$

Any set of probability measures $(\mu_t)_{t>0}$ on $\Omega$ that satisfies (107) for all $\phi \in C_\infty^\infty(\mathbb{R})$ is defined to be a statistical solution of the scalar conservation law—in particular, the flow of measures generated by the entropy solution is a statistical solution. Assuming that all moments of $\mu_t$ are finite, the exponential can be expanded in series (denoting $dx_n = \prod_{j=1}^n dx_j$) as

$$e^{i\langle u, \phi \rangle} = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_{\mathbb{R}^n} \prod_{j=1}^n u(x_j) \phi(x_j) dx_n.$$ 

Let

$$E_{\mu_t} \left[ \prod_{i=1}^n \varphi_i(u(x_i)) \right] = \int_\Omega \prod_{i=1}^n \varphi_i(u(x_i)) \mu_t(du).$$

Substituting the expansion for the exponential into (107) yields the infinite
Re-indexing the l.h.s. of (110) and subtracting the r.h.s. gives
\[ \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_{\mathbb{R}^n} \partial_t E_{\mu t} \left[ \prod_{j=1}^{n} u(x_j) \varphi(x_j) \right] dx_n \]
\[ = \sum_{n=0}^{\infty} \frac{i^{n+1}}{n!} \int_{\mathbb{R}^{n+1}} E_{\mu t} \left[ f(u(x_k)) \varphi'(x_k) \prod_{j \neq k} u(x_j) \varphi(x_j) \right] dx d\mathbf{x}_n. \quad (108) \]

We simplify this hierarchy as follows. If \( g : \mathbb{R}^n \to \mathbb{R} \) satisfies \( g(x_1, \ldots, x_n) = g(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \) for all permutations \( \sigma \) of \( \{1, \ldots, n\} \), then
\[ \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) d\mathbf{x}_n = n! \int_{x_1 < x_2 < \cdots < x_n} g(x_1, \ldots, x_n) d\mathbf{x}_n. \quad (109) \]

We choose
\[ g(x_1, \ldots, x_{n+1}) = \frac{1}{n} \sum_{k=1}^{n+1} E_{\mu t} \left[ f(u(x_k)) \varphi'(x_k) \prod_{j \neq k} u(x_j) \varphi(x_j) \right], \]
in (109) and substitute in (108) to obtain
\[ \sum_{n=0}^{\infty} \frac{i^n}{n!} \int_{x_1 < \cdots < x_n} \partial_t E_{\mu t} \left[ \prod_{j=1}^{n} u(x_j) \varphi(x_j) \right] dx_n \]
\[ = \sum_{n=0}^{\infty} \frac{i^{n+1}}{n!} \int_{x_1 < \cdots < x_{n+1}} \sum_{k=1}^{n+1} E_{\mu t} \left[ f(u(x_k)) \varphi'(x_k) \prod_{j \neq k} u(x_j) \varphi(x_j) \right] dx d\mathbf{x}_n. \quad (110) \]

Re-indexing the l.h.s. of (110) and subtracting the r.h.s. gives
\[ \sum_{n=1}^{\infty} i^n \int_{x_1 < \cdots < x_n} \left\{ \partial_t E_{\mu t} \left[ \prod_{j=1}^{n} u(x_j) \prod_{j=1}^{n} \varphi(x_j) \right] \right. \]
\[ - \sum_{k=1}^{n} E_{\mu t} \left[ f(u(x_k)) \prod_{j \neq k} u(x_j) \varphi'(x_k) \prod_{j \neq k} \varphi(x_j) \right] \left\} \right. \]
\[ \left. \left. dx_n = 0. \quad (111) \right. \right. \]

Finally, the statistical hierarchy (111) can be considerably simplified with the closure assumption that \((\mu_t)_{t \geq 0}\) is the law of a stationary Feller process with the one and two-point operators \(P\) and \((Q_h)_{h \geq 0}\) defined in (80) and (81). With \( h_i = x_i - x_{i-1} \) and \( Q_i = Q_{h_i} \) for \( i = 2, \ldots, n \),
\[ E_{\mu t} \left[ \prod_{i=1}^{n} \varphi_i(u(x_i)) \right] \]
\[ = \int_{\mathbb{R}} p(du_1, t) \varphi_1(u_1) \int_{\mathbb{R}} q_{h_2}(u_1, du_2, t) \varphi_2(u_2) \cdots \int_{\mathbb{R}} q_{h_n}(u_{n-1}, du_n, t) \varphi_n(u_n) \]
\[ = P \varphi_1 Q_2 \varphi_2 \cdots Q_n \varphi_n. \]
Assume the transition measures \( q_h \) are differentiable in \( h \). Transferring the derivative on the test function \( \varphi'(x_k) \) onto \( q_{h_i}(u_{k-1}, du_k) \) in (110) and using integration by parts, the statistical hierarchy is

\[
\sum_{n=1}^{\infty} i^n \int_{x_1 < \cdots < x_n} \left\{ \partial_t \mathbb{E}_{\mu_i} \left[ \prod_{j=1}^{n} u(x_j) \right] \right. \\
+ \sum_{k=1}^{n} \partial_{x_k} \mathbb{E}_{\mu_i} \left[ f(u(x_k)) \prod_{j \neq k} u(x_j) \right] \left. \right\} \prod_{j=1}^{n} \varphi(x_j) dx_n = 0.
\]

Note that the boundary terms in the previous equation vanish due to cancellation of terms and the fact that \( \varphi \in C_c^{\infty}(\mathbb{R}) \) has compact support. The density of tensor products of the form \( \psi(x_1, \ldots, x_n) = \prod_{i=1}^{n} \varphi(x_i) \) in the space of test functions on \( \{ x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 < \cdots < x_n \} \) implies that for \( n \in \mathbb{N} \) and \( x_1 < x_2 < \cdots < x_n \), the following infinite set of equations hold:

\[
\partial_t \mathbb{E}_{\mu_t} \left[ \prod_{j=1}^{n} u(x_j) \right] = -\sum_{k=1}^{n} \partial_{x_k} \mathbb{E}_{\mu_t} \left[ f(u(x_k)) \prod_{j \neq k} u(x_j) \right]. \tag{112}
\]

In terms of the operators \( Q'_i = \partial_{h_i}, Q_i \) (denoting \( h_1 = x_1, Q_1 = \mathcal{P}, Q_{n+1} = I \)) this is

\[
\partial_t \mathbb{E}_{\mu_t} \left[ \prod_{j=1}^{n} u(x_j) \right] = \sum_{k=1}^{n} \left\{ \mathcal{P} u Q_2 u \cdots Q_k f(u) Q_{k+1} u \cdots Q_n u \right. \\
- \left. \mathcal{P} u Q_2 u \cdots Q'_k f(u) Q_{k+1} u \cdots Q_n u \right\},
\]

or equivalently, in terms of the generator \( A \) (using \( Q'_i = A Q_i \))

\[
\partial_t \mathbb{E}_{\mu_t} \left[ \prod_{j=1}^{n} u(x_j) \right] = \sum_{k=1}^{n} \mathcal{P} u Q_2 u \cdots Q_k [f(u), A] Q_{k+1} u \cdots Q_n u.
\]

One obtains the evolution equation for the 1-point function by taking \( h_i \to 0 \) for all \( i \) (i.e., \( Q_i = I \) for all \( i \)), or the equation for the 2-point function by taking \( h_{i+1} = \alpha \) for \( l \in \{1, \ldots, n-1\} \) and \( h_i \to 0 \) for all \( i \neq l+1 \) (i.e., \( Q_{l+1} = Q_\alpha \) and \( Q_i = I \) for all \( i \neq l+1 \)):

\[
\partial_t \mathbb{E}_{\mu_t} \left[ (u(x_1))^n \right] = \sum_{k=1}^{n} \mathcal{P} u^{k-1} [f(u), A] u^{n-k} \tag{113}
\]

\[
\partial_t \mathbb{E}_{\mu_t} \left[ (u(x_1))^l (u(x_1 + \alpha))^{n-l} \right] = \sum_{k=1}^{l} \mathcal{P} y^{k-1} [f(y), A] y^{l-k} Q_\alpha y^{n-l} \\
+ \sum_{k=l+1}^{n} \mathcal{P} y^{k-1} Q_\alpha y^{k-(l+1)} [f(y), A] y^{n-k} \tag{114}
\]
Using the expression (20) for the generator $A$, (113) and (114) simplify:

$$\partial_t P y^n = - \sum_{k=1}^{n} P y^{k-1} \left\{ b(y) f'(y) y^{n-k} + \int_{\mathbb{R}} n(y, dz) (f(z) - f(y)) z^{n-k} \right\}$$

$$= - P \left\{ b(y) f'(y) (y^n)' + \int_{\mathbb{R}} n(y, dz) \frac{f(z) - f(y)}{z - y} (z^n - y^n) \right\}$$

$$\partial_t (P y^l Q y^{n-l})$$

$$= - \sum_{k=1}^{l} P y^{k-1} \left\{ b(y) f'(y) y^{l-k} Q_{\alpha} y^{n-l} + \int_{\mathbb{R}} n(y, dz) (f(z) - f(y)) z^{l-k} Q_{\alpha} z^{n-l} \right\}$$

$$- \sum_{k=l+1}^{n} P y^l Q_{\alpha} y^{k-(l+1)} \left\{ b(y) f'(y) y^{n-k} + \int_{\mathbb{R}} n(y, dz) (f(z) - f(y)) z^{n-k} \right\}$$

$$= - P \left\{ b(y) f'(y) (y^l)' Q_{\alpha} y^{n-l} + \int_{\mathbb{R}} n(y, dz) \frac{f(z) - f(y)}{z - y} (z^l - y^l) Q_{\alpha} z^{n-l} \right\}$$

$$- P y^l Q_{\alpha} \left\{ b(y) f'(y) (y^{n-l})' + \int_{\mathbb{R}} n(y, dz) \frac{f(z) - f(y)}{z - y} (z^{n-l} - y^{n-l}) \right\}.$$ 

In terms of the operator $B$ defined in (21) the above expressions read

$$\partial_t P y^n = P B y^n, \quad \partial_t (P y^l Q_{\alpha} y^{n-l}) = P (B(y) Q_{\alpha} y^{n-l}) - y^l B Q_{\alpha} y^{n-l} + y^l Q_{\alpha} B y^{n-l}).$$

Since $B$ is a linear operator, we approximate $\varphi, \psi \in C_0^\infty(\mathbb{R})$ by polynomials to obtain the evolution equation for the 1- and 2-point operators:

$$\partial_t P \varphi = P B \varphi, \quad \partial_t (P \varphi Q_{\alpha} \psi) = P (B(\varphi Q_{\alpha} \psi) - \varphi B Q_{\alpha} \psi + \varphi Q_{\alpha} B \psi).$$

Therefore we have the evolution $\partial_t A = [A, B]$ exactly as in § 4.

6 Groeneboom’s solution

In this section we verify that the generator

$$A(t) \varphi(y) = \frac{1}{t} \varphi'(y) + \int_{-\infty}^{y} \frac{1}{t^{1/3}} n_*(y^{1/3}, z^{1/3}) (\varphi(z) - \varphi(y)) \ dz$$

(115)

satisfies the Lax equation (22) when $f(u) = u^2/2$. In this case, the evolution of the drift (26) is simply $\dot{b} = -b^2$. It is clear that $b(y, t) = t^{-1}$ is a solution. The equation for the jump density now takes the form

$$\partial_t n(y, z, t) - \frac{1}{2t} (y - z) (\partial_y n - \partial_z n) = Q(n, n)$$

(116)
with the collision kernel
\[ Q(n,n)(y,z,t) = \frac{y-z}{2} \int_z^y n(y,w,t)n(w,z,t) \, dw 
- n(y,z,t) \int_z^y \frac{y-w}{2} n(z,w,t) \, dw 
- n(y,z,t) \int_y^\infty \frac{w-z}{2} n(y,w,t) n(y,w,t) \, dw. \]

We substitute the ansatz
\[ u = yt^{1/3}, \quad v = zt^{1/3}, \quad s = u - v, \quad n(y,z,t) = t^{-1/3} n^\ast(u,v), \]
(118) in (116) and collect terms. We must then verify that
\[ -2 \frac{2}{3} + \left( \frac{u-v}{3} J'(v) / J(v) + \left( \frac{u}{3} - v \right) J'(u) / J(u) - \frac{4}{3} s K'(s) / K(s) \right) \]
\[ = s K * K(s) - \left( s K * J(v) + J * (xK)(v) \right) + \left( J * (xK) / K(u) - s K * J(u) \right). \]
(119)

Here \( x \) plays the role of a dummy variable in the convolutions \( J * (xK) \). Explicitly, \( J * (xK)(u) = \int_x^u J(u-x)xK(x) \, dx \). Note also that the terms in (119) are evaluated at the arguments \( u, v \) or \( s = u - v \) respectively. The functions \( J \) and \( K \) are related by the following identities:
\[ x^2 J = K * J + J' \]
(120)
\[ \frac{3}{2} J * (xK) = x (K * J) + J \]
(121)
\[ x^3 K = 3x(K * K) + 4xK' + 2K. \]
(122)
(The argument of each function in these identities is \( x \)). We substitute these identities in (119), and collect the three terms corresponding to the arguments \( u, v \) and \( s \) and find that (remarkably!) each of them reduces to a polynomial, and the sum of these three polynomials vanishes.

The identities (120)–(122) are proven using the Laplace transform, (18), and the definition \( Ai''(q) = qAi(q) \). First, we observe that \( l = j'/j = -Ai'/Ai \) solves the Riccati equation
\[ l' = -q + l^2. \]
(123)
As a consequence of the definition of \( k \) in (18) we have
\[ l' = k/2. \]
(124)
Therefore, we also have
\[ k' = 2l'' = -2 + 4ll' = -2(1 - lk). \]
(125)
Thus, \( l \) and \( k \) solve an autonomous system. Equation (123) may be rewritten in terms of \( j \) and \( k \) as
\[ j''' - (q + k)j = 0. \]
(126)
This is equivalent to (120). Next, if we rewrite (125) in terms of \( j \) and \( k \), we find immediately that
\[
\frac{3}{2} jk' = (jk)' - j. \tag{127}
\]
This is equivalent to (121). Finally, we differentiate (125) twice and use (123) to eliminate \( l \) to find
\[
k'' = k^2 - 4l(1 - lk) \tag{128}
k''' = 3(k^2)' + 4qk' + 2k, \tag{129}
\]
which is equivalent to (122).

Here is the Painlevé property: differentiate (123) to find \( l'' = 2l^3 - 2lq - 1 \). Now rescale \( \tau = -q^{2/3} \), \( w(\tau) = 2^{-1/3}l(q) \) to see that \( w \) solves the second Painlevé equation with parameter \( 1/2 \) [1].
\[
\frac{d^2 w}{d\tau^2} = 2w^3 + w\tau + \frac{1}{2}. \tag{130}
\]

7 Acknowledgements

This material is based upon work supported by the National Science Foundation under grants DMS 06-05006 and DMS 07-48482. We have benefited greatly from discussions with Percy Deift, Irene Gamba, John Hartigan, David Kinderlehrer, Dave Levermore, Mokshay Madiman, Jonathan Mattingly, Scott McKinley, Irina Nenciu, Christian Pfrang and David Pollard. We thank them all. Most of all, we thank Bob Pego for the pleasure of continued collaboration. Part of this work was completed while the first author was visiting the Indian Statistical Institute, Kolkata.

References

[1] M. J. Ablowitz and P. A. Clarkson, Solitons, nonlinear evolution equations and inverse scattering, vol. 149 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1991.

[2] M. Abramowitz and I. A. Stegun, Handbook of mathematical functions with formulas, graphs, and mathematical tables, vol. 55 of National Bureau of Standards Applied Mathematics Series, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[3] D. F. Andrews, P. J. Bickel, F. R. Hampel, P. J. Huber, W. H. Rogers, and J. W. Tukey, Robust estimates of location: Survey and advances, Princeton University Press, Princeton, N.J., 1972.

[4] M. Avellaneda and W. E, Statistical properties of shocks in Burgers turbulence, Comm. Math. Phys., 172 (1995), pp. 13–38.
[5] J. Bertoin, *The inviscid Burgers equation with Brownian initial velocity*, Comm. Math. Phys., 193 (1998), pp. 397–406.

[6] ——, *Some properties of Burgers turbulence with white or stable noise initial data*, in Lévy processes, Birkhäuser Boston, Boston, MA, 2001, pp. 267–279.

[7] ——, *Some aspects of additive coalescents*, in Proceedings of the International Congress of Mathematicians, Beijing 2002, vol. III, Higher Ed. Press, 2002, pp. 15–23.

[8] J. M. Burgers, *The nonlinear diffusion equation*, Dordrecht: Reidel, 1974.

[9] L. Carraro and J. Duchon, *Solutions statistiques intrinsèques de l’équation de Burgers et processus de Lévy*, C. R. Acad. Sci. Paris Sér. I Math., 319 (1994), pp. 855–858.

[10] ——, *Équation de Burgers avec conditions initiales à accroissements indépendants et homogènes*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 15 (1998), pp. 431–458.

[11] M.-L. Chabanol and J. Duchon, *Markovian solutions of inviscid Burgers equation*, J. Statist. Phys., 114 (2004), pp. 525–534.

[12] H. Chernoff, *Estimation of the mode*, Ann. Inst. Statist. Math., 16 (1964), pp. 31–41.

[13] C. M. Dafermos, *Hyperbolic conservation laws in continuum physics*, Springer-Verlag, New York, 2000.

[14] F. J. Dyson, *A Brownian-motion model for the eigenvalues of a random matrix*, Rev. Modern Phys., 3 (1962), pp. 1191–1198.

[15] W. E and E. Vanden Eijnden, *Statistical theory for the stochastic Burgers equation in the inviscid limit*, Comm. Pure Appl. Math., 53 (2000), pp. 852–901.

[16] L. Frachebourg and P. A. Martin, *Exact statistical properties of the Burgers equation*, J. Fluid Mech., 417 (2000), pp. 323–349.

[17] L. Frachebourg, P. A. Martin, and J. Piasecki, *Ballistic aggregation: a solvable model of irreversible many particles dynamics*, Physica A: Statistical Mechanics and its Applications, 279 (2000), pp. 69 – 99.

[18] R. K. Getoor, *Splitting times and shift functionals*, Z. Wahrsch. Verw. Gebiete, 47 (1979), pp. 69–81.

[19] C. Giraud, *Genealogy of shocks in Burgers turbulence with white noise initial velocity*, Comm. Math. Phys., 223 (2001), pp. 67–86.
[20] P. Groeneboom, *Brownian motion with a parabolic drift and Airy functions*, Probab. Theory Related Fields, 81 (1989), pp. 79–109.

[21] E. Hopf, *The partial differential equation $u_t + uu_x = \mu_{xx}$*, Comm. Pure Appl. Math., 3 (1950), pp. 201–230.

[22] ———, *Statistical hydromechanics and functional calculus*, J. Rational Mech. Anal., 1 (1952), pp. 87–123.

[23] S. V. Kerov, *Asymptotic representation theory of the symmetric group and its applications in analysis*, vol. 219 of Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 2003. Translated from the Russian manuscript by N. V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski.

[24] S. Kida, *Asymptotic properties of Burgers turbulence*, J. Fluid Mech., 93 (1979), pp. 337–377.

[25] J. Kim and D. Pollard, *Cube root asymptotics*, Ann. Statist., 18 (1990), pp. 191–219.

[26] P. D. Lax, *Hyperbolic systems of conservation laws. II*, Comm. Pure Appl. Math., 10 (1957), pp. 537–566.

[27] P. D. Lax, *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 11.

[28] J. C. Mattingly and S. A. McKinley, *Shell models of turbulence*. Work in progress, 2009.

[29] J. C. Mattingly, T. Suidan, and E. Vanden-Eijnden, *Simple systems with anomalous dissipation and energy cascade*, Comm. Math. Phys., 276 (2007), pp. 189–220.

[30] G. Menon and R. L. Pego, *Universality classes in Burgers turbulence*, Comm. Math. Phys., 273 (2007), pp. 177–202.

[31] P. W. Millar, *A path decomposition for Markov processes*, Ann. Probability, 6 (1978), pp. 345–348.

[32] Z.-S. She, E. Aurell, and U. Frisch, *The inviscid Burgers equation with initial data of Brownian type*, Comm. Math. Phys., 148 (1992), pp. 623–641.

[33] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
[34] C. A. TRACY and H. WIDOM, Level-spacing distributions and the Airy kernel, Comm. Math. Phys., 159 (1994), pp. 151–174.

[35] D. VOICULESCU, Addition of certain noncommuting random variables, J. Funct. Anal., 66 (1986), pp. 323–346.

[36] A. I. VOL’PERT, Spaces BV and quasilinear equations, Mat. Sb. (N.S.), 73 (115) (1967), pp. 255–302.

[37] J. VON NEUMANN, Collected works. Vol. VI: Theory of games, astrophysics, hydrodynamics and meteorology, General editor: A. H. Taub. A Pergamon Press Book, The Macmillan Co., New York, 1963.

[38] M. WINKEL, Limit clusters in the inviscid Burgers turbulence with certain random initial velocities, J. Statist. Phys., 107 (2002), pp. 893–917.

[39] W. A. WOYCZYŃSKI, Burgers–KPZ turbulence, vol. 1700 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1998. Göttingen lectures.