Available online at www.sciencedirect.com
ScienceDirect
Acta Mathematica Scientia 2012,**32B**(1):281–294
Mathematica Acta Scientia 数学物理学报
http://actams.wipm.ac.cn

LESSER KNOWN MIRACLES OF BURGERS EQUATION*

Dedicated to Professor Constantine M. Dafermos on the occasion of his 70th birthday

Govind Menon
Division of Applied Mathematics, Box F, Brown University, Providence, RI 02912, USA
E-mail: menon@dam.brown.edu

Abstract This article is a short introduction to the surprising appearance of Burgers equation in some basic probabilistic models.

Key words Burgers equation; random matrix theory; kinetic theory; Dyson's Brownian motion; Kerov's kinetic equation; shock clustering; integrable systems

2000 MR Subject Classification 35R60; 35L67; 37K10; 60J75; 82C99

1 Introduction

1.1 Burgers Equation with a Simple Pole as Initial Data

The viscous Burgers equation and its inviscid limit appear in many textbooks on applied mathematics as a fundamental model of nonlinear phenomenon. In his pioneering analysis of Burgers equation, Hopf established the importance of singularity formation, weak solutions, and a vanishing viscosity limit as basic themes in the analysis of nonlinear partial differential equations [17]. When considering the modern theory of hyperbolic conservation laws, it is almost miraculous to see the ideas that have sprung from so simple a beginning.1) It is tempting to believe that we know the whole story, but there seem to be many new vistas ahead as conservation laws continue to arise in unexpected ways in apparently unrelated areas of mathematics.

Let me illustrate this point with a question. What is the right solution to the initial value problem

$$\partial_t g + g\partial_x g = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1}$$

$$g(x,0) = \frac{1}{x}? \tag{2}$$

*Received November 7, 2011. This material is based upon work supported by the National Science Foundation under grants DMS 07-04842 (CAREER) and EFRI–10-22730.

1) There are no miracles in science, and as Feynman writes about the gyroscope, 'it is a wonderful thing– but it is not a miracle' [13]. I hope the reader will forgive my poetic license. Lax calls Hopf's work 'epoch-making' in his commentary in [26], an exuberant endorsement in its own right.

The issue here is the singular initial data which is integrable neither at zero nor at infinity. As a consequence, we cannot naively use the Cole-Hopf solution formula to define an entropy solution for this initial value problem.

1.2 A Self-similar Solution and Wigner's Semicircle Law

This problem is not an isolated curiosity. It appears in various guises as a limit problem in algebra, combinatorics and probability theory. In operator theory and random matrix theory, the appropriate solution to (1)–(2) is the self-similar solution with diffusive scaling

$$g(t,z) = \frac{1}{\sqrt{t}} g_*(z\sqrt{t}), \tag{3}$$

where g_* is given explicitly in a slit plane

$$g_*(z) = \frac{1}{2}\left(z - \sqrt{z^2-4}\right), \quad z \in \mathbb{C}\backslash[-2,2]. \tag{4}$$

g_* is a Pick function [10]. It is the Cauchy transform of a probability measure μ_* supported in the interval $[-2,2]$

$$g_*(z) = \int_{-2}^{2} \frac{1}{z-x}\, \mu_*(\mathrm{d}x), \quad z \in \mathbb{C}\backslash[-2,2]. \tag{5}$$

In fact, μ_* is Wigner's celebrated semicircle law with density

$$\mu_*(\mathrm{d}x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathrm{d}x, \quad x \in [-2,2]. \tag{6}$$

g_* may be extended to the slit $x \in [-2,2]$ by computing its principal value. This is simply the average $g_*(x) = (g_*(x_+) + g_*(x_-))/2$ where the subscripts denote the limits taken from above and below in $\mathbb{C}\backslash[-2,2]$. We then find

$$g(x,t) = \frac{x}{2t}, \quad x \in [-2\sqrt{t}, 2\sqrt{t}]. \tag{7}$$

Equations (3), (4) and (7) complete the prescription of $g(x,t)$ for $x \in \mathbb{R}$ and $t > 0$. The factor of 2 in equation (7) is crucial: it shows that g not an entropy solution to Burgers equation for $x \in \mathbb{R}$, $t > 0$ although it does define a solution to the initial value problem (1)–(2) analytic in a moving domain $\mathbb{C}\backslash[-2\sqrt{t}, 2\sqrt{t}]$.

That g is not an entropy solution may appear disconcerting at first sight. But this appears less arbitrary when one realizes that the solution (3) is the limit as $n \to \infty$ of a sequence $g_n(t,z)$ that satisfy the stochastic partial differential equation

$$\partial_t g_n + g_n \partial_z g_n = \frac{1}{n}\left(\frac{1}{\beta} - \frac{1}{2}\right)\partial_z^2 g_n + \sqrt{\frac{2}{\beta n^3}} \sum_{k=1}^{n} \frac{\dot{B}_k}{(z-x_k)^2}, \quad z \in \mathbb{C}_+, \ \ t > 0. \tag{8}$$

Here n is a positive integer, $\beta = 1$, 2 or 4, and B_k, $k = 1, \cdots, n$ are standard, independent, Brownian motions. It is not obvious at all that this stochastic partial differential equation is well-posed, particularly since the viscous term vanishes when $\beta = 2$ and has the wrong sign for $\beta = 4$. But this will follow naturally from the discussion below.

1.3 Some Context

Wigner's semicircle law is of the first importance in the theory of random matrices and it is to to my mind surprising and delightful that Burgers equation should reappear here. This connection is just one of several fascinating links between Burgers equation and some basic probabilistic models. Many of these links are known to experts, but they are not as well known to the working applied mathematician as they should be, especially since they constitute a beautiful class of exact solutions with wide appeal. I stumbled upon these links in connection with the problem of Burgers turbulence (as explained in Section 5 below) and was a little dismayed to realize when I dug into the literature that my calculations were "well-known". Of course, what is well-known in one community is seldom well-known in another, and my intention here is to explain a few such links in a simple and transparent manner. To this end, I highlight the main calculations avoiding all technicalities. In keeping with the informal tone of this article, the references are representative, not exhaustive.(But all assertions can be proved rigorously with a little work and reference to [3, 16, 19] when needed. More complete references can be found in [23, 25]). I hope the exposition will be of value to applied mathematicians with an interest in differential equations, integrable systems, probability theory and kinetic theory as an introduction to the fascinating interplay between these areas.

The rest of the article is organized as follows. Section 2 and Section 3 provide an overview of the role of Burgers equation in random matrix theory. We begin with Brownian motion in the space of $n \times n$ self-adjoint random matrices and show how it naturally leads to the approximation (8) and the initial value problem (1)–(2) in the limit $n \to \infty$. Random matrix theory leads to kinetic theory and growth processes in Section 4. Finally, we turn to a particular statistical theory of turbulence–the motivation for both Burgers and Hopf– and connect this to random matrix theory and growth processes. We conclude with some speculation about what appear to be deep connections between a family of integrable systems and probability theory in Section 5.

1.4 Dedication

The article is dedicated to Constantine Dafermos in admiration and respect. He has been an exemplary teacher and colleague. More than ten years ago, as a graduate student I looked to Costas to clear the fog of my confusion. Today I look back and see that nothing has changed. It is a pleasure to dedicate this article to him on the occasion of his 70th birthday.

2 Brownian Motion of Self-adjoint Matrices

2.1 Dyson's Brownian Motion

We first introduce Brownian motion in the space of self-adjoint $n \times n$ matrices. A standard Brownian motion $B(t)$, $t \geq 0$ is a Gaussian process with continuous paths and independent increments normalized so that $\mathbb{E}(B_t^2) = t$. It may be used to construct Brownian motion in the space of self-adjoint matrices. Here the term self-adjoint is used to encompass real symmetric, complex Hermitian, and quaternion self-dual matrices. We consider real symmetric ($\beta = 1$) and complex Hermitian ($\beta = 2$) matrices $M(t)$ whose entries M_{jk} are Brownian motions normalized as follows:

$$M_{jk}(t) = \frac{1}{\sqrt{\beta}}\left(B^1_{jk}(t) + i(\beta-1)B^2_{jk}(t)\right) \quad 1 \leq j < k \leq n, \tag{9}$$

$$M_{jj}(t) = \sqrt{\frac{2}{\beta}} B_j(t), \quad 1 \leq j \leq n. \tag{10}$$

Here B^l_{jk} and B_j are standard, independent Brownian motions. The asymmetry between diagonal and off-diagonal terms is introduced in order that the distribution of $M(t)$ is unitarily invariant. Since the M_{jk} are independent Gaussian random variables, their joint density is a product, and the normalization above ensures it is of the form $c_n t^{-n(n+1)/4}\exp(-\beta \mathrm{Tr}(M^*M))$. The distribution is therefore invariant under the natural transformation $M\mapsto Q^*MQ$ for an orthogonal ($\beta=1$) or unitary ($\beta=2$) matrix Q. We have not defined the case $\beta=4$ of self-dual quaternion matrices for simplicity. At the snapshot $t=1$, these stochastic processes correspond to the fundamental ensembles of random matrix theory: the Gaussian Orthogonal Ensemble (GOE), Gaussian Unitary Ensemble (GUE), and Gaussian Symplectic Ensemble (GSE) [22].

While the entries of $M(t)$ are independent, its eigenvalues $\lambda_k(t)$, $1\le k\le n$ are not. Yet, Dyson showed that the eigenvalues satisfy the stochastic differential equation [11]

$$\mathrm{d}\lambda_k=\sum_{j\neq k}\frac{\mathrm{d}t}{\lambda_k-\lambda_j}+\sqrt{\frac{2}{\beta}}\,\mathrm{d}B_k,\quad 1\le k\le n, \tag{11}$$

where B_k, $1\le k\le n$ are independent Brownian motions. That is, the eigenvalues behave like repulsive unit charges on the line perturbed by independent white noise! More generally, we may consider Brownian motion $N(t)$ that starts at a fixed self-adjoint matrix N_0

$$N(t)=N_0+M(t), \tag{12}$$

where M is as above. The eigenvalues of $N(t)$ continue to satisfy (11) but now with initial conditions $\lambda_j(0)$ that are the eigenvalues of N_0.

2.2 Itô's Formula and Dyson's Equation

We first derive Dyson's equation (11) as an interesting application of Itô's formula. Here is a quick summary of Itô's formula [21]: If B is a one-dimensional Brownian motion and $f:\mathbb{R}\to\mathbb{C}$ a twice differentiable function, we replace the usual chain rule for the composition $f(B(t))$ with

$$\mathrm{d}f(B)=f'(B)\mathrm{d}B+\frac{1}{2}f''(B)(\mathrm{d}B)^2,\quad (\mathrm{d}B)^2=\mathrm{d}t. \tag{13}$$

The extension to an N-dimensional Brownian motion $B(t)=(B_1(t),\cdots,B_N(t))$ is natural. If $f:\mathbb{R}^N\to\mathbb{C}$ is twice differentiable, Itô's formula is

$$\mathrm{d}f(B)=Df(B)\mathrm{d}B+\frac{1}{2}D^2f(B)[\mathrm{d}B,\mathrm{d}B]. \tag{14}$$

The first term is computed as usual, and the second term is given by

$$D^2f(B)[\mathrm{d}B,\mathrm{d}B]=\sum_{j,k=1}^{N}\partial_j\partial_k f(B)\mathrm{d}B_j\mathrm{d}B_k,\quad \mathrm{d}B_j\mathrm{d}B_k=\delta_{jk}\mathrm{d}t, \tag{15}$$

with $\delta_{jk}=1$ if $j=k$ and 0 otherwise. The rule $\mathrm{d}B_j\mathrm{d}B_k=\delta_{jk}\mathrm{d}t$ is based on the quadratic variation and independence of each B_k.

We will derive Dyson's equation (11) by choosing f to be $\log\det(zI-M)$ in Itô's formula. Here we fix z in the upper half plane $\mathbb{C}_+$. Since the eigenvalues are real, $\det(zI-M)$ never vanishes in $\mathbb{C}_+$ and $\log\det(zI-M)$ is well defined. To apply Itô's formula we must compute the first and second derivatives of the determinant of a matrix. One way to do this is to expand the determinant. For small $A \in \mathbb{M}^{n\times n}$ we use the definition of the determinant to obtain

$$\det(I-A) = 1 - \mathrm{Tr}(A) + \frac{1}{2}\sum_{j,k=1}^{n} \det\begin{pmatrix} A_{jj} & A_{jk} \\ A_{kj} & A_{kk}\end{pmatrix} + \cdots. \tag{16}$$

Let us apply this expansion to $\det(zI-M)$. Let $R=(zI-M)^{-1}$ and $A = R\mathrm{d}M$. We write $zI-(M+\mathrm{d}M) = (zI-M)(I-A)$ and use (16) to obtain

$$\begin{aligned}&\det(zI-(M+\mathrm{d}M))\\ &= \det(zI-M)\left(1-\mathrm{Tr}(A) + \frac{1}{2}\sum_{j,k=1}^{n}\det\begin{pmatrix} A_{jj} & A_{jk} \\ A_{kj} & A_{kk}\end{pmatrix} + \cdots\right).\end{aligned} \tag{17}$$

If M is a diagonal matrix with entries $\lambda_1,\cdots,\lambda_n$, then R is also diagonal with entries $(z-\lambda_1)^{-1},\cdots,(z-\lambda_n)^{-1}$. In this case,

$$\mathrm{Tr}(A) = \mathrm{Tr}(R\mathrm{d}M) = \sum_{j=1}^{n}\frac{\mathrm{d}M_{jj}}{z-\lambda_j} = \sqrt{\frac{2}{\beta}}\sum_{j=1}^{n}\frac{\mathrm{d}B_j}{z-\lambda_j}. \tag{18}$$

Similarly, by the definition of M_{jk} in (9) we find

$$\det\begin{pmatrix} A_{jj} & A_{jk} \\ A_{kj} & A_{kk}\end{pmatrix} = \frac{\mathrm{d}M_{jj}\mathrm{d}M_{kk} - \mathrm{d}M_{jk}\mathrm{d}M_{kj}}{(z-\lambda_j)(z-\lambda_k)} = -\frac{(1-\delta_{jk})\mathrm{d}t}{(z-\lambda_j)(z-\lambda_k)}. \tag{19}$$

The general calculation can be reduced to this one by the unitary invariance of the law of $M(t)$. We factorize $M = Q\Lambda Q^*$ with $\Lambda = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ and Q unitary. The matrix $\mathrm{d}M$ in (18) and (19) is then replaced with $Q\mathrm{d}MQ^*$ which has the same law. In order to derive (11), we first note that

$$\mathrm{d}(\log\det(zI-M)) = -\sum_{j=1}^{n}\frac{\mathrm{d}\lambda_j}{z-\lambda_j}. \tag{20}$$

We combine this observation with (14), (17), (18) and (19) to obtain

$$\sum_{j=1}^{n}\frac{\mathrm{d}\lambda_j}{z-\lambda_j} = \frac{1}{2}\sum_{j=1}^{n}\sum_{k\neq j}\frac{\mathrm{d}t}{(z-\lambda_j)(z-\lambda_k)} + \sqrt{\frac{2}{\beta}}\sum_{j=1}^{n}\frac{\mathrm{d}B_j}{z-\lambda_j}. \tag{21}$$

We now let $z\to\lambda_k$ for each k to obtain (11).

3 Burgers Equation in the Spectral Plane

3.1 The Cauchy Transform of the Empirical Distribution

The empirical spectral measure of $N^{(n)}$ is defined to be the probability distribution with atoms of mass $1/n$ at each eigenvalue of $N^{(n)}$ (the superscript is introduced to make the size of the matrices explicit). The empirical spectral measure is random and it is natural to seek a law of large numbers after suitable rescaling as $n\to\infty$. We can guess the form of the scaling as follows. Suppose $x = \lambda/c$, where c is a scale parameter to be determined. We use (11) to obtain

$$\mathrm{d}x_k = \frac{1}{c^2}\sum_{j\neq k}\frac{\mathrm{d}t}{x_k - x_j} + \sqrt{\frac{2}{c^2\beta}}\mathrm{d}B_k, \quad 1 \le k \le n. \tag{22}$$

The choice $c^2 = n$ makes the first term a functional of the empirical spectral measure. It also ensures that the stochastic forcing vanishes as $n \to \infty$. Therefore, we rescale $\lambda = x\sqrt{n}$ and define $\nu_t^{(n)}$ to be the empirical spectral measure of the rescaled matrices $n^{-1/2}N^{(n)}(t)$.

Some care is needed though. As $n \to \infty$ the eigenvalues are bunched closer together and we must expect a singularity in the first term in (22). As in many probability limit theorems, a suitable integral transform clarifies the analysis. Here it is the Cauchy (or Cauchy-Stieltjes) transform

$$g_n(t,z) = \int_{\mathbb{R}} \frac{1}{z-x}\,\nu_t^{(n)}(\mathrm{d}x) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{z-x_j}, \quad z \in \mathbb{C}_+. \tag{23}$$

We show below that g_n satisfies the stochastic partial differential equation (8). The initial condition $g_n(0,z)$ for (8) is the Cauchy transform of $\nu_0^{(n)}$, the empirical spectral measure of the random matrix $n^{-1/2}N_0^{(n)}$. In the simplest situation $N_0 = 0$ and $\nu_0^{(n)}$ is a unit atom at the origin with Cauchy transform

$$g_n(z,0) = \frac{1}{z}. \tag{24}$$

In particular, as $n \to \infty$, we see that the limit g of g_n formally satisfies the initial value problem (1)–(2)

$$\partial_t g + g\partial_x g = 0, \;\; x \in \mathbb{R}, \;\; t > 0, \qquad g(x,0) = \frac{1}{x}.$$

More generally, we must assume that the spectra of the initial matrices $n^{-1/2}N_0^{(n)}$ converge to a probability measure ν_0 with compact support. In this case, $g_n(z,t)$ is the spectral measure of $n^{-1/2}N^{(n)}(t)$, and as $n \to \infty$, g_n converges to the solution of (1) with initial condition

$$g(z,0) = \int_{\mathbb{R}} \frac{1}{z-x}\,\nu_0(\mathrm{d}x). \tag{25}$$

Well-posedness of (8) with initial condition (25) follows from well-posedness of Dyson's stochastic differential equation (11). Precise statements and a rigorous proof of convergence in the framework of large deviations theory may be found in [16]. For our purposes, we simply highlight how (8) can be derived using Itô's formula and a clever identity from [19].

3.2 Itô's Formula and the Cauchy Transform

We first derive (8) using Itô's formula. By Itô's formula

$$\mathrm{d}g_n = \sum_{k=1}^{n}\frac{\partial g_n}{\partial x_k}\mathrm{d}x_k + \frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 g_n}{\partial x_j \partial x_k}\mathrm{d}x_j\mathrm{d}x_k \tag{26}$$

The second term simplifies to

$$\frac{1}{2n}\sum_{k=1}^{n}\frac{2\mathrm{d}x_k^2}{(z-x_k)^3} = \frac{1}{\beta n^2}\sum_{k=1}^{n}\frac{2\mathrm{d}t}{(z-x_k)^3} = \frac{1}{\beta n}\partial_z^2 g_n \mathrm{d}t, \tag{27}$$

where we evaluate $\mathrm{d}x_k^2$ using (22) with $c^2 = n$ and the rules $\mathrm{d}B_k^2 = \mathrm{d}t$ and $\mathrm{d}t^2 = \mathrm{d}B_k \mathrm{d}t = 0$. The first term in (26) is given by the chain rule and is

$$\frac{1}{n}\sum_{k=1}^{n}\frac{1}{(z-x_k)^2}\left(\frac{1}{n}\sum_{j\neq k}\frac{\mathrm{d}t}{x_k-x_j}+\sqrt{\frac{2}{\beta n}}\mathrm{d}B_k\right). \tag{28}$$

The stochastic forcing term is as in (8). Therefore, we focus on the deterministic part. Here we use the interesting identity

$$\frac{1}{n^2}\sum_{k=1}^{n}\sum_{j\neq k}\frac{1}{(z-x_k)^2(x_k-x_j)} = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{j\neq k}\frac{1}{(z-x_k)^2(z-x_j)}. \tag{29}$$

The proof is elementary. A couple of lines of algebra yields

$$\frac{1}{(z-x_k)^2(x_k-x_j)} = \frac{1}{(z-x_k)^2(z-x_j)} - \frac{1}{(x_k-x_j)^2}\left(\frac{1}{z-x_k}-\frac{1}{z-x_j}\right),$$

and the second term vanishes when we sum over j and k. Finally, we note that the right hand side of (29) can be expressed in terms of g_n as

$$-g_n\partial_z g_n - \frac{1}{2n}\partial_z^2 g_n. \tag{30}$$

We combine (26), (27) and (30) to obtain (8).

3.3 The Semicircle Law Again

Let us now solve the initial value problem (1). Since the initial data has no intrinsic scale, it is natural to expect a self-similar solution $g(z,t) = l_1(t)f(z/l_2(t))$. In order to guess the time-dependent scale factors we reason as follows. First, for any $t \geq 0$, g is the Cauchy transform of the probability distribution function μ_t (we write μ_t instead of ν_t since this initial condition is special). Therefore, $g(z,t) \sim z^{-1}$ as $z \to \infty$. Thus, $l_1(t) = 1/l_2(t)$ and we may write $g(z,t) = l(t)^{-1}g_*(z/l(t))$. Next, the scale factor $l(t)$ reflects the length of the support of μ_t. This length may be determined by focusing on the largest eigenvalue. Dyson's calculation tells us that the largest eigenvalue is repelled by all of the eigenvalues to the left, and we expect based on (11) that $\dot{l} = O(l^{-1})$. Therefore, $l = O(\sqrt{t})$.

We thus define the similarity variable $\xi = z/\sqrt{t}$ and substitute the ansatz $g(z,t) = g_*(\xi)/\sqrt{t}$ in (1). This ansatz is 'diffusive'. That is, it is what one would use for the heat equation on the line. It is very unusual in the context of Burgers equation. The point of the discussion above is to explain why it is the right guess here. A routine calculation then yields

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left(g_*^2 - \xi g_*\right) = 0. \tag{31}$$

Thus, $g_*^2 - \xi g_*$ is a constant. We can choose this constant to be 1– every other choice can be recovered by scaling ξ. Thus our solution is given implicitly by $g_*^2 - \xi g_* = 1$ and explicitly by (4). We now see that the self-similar solution is defined in the exterior of a growing slit in the complex plane, and we have

$$g(t,z) = \frac{1}{2t}\left(z - \sqrt{z^2-4t}\right), \quad z \in \mathbb{C}\backslash[-2\sqrt{t}, 2\sqrt{t}]. \tag{32}$$

g is the Cauchy transform of the probability measure μ_t. This measure is

$$\mu_t(\mathrm{d}x) = \frac{1}{2\pi t}\sqrt{4t - x^2}\,\mathrm{d}x, \quad x \in [-2\sqrt{t}, 2\sqrt{t}]. \tag{33}$$

(It takes some work to invert the Cauchy transform, but not too much work to verify that this is the inverse). μ_t is Wigner's semicircle law with a scale parameter t.

4 Kinetic Theory and Growth Processes

4.1 Kerov's Kinetic Equation

Our calculations so far have been eased by the Cauchy transform. There is a bijective correspondence between Pick functions and certain positive measures and the pointwise convergence of $g_n(t, z)$ to $g(t, z)$ for fixed z in the upper-half plane is enough to guarantee the weak convergence of the empirical measures $\nu_t^{(n)}$ to ν_t [10]. In addition, we have the representation formula

$$g(z, t) = \int \frac{1}{z - x}\nu_t(\mathrm{d}x), \quad z \in \mathbb{C}_+. \tag{34}$$

Thus, as one may expect, (1) describes completely not just the evolution of $g(t, \dot{})$, but also the evolution of a measure ν_t such that $g(\cdot, t)$ is its Cauchy transform. If g satisfies (1), what evolution equation does ν satisfy?

An explicit description of the evolution of ν_t (as opposed to the implicit prescription of (1)) seems to have been first found by Kerov [19, §4.6.5]. For finite n, the eigenvalues repel one another and are perturbed by noise as described by (22). In the scaling limit, the perturbative noise is washed out, and ν satisfies a kinetic equation that may be written in the weak form

$$\frac{\partial}{\partial t}\int_{\mathbb{R}} \varphi(x)\nu_t(\mathrm{d}x) = \int_{\mathbb{R}}\int_{\mathbb{R}} \frac{\varphi'(x) - \varphi'(y)}{x - y}\nu_t(\mathrm{d}x)\nu_t(\mathrm{d}y). \tag{35}$$

Here φ denotes a complex-valued test function that is continuous in x and decays sufficiently fast at infinity. We choose $\varphi(x) = 1/(z - x)$, $z \in \mathbb{C}_+$, to recover (1).

It is easier to solve (1) than (35), but the solutions are equivalent. The general solution to (1) with initial condition (25) is of particular importance in the theory of free probability and it is common to say that ν_t solves the with initial condition ν_0 or that ν_t is given by the free convolution of ν_0 with the scaled semicircle law μ_t [28]. While it would take us too far afield to describe these results, it is not that hard to describe the underlying motivation. The solution to (1) with the general initial condition (25) describes the evolution of the limiting spectral measure of the matrices $n^{-1/2}N^{(n)}(t)$ when the spectral measure of the matrices $n^{-1/2}N_0^{(n)}$ converges to ν_0 as $n \to \infty$. To be concrete, suppose $n^{-1/2}N_0^{(n)}$ are discretizations of a bounded linear operator with spectrum ν_0. In this case, the discretizations $n^{-1/2}N^{(n)}(t)$ help us formulate a notion of an operator-valued Brownian motion starting at N_0. Roughly speaking, the semicircle law μ_t plays a role analogous to the heat kernel and Burgers equation plays the role of the heat equation in making these notions precise.

4.2 Plancherel Growth and its Scaling Limit

Kerov's kinetic equation is also the limit of an important Markov process called Plancherel growth. The full sweep of ideas involved can only be appreciated by a careful reading of his beautiful thesis [19], but here is a quick summary. Fix a positive integer n, and consider an interlacing sequence of real numbers $x_1 < y_1 < x_2 < y_2 \cdots < y_{n-1} < x_n$. We consider a 1-Lipschitz function $\omega(x)$ that has minima at $x_1 < \cdots < x_n$, maxima at $y_1 < \cdots < y_{n-1}$ and is asymptotic to $|x-\alpha|$ where $\alpha = \sum x_j - \sum y_j$. Kerov and Vershik call such a function a diagram and α its center. The graph of such a diagram is a sawtooth with slope ± 1, upward corners at x_j and downward corners at y_j. More generally, a diagram is a 1-Lipschitz function ω that (i) oscillates within an interval $[a, b]$; (ii) and such that $\omega(x) = |x - \alpha|$ for all x outside the interval $[a, b]$ for some $\alpha \in \mathbb{R}$ that depends only on ω. $\mathcal{D}_{[a,b]}$ denotes the set of such diagrams.

We now introduce a bijective correspondence between diagrams and probability measures. We consider positive numbers p_j, $1 \le j \le n$ such that $\sum\limits_{j=1}^{n} p_j = 1$ and x_j as above. The bijective correspondence between diagrams with n minima at x_j and the probability measure $\mu = \sum\limits_{j=1}^{n} p_j \delta_{x_j}$ is given by the partial fractions expansion

$$\sum_{j=1}^{n} \frac{p_j}{x - x_j} = \frac{\prod\limits_{i=1}^{n-1} (x - y_i)}{\prod\limits_{j=1}^{n} (x - x_j)}. \tag{36}$$

y_j are determined once p_k are prescribed and vice versa. This correspondence is of particular importance in combinatorics and representation theory. Every permutation of n can be decomposed into increasing sequences, and the length of these sequences is conveniently represented by a Young diagram. Tilt the Young tableaux by 45^o and observe that its boundary is a diagram in the sense of Kerov and Vershik. The Young diagram of each partition of $n+1$ can be obtained from that of a partition of n by adding a block to it. This gives rise to a growth process on diagrams. As n increases to $n+1$ we form a new diagram from an old one by the attachment of a unit square at a minimum.

In order to describe a Markov process we must prescribe the states of the process and the transition rates between states. Plancherel growth is a discrete Markov process whose states are Young diagram $(\omega_1, \omega_2, \cdots, \omega_n, \cdots)$. The transition probability between diagram ω_n and ω_{n+1} is given by μ: with probability p_j a unit square is attached at x_j. The importance of this transition measure is that ω_n is distributed according to the Plancherel measure on Young diagram, which in turn is the push-forward of the uniform measure on the symmetric group S_n onto its irreducible representations.

The connection to kinetic theory is as follows. (i) It is true (and certainly not obvious!) that the above bijective correspondence between discrete diagrams and probability measures can be extended to a bijective correspondence between $\mathcal{D}_{[a,b]}$ and probability measures with support $[a, b]$. (ii) Plancherel growth can be replaced by a discrete deterministic evolution where a diagram grows rougher by the attachment of a square of size p_k at each minima x_k in unit time. The continuum limit of this process is Kerov's kinetic equation (35).

5 Scalar Conservation Laws with Random Data

5.1 Shock Clustering and Burgers Turbulence

The goal of the discussion so far has been to introduce the reader to the fact that Burgers equation plays an important role in random matrix theory and representation theory. Let us now return to nonlinear science. Both Burgers and Hopf spent substantial parts of their careers working on turbulence. Burgers worked on flow instabilities, (re)-introduced the equation that bears his name, analyzed shock formation and the interaction of shocks, and initiated the study of Burgers equation with random initial data. In addition to [17], Hopf introduced the idea that the transition to turbulence occurs through the loss of stability of quasiperiodic solutions to the Navier-Stokes equation in a sequence of Hopf(!)-bifurcations. He was also the first to precisely formulate a statistical theory of turbulence as the flow of a probability measure on the space of solutions to the Navier-Stokes equations [26]. While complete results on these problems continue to elude us, these ideas form the foundation for statistical theories of turbulence. Von Neumann's 1950's survey [29] is a wonderful historical document of these ideas in gestation. See also [12] for a recent survey.

A modern formulation of Burgers' model of turbulence is the following. What are the statistics of the entropy solution to Burgers equation

$$\partial_t u + \partial_x \left(\frac{u^2}{2}\right) = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{37}$$

when the initial data $u(x,0) = u_0(x)$ is random? Two important cases of random initial data are: (i) u_0 is Brownian motion; (ii) u_0 is white noise (the "derivative" of Brownian motion).

Both problems are exactly solvable! The solutions rely on closure theorems for the entropy solution to (37) that assert that a class of stochastic processes is left invariant by the entropy solution to (37). The main underlying insight here is that several classes of stochastic processes are "rigid". For example, Markov processes that are sufficiently regular (Feller processes) are completely characterized by their generators. The characterization of Lévy processes by the Lévy-Khintchine formula is an important special case of this general characterization.

For case (i) the relevant closure theorem is the following [4]: if $u_0(x)$, $x > 0$ is a one-sided Lévy process in x with only downward jumps, then so is $u(x,t) - u(0,t)$, $x \geq 0$ for every $t > 0$. Thus, the entropy solution to Burgers equation preserves the class of Lévy processes with one-sided jumps. This theorem is quite unexpected: a priori the notion of Lévy processes is probabilistic and has nothing to do with conservation laws, so it is surprising that it should be compatible with the notion of an entropy solution as long as we impose the requirement that the jumps are one-sided.

The closure theorem may be applied to solve explicitly for the statistics of u as follow. The statistics of Lévy processes with only downward jumps are completely characterized by their Laplace exponent ψ defined as follows:

$$\psi(q,t) = \frac{1}{x} \log \mathbb{E}\left(\mathrm{e}^{q(u(x,t)-u(0,t))}\right), \quad q \in \mathbb{C}_+, \quad x > 0, \quad t \geq 0. \tag{38}$$

Here q denotes a spectral variable. As a consequence of the closure theorem, the problem of determining the evolution of shock statistics reduces to determining an equation for the evolution of ψ. Astonishingly, the evolution of ψ is described by Burgers equation again, but now in the spectral plane.

$$\partial_t \psi + \psi \partial_q \psi = 0, \quad q \in \mathbb{C}_+, \quad t > 0. \tag{39}$$

These remarkable calculations were first discovered by Carraro and Duchon [7] and made rigorous shortly thereafter by Bertoin [4]. The initial data $\psi(q,0)=\psi_0(q)$ is given by the Laplace exponent of u_0. For example, when u_0 is a Brownian motion, this is $\psi_0(q)=q^2$. Equation (39) is easily solved in this case and we obtain the self-similar solution

$$\psi(q,t)=\frac{1}{t^2}\psi_*(qt), \quad \psi_*(q)=q+\frac{1}{2}-\sqrt{q+\frac{1}{4}}. \tag{40}$$

Equation (39) also has a kinetic interpretation: the Laplace exponent ψ is determined through the Lévy-Khintchine formula

$$\psi(q,t)=\int_0^\infty \left(\mathrm{e}^{qs}-1+qs\right)\Lambda_t(\mathrm{d}s). \tag{41}$$

Roughly speaking $\Lambda_t(\mathrm{d}s)$ denotes the number density of shocks of size s. The shocks evolve through ballistic aggregation and this induces an evolution of Λ_t. This process is described by a kinetic equation which may be rescaled to Smoluchowsk's coagulation equation with additive kernel. In analogy with Kerov's kinetic equation (35) we write Smoluchowski's equation in the weak form

$$\partial_\tau \int_0^\infty \varphi(s)\lambda_\tau(\mathrm{d}s)=\frac{1}{2}\int_0^\infty\int_0^\infty \left(\varphi(r+s)-\varphi(s)-\varphi(r)\right)(r+s)\lambda_\tau(\mathrm{d}r)\lambda_\tau(\mathrm{d}s). \tag{42}$$

(here λ_τ is a rescaling of Λ_t, and τ denotes a time scale related to the total variation of $u(x,t)$. See [24, §2.5]).

We now turn to case (ii). The statistics of u were determined explicitly by Groeneboom [15] and Frachebourg and Martin [14] following earlier work of Burgers [6]. We developed a different perspective on this solution by establishing a closure theorem of greater generality suggested by work of Chabanol and Duchon [8]. Our theorem applies to the entropy solution to a scalar conservation law with a C^1 convex flux f

$$\partial_t u+\partial_x f(u)=0, \quad x\in\mathbb{R}, \quad t>0. \tag{43}$$

Unlike Burgers equation, rarefactions waves are no longer straight lines for general f, and therefore cannot be sample paths of a Lévy process. However, the entropy condition *is* compatible with the broader class of Markov processes. Here is the general closure theorem [25, §3]: assume $u_0(x)$, $x\in\mathbb{R}$ is a Markov process with only downward jumps, then so is the entropy solution $u(x,t)$, $x\in\mathbb{R}$, to (43).

If $u(x,t)$, $x\in\mathbb{R}$ is a stationary Feller process whose sample paths have bounded variation then its generator is an integro-differential operator

$$\mathcal{A}(t)\varphi(u)=b_t(u)\varphi'(u)+\int_{-\infty}^u \left(\varphi(v)-\varphi(u)\right)n_t(u,\mathrm{d}v). \tag{44}$$

The characteristics of $\mathcal{A}(t)$ are the drift term b_t and the jump measure $n_t(u,\mathrm{d}v)$ and φ denotes a C_c^1 test function. The closure theorem reduces the problem to determining an evolution equation for the generator $\mathcal{A}$. We showed in [25] that $\mathcal{A}$ satisfies the Lax equation

$$\partial_t\mathcal{A}=[\mathcal{A},\mathcal{B}]. \tag{45}$$

Here the operator $\mathcal{B}(t)$ is defined by its action on a test function φ as follows

$$\mathcal{B}(t)\varphi(u) = -f'(u)b_t(u)\varphi'(u) - \int_{-\infty}^{u} \frac{f(v)-f(u)}{v-u}\left(\varphi(v)-\varphi(u)\right) n_t(u,\mathrm{d}v). \tag{46}$$

Further, since $\mathcal{A}(t)$ is characterized by (b_t, n_t), the Lax equation (45) expands into evolutions equations for b_t and n_t. These are a family of kinetic equations of Vlasov-Boltzmann type that describe the statistics of the clustering of shocks (see [25, §1.4]). Equation (42) is a special case of these kinetic equations.

The exact solution to Burgers equation with white noise is a stationary Markov process whose generator $\mathcal{A}(t)$ has characteristics

$$b_t(u) = \frac{1}{t}, \quad n_t(u,v) = \frac{1}{t^{1/3}} n_*(ut^{1/3}, vt^{1/3}). \tag{47}$$

The jump density of the integral operator is given explicitly as follows:

$$n_*(u,v) = \frac{J(v)}{J(u)}K(u-v), \quad u > v, \quad n_*(u,v) = 0, \quad u \le v. \tag{48}$$

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

$$j(q) = \int_{-\infty}^{\infty} \mathrm{e}^{-qu} J(u)\mathrm{d}u, \quad k(q) = \int_0^{\infty} \mathrm{e}^{-qs} K(s)\mathrm{d}s, \tag{49}$$

are given explicitly by

$$j(q) = \frac{1}{\mathrm{Ai}(q)}, \quad k(q) = -2\frac{\mathrm{d}^2}{\mathrm{d}q^2}\log \mathrm{Ai}(q). \tag{50}$$

The formula for $k(q)$ is strongly suggestive of determinantal formulas in the theory of integrable systems, in particular Dyson's formula for the Gelfand-Levitan-Marchenko equation [20].

5.2 Random Matrix Theory Revisited

Let us now connect Burgers turbulence and random matrix theory. Here are some features common to both models.

1. Both yield Burgers equation in the spectral plane (equations (1) and (39) respectively).

2. Both g and ψ are characterized by positive measures ν_t and Λ_t respectively: g through the Pick (or Herglotz) representation (34) and ψ through the Lévy-Khintchine formula (41).

3. These measures evolve through kinetic equations – Kerov's kinetic equation for ν_t and Smoluchowski's coagulation equation for λ_τ.

4. These kinetic equations are scaling limits of fundamental stochastic process: Plancherel growth and the additive coalescent respectively [2, 5, 19].

5. We have found a self-similar solution (3)–(4) to the initial value problem (1)–(2) and the self-similar solution (40) to equation (39) with $\psi_0(q) = q^2$. These correspond to Wigner's semicircle law and the solution to Burgers equation with Brownian motion as initial velocity. These solutions are related by a simple transformation in spectral variables.

$$g_*(z) = \frac{\psi_*(q)}{q}, \quad z = 2 + \frac{1}{q}, \quad \text{or} \quad \frac{g(z\sqrt{t},t)}{\sqrt{t}} = \frac{\psi(q/t,t)}{q/t}. \tag{51}$$

Is this coincidence or something deeper? My view is that the common aspects of the models suggest deep links between probability theory and integrable systems that remain to be explored. The methods of integrable systems have played a critical role in the description of new limit laws in the theory of random matrices, particularly important results being the work of the Kyoto school, followed by Tracy and Widom [18, 27]. Yet there still seems to be room for new perspectives that provide an intrinsic probabilistic explanation for the utility of methods from integrable systems in these models.

By an intrinsically probabilistic perspective we mean the following: we seek natural symplectic structures on spaces of probability measures, completely integrable flows of probability measures with respect to this symplectic structure, and symplectic transformations that link different probabilistic models. The deformation of probability measures by the deterministic transformation of sample paths is of fundamental importance in probability theory. For example, the Girsanov theorem describes the Jacobian of the transformation of Wiener measure under a deterministic shift of each sample path. Other important examples of path transformation such as Vervaat's transformation arise in the theory of Lévy processes with one-sided jumps and the study of random trees [5]. Observe now that the scalar conservation law (43) induces a deformation of the law of u_0 through a nonlinear transformation of each sample path. Our recent work suggests that the Lax equation (45) is a completely integrable Hamiltonian system with respect to a probabilistically natural symplectic structure [23]. Moreover, it is intimately tied to several classical integrable systems such as geodesic flow on a Lie group with a Manakov metric. This is still not fully understood: what we have shown in [23] is that a natural discretization of (45) carries the natural symplectic structure of a coadjoint orbit of a Lie group and that the discrete system can be linearized by the Adler-Kostant-Symes theorem [1]. Nevertheless, these results strongly suggest that the path space of Markov processes with bounded variation and downward jumps on the line carries a natural symplectic structure and deformations of the probability measure induced by the scalar conservation law (43) commute for different f.

The coincidences enumerated above suggest to us quite strongly that the flow of probability measures induced by the limit of Plancherel growth is also completely integrable. Moreover, we expect that there is a symplectic transformation of probability measures that links Burgers turbulence and Plancherel growth. Partial evidence is provided by the following: formally one can introduce a symplectic structure for (1) using the inverse spectral theorem of Stieltjes. Roughly, the Pick measure ν_t is in bijective correspondence with a Jacobi matrix, and the space of Jacobi matrices carries a natural symplectic structure [9]. We conjecture that (i) Kerov's kinetic equation is Hamiltonian with this symplectic structure; (ii) the transformation (51) is a particular case of a symplectic transformation that carries solutions of (35) to (42); (iii) the underlying deformation of stochastic processes are itself related through symplectic transformations. It remains to be seen if these conjectures are true.

Acknowledgements The author thanks the Department of Mathematics at the University of Auckland for hospitality during the completion of this manuscript.

References

[1] Adler M, Van Moerbeke P, Vanhaecke P. Algebraic integrability, Painlevé geometry and Lie algebras. Springer-Verlag, 2004

[2] Aldous D J. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. Bernoulli, 1999, **5**: 3–48

[3] Anderson G, Guionnet A, Zeitouni O. An Introduction to Random Matrices. Cambridge University Press, 2010

[4] Bertoin J. The inviscid Burgers equation with Brownian initial velocity. Comm Math Phys, 1998, **193**: 397–406

[5] Bertoin J. Some aspects of additive coalescents//Proceedings of the International Congress of Mathematicians, Beijing 2002, vol III. Beijing Higher Edu Press, 2002: 15–23

[6] Burgers J M. The Nonlinear Diffusion Equation. Dordrecht: Reidel, 1974

[7] Carraro L, Duchon J. Solutions statistiques intrinsèques de l'équation de Burgers et processus de Lévy. C R Acad Sci Paris Sér I Math, 1994, **319**: 855–858

[8] Chabanol M -L, Duchon J. Markovian solutions of inviscid Burgers equation. J Statist Phys, 2004, **114**: 525–534

[9] Deift P. Orthogonal Polynomials and Random Matrices: a Riemann–Hilbert approach, vol 3. New York Univ Courant Inst, 1999

[10] Donoghue W. Monotone Matrix Functions and Analytic Continuation. Berlin, New York: Springer, 1974

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

[12] E W, Sinaĭ Y G. New results in mathematical and statistical hydrodynamics. Uspekhi Mat Nauk, 2000, **55**: 25–58

[13] Feynman R, Leighton R, Sands M, et al. The Feynman Lectures on Physics, vol 1. Reading, MA: Addison-Wesley, 1964

[14] Frachebourg L, Martin P A. Exact statistical properties of the Burgers equation. J Fluid Mech, 2000, **417**: 323–349

[15] Groeneboom P. Brownian motion with a parabolic drift and Airy functions. Probab Theory Related Fields, 1989, **81**: 79–109

[16] Guionnet A. Large random matrices: Lectures on macroscopic asymptotics//École d'Eté Probabilités de Saint-Flour XXXVI-2006, vol 1957 of Lecture Notes in Math. Berlin: Springer, 2009: 1–310

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

[18] Jimbo M, Miwa T, Môri Y, Sato M. Density matrix of an impenetrable Bose gas and the fifth Painleve transcendent. Physica D, 1980, **1**: 80–158

[19] Kerov S V. Asymptotic Representation Theory of the Symmetric Group and Its Applications in Analysis. Providence RI: American Mathematical Society, 2003

[20] Lax P. Functional Analysis. John Wiley and Sons, 2002

[21] McKean Jr H P. Stochastic Integrals. New York: Academic Press, 1969

[22] Mehta M. Random Matrices, vol 142. Academic Press, 2004

[23] Menon G. Complete integrability of shock clustering and burgers turbulence. Arch Ration Mech Anal, 2011, doi 10.1007/S00205-011-0461-8

[24] Menon G, Pego R L. Universality classes in Burgers turbulence. Comm Math Phys, 2007, **273**: 177–202

[25] Menon G, Srinivasan R. Kinetic theory and Lax equations for shock clustering and Burgers turbulence. J Stat Phys, 2010, **140**: 1195–1223

[26] Morawetz C, Serrin J, Sinaĭ Y. Selected Works of Eberhard Hopf: With Commentaries. American Mathematical Society, 2002

[27] Tracy C A, Widom H. Distribution functions for largest eigenvalues and their applications//Proceedings of the International Congress of Mathematicians, Vol I (Beijing, 2002). Beijing: Higher Edu Press, 2002: 587–596

[28] Voiculescu D. Addition of certain noncommuting random variables. J Funct Anal, 1986, **66**: 323–346

[29] Taub A H, ed. John von Neumann. Collected Works. Vol VI: Theory of Games, Astrophysics, Hydrodynamics and Meteorology. New York: Pergamon Press, Macmillan Co, 1963