Arch. Rational Mech. Anal.
Digital Object Identifier (DOI) 10.1007/s00205-011-0461-8

Complete Integrability of Shock Clustering and Burgers Turbulence

Govind Menon

Communicated by W. E

Abstract

We consider scalar conservation laws with convex flux and random initial data. The Hopf–Lax formula induces a deterministic evolution of the law of the initial data. In a recent article, we derived a kinetic theory and Lax equations to describe the evolution of the law under the assumption that the initial datum is a spectrally negative Markov process. Here we show that: (i) the Lax equations are Hamiltonian and describe a principle of least action on the Markov group; (ii) the Lax equations are completely integrable and linearized via a loop-group factorization of operators; (iii) the associated zero-curvature equations can be solved via inverse scattering. Our results are rigorous for N-dimensional approximations of the Lax equations, and yield formulas for the limit $N \to \infty$. The main observation is that the Lax equations and zero-curvature equations are a Markovian analog of known integrable systems (geodesic flow on Lie groups and the N-wave model respectively). This allows us to introduce a variety of methods from the theory of integrable systems.

1. Introduction

1.1. Turbulence and Flows of Probability Measures

A fundamental problem in the statistical theory of turbulence is to construct random incompressible velocity fields that model isotropic homogeneous turbulence. Such random fields must be supported on weak solutions to the Euler equations that dissipate kinetic energy in accordance with the criterion of Kolmogorov and Onsager [22,25,35,41]. This problem is currently out of reach. Much of our understanding is based, instead, on vastly simplified models. One such model, proposed by Burgers, is to understand the statistics of the Cole–Hopf (or entropy) solution to Burgers' equation

Supported by NSF grant DMS 07-48482.

Published online: 13 September 2011

$$\partial_t u + \partial_x \left(\frac{u^2}{2}\right) = 0, \qquad x \in \mathbb{R}, \quad t > 0, \tag{1}$$

with random initial data u_0 such as white noise [18,43]. A closely related problem is to consider (1) with random forcing, and to understand the associated equilibrium measure. While we focus on the unforced equation in this article, some of our methods also apply to (1) with random forcing [38]. This class of problems is called Burgers-KPZ turbulence. An explanation of its role in statistical hydrodynamics may be found in [23].

Random initial data also arise in applications unrelated to turbulence. For example, the coarsening of domains in the kinetics of phase transitions is modeled by solutions to the Allen-Cahn and Cahn-Hilliard equations that emerge from disorder [37]. In such problems, the equations of continuum physics induce an evolution of the law of the initial data, and it is of basic interest to understand this evolution. The first clear formulation of an evolution equation for the law of solutions is due to Hopf [30]. In this article, we show that this problem is surprisingly rich, even in the setting of perhaps the simplest nonlinear equations.

We consider the scalar conservation law

$$\partial_t u + \partial_x f(u) = 0, \quad x \in \mathbb{R},\ t > 0, \quad u(x,0) = u_0(x), \tag{2}$$

with a strictly convex, C^1 flux f. When the initial data is random, the Hopf–Lax formula defines a unique entropy solution to (2) for each sample path u_0. The overall effect is a nonlinear evolution of the law of u_0. Our main contribution is to show that if u_0 is a Markov process in x with only downward jumps (or a limit of such processes), then the *evolution of its law is completely integrable*.

Our results include Burgers model, but the assumption $f(u) = u^2/2$ is not necessary. Neither do we require a special choice of initial condition such as white noise or Brownian motion, though these yield important exact solutions. For all convex, $C^1 f$ and a broad class of random initial data, the evolution of the law of $u(\cdot,t)$ is given by kinetic equations for shock clustering that are a continuum limit of a Markovian variant of the N-wave model. The N-wave model generalizes the three-wave model of Manakov and Zakharov in nonlinear optics [47] and is well-known to be completely integrable [3, p. 55], [48, III.4]. We show that this model also underlies Burgers turbulence and shock clustering for (2) with arbitrary convex f. We stress that it is the evolution of the *law* of $u(\cdot,t)$ that is integrable, and this is completely distinct from the integrability of (1).

This is surprising enough, but more is true. Our kinetic equations sit at a rich juncture of problems: (i) geodesic flows on Lie groups [36]; (ii) the integrable systems of the 19th century, as recast by Moser [39]; (iii) the completely integrable systems of random matrix theory discovered by the Kyoto school [31]; (iv) integrable hierarchies on groups [45]; and (v) asymptotic problems in representation theory [33]. This reveals a close and unexpected relation between the theory of Markov processes, kinetic theory and integrable systems.

This confluence of ideas is quite bewildering, and a full explanation for their role in what should be a purely probabilistic problem still eludes us. Some heuristic explanation is perhaps the following: at its heart, complete integrability is an explicit understanding of the hidden symmetries of a Hamiltonian system. On the other hand, understanding continuous deformations of the law of a stochastic process induced by a transformation of each sample path of the process is a basic problem in probability theory. (For example, the Girsanov theorem may be viewed in this light). The deeper principles here seem to be that: (i) the space of Feller processes with bounded variation on the Skorokhod space can be given a natural symplectic structure; (ii) for every convex, C^1 flux f (2) induces a Hamiltonian flow on this space with respect to this symplectic structure; (iii) these flows commute for distinct f. A precise formulation of these ideas is subtle, and we have been unable to develop this viewpoint completely, even though we obtain partial results. In order to state precisely what we prove, and what is mere conjecture, we first review some recent work.

1.2. Lax Equations for Shock Clustering with Markov Data

We always assume that random initial data u_0 for (2) is a Markov process in x with only downward jumps (a *spectrally negative Markov process*) or a limit of such processes. This assumption is motivated by two considerations. First, in order to obtain a detailed understanding of the evolution of the law of u_0 under (2), it is necessary to work with a class of well-understood random processes on the line and it is natural to choose Markov processes. Second, it is a surprising fact that for every $t > 0$ the solution to Burgers equation with white noise is a stationary, spectrally negative Markov process in x. The Markov property of this solution was assumed by Burgers, and first proved (in a completely different context) by Groeneboom [29]. The importance of the Markov property in the context of Burgers turbulence was first noted by Avallaneda and E [11].

In a recent article, we proved the following *closure theorem* for the entropy solution to (2): assume the initial data $u_0(x)$ is a spectrally negative strong Markov process in x. Then for every $t > 0$ the entropy solution to (2) remains a spectrally negative Markov process in x [38, Thms.2,3]. This shows that the entropy solution to (2) leaves this class of stochastic processes invariant. It is not necessary to assume that $f(u) = u^2/2$, only that f is convex and C^1. There is also no need to assume that u_0 is stationary in x. Under an additional assumption of regularity (preservation of the Feller property), the closure theorem forms the basis for a kinetic theory of shock clustering as follows.

Feller processes are characterized by their generators. The simplest such characterization is the Lévy-Khintchine formula for Lévy processes. A general characterization, attributed to Courrége in [8, Thm 3.5.3] builds on the Lévy-Khintchine formula. Assume $u(x)$, $x \in \mathbb{R}$ is a stationary, Feller process. Then its generator $\mathcal{A}$ is an integro-differential operator that acts on C_c^∞ test functions in its domain as follows:

$$\begin{aligned}\mathcal{A}\varphi(u) &= a(u)\varphi''(u) + b(u)\varphi'(u) + c(u)\varphi(u)\\ &\qquad + \int_{\mathbb{R}\backslash\{u\}} \big(\varphi(v) - \varphi(u) - \psi(u,v)\varphi'(u)\big)\, n(u, \mathrm{d}v).\end{aligned} \tag{3}$$

Here a, b, c are functions on the line satisfying a certain continuity criterion, $n(u, \mathrm{d}v)$ is a Lévy measure that describes the jumps of the process, and $\psi(u, v)$ is a local unit (in the simplest situation, we have $\psi(u, v) = v - u$). The precise assumptions are stated in [8, Thm. 3.5.3]. There is an intimate relation between this characterization and the sample paths of the Feller process: a describes the diffusion of the Feller process, so that $a \geqq 0$; c describes killing, so that $c \leqq 0$; b describes the drift, and its sign is not restricted.

Our basic idea, following [21], is to study the shock statistics through an evolution equation for the generator. The general formula (3) simplifies because for any $t > 0$, the entropy solution to (2) has bounded variation, only downward jumps, and no killing. Thus a and c vanish, and the support of $n(u, \mathrm{d}v)$ is $(-\infty, u)$. For fixed $t > 0$, if $u(x, t)$, $x \in \mathbb{R}$ is a stationary Feller process, its generator $\mathcal{A}(t)$ is an integro-differential operator of the form

$$\mathcal{A}(t)\varphi(u) = b(u, t)\varphi'(u) + \int_{-\infty}^{u} (\varphi(v) - \varphi(u))\, n(u, \mathrm{d}v, t). \tag{4}$$

One of the main results in [38] is that $\mathcal{A}(t)$ satisfies the Lax equation

$$\partial_t \mathcal{A} = [\mathcal{A}, \mathcal{B}]. \tag{5}$$

Here $[\mathcal{A}, \mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}$ denotes the Lie bracket, and the operator $\mathcal{B}$ is defined by its action on test functions as follows:

$$\mathcal{B}\varphi(u) = -f'(u)b(u, t)\varphi'(u) - \int_{-\infty}^{u} [f]_{u,v}\, (\varphi(v) - \varphi(u))\, n(u, \mathrm{d}v, t). \tag{6}$$

$[f]_{u,v}$ is abbreviated notation for the Rankine–Hugoniot speed of a shock connecting states u and v.

$$[f]_{u,v} := \frac{f(v) - f(u)}{v - u}. \tag{7}$$

We do not need to assume that the process $u(\cdot, t)$ is stationary in x. Non-stationary data arise when we wish to model the spread of 'turbulent bursts' (that is localized data $A_0(x)$) or Riemann data (for example one-sided initial conditions as in [46]). These situations are described by the zero-curvature equation

$$\partial_t \mathcal{A} - \partial_x \mathcal{B} = [\mathcal{A}, \mathcal{B}]. \tag{8}$$

It requires considerable insight to realize that this approach is fruitful, and our work was greatly inspired by Duchon et al. [19–21]. Several open questions remain: in particular, it is still necessary to justify the connection between (2) with random data and (8) in full generality [38, §1.6]. Our earlier work showed only that the entropy solution preserves the strong Markov property, whereas we really need to show that it preserves the Feller property. This requires a well-posedness theory and uniform estimates for (8). One of our goals in this article is to lay a foundation for the analysis of (8) that also allows us to address this question.

1.3. Kinetic Equations

To convince the reader of the merit of this approach, let us briefly explain how it describes the evolution of shock statistics. Since the generator $\mathcal{A}$ is characterized by the drift, b, and jump measure, n, the Lax equation can be expanded using (4) and (6) to yield evolution equations for b and n. The drift satisfies the simple differential equation

$$\partial_t b(u,t) = -f''(u)b^2(u,t). \tag{9}$$

The evolution of n is more interesting, since the shock statistics evolve by decay of rarefaction waves and growth in binary collisions. To state the evolution equations, we assume for simplicity that the jump measure has a density, say $n(u, \mathrm{d}v, t) = n(u, v, t)\,\mathrm{d}v$. Then n satisfies a kinetic equation of Vlasov–Boltzmann type

$$\begin{aligned}&\partial_t n(u,v,t) + \partial_u \left(nV_u(u,v,t)\right) + \partial_v \left(nV_v(u,v,t)\right)\\ &\quad = Q(n,n) + n\left(\left([f]_{u,v} - f'(u)\right)\partial_u b - bf''(u)\right).\end{aligned} \tag{10}$$

Here the drift velocities V_u and V_v are defined by

$$V_u(u,v,t) = \left([f]_{u,v} - f'(u)\right)b(u,t), \quad V_v(u,v,t) = \left([f]_{u,v} - f'(v)\right)b(v,t), \tag{11}$$

and the collision kernel Q counts growth and loss in binary clustering

$$\begin{aligned}Q(n,n)(u,v,t) = &\int_v^u \left([f]_{u,w} - [f]_{w,v}\right) n(u,w,t)n(w,v,t)\,\mathrm{d}w\\ &- \int_{-\infty}^v \left([f]_{u,v} - [f]_{v,w}\right) n(u,v,t)n(v,w,t)\,\mathrm{d}w\\ &- \int_{-\infty}^u \left([f]_{u,w} - [f]_{u,v}\right) n(u,v,t)n(u,w,t)\,\mathrm{d}w.\end{aligned} \tag{12}$$

The distinction with earlier work on kinetics of shock clustering [24,34] is summarized in [38].

1.4. Exact Solutions

Despite their formidable appearance, these kinetic equations admit surprising exact solutions for Burgers' equation. The first class of solutions correspond to spectrally negative Lévy processes (for example when u_0 is a Brownian motion as in [44]). We then have $\mathcal{A}\mathrm{e}^{qy} = \psi(q,t)\mathrm{e}^{qy}$ and $[\mathcal{A},\mathcal{B}]\mathrm{e}^{qy} = -\psi\partial_q\psi(q,t)$, where $\psi(q,t)$ denotes the Laplace exponent of the process $u(x,t) - u(0,t)$, $x \geqq 0$, for fixed $t > 0$. Then the Lax equation (5) yields

$$\partial_t\psi + \psi\partial_q\psi = 0, \quad t > 0,\ q > 0. \tag{13}$$

The beautiful fact that the Laplace exponent itself evolves by Burgers' equation was discovered by Carraro and Duchon [19], and made rigorous by Bertoin

[16]. It is reminiscent of the inverse scattering method in integrable systems. In addition, (10) reduces to Smoluchowski's coagulation equations with additive kernel. It has been known for some time that Smoluchowski's equation can be solved explicitly by the Laplace transform [28]. But this solution takes on new meaning when we recognize that it describes exactly the clustering of shocks in Burgers' equations with Lévy process data. This connection is the basis for several deeper results connecting stochastic coalescence and Burgers' turbulence [17].

Another remarkable solution to (5) corresponds to the shock statistics in Burgers' equation with white noise initial data [27,29]. This corresponds to a self-similar solution to (9) and (10) of the form

$$b(u,t) = \frac{1}{t}, \quad n(u,v,t) = \frac{1}{t^{1/3}} n_*(ut^{1/3}, vt^{1/3}). \tag{14}$$

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

$$n_*(u,v) = \frac{J(v)}{J(u)} K(u-v), \quad u > v, \tag{15}$$

and vanishes if $u \geqq v$. Here J and K are positive functions defined on the line and positive half-line respectively, whose Laplace transforms

$$j(q) = \int_{-\infty}^{\infty} \mathrm{e}^{-qy} J(y)\,\mathrm{d}y, \quad k(q) = \int_0^{\infty} \mathrm{e}^{-qy} K(y)\,\mathrm{d}y, \tag{16}$$

are meromorphic functions on $\mathbb{C}$ given by

$$j(q) = \frac{1}{\mathrm{Ai}(q)}, \quad k(q) = -2\frac{d^2}{dq^2} \log \mathrm{Ai}(q). \tag{17}$$

Ai denotes the Airy function [4, 10.4]. This generator was computed by GROENEBOOM (but not as a solution to (5)!) [29]. When written this way, the formula for k is reminiscent of determinantal formulas in soliton theory.

1.5. Discrete Lax Equations and the Markov N-Wave Model

The most intriguing aspect of [38] is that the kinetic theory of shock clustering has many features reminiscent of completely integrable systems. These include the formulation as a Lax pair, explicit exact solutions as above, a Painlevé property for the self-similar solution, and links with random matrix theory (see [38, §1.5.1]). Our main contribution here is to understand the origin of these coincidences.

In order to show that (5) is completely integrable, we must show that it defines a Hamiltonian system with respect to a suitable Poisson or symplectic structure and construct infinitely many commuting integrals. In addition, what is required is an explicit solution via inverse scattering or a Riemann–Hilbert problem. In order to address these questions, we first study exact finite-dimensional discretizations of (5) (this terminology is explained below). Here the definition of Hamiltonian structure and complete integrability are unambiguous, and it has to be shown that (5) actually has this structure. This lays a foundation for the analysis of (5) and (8).

We discretize the system as follows. Fix a positive integer N and discrete velocities $-\infty < u_1 < u_2 < \cdots < u_N < \infty$. We consider a continuous x Markov processes that takes the values u_k, $1 \leqq k \leqq N$. The sample paths of this Markov process are piecewise constant paths. Let $\mathbb{M}_N$ denote the space of $N \times N$ matrices. The generator $A \in \mathbb{M}_N$ satisfies

$$A_{ij} \geqq 0, \quad i \neq j, \quad \sum_{j=1}^{N} A_{ij} = 0. \tag{18}$$

More formally, A corresponds to an operator $\mathcal{A}$ with drift $b \equiv 0$ and a jump measure $n(u_j, \mathrm{d}v) = \sum_{k \neq j} A_{jk} \delta_{u_k}(\mathrm{d}z)$. The jump measure vanishes if u is not one of the discrete states u_j. In order to define the discretization B of $\mathcal{B}$ we introduce the symmetric matrix

$$F_{ij} = -\frac{f(u_i) - f(u_j)}{u_i - u_j}, \quad i \neq j, \quad F_{ii} = -f'(u_i). \tag{19}$$

Here f is the flux in the conservation law (2), so that $-F_{ij}$ is the Rankine–Hugoniot speed of the shock connecting states u_i and u_j. We then set

$$B_{ij} = F_{ij} A_{ij}, \quad i \neq j, \quad B_{ii} = -\sum_{j \neq i} B_{ij}. \tag{20}$$

The discrete zero-curvature equation is

$$\partial_t A - \partial_x B = [A, B], \tag{21}$$

and the discrete Lax equation is

$$\dot{A} = [A, B]. \tag{22}$$

Let us briefly comment on why this discretization is natural. Initial data to (21) correspond to random initial data u_0 to (2). In particular, initial data to (21) that are lower triangular corresponds to u_0 that are spectrally negative Markov processes with a finite number of velocities. For such initial data, the Feller property is preserved in time and the evolution of statistics of random initial data by the conservation law (2) is described exactly by (21). Thus, solutions to (21) that are lower triangular and generators of Markov processes correspond to exact solutions to (8). This is what we mean when we say that the discretization is exact.

When we consider (22), complete integrability immediately becomes plausible. Indeed, ordinary differential equations of the form (22) arise in basic examples in the theory of integrable systems: these include geodesic flows on $SO(N)$ [36], the integrable flows of Neumann and Jacobi [39], and the integrable flows of random matrix theory [31]. Moreover, the zero-curvature equations (21) are very similar to the N-wave model [3,48] with an important difference. For the N-wave model we typically assume $A \in \mathfrak{u}$, the algebra of the unitary group. (Perhaps the only study where A is *not* assumed to lie in $\mathfrak{u}$ is [1]). But $A \in \mathfrak{u}$ is incompatible with the Markov property. Motivated by these considerations, we call (21) the Markov N-wave model, or mN-wave model for short.

1.6. Statement of Results

A complete study of (5) and (8) requires a combination of methods from the theory of Markov processes, integrable systems and spectral theory. Our approach is to prove complete results in the finite-dimensional setting in this article, and to study the limit $N \to \infty$ in a sequel. Once one has recognized the structure of the problem in the discrete setting, the proofs only rely on well-established techniques from integrable systems. (Of course, the main difficulty in our work was to recognize this structure!) Our main results are:

1. (22) is Hamiltonian and is associated to a principle of least action (see Section 2 and Theorem 1).
2. (22) is completely integrable and linearized by a loop-group factorization (see Section 3 and Theorem 3).
3. an inverse scattering theory and well-posedness for (21) (see Section 4, in particular Theorem 9).

Though obviously related, the method of solution for (5) and (8) turn out to be quite distinct. The Lax equation (5) can be studied through an operator factorization problem, while the zero-curvature equation (8) can be attacked by inverse scattering.

Let us briefly explain how these results are proved. Hamiltonian structures and complete integrability are deeply linked to group actions. In seeking an algebraic, but probabilistically natural, approach to (5), we find the following simple structure. We observed in [38, §2.7] that the space $\mathfrak{m}_\infty$ of integro-differential operators $\mathcal{C}$ of the form

$$\mathcal{C}\varphi(u) = \beta(u)\varphi'(u) + \int_{\mathbb{R}} (\varphi(v) - \varphi(u))\, \nu(u,v)\, \mathrm{d}v, \quad \beta \in C_c^\infty(\mathbb{R}),\, \nu \in C_c^\infty(\mathbb{R}^2), \tag{23}$$

formally constitutes a Lie algebra. Of course, it is not clear that these operators generate an infinite-dimensional Lie group. But we may use this insight at the discrete level quite easily. The set of generators of Markov processes defined by (18) forms a cone $\mathfrak{q}_N \subset \mathbb{M}_N$. Each element of $\mathfrak{q}_N$ generates a one-parameter Markov semigroup. Let us define

$$\mathfrak{Q}_N = \{g \in GL(N,\mathbb{R}) | \quad g = \mathrm{e}^{xA},\ A \in \mathfrak{q}_N,\ x \geqq 0\}. \tag{24}$$

$\mathfrak{Q}_N$ can be naturally embedded in a Lie group as follows. Define the *Markov algebra*

$$\mathfrak{m}_N = \left\{ A \in \mathbb{M}_N | \quad \sum_{j=1}^N A_{ij} = 0, \quad 1 \leqq i \leqq N \right\}. \tag{25}$$

$\mathfrak{m}_N$ is a Lie algebra. It generates the *Markov group*

$$\mathfrak{M}_N = \{g \in GL(N,\mathbb{R}) | \quad g = \mathrm{e}^{xA}, \quad A \in \mathfrak{m}_N,\ x \in \mathbb{R}\}. \tag{26}$$

Since N will be fixed in our results, we will mostly suppress the subscript N in what follows. Clearly $\mathfrak{q} \subset \mathfrak{m}$, and $\mathfrak{Q} \subset \mathfrak{M}$. Some basic properties of $\mathfrak{m}$ and $\mathfrak{M}$ may be found in [32].

The first result is that (22) is a Hamiltonian flow on $\mathfrak{m}$. In addition, when f is convex and A lower triangular, this Hamiltonian flow leaves $\mathfrak{q}$ invariant. These results are seen as follows. Co-adjoint orbits of Lie groups carry a natural (Kirillov-Kostant) symplectic structure. In addition, Lie algebras that admit direct sum decompositions into subalgebras carry more than one symplectic structure. This idea has been formalized by the notion of an r-matrix [42]. The Lie algebra $\mathfrak{g} = gl(N, \mathbb{R}) = \mathbb{M}_N$ admits a natural splitting induced by $\mathfrak{m}$ (see (31) below). We show that (22) is a Hamiltonian system on the algebra $\mathfrak{g}$ with a Lie-Poisson bracket induced by this splitting. This is stated precisely in Theorem 1 below. Once we have established the Hamiltonian structure of (22) we treat the probabilistically important case ($A \in \mathfrak{q}$) in Theorem 2. The proof shows clearly the role of convexity of f and spectral negativity at the level of the Lax equations.

We then formulate an associated principle of least action for (22). This is in precise analogy with geodesic flow on $SO(N)$ and in particular, with Euler's equation for a free rigid body. When F_{ij} is positive, the role of the flux f is to define a (degenerate) metric on $\mathfrak{M}$. More generally, f defines a quadratic action through the multiplier F_{ij}. This principle of least action describes the evolution of a probability measure on path space through (22), and is completely distinct from the usual principle of least action for the Hopf–Lax solution to (2).

Complete integrability is based on an elegant observation of Manakov [36]. Define the diagonal matrices

$$\mathcal{M} = \mathrm{diag}(u_1, \ldots, u_N), \quad \mathcal{N} = \mathrm{diag}(f(u_1), \ldots, f(u_N)), \tag{27}$$

and observe that A and B are related through the algebraic relation

$$[A, \mathcal{N}] - [\mathcal{M}, B] = 0. \tag{28}$$

This allows us to introduce a spectral parameter $z \in \mathbb{C}$ in (22) and embed the flow in a loop-algebra

$$\frac{\mathrm{d}}{\mathrm{d}t}(A - z\mathcal{M}) = [A - z\mathcal{M}, B + z\mathcal{N}], \quad z \in \mathbb{C}. \tag{29}$$

Thus, the spectral curve $\{(z, \lambda) \in \mathbb{C}^2 | \det(A - z\mathcal{M} - \lambda I) = 0\}$ is invariant, and its coefficients are integrals. Rather than verify explicitly that these integrals are in involution as in [36], it is simpler to apply the Adler–Kostant–Symes (AKS) theorem to show that (29) is completely integrable. Again we need to find a suitable r-matrix, now of the loop algebra. This is stated precisely in Theorem 3. The matrix factorization in the AKS theorem also yields a Riemann–Hilbert problem. Once this is formulated for finite N, it also yields a limiting factorization problem for the Lax equation (5).

Scattering and inverse scattering for a class of integrable systems including the N-wave model on $\mathfrak{u}$ was established by Beals and Coifman [13,14] and formulated in a general Lie algebraic setting by Beals and Sattinger [15]. This theory does not directly apply to (21) since $\mathfrak{m}$ is not semi-simple. We modify [13,14] to obtain a scattering and inverse scattering theory for (21). Among other results, we obtain a hierarchy of integrable flows and global well-posedness theorems for (21) including the probabilistically natural case. These results are contained in Section 4. In addition to these rigorous results, the method also yields a formal solution procedure for (8) via inverse scattering.

In the limit $N \to \infty$, $\mathcal{A}$ is an integro-differential operator of the form (4). Subtle problems arise in the approximation: under suitable assumptions, an operator of the form (3) can be approximated by operators in $\mathfrak{m}_N$. However, operators of the general form (3) are not closed under the commutator! In addition, while operators of the form (23) form a Lie algebra, it is not clear that they correspond to an infinite-dimensional Lie group. Nevertheless, our work yields a formal understanding of (5) and (8), and thus (2) with spectrally negative Markov data. Theorem 1 suggests that (5) is Hamiltonian with an associated principle of least action on the semigroup of Markov operators. Theorem 3 suggests that (5) is completely integrable, and yields an operator factorization problem for (5). Finally, the inverse scattering problem for (8) extends formally to unbounded operators with little change. We hope that these natural conjectures will stimulate rigorous results of full generality.

2. Hamiltonian Structure

In this section, we show that the Lax equation (22) defines a Hamiltonian system. As is well-known, co-adjoint orbits of a Lie group carry a natural symplectic structure [7,10]. If $\mathfrak{g}$ denotes a finite-dimensional Lie algebra, and H a smooth Hamiltonian defined on the dual space $\mathfrak{g}^*$, then Hamilton's equations may be rewritten as Kirillov's equation

$$\dot{\alpha} = \mathrm{ad}^*_{dH(\alpha)}(\alpha), \quad \alpha \in \mathfrak{g}^*. \tag{30}$$

Our main observation is that the Lax equation (22) takes this form on the algebra $\mathfrak{g}$ with a probabilistically natural bracket defined in (34) below.

2.1. Algebraic Preliminaries

Let $\mathfrak{g} = gl(N, \mathbb{R})$ denote the Lie algebra of real, $N \times N$ matrices equipped with the bracket $[A, B] = AB - BA$. We have already defined the subspace $\mathfrak{m} \subset \mathfrak{g}$ in (25). It is easily checked that $\mathfrak{m}$ is a subalgebra of $\mathfrak{g}$. Let $\mathfrak{d}$ denote the subspace of diagonal matrices. Since diagonal matrices commute, $\mathfrak{d}$ is trivially a subalgebra. Its importance here lies in the fact that $\mathfrak{g}$ admits a direct sum (vector space) decomposition

$$\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{d}. \tag{31}$$

The decomposition (31) is obtained as follows. Recall that $e = (1, 1, \ldots, 1)^T$. Given $A \in \mathfrak{g}$, we define the projections

$$P_{\mathfrak{m}} A = A - \mathrm{diag}(Ae), \qquad P_{\mathfrak{d}} A = \mathrm{diag}(Ae). \tag{32}$$

Then $P_{\mathfrak{m}}^2 = P_{\mathfrak{m}}$, $P_{\mathfrak{d}}^2 = P_{\mathfrak{d}}$, $P_{\mathfrak{m}}A \in \mathfrak{m}$ and $P_{\mathfrak{d}}A \in \mathfrak{d}$, and every matrix $A \in \mathfrak{g}$ may be written as

$$A = P_{\mathfrak{m}}A + P_{\mathfrak{d}}A. \tag{33}$$

Associated to a splitting such as (31) is an r-matrix [42]. This allows us to introduce a new Lie bracket on $\mathfrak{g}$. If $A, B \in \mathfrak{g}$, then we define a new ad-action

$$\mathrm{ad}^r_B A = [B, A]_r = [P_{\mathfrak{m}}B, P_{\mathfrak{m}}A] - [P_{\mathfrak{d}}B, P_{\mathfrak{d}}A] = [P_{\mathfrak{m}}B, P_{\mathfrak{m}}A]. \tag{34}$$

The last equality holds because diagonal matrices commute. $\mathfrak{g}$ remains a Lie algebra with the new bracket $[\cdot, \cdot]_r$.

We identify $\mathfrak{g}$ with its dual space $\mathfrak{g}^*$ through the non-degenerate Ad-invariant pairing

$$(\alpha, A) = \mathrm{Tr}(\alpha A), \quad \alpha \in \mathfrak{g}^*, \ A \in \mathfrak{g}. \tag{35}$$

The dual spaces $\mathfrak{m}^*$ and $\mathfrak{d}^*$ are naturally identified with the orthogonal complements $\mathfrak{d}^\perp$ and $\mathfrak{m}^\perp$ under $(\cdot, \cdot)$. It is easy to compute

$$\mathfrak{m}^* \cong \mathfrak{d}^\perp = \{\alpha |\, \mathrm{diag}(\alpha) = 0\}. \tag{36}$$

Since $Ae = 0$ for every $A \in \mathfrak{m}$, where the dimension of $\mathfrak{m}$ is $N^2 - N$ we also find

$$\mathfrak{d}^* \cong \mathfrak{m}^\perp = \left\{ \alpha |\, \alpha = \sum_{j=1}^N c_j E_j = 0 \right\}, \tag{37}$$

where $c_j \in \mathbb{R}$, $1 \leqq j \leqq N$ and E_j is the matrix obtained from the zero matrix by replacing the jth column with e. The projections $P_{\mathfrak{m}}$ and $P_{\mathfrak{d}}$ induce dual projections $P_{\mathfrak{m}^\perp}$ and $P_{\mathfrak{d}^\perp}$ in the natural manner: for $A \in \mathfrak{g}$ and $\alpha \in \mathfrak{g}^*$

$$(P_{\mathfrak{d}^\perp}\alpha, A) = (\alpha, P_{\mathfrak{m}}A), \quad (P_{\mathfrak{m}^\perp}\alpha, A) = (\alpha, P_{\mathfrak{d}}A). \tag{38}$$

We now compute the ad* action with the bracket (34) and the non-degenerate pairing (35). For every $A, B \in \mathfrak{g}$ and $\alpha \in \mathfrak{g}^*$ we have

$$\begin{aligned}
\mathrm{ad}^{*r}_B \alpha(A) &= (\alpha, \mathrm{ad}^r_B A) = (\alpha, [P_{\mathfrak{m}}B, P_{\mathfrak{m}}A]) \\
&= \mathrm{Tr}\,(\alpha[P_{\mathfrak{m}}B, P_{\mathfrak{m}}A]) = \mathrm{Tr}\,([\alpha, P_{\mathfrak{m}}B]P_{\mathfrak{m}}A) \\
&= ([\alpha, P_{\mathfrak{m}}B], P_{\mathfrak{m}}A) = (P_{\mathfrak{d}^\perp}[\alpha, P_{\mathfrak{m}}B], A).
\end{aligned}$$

Since this holds for every $A \in \mathfrak{g}$, we find

$$\mathrm{ad}^{*r}_B \alpha = P_{\mathfrak{d}^\perp}[\alpha, P_{\mathfrak{m}}B]. \tag{39}$$

2.2. *Quadratic Hamiltonians and Kirillov's Equation*

The calculation so far has been purely algebraic and we have been careful to distinguish $\alpha \in \mathfrak{g}^*$ from $A \in \mathfrak{g}$. Since we have now computed the ad^{*r} action and $\mathfrak{g}$ has been identified with $\mathfrak{g}^*$ via $(,)$ we may drop this notation. We assume $H : \mathfrak{g} \to \mathbb{R}$ is C^1, replace α by A and B by $dH(A)$ in (39) to obtain Kirillov's equation with the bracket (34)

$$\dot{A} = \mathrm{ad}^{*r}_{dH(A)} A = P_{\mathfrak{d}^\perp}[A, P_{\mathfrak{m}} dH(A)] \tag{40}$$

These calculations have made the Hamiltonian structure precise. We now show that the Lax equation (22) is Hamiltonian with this symplectic structure.

Theorem 1. *Assume given a real symmetric matrix F, and let $F \circ A$ denote the Hadamard product $(F \circ A)_{ij} = F_{ij} A_{ij}$. Define the quadratic Hamiltonian $H : \mathfrak{g} \to \mathbb{R}$*

$$H(A) = \frac{1}{2}\mathrm{Tr}(AF \circ A) = \frac{1}{2}\sum_{i,j=1}^{N} F_{ij} A_{ij} A_{ji}. \tag{41}$$

(a) *The associated Hamiltonian vector field on $\mathfrak{g}$ with the bracket* (34) *is*

$$\dot{A} = [A, P_{\mathfrak{m}} F \circ A]. \tag{42}$$

(b) $\mathfrak{m}^\perp, \mathfrak{d}^\perp, \mathfrak{m}$, *and $\mathfrak{d}$ are invariant under* (42).

(c) *When F is defined by* (19)*, the vector field* (42) *is identical to* (22).

Proof. (a) Kirillov's equation on $\mathfrak{g}^*$ with the bracket (34) takes the form (40). We need only to show that

$$P_{\mathfrak{d}^\perp}[A, P_{\mathfrak{m}} dH(A)] = [A, P_{\mathfrak{m}} dH(A)] = [A, P_{\mathfrak{m}} F \circ A]. \tag{43}$$

The identity (43) is seen as follows. By (41), $dH(A) = F \circ A$ and

$$[A, dH(A)]_{ij} = \sum_{k=1}^{N} \left(F_{kj} - F_{ik}\right) A_{ik} A_{kj}. \tag{44}$$

Since F is symmetric, $[A, dH(A)]$ vanishes on the diagonal. In addition, since $P_{\mathfrak{d}} dH(A)$ is a diagonal matrix, $[A, P_{\mathfrak{d}} dH(A)]$ also vanishes on the diagonal. Thus, by (36)

$$P_{\mathfrak{d}^\perp}[A, dH(A)] = [A, dH(A)], \quad P_{\mathfrak{d}^\perp}[A, P_{\mathfrak{d}} dH(A)] = [A, P_{\mathfrak{d}} dH(A)],$$

which implies (43).

(b) Every vector field of the form (40) vanishes on $\mathfrak{m}^\perp$. Thus, all such vector fields (not just quadratic Hamiltonians) leave $\mathfrak{d}^\perp$ and $\mathfrak{m}^\perp$ invariant. If $A \in \mathfrak{d}$, then $P_{\mathfrak{m}} dH(A)$ vanishes. Thus, $\mathfrak{d}$ is invariant. Next, $\mathfrak{m}$ is a subalgebra. So if $A \in \mathfrak{m}$, then $[A, P_{\mathfrak{m}} dH(A)] \in \mathfrak{m}$ and $\mathfrak{m}$ is also invariant. □

Remark 1. The following subtlety should be noted in (b). $\mathfrak{d}^\perp$ and $\mathfrak{m}^\perp$ are Poisson subspaces, but $\mathfrak{d}$ and $\mathfrak{m}$ are not. A subspace $V \subset \mathfrak{g}$ is a Poisson subspace if and only if the restriction of every Hamiltonian vector field on $\mathfrak{g}$ to V is tangent to V [7, Prop. 3.33]. However, it is only vector fields corresponding to the quadratic Hamiltonians (41) that vanish on $\mathfrak{d}$ and $\mathfrak{m}$.

2.3. Convexity of f and Spectral Negativity

We defined the Markov group in order to give the Lax equations (22) the Hamiltonian structure stated in Theorem 1. We now return to the probabilistically interesting case. We consider the initial value problem for (22) with initial data that are generators of spectrally negative Markov processes with discrete states $-\infty < u_1 < \cdots < u_N < \infty$. In this discrete setting, spectral negativity simply means that the generator is a lower-triangular matrix. For brevity, let $\mathfrak{l}$ denote the algebra of lower-triangular matrices. We then have

Theorem 2. *Assume f is a C^1 convex flux and F is defined by (19). Then*

(a) *The vector field (22) leaves $\mathfrak{l}$ invariant.*
(b) *The vector field (22) leaves $\mathfrak{q} \cap \mathfrak{l}$ positively invariant.*
(c) *For every $A_0 \in \mathfrak{q} \cap \mathfrak{l}$, there is a unique, global, C^∞ solution $A : [0, \infty)$ with initial condition $A(0) = A_0$.*

Proof. (a) is easy. If $A \in \mathfrak{l}$, then $B \in \mathfrak{l}$ and $[A, B] \in \mathfrak{l}$. (This does not require convexity of f).

(b) Since we have already shown that the lower-triangular form is preserved, we need to show only that the flow preserves positivity. In order to establish positive invariance, we show that the vector field (22) 'points into' $\mathfrak{q} \cap \mathfrak{l}$ for a point on its boundary. A point on the boundary has $A \in \mathfrak{q} \cap \mathfrak{l}$ and $A_{ij} = 0$ for some indices (i, j) in the lower-triangular region (that is $i > j$). We claim that $\dot{A}_{ij} \geqq 0$ so that positivity is preserved. Indeed, by (22)

$$\dot{A}_{ij} = \sum_{k \neq i,j} \left(F_{kj} - F_{ik}\right) A_{ik} A_{kj} + \left(A_{ii} - A_{jj}\right) B_{ij} + \left(B_{jj} - B_{ii}\right) A_{ij}. \tag{45}$$

The second and third term vanish since $A_{ij} = 0$ and $B_{ij} = F_{ij} A_{ij} = 0$. We now consider the first term. Since A is lower-triangular, the sum extends only over $j < k < i$. For such k, $u_j < u_k < u_i$, and the convexity of f and (19) imply

$$F_{kj} - F_{ik} = \frac{f(u_i) - f(u_k)}{u_i - u_k} - \frac{f(u_k) - f(u_j)}{u_k - u_j} \geqq 0. \tag{46}$$

Finally, since $A \in \mathfrak{q} \cap \mathfrak{l}$ we also have $A_{ik} \geqq 0$ and $A_{kj} \geqq 0$. Thus, $\dot{A}_{ij} \geqq 0$.

(c) The vector field (22) vanishes on the diagonal, so that the diagonal terms are conserved. The diagonal terms control the off-diagonal terms since $A \in \mathfrak{q}$. Thus, we have the uniform bound

$$0 \leqq A_{ij}(t) \leqq |A_{ii}(t)| = |A_{ii}(0)|, \quad i > j, \quad t \geqq 0. \tag{47}$$

□

Remark 2. As in Remark 1, $\mathfrak{l}$ constitutes an invariant subspace, but not a Poisson subspace. In fact, if we set $F_{ii} = 0$ (which does not affect the flow), we see that the Hamiltonian $H = \mathrm{Tr}(AF \circ A)/2$ vanishes on the subalgebra of lower triangular Markov matrices.

2.4. A Principle of Least Action

In this section, we assume that F_{ij} does not vanish. Equation (20) then defines an endomorphism of $\mathfrak{m}$, and we have the following principle of least action. Let $g : [0, 1] \to \mathfrak{M}$ denote a C^1 path with given endpoints. Assume $F_{ij} \neq 0$, define the endomorphism $I : \mathfrak{m} \to \mathfrak{m}$, $B \mapsto A$ given by (20), and define the action

$$S[g] = \frac{1}{2} \int_0^1 \mathrm{Tr} \left(I \left(g^{-1} \dot{g} \right) g^{-1} \dot{g} \right) \, \mathrm{d}t. \tag{48}$$

Then the Euler–Lagrange equations for minimizing this action are

$$\dot{A} = [A, B], \quad \dot{g} = gB. \tag{49}$$

The second equation is a linear non-autonomous equation and can be integrated once we have solved the first, which is of course, (22).

It is easy to check this assertion, but the main point to observe is that when $F_{ij} > 0$, a similar endomorphism of $\mathfrak{so}(N)$ is used to define geodesic flow with respect to a left-invariant metric on $SO(N)$. Let us first recall these ideas. This will immediately explain the origin of (48).

We follow the notation of [7, Ch. 8, p. 265] (see also [9, Appendix 2]). Assume F is a symmetric matrix with strictly positive entries and let $I : \mathfrak{so}(N) \to \mathfrak{so}(N)$ denote the endomorphism $\omega \mapsto I(\omega)$ with $I(\omega)_{ij} = \omega_{ij}/F_{ij}$. In physical terms, ω is the angular velocity in the body frame and $X = I(\omega)$ is the angular momentum in the body frame. The letter I stands for the inertia tensor. Since $F_{ij} > 0$ and $\mathfrak{so}(N)$ is semi-simple, the quadratic form

$$(\omega_1, \omega_2)_F := \mathrm{Tr} \left(I(\omega_1) \omega_2 \right), \quad \omega_1, \omega_2 \in \mathfrak{so}(N) \tag{50}$$

is an inner-product on $\mathfrak{so}(N)$. The inner-product $(\cdot, \cdot)_F$ then defines a left-invariant metric on $SO(N)$ by left-translation. The length of a C^1 path $g : [0, 1] \to SO(N)$ with respect to this metric is given by

$$L[g] = \int_0^1 \sqrt{\left(g^{-1} \dot{g}, \, g^{-1} \dot{g} \right)_F} \, \mathrm{d}t. \tag{51}$$

The problem of minimizing the length is the same as that of minimizing the action

$$S[g] = \frac{1}{2} \int_0^1 \left(g^{-1} \dot{g}, \, g^{-1} \dot{g} \right)_F \, \mathrm{d}t. \tag{52}$$

The Euler–Lagrange equations for this variational principle are precisely the Euler equations:

$$\dot{X} = [X, \omega], \quad \dot{g} = g\omega. \tag{53}$$

If we replace the group $SO(N)$ with the Markov group $\mathfrak{M}$, the angular momentum X with A, and the angular velocity ω with B, then we have a flow on $\mathfrak{M}$ given by precisely (49). The analogy with geodesic flow on $\mathfrak{so}(N)$ is now clear.

In the calculations above, we did not need to assume that F_{ij} is of the form (19). Indeed, every symmetric matrix F with positive entries defines a diagonal left-invariant metric on $SO(N)$. Geodesic flow with this metric is Hamiltonian, but not necessarily completely integrable for $N \geqq 4$. However, Manakov discovered that when

$$F_{ij} = \frac{f_i - f_j}{u_i - u_j} \tag{54}$$

for two vectors $(f_1, \ldots, f_N)$, and $(u_1, \ldots, u_N)$, the geodesic flow is integrable. It is quite remarkable that we find ourselves in exactly this situation with the flux function f in (2) defining F as in (19).

The analogy with geodesic flow is incomplete in the following respect. First, we do not assume that $F_{ij} > 0$. Moreover, even when $F_{ij} > 0$, the metric on $\mathfrak{m}$ is degenerate since $\mathfrak{m}$ is not semi-simple. (The Killing form of $A, B \in \mathfrak{g}$ is $2N\mathrm{Tr}(AB) - 2\mathrm{Tr}(A)\mathrm{Tr}(B)$. This vanishes on the identity in $\mathfrak{g}$, and on $\mathfrak{m}$ if $A = \sum_{j=1}^{N} c_j E_j$ with $\sum_{j=1}^{N} c_j = 0$.)

3. Complete Integrability

Complete integrability of all the systems alluded to in Section 1.1 can be established in the unified framework of [5,6,40]. There are two distinct aspects to these studies: the first is to establish complete integrability via a suitable loop algebra splitting. The second is to explicitly linearize the flow on a Jacobi variety. Here we consider only the first aspect of the problem. We show that the Lax equation (22) defines a completely integrable Hamiltonian system. The proof is almost a textbook application of the Adler–Kostant–Symes (AKS) theorem and we follow the treatment in [7, §4.4]. Construction of the linearizing transformation is more difficult and will be considered in a separate article.

3.1. Integrability via the AKS Theorem

We introduce the loop algebra of formal finite Laurent expansions valued in $\mathfrak{g}$

$$L(\mathfrak{g}) = \left\{ X(z) = \sum_{m}^{n} A_k z^k, \, m, n \in \mathbb{Z} \;\; A_k \in \mathfrak{g} \right\}. \tag{55}$$

The natural Lie bracket on $L(\mathfrak{g})$ is given by

$$\left[\sum_{i \leqq n} A_i z^i, \sum_{j \leqq m} B_j z^j \right] = \sum_{k \leqq m+n} z^k \left(\sum_{i+j=k} [A_i, B_j] \right). \tag{56}$$

The sum includes only a finite number of terms by (55). We pair $\mathfrak{g}$ with $\mathfrak{g}^*$ via the non-degenerate, Ad-invariant pairing (35). There are then various Ad-invariant pairings that one may introduce on $L(\mathfrak{g})$. We use the pairing

$$\langle X | Y \rangle = \sum_{i+j=0} (X_i, Y_j) = \frac{1}{2\pi i} \oint_{|z|=1} \mathrm{Tr}(X(z)Y(z)) \frac{\mathrm{d}z}{z}. \tag{57}$$

It is easily checked that this pairing is non-degenerate and Ad-invariant.

The direct sum decomposition (31) also induces a decomposition of $L(\mathfrak{g})$. We define the subalgebras

$$L(\mathfrak{g})_+ = \left\{ X(z) = \sum_{k \geqq 0} A_k z^k, \quad A_0 \in \mathfrak{m}, \;\; A_k \in \mathfrak{g}, k \geqq 1 \right\}, \tag{58}$$

$$L(\mathfrak{g})_- = \left\{ X(z) = \sum_{k \leqq 0} A_k z^k, \quad A_0 \in \mathfrak{d}, \;\; A_k \in \mathfrak{g}, k \leqq -1 \right\}. \tag{59}$$

It is immediate from the calculations of Section 2.1 that

$$L(\mathfrak{g}) = L(\mathfrak{g})_+ \oplus L(\mathfrak{g})_-. \tag{60}$$

The respective projections are given by

$$X(z)_+ = P_{\mathfrak{m}} A_0 + \sum_{k \geqq 1} A_k z^k, \quad X(z)_- = P_{\mathfrak{d}} A_0 + \sum_{k \leqq -1} A_k z^k. \tag{61}$$

The orthogonal complements with respect to the pairing (57) are given by

$$L(\mathfrak{g})_+^\perp = \left\{ Y(z) = \sum_{k \geqq 0} Y_k z^k, \quad Y_0 \in \mathfrak{m}^\perp, \; Y_k \in \mathfrak{g}, k \geqq 1 \right\}, \tag{62}$$

$$L(\mathfrak{g})_-^\perp = \left\{ Y(z) = \sum_{k \leqq 0} Y_k z^k, \quad Y_0 \in \mathfrak{d}^\perp, \; Y_k \in \mathfrak{g}, k \leqq -1 \right\}. \tag{63}$$

The gradient of a function $H : L(\mathfrak{g}) \to \mathbb{C}$ is defined through the pairing (57). For $X, Y \in L(\mathfrak{g})$

$$\langle \nabla H(X) | Y \rangle = \frac{\mathrm{d}}{\mathrm{d}\tau} H(X + \tau Y)|_{\tau=0}. \tag{64}$$

Hamiltonian flows on $L(\mathfrak{g})$ correspond to the Lax equation

$$\dot{X} = [\nabla H(X), X]. \tag{65}$$

If H is Ad-invariant then the vector-field (64) vanishes. On the other hand, Ad-invariant Hamiltonians define non-trivial vector fields through the r-matrix induced by the splitting (60). By the Adler–Kostant–Symes theorem, these vector fields correspond to the Lax equation

$$\dot{X} = \pm [X, \nabla H(X)_\mp]. \tag{66}$$

We now show that (22) is of the form (66) for a suitable Ad-invariant Hamiltonian on $L(\mathfrak{g})$. Let f denote the flux in the scalar conservation law (2), and consider its antiderivative

$$F(s) = \int_0^s f(r) \mathrm{d}r. \tag{67}$$

Define the Hamiltonian $H_F : L(\mathfrak{g}) \to \mathbb{C}$

$$H_F(X(z)) = \left\langle F(X(z)z^{-1})|z^2 \right\rangle = \frac{1}{2\pi i} \oint_{|z|=1} F(X(z)z^{-1}) z \, \mathrm{d}z. \tag{68}$$

If f is a polynomial, the second equality follows from Cauchy's theorem. The general case follows by approximation. The Hamiltonian H_F is *distinct* from the Hamiltonian of Theorem 1, and in some sense is more natural. A few calculations (see [7, p. 94]) then yield that H_F is Ad-invariant and

$$\nabla H_F(X) = f(Xz^{-1})z. \tag{69}$$

Now consider the diagonal matrices $\mathcal{M}$ and $\mathcal{N}$ as in (27), and consider the finite-dimensional subspace V of $L(\mathfrak{g})_+$ consisting of linear polynomials of the form

$$V = \{X \in L(\mathfrak{g})_+ | X(z) = z\mathcal{M} - A\}. \tag{70}$$

A direct computation based on the definition of H_F then yields

$$(\nabla H_F(X))_+ = z\mathcal{N} + B, \tag{71}$$

where B is as in (20). Thus, the Hamiltonian flow defined by H_F on V is exactly (29), which is of course, identical to (22).

We now see that every C^1 flux f gives rise to a Hamiltonian flow as in (29). Since each of these Hamiltonians is Ad-invariant, they are all in involution. One may now count the number of integrals and invoke the Liouville theorem. Here we linearize the flow via the AKS theorem.

Theorem 3. *Let $A_0 \in \mathfrak{m}$ and $X_0(z) = z\mathcal{M} - A_0$. Let $g_\pm(t)$ denote the smooth curves in $\mathcal{G}_\pm$ which solve the factorization problem*

$$\exp(-t\nabla H_F(X_0)) = g_+(t)^{-1} g_-(t), \quad g_\pm(0) = \mathrm{Id}, \tag{72}$$

for t in a maximal open interval I containing 0*. Then the solution X_t, $t \in I$ to* (29) *with initial condition X_0 is given by*

$$X_t = \mathrm{Ad}_{g_+(t)} X_0 = \mathrm{Ad}_{g_-(t)} X_0. \tag{73}$$

Note that a solution to the factorization problem always exists for $|t|$ small, thus a maximal interval of existence for a smooth solution to the factorization problem is well-defined. It is also well-known that this factorization problem is equivalent to a Riemann–Hilbert problem (see [12, Ch. 3.5]).

3.2. An Operator Factorization Problem

While Theorem 3 applies only to the discrete Lax equations (22), one may easily guess the associated factorization problem for (5). Recall that the generators $\mathcal{A}$ and $\mathcal{B}$ are integro-differential operators defined by (4) and (6). As $N \to \infty$, the $N \times N$ diagonal matrices $\mathcal{M}$ and $\mathcal{N}$ of Theorem 3 are replaced by multiplication operators that act on test functions via

$$\mathcal{M}\varphi(u) = u\varphi(u), \quad \mathcal{N}\varphi(u) = f(u)\varphi(u). \tag{74}$$

The crucial algebraic relation (28) continues to hold.

$$[\mathcal{A}, \mathcal{N}] - [\mathcal{M}, \mathcal{B}] = 0. \tag{75}$$

Formally, this is all that is required to embed (5) in a loop-group and we now find the factorization problem

$$\exp(-t\nabla H_F(X_0)) = g_+(t)^{-1} g_-(t), \quad g_\pm(0) = \mathrm{Id}, \tag{76}$$

with $X_0 = z\mathcal{M} - \mathcal{A}$. Rather than develop these ideas in formal generality, let us mention one interesting example. Assume we consider Burgers' equation with Brownian motion initial data. Then the Hamiltonian is $H_F(s) = s^3/6$ and the generator of initial data is $\mathcal{A}_0\varphi(u) = -\varphi''(u)/2$. Then we find

$$\nabla H_F(X_0) = \frac{1}{2}\left(z\mathcal{M} - \frac{1}{2}\frac{d^2}{du^2}\right)^2. \tag{77}$$

For $z \in \mathbb{R}$, this is the square of the Airy operator. To the best of our knowledge, the factorization suggested by (76) is new.

4. Scattering and Inverse Scattering Theory for the *mN*-Wave Model

In this section we develop a scattering and inverse scattering theory for the discrete zero-curvature equations (21).

$$\partial_t A - \partial_x B = [A, B]. \tag{78}$$

The linear problem that underlies the scattering theory of (21) is as follows. Assume that $\mathcal{M}$ is a fixed diagonal matrix as in (27) and $A \in L^1(\mathbb{R}, \mathfrak{m})$. We consider a fundamental matrix for the linear equation

$$\psi_x = \psi\left(z\mathcal{M} + A\right), \quad x \in \mathbb{R}, \tag{79}$$

such that $\psi(x, z) \sim \mathrm{e}^{zx\mathcal{M}}$ as $x \to -\infty$. Such solutions are called *wave functions* (we use the terminology of [45]). The scattering theory for this equation when A is a matrix that vanishes on the diagonal was considered by ZAKHAROV ET AL. [48] and by BEALS and COIFMAN [13,14]. In our work, $A \in \mathfrak{m}$. As a consequence, even though A is completely determined by its off-diagonal elements, its diagonal entries do not vanish. Thus, the scattering theory of [13,14] does not immediately apply to our model and we have to rederive some results. For the most part, this is straightforward. To prevent too much repetition, we state the results we need, and present the main calculations that explain how the results of [13,14] are to be modified in Sections 4.5 and 4.6.

A short outline of the results of this section is as follows. The scattering theory is addressed in Section 4.1. Theorems 4 and 5 associate spectral data to $A \in L^1(\mathbb{R}, \mathfrak{m})$. We consider the inverse scattering theory in Section 4.2 and Theorem 6. The time evolution of spectral data and the Cauchy problem is considered in Section 4.3. The linear evolution of spectral data is stated in (89). The combination of inverse scattering theory and evolution yields several well-posedness theorems for (21). Finally, we construct a hierarchy of integrable models as in the ZS-AKNS hierarchy in Section 4.4. It is not apparent to us that these have intrinsic probabilistic significance, but it is interesting to note that one may construct other integrable flows on $\mathfrak{M}$ with little effort.

All results are rigorous for $N \times N$ matrices and have a natural, but formal, extension to the integro-differential operators $\mathcal{A}$ and $\mathcal{B}$. The linear evolution of the spectral data for these operators remains (89) with $\mathcal{M}$ and $\mathcal{N}$ replaced by the multiplication operators (74). However, we are unaware of rigorous results on the inverse spectral problem for such operators, and a full well-posedness theorem for (8) via inverse scattering requires further study.

4.1. Scattering Theory

It is more convenient to work with the new variable

$$m(x, z) = \mathrm{e}^{-z\mathcal{M}x} \psi(x, z). \tag{80}$$

m satisfies the linear equation

$$m_x = z[m, \mathcal{M}] + mA \tag{81}$$

Solutions to (81) such that $\|m(\cdot, z)\|_{L^\infty(\mathbb{R})} < \infty$ and $m(x, z) \to I$ as $x \to -\infty$ are called *global reduced wave functions*.

Theorems 4 and 5 below closely follow Beals and Coifman [13]. Let $\Sigma = i\mathbb{R}$ denote the imaginary axis in the complex z-plane, and let $\mathbf{P}$ denote the set of maps $A \in L^1(\mathbb{R}, \mathfrak{m})$. We call these maps *potentials* and the subclass $\mathbf{P}_0$ below *generic potentials*.

Theorem 4. ([13, Thm. A])

(a) *Suppose $A \in \mathbf{P}$. There is a bounded discrete set $Z \subset \mathbb{C}\backslash\Sigma$ such that $m(\cdot, z)$ is a unique global reduced wave function for every $z \in \mathbb{C}\backslash(\Sigma \cup Z)$. Moreover, $m(x, \cdot)$ is meromorphic in $\mathbb{C}\backslash\Sigma$ with poles precisely at the points of Z and* $\lim_{z\to\infty} m(x, z) = I$.

(b) *There is a dense open set $\mathbf{P}_0 \subset \mathbf{P}$ such that if $A \in \mathbf{P}_0$ then*
   1. *Z is finite.*
   2. *The poles of $m(x, \cdot)$ are simple.*
   3. *Distinct columns of $m(x, \cdot)$ have distinct poles.*

4. *$m(x,\cdot)$ admits limits $m^{\pm}(x,)$ as $z \to \Sigma\backslash Z$ from the left and right half-planes.*

Theorem 5. ([13, Thm. B])

(a) *Suppose $A \in \mathbf{P}_0$. For $z \in \Sigma$ there is a unique matrix $v(z)$ such that for every $x \in \mathbb{R}$*

$$m^{+}(x,z) = \mathrm{e}^{-xz\mathcal{M}} v(z) \mathrm{e}^{xz\mathcal{M}} m^{-}(x,z). \tag{82}$$

(b) *For each pole $z_j \in Z$, there is a matrix $v(z_j)$ such that the residue of m satisfies*

$$\operatorname{Res}(m(x,\cdot); z_j) = \lim_{z\to z_j} \mathrm{e}^{-xz\mathcal{M}} v(z) \mathrm{e}^{xz\mathcal{M}} m(x,z). \tag{83}$$

(c) *The generic potential A is uniquely determined by the jump matrix $v(z)$, $z \in \Sigma$ and the residues* $\operatorname{Res}(m(x,\cdot); z_j)$, *$z_j \in Z$.*

4.2. Inverse Scattering Theory

The jump matrix $v(z)$, the poles $Z = \{z_1, \ldots, z_M\}$ and the residues $v(z_j)$ constitute the *scattering data*. The reconstruction of A from the scattering data is the *inverse spectral problem*. Though Theorem 5 guarantees that the scattering data associated to a generic potential is unique, this assertion is proved via an application of Liouville's theorem and is not constructive. What is required is a constructive procedure to obtain A given the scattering data.

In order to state the inverse spectral theorems, we work with potentials that lie in the Schwartz class $\mathcal{S}(\mathbb{R}, \mathfrak{m})$. This assumption is not necessary, but it simplifies the exposition. Analogous finite regularity results can also be obtained as in [13].

Theorem 6. (a) *Suppose $A \in \mathcal{S}(\mathbb{R}, \mathfrak{m})$. Then there is $R > 0$ and C^{∞} functions $m^{(k)} : \mathbb{R} \to \mathfrak{g}$, $k = 0, 1, \ldots$ such that*

$$m(x,z) = \sum_{k=0}^{\infty} z^{-k} m^{(k)}(x), \quad x \in \mathbb{R}, |z| > R, \tag{84}$$

and the series converges uniformly in x and z.

(b) *The coefficients $m^{(k)}$ may be determined recursively. In particular, $m^{(0)}$ is a diagonal matrix with entries*

$$m^{(0)}_{ii}(x) = \exp\left(\int_{-\infty}^{x} A_{ii}(s)\,\mathrm{d}s\right), \quad i = 1, \ldots, N, \tag{85}$$

and the off-diagonal entries of $m^{(1)}$ are given by

$$m^{(1)}_{ij} = \frac{m^{(0)}_{ii}}{u_j - u_i} A_{ij}(x), \quad i \neq j. \tag{86}$$

(c) *The asymptotic expansion* (84) *may also be written*

$$m(x,z) = m^{(0)}(x)h(x,z), \quad h(x,z) = \sum_{k=0}^{\infty} z^{-k}h^{(k)}(x), \quad x \in \mathbb{R}, |z| > R, \tag{87}$$

where $h^{(0)}(x) \equiv I$, *and* $h^{(k)}$, $k \geqq 1$, *are in the Schwartz class* $\mathcal{S}(\mathbb{R}, \mathfrak{g})$.

Recall that here $u_1 < u_2 < \cdots < u_N$ are the diagonal entries of $\mathcal{M}$. Part (b) of the theorem allows us to uniquely reconstruct the potential. Assume given a global reduced wave function m with the asymptotic expansion (84). Then the off-diagonal terms of A are given by (86), and the diagonal terms are given by the relation $A_{jj} = -\sum_{k\neq j} A_{jk}$. It is necessary to assume that m is a reduced wave function: an arbitrary set of functions $m^{(k)}(x)$ is not admissible. Indeed, the constraint $A_{jj} = -\sum_{k\neq j} A_{jk}$ implies many relations between the coefficients $m^{(k)}$. For example, we have

$$m_{ii}^{(0)}(x) = \sum_{j\neq i} \int_{-\infty}^{x} m_{ij}^{(1)}(s)\, \mathrm{d}s. \tag{88}$$

The full inverse scattering problem relates the scattering data to the potential A. In light of Theorem 6, it suffices to reconstruct m from the scattering data. Since m is holomorphic in $\mathbb{C}\backslash(\Sigma \cup Z)$ it is expressed in terms of the scattering data by Cauchy integrals. The associated integral equations are independent of our assumption that $A \in \mathfrak{m}$, and the results of [13] relating m and the scattering data apply directly. The subtlety is that the scattering data satisfy algebraic, analytic and topological constraints. For example, these may be constraints involving the zeros and winding numbers of principal minors of v [13, Thm D]. For generic potentials that satisfy these constraints, m and the scattering data are related by Cauchy integrals that preserve the Schwartz class. For such scattering data, the inverse scattering problem is solved by mapping the scattering data to m via Cauchy integrals, and then m to A via Theorem 6. In the simplest situation, Z is empty, and the wave function is reconstructed from the jump on Σ alone.

4.3. Evolution of Scattering Data and the Cauchy Problem

We now combine the x and t dependence, and consider the Cauchy problem for the discrete zero-curvature equations (21) with initial data $A(x,0) = A_0(x) \in \mathfrak{m}$. Let $v_0(z)$, $z \in \Sigma$ and $v_0(z_j)$ denote the scattering data of A_0. The scattering data evolve by the simple linear equations

$$\frac{\partial v(z)}{\partial t} = [z\mathcal{N}, v(z)], \quad z \in \Sigma \tag{89}$$

$$\frac{\partial v(z_j)}{\partial t} = [z\mathcal{N}, v(z_j)], \quad z_j \in Z, \tag{90}$$

with the unique solution

$$v(z,t) = \mathrm{e}^{tz\mathcal{N}} v_0(z) \mathrm{e}^{-tz\mathcal{N}}, \quad v(z_j,t) = \mathrm{e}^{tz_j\mathcal{N}} v_0(z_j) \mathrm{e}^{-tz_j\mathcal{N}}. \tag{91}$$

The evolution is *formally stable* in the terminology of [14, §3.12]). We may now combine Theorem A and Theorem C of [14] to obtain the following basic well-posedness theorem for (21).

Theorem 7. *Assume A_0 is a generic potential in $\mathcal{S}(\mathbb{R}, \mathfrak{m})$ with associated scattering data $v_0(z)$, $z \in \Sigma$ and $v_0(z_j)$, $z \in Z$. Then there is $T > 0$ and a unique smooth map $[0, T) \to \mathcal{S}(\mathbb{R}, \mathfrak{m})$, $t \mapsto A(\cdot, t)$, such that the scattering data of $A(\cdot, t)$ is given by (91) and $A(x, t)$ solves the Cauchy problem for the discrete zero-curvature equations (21) with initial data A_0.*

In addition, the following dichotomy holds [14, Thm. B].

Theorem 8. *Under the hypotheses of Theorem 7, suppose $T \in (0, \infty]$ is maximal. Then either $T = \infty$ or $\lim_{t \to T} \|A(\cdot, t)\|_{L^2(\mathbb{R})} = \infty$.*

In general, the maximal time interval is finite. However, global existence is guaranteed if A_0 is triangular. If A is triangular, so are B and $[A, B]$, and $[A, B]$ vanishes on the diagonal. Thus

$$\operatorname{Tr}(A^T [A, B]) = 0. \tag{92}$$

In addition, for $A \in \mathcal{S}(\mathbb{R}, \mathfrak{m})$

$$\int_{\mathbb{R}} \operatorname{Tr}\left(A^T(x) \partial_x B\right) \,\mathrm{d}x = \sum_{i,j} F_{ij} \int_{\mathbb{R}} A_{ij} \partial_x A_{ij} \,\mathrm{d}x = 0. \tag{93}$$

Thus, the evolution of (21) is *dissipative* (see [14, §1.11]) and we have global existence.

Corollary 1. *Assume the hypotheses of Theorem 7 and assume in addition that $A_0(x)$ is triangular for every $x \in \mathbb{R}$. Then $T = \infty$.*

It is surprising that we do not need to assume that A is everywhere lower triangular or everywhere upper triangular. The assumption that A is triangular pointwise is enough to ensure (92) for every $x \in \mathbb{R}$, which in turn implies dissipativity.

Finally, let us connect these results with the probabilistic context that motivated us. In order to ensure that A is truly a generator, we must ensure that the off-diagonal terms are positive. Since smooth solutions exist, this is preserved at least for a short time. However, it is more subtle to ensure global existence. Here the convexity of f plays an important role.

Theorem 9. *Assume the hypotheses of Theorem 7. In addition, assume that f is convex, and that $A_0(x)$ is the generator of a spectrally negative Markov process. Then $T = \infty$ and $A(x, t)$ remains the generator of a spectrally negative Markov process for every $t > 0$.*

Proof. Equation (21) may also be solved by the method of characteristics. Indeed, $B_{ij} = F_{ij} A_{ij}$, thus each entry A_{ij} evolves on a characteristic with speed $-F_{ij}$.

The characteristic speed is simply the Rankine–Hugoniot condition associated to the shock connecting states u_i and u_j. We integrate (21) on characteristics to find

$$A_{ij}(x,t) = (A_0)_{ij}(x + F_{ij}t) + \int_0^t [A, B]_{ij}(x + F_{ij}s, s)\, \mathrm{d}s. \tag{94}$$

Since $u_1 < \cdots < u_M$ and f is convex, we now find exactly as in the proof of Theorem 2 that $[A, B]_{ij} \geqq 0$, $i > j$. The diagonal terms are conserved on characteristics since $[A, B]_{ii} = 0$ since $B = F \circ A$. Similarly, the upper-triangular part of $[A, B]$ vanishes if A and B are lower-triangular. A simple maximum principle argument shows that A remains lower-triangular. $\square$

The integral equations (94) can also be used to give a direct proof of global existence of solutions without the assumption that A_0 is generic. Since $A_{ij} \geqq 0$ on the off-diagonal, and A_{ii} is conserved, we have the bound $\|A_{ij}(\cdot, t)\|_{L^\infty(\mathbb{R})} \leqq \|(A_0)_{ii}\|_{L^\infty(\mathbb{R})}$, which ensures global existence.

4.4. The ZS-AKNS Hierarchy

The discrete zero-curvature equations (21) are part of a hierarchy of commuting Hamiltonian flows. The existence of such hierarchies was established in the pioneering work of Zakharov and Shabat [49] and Ablowitz et al. [2]. We now derive the associated hierarchy for (21): as expected this is a modification of the hierarchy for the N-wave model. Our calculations and notation follow [45].

We fix a diagonal matrix $\mathcal{N}$ and consider the asymptotic behavior of $Q(x, z) = m^{-1}\mathcal{N}m$ as $z \to \infty$. By Theorem 6(c), we have $m^{-1}\mathcal{N}m = h^{-1}\mathcal{N}h$, and the expansion (87) yields

$$Q(x, z) = h^{-1}\mathcal{N}h \sim \sum_{k=0}^{\infty} Q^{(k)} z^{-k}, \quad z \to \infty. \tag{95}$$

We call $Q^{(k)}$ the kth flux. It admits an expansion

$$Q^{(k)} = z^k\mathcal{N} + z^{k-1}B_1 + z^{k-2}B_2 + \cdots + B_k, \quad B_k \in \mathfrak{m}. \tag{96}$$

Definition 1. The kth flow in the hierarchy is given by the equation

$$\partial_t A - \partial_x Q^{(k)} = [A, Q^{(k)}], \quad k \geqq 0. \tag{97}$$

The zero-curvature equation (97) may also be written in the Lax form

$$\left[\partial_x + z\mathcal{M} + A, \partial_t + Q^{(k)}\right] = 0. \tag{98}$$

The kth flux is obtained as follows. We show below that Q satisfies the linear equation

$$Q_x = [Q, z\mathcal{M} + A]. \tag{99}$$

The asymptotic expansion (95) now yields the hierarchy of linear equations

$$0 = [Q^{(0)}, \mathcal{M}], \tag{100}$$

and

$$Q_x^{(k)} - [Q^{(k)}, A] = [Q^{(k+1)}, \mathcal{M}], \quad k \geqq 0. \tag{101}$$

Since $h(x, z) \sim I + z^{-1}h^{(1)}$ as $z \to \infty$, we have

$$Q(x, z) \sim \mathcal{N} + \frac{[\mathcal{N}, h^{(1)}]}{z} + \cdots, \quad z \to \infty. \tag{102}$$

Thus, $Q^{(0)} = \mathcal{N}$ is the solution to (100). For $k \geqq 1$ the ansatz (96) yields $k + 2$ linear equations for B_j. We use (96) and (98) to obtain the equations

$$O(z^{k+1}) : \ [\mathcal{M}, \mathcal{N}] = 0, \tag{103}$$

$$O(z^k) : \ \partial_x \mathcal{N} + [A, \mathcal{N}] + [\mathcal{M}, B_1] = 0, \tag{104}$$

$$O(z^j), k - 1 \geqq j \geqq 1 : \ \partial_x B_{k-j} + [A, B_{k-j}] + [\mathcal{M}, B_{k+1-j}] = 0, \tag{105}$$

$$O(1) : \ \partial_x B_k - \partial_t A + [A, B_k] = 0. \tag{106}$$

The $O(z^{k+1})$ equation is trivially satisfied since $\mathcal{N}$ is diagonal. Since $\mathcal{N}$ is independent of x, the $O(z^k)$ equation generalizes equation (28). When $k = 1$ this yields the off-diagonal terms of $B_1 = B$ in accordance with (20). For $k > 1$ we recursively solve (103) until we obtain B_k. This is very similar to the recursion for the classical N-wave model with one important difference. When solving (103) recursively, we realize that the $O(z)$ term yields only the off-diagonal terms of B_k. In the earlier work of Zakharov and Manakov, the diagonal terms of B_k vanished because of the assumption that $B_k \in \mathfrak{u}$. Here these terms suffice to determine B_k, since $B_k \in \mathfrak{m}$.

The kth flow may be solved by the inverse scattering method. The scattering data evolve by

$$v(z, t) = \mathrm{e}^{tz^k\mathcal{N}} v_0(z)\mathrm{e}^{-tz^k\mathcal{N}}, \quad v(z_j, t) = \mathrm{e}^{tz_j^k\mathcal{N}} v_0(z_j)\mathrm{e}^{-tz_j^k\mathcal{N}}. \tag{107}$$

Well-posedness of the kth flow requires that the evolution is formally stable [14, 3.12]. In our case, this requires $\mathrm{Re}(z^k\mathcal{N}_{jj}) = 0$ for $z \in \Sigma$ and each diagonal entry of $\mathcal{N}$. Since Σ is the imaginary axis, this is satisfied for all odd k when $\mathcal{N}$ is real (in particular, for $\mathcal{N}$ given by $\mathrm{diag}(f(u_1), \ldots, f(u_M))$). For even k, we need to choose $\mathcal{N}$ purely imaginary, and we find that B_k is purely imaginary if A is real. This is incompatible with $A \in \mathfrak{m}$. For odd k, Theorems 7 and 8 hold with (91) replaced by (107) and (21) replaced by (97). It is not clear if these equations have a true probabilistic interpretation.

4.5. Proofs of Theorem 4 and Theorem 5

We now present the calculations and matrix factorization theorem that underlie Theorems 4 and 5. To this end, it is enough to assume that A is C^∞ with compact support in x. A density argument as in [13] yields the conclusions for $A \in L^1(\mathbb{R}, \mathfrak{m})$. To construct a globally bounded solution $m(x,z)$ we assume that A has compact support and solve the initial value problem

$$\tilde{m}_x = z[\tilde{m}, \mathcal{M}] + \tilde{m}A, \quad \tilde{m}(x,z) = I, \ x \ll 0. \tag{108}$$

It is clear that (108) has a unique solution that is holomorphic in z. In addition, since $\tilde{m}_x = z[\tilde{m}, \mathcal{M}]$, $x \gg 0$, there exists a holomorphic matrix $s(z)$ such that

$$\tilde{m}(x,z) = \begin{cases} I, & x \ll 0, \\ \mathrm{e}^{-zx\mathcal{M}} s(z) \mathrm{e}^{zx\mathcal{M}}, & x \gg 0. \end{cases} \tag{109}$$

Observe also that $\tilde{m}$ is always invertible because

$$\det(\tilde{m})(x,z) = \exp\left(\int_{-\infty}^{x} \mathrm{Tr}(A(s))\, \mathrm{d}s\right). \tag{110}$$

We seek a bounded solution to (81) of the form $m(x,z) = \tilde{g}(x,z)\tilde{m}(x,z)$. Since m solves (81) and $\tilde{m}$ solves (108) we find that $\tilde{g}(x,z) = \mathrm{e}^{-xz\mathcal{M}} g(z) \mathrm{e}^{xz\mathcal{M}}$ for some matrix $g(z)$. The asymptotic behavior of $\tilde{m}$ in (109) then implies

$$m(x,z) = \begin{cases} \mathrm{e}^{-xz\mathcal{M}} g(z) \mathrm{e}^{xz\mathcal{M}}, & x \ll 0, \\ \mathrm{e}^{-zx\mathcal{M}} g(z) s(z) \mathrm{e}^{zx\mathcal{M}}, & x \gg 0. \end{cases} \tag{111}$$

It remains only to choose g so that m is globally bounded in x for fixed z.

First assume z is in the left half-plane. Recall that $\mathcal{M} = \mathrm{diag}(u_1, \ldots, u_m)$ with $u_1 < u_2 < \cdots < u_m$. Thus, $m_{jk}(x,z) = g_{jk}(z)\mathrm{e}^{xz(u_k - u_j)}$ for $x \ll 0$. Since $\mathrm{Re}(z) > 0$, m_{jk} is bounded only if $g_{jk} = 0$ for $j < k$. Thus, g is *lower-triangular*. We next find that m is bounded as $x \to \infty$ only if $g(z)s(z)$ is *upper-triangular*. Finally, since $m \to I$ as $x \to -\infty$, we see that the diagonal entries of g are all 1. In order to choose g in accordance with these constraints, recall that Gaussian elimination may be written as the matrix factorization

$$s(z) = L(z)D(z)U(z), \tag{112}$$

where L and U are lower and upper triangular matrices that are 1 on the diagonal and $D(z) = \mathrm{diag}(\det(s_1(z), \ldots, s_m(z))$ where $s_k(z)$ denotes the $k \times k$ upper-block of $s(z)$ [26, Thm 1.1]. By construction, L, D and U are unique except at the zeros of $\det(s_k(z))$, $k = 1, \ldots, m$. Since s is entire, this set is discrete. We then choose $g^{-}-(z) = L(z)^{-1}$, the superscript denoting the left half-plane. By construction, g is meromorphic in the left half-plane.

A similar calculation in the right half-plane $\mathrm{Re}(z) > 0$ reveals that $g(z)$ must be upper-triangular and $g(z)s(z)$ must be lower-triangular. In this case, we factorize

$$s(z) = \tilde{L}(z)\tilde{D}(z)\tilde{U}(z), \tag{113}$$

where now $\tilde{D} = \mathrm{diag}(\det(\tilde{s}_1(z), \ldots, \tilde{s}_m(z))$ and $\tilde{s}_k(z)$ denotes the $k \times k$ lower-block of $s(z)$. We now find $g^+(z) = \tilde{U(z)}^{-1}$.

Let Z denotes the set of zeros of $\det(s_k(z))$ and $\det(\tilde{s_k}(z))$, $k = 1, \ldots, m$. The factorizations $g^{\pm}(z)$ are continuous on $z \in \Sigma \backslash Z$. To summarize, we have

$$m^{\pm}(x, z) = \mathrm{e}^{-xz\mathcal{M}} g^{\pm}(z) \mathrm{e}^{xz\mathcal{M}} \tilde{m}(x, z), \quad \in \mathbb{R}, z \in \mathbb{C} \backslash Z, \tag{114}$$

where $\tilde{m}(x, z)$ is entire in z, $g^{\pm}$ are obtained by factorizing s as in (112) and (113). In order to obtain the scattering data, we isolate the jump in m on Σ. If we set $\varphi(x, z) = m^+(m^-)^{-1}$, $x \in \mathbb{R}$, $z \in \Sigma$, we find $\varphi_x = z[\varphi, \mathcal{M}]$. Thus, there exists a matrix $v(z)$ such that $\varphi(x, z) = \mathrm{e}^{-xz\mathcal{M}} v(z) \mathrm{e}^{xz\mathcal{M}}$ and we have

$$m^+(x, z) = \mathrm{e}^{-xz\mathcal{M}} v(z) \mathrm{e}^{xz\mathcal{M}} m^-(x, z), \quad x \in \mathbb{R}, z \in \Sigma \backslash Z. \tag{115}$$

These are the main calculations needed to establish the existence of the jump measure, and it is clear that the assumption $A \in \mathfrak{m}$ (as opposed to $A_{jj} = 0$) has played only a minor role (for example (110) has replaced $\det(\tilde{m}) \equiv 1$). The arguments in [13, pp. 48–49] are similarly modified to yield Theorem 4 and Theorem 5.

4.6. *Proof of Theorem 6*

Theorem 6 is a modification of [13, Thm 6.1]. The main differences are that m_0 is no longer the identity, and we have to solve separately for diagonal and off-diagonal terms. It is simplest to postulate an expansion of the form (84) and solve for the terms m_j. One may then justify the expansion for A that is suitably regular as in [13, §6].

Assume (84) holds and $m(x, z) \to I$ as $x \to -\infty$. We substitute this ansatz in (81) to find the hierarchy of equations

$$0 = [m^{(0)}(x), \mathcal{M}], \tag{116}$$

and

$$\frac{\mathrm{d}m^{(k)}}{\mathrm{d}x} - m_k A = [m^{(k+1)}(x), \mathcal{M}], \quad k = 0, 1, \ldots \tag{117}$$

Equation (117) implies that on the diagonal

$$\frac{\mathrm{d}m_{ii}^{(k)}}{\mathrm{d}x} - \left(m^{(k)} A\right)_{ii} = 0. \tag{118}$$

On the off-diagonal

$$\frac{\mathrm{d}m_{ij}^{(k)}}{\mathrm{d}x} - \left(m^{(k)} A\right)_{ij} = \left(u_j - u_i\right) m_{ij}^{(k+1)}, \quad i \neq j. \tag{119}$$

Equation (116) implies that $m^{(0)}$ is a diagonal matrix. We then solve (118) with $k = 0$ to obtain (85). It is simplest to solve the rest of the hierarchy by making the ansatz

$$m(x, z) = m^{(0)} \sum_{k=0}^{\infty} z^{-k} h^{(k)}(x, z) = m^{(0)} h(x, z). \tag{120}$$

Let A_d and A_o denote the diagonal and off-diagonal terms of A. We substitute (120) in (81) to find

$$h_x = [h, z\mathcal{M} + A_d] + hA_o. \tag{121}$$

Then we have the hierarchy of linear equations

$$0 = [h^{(0)}, \mathcal{M}], \tag{122}$$

and

$$h_x^{(k)} - [h^{(k)}, A_d] - h^{(k)}A_o = [h^{(k+1)}, \mathcal{M}], \quad k \geqq 0. \tag{123}$$

Equation (122) implies that $h^{(0)}$ is diagonal. We then consider the diagonal terms of (123) with $k = 0$ to find $h_x^{(0)} = 0$. Since $\lim_{x\to-\infty} h(x, z) = I$, this implies $h^{(0)} \equiv I$ as expected. The off-diagonal terms of (123) with $k = 0$ may be solved algebraically and yield

$$h_{ij}^{(1)} = \frac{A_{ij}}{u_i - u_j}, \quad i \neq j. \tag{124}$$

This process can be continued indefinitely. At each step, we first solve a differential equation that yields the diagonal terms of $h^{(k)}$, and then an algebraic equation that yields the off-diagonal terms of $h^{(k+1)}$. For example, we find

$$h_{ii}^{(1)}(x) = \sum_{j\neq i} \frac{1}{a_i - a_j} \int_{-\infty}^{x} A_{ij}(s)A_{ji}(s)\,\mathrm{d}s, \tag{125}$$

and for the off-diagonal terms of $h^{(2)}$

$$h_{ij}^{(2)} = \frac{1}{u_i - u_j}\left(\sum_{k\neq i,j} \frac{A_{ik}A_{kj}}{u_i - u_k} + h_{ii}^{(1)}A_{ij} + \frac{A_{ij}(A_{ii} - A_{jj})}{u_i - u_j} - \frac{\partial_x A_{ij}}{u_i - u_j})\right). \tag{126}$$

This process becomes increasingly unwieldy, but at every step $h^{(k)}$ is expressed as a finite number of integro-differential terms of A. If A is in the Schwartz class, so is $h^{(k)}$ for $k \geqq 1$. This is enough to establish the uniform convergence of Theorem 6.

Acknowledgments. I thank Mark Ablowitz, Mark Adler, Percy Deift, Luen-Chau Li, David Mumford, Bob Pego, and Fraydoun Rezakhanlou for stimulating discussions. This work is part of a general program developed with Ravi Srinivasan and I am particularly indebted to him. Thanks also to Michele Benzi for informing me of [32].

References

1. Ablowitz, M., Haberman, R.: Resonantly coupled nonlinear evolution equations. *J. Math. Phys.* **16**, 2301 (1975)
2. Ablowitz, M., Kaup, D., Newell, A., Segur, H.: The inverse scattering transform-Fourier analysis for nonlinear problems. *Stud. App. Math.* **53**, 249–315 (1974)
3. Ablowitz, M.J., Clarkson, P.A.: Solitons, nonlinear evolution equations and inverse scattering. *London Mathematical Society Lecture Note Series*, Vol. 149. Cambridge University Press, 1991
4. Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions with formulas, graphs, and mathematical tables. *National Bureau of Standards Applied Mathematics Series*, Vol. 55. U.S. Government Printing Office, Washington, D.C., 1964
5. Adler, M., van Moerbeke, P.: Linearization of Hamiltonian systems, Jacobi varieties and representation theory. *Adv. Math.* **38**(3), 318–379 (1980)
6. Adler, M., Van Moerbeke, P.: Completely integrable systems, Euclidean Lie algebras, and curves. *Adv. Math.* **38**(3), 267–317 (1980)
7. Adler, M., Van Moerbeke, P., Vanhaecke, P.: *Algebraic integrability, Painlevé geometry and Lie algebras*. Springer, 2004
8. Applebaum, D.: *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2004
9. Arnold, V.: *Mathematical Methods of Classical Mechanics*. Springer, Berlin, 1989
10. Arnold, V., Khesin, B.: *Topological Methods in Hydrodynamics*. Springer, Berlin, 1998
11. Avallaneda, M., E, W.: Statistical properties of shocks in Burgers turbulence. *Commun. Math. Phys.* **172**, 13–38 (1995)
12. Babelon, O., Bernard, D., Talon, M.: *Introduction to Classical Integrable Systems*. Cambridge University Press, 2003
13. Beals, R., Coifman, R.: Scattering and inverse scattering for first order systems. *Commun. Pure App. Math.* **37**(1), 39–90 (1984)
14. Beals, R., Coifman, R.: Inverse scattering and evolution equations. *Commun. Pure App. Math.* **38**(1), 29–42 (1985)
15. Beals, R., Sattinger, D.: On the complete integrability of completely integrable systems. *Commun. Math. Phys.* **138**(3), 409–436 (1991)
16. Bertoin, J.: The inviscid Burgers equation with Brownian initial velocity. *Commun. Math. Phys.* **193**, 397–406 (1998)
17. Bertoin, J.: Some aspects of additive coalescents. In: *Proceedings of the International Congress of Mathematicians*, Beijing 2002, Vol. III, 15–23. Higher Ed. Press, 2002
18. Burgers, J.M.: *The Nonlinear Diffusion Equation*. Reidel, Dordrecht, 1974
19. Carraro, L., Duchon, J.: Solutions statistiques intrinsèques de l'équation de Burgers et processus de Lévy. *C. R. Acad. Sci. Paris Sér. I Math.*, **319**(8), 855–858 (1994)
20. Carraro, L., Duchon, J.: Équation de Burgers avec conditions initiales à accroissements indépendants et homogènes. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **15**(4), 431–458 (1998)
21. Chabanol, M.L., Duchon, J.: Markovian solutions of inviscid Burgers equation. *J. Stat. Phys.* **114**(1–2), 525–534 (2004)
22. Constantin, P., E, W., Titi, E.: Onsager's conjecture on the energy conservation for solutions of Euler's equation. *Commun. Math. Phys.* **165**(1), 207–209 (1994)
23. E, W., Sinai, Y.: New results in mathematical and statistical hydrodynamics. *Russian Math. Surv.* **55**(4), 635–666 (2000)
24. E, W., Vanden Eijnden, E.: Statistical theory for the stochastic Burgers equation in the inviscid limit. *Commun. Pure Appl. Math.* **53**(7), 852–901 (2000)
25. Eyink, G.: Energy dissipation without viscosity in ideal hydrodynamics I. Fourier analysis and local energy transfer* 1. *Physica D*, **78**(3–4), 222–240 (1994)

26. Faddeev, D., Faddeeva, V.: *Computational Methods of Linear Algebra*. WH Freeman, San Francisco, 1963
27. Frachebourg, L., Martin, P.A.: Exact statistical properties of the Burgers equation. *J. Fluid Mech.* **417**, 323–349 (2000)
28. Golovin, A.M.: The solution of the coagulating equation for cloud droplets in a rising air current. *Izv. Geophys. Ser.* 482–487 (1963)
29. Groeneboom, P.: Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Relat. Fields*, **81**(1), 79–109 (1989)
30. Hopf, E.: Statistical hydromechanics and functional calculus. *J. Rational Mech. Anal.* **1**, 87–123 (1952)
31. Jimbo, M., Miwa, T., Môri, Y., Sato, M.: Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent. *Physica D* **1**(1), 80–158 (1980)
32. Johnson, J.: Markov-type Lie groups in GL (n, R). *J. Math. Phys.* **26**(2), 252 (1985)
33. Kerov, S.V.: Asymptotic representation theory of the symmetric group and its applications in analysis, *Translations of Mathematical Monographs*, Vol. 219. American Mathematical Society, Providence, RI, 2003
34. Kida, S.: Asymptotic properties of Burgers turbulence. *J. Fluid Mech.* **93**(2), 337–377 (1979)
35. Kolmogorov, A.N.: Dissipation of energy in the locally isotropic turbulence. *Proc. R. Soc. London Ser. A* **434**(1890), 15–17 (1991). Translated from the Russian by V. Levin, Turbulence and stochastic processes: Kolmogorov's ideas 50 years on
36. Manakov, S.: Note on the integration of Euler's equations of the dynamics of an n-dimensional rigid body. *Funct. Anal. Appl.* **10**(4), 328–329 (1976)
37. Mendoza, R., Savin, I., Thornton, K., Voorhees, P.: Topological complexity and the dynamics of coarsening. *Nat. Mater.* **3**(6), 385–388 (2004)
38. Menon, G., Srinivasan, R.: Kinetic theory and Lax equations for shock clustering and Burgers turbulence. *J. Stat. Phys.* **140**(6), 1195–1223 (2010)
39. Moser, J.: Geometry of quadrics and spectral theory. In: *The Chern Symposium, 1979: Proceedings of the International Symposium on Differential Geometry in Honor of S.-S. Chern*, held in Berkeley, California, June 1979, p. 147. Springer, 1980
40. Novikov, S.P. (ed.): Dynamical systems, VII. *Encyclopaedia of Mathematical Sciences*, Vol. 16. Springer, Berlin, 1994
41. Onsager, L.: Statistical hydrodynamics. *Il Nuovo Cimento* (1943–1954), **6**, 279–287 (1949)
42. Semenov-Tian-Shansky, M.: What is a classical r-matrix? *Funct. Anal. Appl.* **17**(4), 259–272 (1983)
43. She, Z.S., Aurell, E., Frisch, U.: The inviscid Burgers equation with initial data of Brownian type. *Commun. Math. Phys.* **148**(3), 623–641 (1992)
44. Sinai, Y.G.: Statistics of shocks in solutions of inviscid Burgers equation. *Commun. Math. Phys.* **148**(3), 601–621 (1992)
45. Terng, C., Uhlenbeck, K.: Bäcklund transformations and loop group actions. *Commun. Pure Appl. Math.* **53**(1), 1–75 (2000)
46. Valageas, P.: Statistical properties of the Burgers equation with Brownian initial velocity. *J. Stat. Phys.* **134**(3), 589–640 (2009)
47. Zakharov, V., Manakov, S.: Resonant interaction of wave packets in nonlinear media. *JETP Lett. (USSR)* (Engl. Transl.) **18**(7), 243–245 (1973)
48. Zakharov, V., Manakov, S., Novikov, S., Pitaevskii, L.: *Theory of Solitons: The Inverse Scattering Method*. Plenum, New York, 1984
49. Zakharov, V., Shabat, A.: Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *JETP* **34**, 62–69 (1972)

Govind Menon

Division of Applied Mathematics,
Brown University,
182 George St., Providence,
RI 02912, USA.
e-mail: menon@dam.brown.edu

(*Received February 28, 2011 / Accepted August 5, 2011*)