Commun. Math. Phys. 179, 417–466 (1996)

Communications in Mathematical Physics

The Large Time Stability of Sound Waves

B. Temple[1,*], **R. Young**[2,**]

[1] Department of Mathematics, University of California, Davis CA 95616-8618, USA
[2] Department of Applied Mathematics, State University of New York, Stony Brook, NY 11794, USA

Received: 4 August 1995 / Accepted: 21 November 1995

Abstract: We demonstrate the existence of solutions to the full 3×3 system of compressible Euler equations in one space dimension, up to an arbitrary time $T > 0$, in the case when the initial data has arbitrarily large total variation, and sufficiently small supnorm. The result applies to periodic solutions of the Euler equations, a nonlinear model for sound wave propagation in gas dynamics. Our analysis establishes a growth rate for the total variation that depends on a new length scale d that we identify in the problem. This length scale plays no role in 2×2 systems, (or any system possessing a full set of Riemann coordinates), nor in the small total variation problem for $n \times n$ systems, the cases originally addressed by Glimm in 1965. Recent work by a number of authors has demonstrated that when the total variation is sufficiently large, solutions of 3×3 systems of conservation laws can in general blow up in finite time, (independent of the supnorm), due to amplifying instabilities created by the non-trivial Lie algebra of the vector fields that define the elementary waves. For the large total variation problem, there is an interaction between large scale effects that amplify and small scale effects that are stable, and we show that the length scale on which this interaction occurs is d. In the limit $d \to \infty$, we recover Glimm's theorem, and we observe that there exist linearly degenerate systems within the class considered for which the growth rate we obtain is sharp.

1. Introduction

We consider the initial value problem for the system of compressible Euler equations in one space dimension,

* Supported in part by NSF Applied Mathematics grant numbers DMS-92-06631, DMS-95000694, in part by ONR, US Navy grant number N00014-94-1-0691, A Guggenheim fellowship, and by the Institute of Theoretical Dynamics, UC-Davis.
** Partially supported by DOE grant number DE-FG02-88ER25053 while at the Courant Institute, and by NSF grant number DMS-9201581 and DOE grant number DE-FG02-90ER25084.

$$\begin{aligned} \rho_t + (\rho v)_x &= 0\,, \\ (\rho v)_t + (\rho v^2 + p)_x &= 0\,, \\ E_t + ((E+p)v)_x &= 0\,, \end{aligned} \tag{1.1}$$

where ρ is the density, v the velocity, p the pressure, and E the energy of the fluid. This is the special case of the general initial value problem for a system of conservation laws,

$$u_t + f(u)_x = 0, \qquad u(x,0) = u_0(x)\,, \tag{1.2}$$

where for (1.1), $u = (\rho, \rho v, E)$. We let r^i denote the i^{th} (normalized) right eigenvector and λ_i the i^{th} eigenvalue of the Jacobian matrix Df,

$$Df \cdot r^i = \lambda_i r^i, \quad i = 1,2,3\,,$$

and refer to the eigenpair (λ_i, r^i) as the i^{th} characteristic field, or i^{th} wave family. When discussing the general system (1.2), we always assume, as is the case for (1.1), that the system is strictly hyperbolic ($\lambda_i \neq \lambda_j$) and either genuinely nonlinear ($r^i \cdot \nabla\lambda_i \neq 0$) or linearly degenerate ($r^i \cdot \nabla\lambda_i \equiv 0$) in each characteristic field, and that solutions take values in a small neighborhood $\mathcal{U}$ of a reference state $\tilde{u}$, cf. [13].

System (1.1) represents the zero dissipation limit of the compressible Navier–Stokes equations. It is well known that shock-waves form in solutions of (1.1) even in the presence of smooth data, and the shock waves encode the dissipation that carries over to the zero dissipation limit. Shock waves introduce time-irreversibility, loss of information, and the increase of entropy (in a generalized sense), and this leads to the decay of solutions for general systems of type (1.2), cf. [2, 13].

For convenience, we study the Lagrangian version of the Euler system [13],

$$\begin{aligned} (1/\rho)_t - v_x &= 0\,, \\ v_t + p_x &= 0\,, \\ E_t + (pv)_x &= 0\,, \end{aligned} \tag{1.3}$$

where the conserved quantities are $(1/\rho, v, E)$. For smooth solutions, (1.3) implies

$$S_t = 0\,,$$

where S is the entropy, [13]. This "entropy equation" implies that the entropy decouples from the solution in smooth regions, which is characteristic of a *Riemann coordinate*. A wave (or eigen-) family has a Riemann coordinate $w : \mathcal{U} \to R$ if the left eigenvector of the flux matrix can be written as the gradient of that function,

$$\ell \cdot Df = \lambda\ell, \quad \text{where } \ell = \nabla_u w\,.$$

For system (1.3), the second eigenvalue (corresponding to the entropy) satisfies $\lambda_2 \equiv 0$. We thus refer to 2-waves as *contact discontinuities* or *entropy jumps.*

More generally, and with the goal in mind of isolating a particular nonlinear aspect of the Euler equations, we consider in this paper any 3×3 system within the class of conservation laws that possess a Riemann coordinate for one of the families. In analogy with (1.1), we assume for convenience that the second family has a Riemann coordinate, and is also linearly degenerate. After a suitable normalization, this class has the same (unsigned) Lie bracket components as the Euler system. For any system in this class, we obtain a large time existence theorem for initial data

of arbitrarily large total variation and sufficiently small supnorm. The main estimate involves a new length scale that must be accounted for, and to motivate this, we first recall the earlier work of Glimm.

Note that when the equation of state is of the form $p = p(\rho)$, (the case of isothermal and isentropic flow, [13]), the first two equations in (1.3) uncouple from the third, and system (1.1) reduces to the 2×2 system that we refer to as the p-system. There is a well developed existence theory for 2×2 systems of conservation laws, but for three equations, the only general existence theorems that apply to the full nonlinear system (1.3) are based essentially on methods of analysis first introduced by Glimm in 1965: methods that only apply when the initial total variation is sufficiently small, or when the system possesses a full set of Riemann coordinates (like 2×2 systems), cf. [13]. A full set of Riemann coordinates determines a coordinate system in which the coordinate vector-fields are eigenvectors of the flux, and thus have pairwise vanishing Lie brackets $[r^i, r^j] \equiv 0$. These systems are referred to as *rich* systems by Serre, who has shown the applicability of the methods of compensated compactness to these systems, cf. [11, 12]. In the small variation case for general systems, Glimm's theorem can be stated as follows, [3]:

Theorem (Glimm 1965). *If the total variation of the initial data $u_0(x)$ is smaller than a threshold value V_{crit},*

$$TV\{u_0(\cdot)\} < V_{\text{crit}} ,$$

then a global weak solution with shocks exists for all time and

$$TV\{u(\cdot, t)\} < CTV\{u_0(\cdot)\} ,$$

where V_{crit} and C depend only on the flux function f in the neighborhood of the solution.

In [17], it was shown that

$$V_{\text{crit}} = O(1/\Lambda) , \tag{1.4}$$

where Λ is a measure of the strength of the geometric coupling of the wave families, which is determined by the flux:

$$\Lambda = \max\{|\Lambda_i^{jk}|\} , \tag{1.5}$$

where

$$[r^i, r^j] = \sum_i \Lambda_i^{jk} r^i . \tag{1.6}$$

In 1970, Glimm and Lax went on to prove that for 2×2 systems like the p-system, periodic solutions decay at a rate $O(1/t)$ in the total variation, so long as the oscillation of the initial data is sufficiently small, see [4]. (The oscillation is equivalent to the supnorm $\text{Sup}|u(\cdot) - \tilde{u}|$, once an origin $\tilde{u}$ is chosen.) This was a triumph for the mathematical theory of shock-waves because it provides a quantitative estimate of the dissipation present in the zero dissipation limit of gas dynamics. However, the methods of Glimm and Lax give only a *short time* existence theorem for periodic solutions of the 3×3 system (1.1). Little is known about the

[1] See e.g., [10], where large total variation is allowed among components that are *almost planar.*

[2] The result was extended by Zumbrun to $n \times n$ systems which possess a full set of Riemann coordinates, [20].

3×3 case, the simplest setting in which the true physical entropy effects the time-irreversibility. *The long time existence problem for periodic solutions of the full Euler system* (1.1)–(1.2) *has remained open since that time.*

In this paper we prove the following theorem which demonstrates the large time stability of solutions for systems which possess a Riemann coordinate in one of the characteristic families, as does the Euler system. The theorem applies to solutions defined in a small enough neighborhood $\mathcal{U}$ of a reference state $u = \tilde{u}$:

Theorem 1. *Consider any* 3×3 *strictly hyperbolic system* (1.2) *that has a Riemann coordinate for one wave family in a neighborhood* $\mathcal{U}$. *Let* $V_0 > 0, d > 0$ *and* $T > 0$ *be given arbitrarily. Then there exists an* $\varepsilon \equiv \varepsilon(V_0, d, T)$, *such that, if the initial data* $u_0(\cdot)$ *of the Cauchy problem* (1.2) *satisfies*

$$\operatorname{Sup}|u_0(\cdot) - \tilde{u}| \leqq \varepsilon\,, \tag{1.7}$$

$$TV(u_0(\cdot)) \leqq V_0\,, \tag{1.8}$$

and

$$\|u_0(\cdot)\|_d < V_{\text{crit}}\,, \tag{1.9}$$

then the conservation law admits a weak solution up to time T *with bounded supnorm and bounded total variation, and*

$$TV\{u(\cdot, t)\} \leqq V_0 \exp(KT/d)\,, \tag{1.10}$$

where K *is a constant depending on the equations, and* $\|u_0(\cdot)\|_d$ *denotes the maximum total variation of the function* u *over intervals of length* d.

Theorem 1 identifies a new length scale d in the Cauchy problem, and determines a corresponding growth rate for the total variation of the solution, cf. [6]. It gives the existence of solutions up to an arbitrary time T in the case when the initial data has arbitrarily large total variation and sufficiently small supnorm. Indeed, the theorem really is a large total variation result for the Euler equations, because by taking the initial sup-norm sufficiently small, it allows for the case in which waves of arbitrarily large total variation all interact before time T. Using finite speed of propagation, Theorem 1 directly implies the large time existence of periodic solutions of the Euler equations, a nonlinear model for sound wave propagation in gas dynamics.

Our proof is based on new functionals and new estimates for the Glimm scheme, the identification of a new length scale in the problem, and the introduction of a new norm $\|\cdot\|_d$, which we call the "d-norm," that is natural for estimating the nonlinearities in the problem. There is no corresponding finite length scale that plays a similar role in 2×2 systems, or in $n \times n$ systems when the total variation of the initial data is sufficiently small.

The difficulty when the total variation is large and the system does not possess a full set of Riemann coordinates, (e.g., periodic solutions of (1.1) or (1.3)), is that when $TV\{u_0(\cdot)\} > V_{\text{crit}}$, there is a de-stabilizing, amplification effect due to the non-trivial Lie algebra of eigenvector fields $\{r^i\}$. For 2×2 systems, this geometric effect is not present because vector fields in the plane can always be rescaled to have pairwise vanishing Lie brackets. For 3×3 systems like (1.1), the Lie brackets play a dominant role, and recent studies of the geometrical optics approximation of (1.2) have demonstrated that certain 3×3 systems are *resonant*, and solutions can *blow up* in a finite time, independent of the supnorm, when the Lie algebra structure

of $\{r^i\}$ has a special form. (See [5, 9, 6].) *Thus Theorem* 1 *demonstrates the large time stability of periodic solutions of* (1.3), *thus ruling out resonant blowup in the Euler equations.*

To understand this de-stabilizing effect when the Lie algebra of $\{r^i\}$ is nontrivial, consider that for general systems, the nonlinear coupling of characteristic fields is determined by the nonlinear geometry of the eigen-fields (λ_i, r^i). The nonlinear coupling of different fields is manifested in the scattering of new waves when two waves from different families interact, and the Lie algebra of the eigenvector fields r^i encodes the description of the scattering process to leading (quadratic) order. The effect of interaction on a particular wave family is given by components of the Lie bracket of the eigenvector fields associated with the incident waves, times the product of the strengths of the incident waves, (an inherently quadratic effect), with an error that is at most *cubic* in incident wave strengths, [17]. The quadratic interaction effects become destabilizing when the total variation is large, even when the supnorm of the solution stays small, due to the new waves that are generated as a result of interactions. Indeed, when the supnorm is small, the waves in the solution are weak, and in this case the combined quadratic effects of a *small number* of interactions are dominated by the incident waves. For example, if the incident waves have strength $O(\varepsilon)$, then the cumulative effect of $O(1)$ interactions is $O(\varepsilon^2)$, which is small. This corresponds to a total variation of $O(\varepsilon)$, namely $O(1)$ waves of strength ε. Increasing the total variation to $O(1)$ while wave strengths remain small requires $O(1/\varepsilon)$ waves, so that there will be $O(1/\varepsilon^2)$ interactions. The cumulative effect of these interactions is then $O(1)$, and this drives the growth of the total variation of the solution when the initial total variation exceeds a critical value. This explains, in principle, why the large total variation, small oscillation problem is inherently different from the small total variation problem originally addressed by Glimm: as long as the total variation remains small, the quadratic effects can be estimated by the strengths of the initial waves. However, as the initial total variation increases, cumulative quadratic effects become dominant. These quadratic effects cause solutions to grow in amplitude, and in some cases this leads to resonance, and to the finite time blow-up of solutions.

This growth of solutions for general systems leads us to look for time-dependent bounds for the total variation of the solution, when the initial total variation is large. Thus, suppose that the total variation at time t in a solution of (1.2) is given by $V(t)$, and suppose that $V(t) \to \infty$ as t increases. The observation that system (1.2) is invariant under the scaling $(x, t) \to (\alpha x, \alpha t)$ means that by scaling, we can construct a solution that grows without bound in arbitrarily short finite times. This leads us to conclude that the norm on the initial data that controls the growth of the solution cannot be invariant under the scaling $(x, t) \to (\alpha x, \alpha t)$, as are the supnorm and the total variation norm. Said differently, by this argument there must be another length scale in the problem, call it d, which should determine the rate of growth of solutions when the total variation of the initial data is large and the supnorm is small.

With this in mind, we define the length scale d, which appears in the growth rate estimate $\exp(KT/d)$ in Theorem 1, as the length of the largest interval over which the total variation of the initial data is smaller than the critical total variation V_{crit} required for Glimm's method. For 3×3 systems with a Riemann coordinate, we relax this to require only the total variation in the Riemann coordinate be less than V_{crit} over intervals of length d. We define the d-norm of a function of bounded variation to be the supremum of the total variation over

intervals of length d,

$$\|u_0\|_d = \operatorname*{Sup}_a TV_{[a,a+d]}(u_0)\,.$$

The d-norm scales like a length under rescalings $(x,t) \to (\alpha x, \alpha t)$. Thus Theorem 1 states that, for systems with the same (unsigned) Lie bracket components as the Euler system (1.3), the total variation grows no worse than $O(1)\exp(KT/d)$, so long as $\|u_0(\cdot)\|_d < V_{\text{crit}}$. Thus the d-norm controls the growth rate of solutions in the class of systems considered here. Moreover, the examples given in [18, 15] show that this growth rate is sharp for the system

$$\begin{aligned} u_t + (w + 2uv)_x &= 0\,, \\ v_t &= 0\,, \\ w_t + (u(1-4v^2) - 2vw)_x &= 0\,, \end{aligned} \tag{1.11}$$

a linearly degenerate system contained within the class of systems to which Theorem 1 applies.

The length scale d plays no role in 2×2 systems, nor in $n \times n$ systems possessing a full set of Riemann coordinates, nor when the initial total variation is less than V_{crit}. Indeed, $\exp(KT/d) \to 1$ as $d \to \infty$, and our growth rate estimate reduces to the time independent estimate obtained by Glimm in the limit $V_0 < V_{\text{crit}}$. Moreover, in the limit $d \to \infty$, the d-norm and total variation coincide, and so our restriction on the initial data reduces to Glimm's. Thus we conclude that d represents a new length scale that is relevant to the Euler system (1.3) when the initial total variation is large.

In Sect. 2 we show that for any $\varepsilon > 0$, and any function f of bounded variation over $x \in \mathbf{R}$, there is a length d such that

$$\|f\|_d < \varepsilon\,.$$

Thus, for fixed initial data of bounded total variation and small supnorm, the length scale d is determined by the data. Our new length scale d is the small scale on which interaction effects *do not* accumulate, so that the growth of the solution is apparent only on the larger scale of the support of the initial data. Because there are $O(1)$ waves on this small length scale, the heuristic argument above indicates that interactions on this scale do not drive the growth of the solution. Rather, the growth of the solution is driven by the long-range effects of multiple interactions which occur over the longer scale of the support of the solution, which is the larger length scale. This heuristic point of view motivates our method of analysis, and clarifies the principles used in obtaining the equations of weakly nonlinear geometric optics, [5, 9, 6].

To prove Theorem 1, we use an extension of the *method of re-orderings* which was introduced by Young in 1991 [17]. The idea in the method of re-orderings is to sup the value of functionals over all possible "future" re-orderings of waves, each re-ordering representing a possible order in which waves could interact in the actual solution. By sup'ing over *all possible* re-orderings, one can obtain sharp estimates for functionals without having to determine the true ordering of interactions in an actual solution, a determination that would require refined estimates for the wave speeds as they evolve (nonlinearly) in the problem. In [17], Young used the method of re-orderings to account for a cancellation in the quadratic effects arising from the Lie brackets of the eigenvector fields – a cancellation that is based on the

bilinearity of the bracket. The idea is that, when one computes the sup, over sequences of consecutive waves of the same family at a given time $t > 0$, of the sums of the (signed) strengths of the waves in the sequence, one has an estimate for the supnorm of the solution at time t. Then, if one estimates this value by sup'ing over all possible reorderings of the initial waves with cumulative quadratic effects included, one obtains an estimate for the supnorm that increases only *third* order, instead of second order, at each interaction. The point is that this measure of the supnorm accounts for a cancellation in quadratic effects based on the bilinearity of the bracket – and by accounting for the cancellation, the improved *third order* estimate for the increase in the supnorm at interactions leads to an improvement in Glimm's original estimate of the supnorm at $t > 0$. Based on this better estimate, Young used a strategy of proof set out in [14] to show that any Glimm solution satisfying $TV\{u_0(\cdot)\} < V_{\text{crit}}$ is stable in the *supnorm*, while Glimm's original argument only demonstrated bounds in the *total variation norm*:

Theorem (Young 1991). *Let $u(x,t)$ denote any Glimm solution to which Glimm's original assumptions apply, so that*

$$TV\{u_0(\cdot)\} < V_{\text{crit}} . \tag{1.12}$$

Then

$$\text{Sup}|u(\cdot,t) - \tilde{u}| < C\,\text{Sup}|u_0(\cdot) - \tilde{u}| , \tag{1.13}$$

where C depends only on values of the flux f in a neighborhood of the solution.

In this paper, we use the method of re-orderings to define a functional V^* to bound the total variation, a functional Q^* to account for potential interactions (analogous to Glimm's original Q), as well as the functional P from [17] to bound the supnorm. The functionals V^* and Q^* are evaluated on a sequence of waves by sup'ing over all possible re-orderings of the waves up to time T, while accounting for *all* cumulative quadratic effects of interactions. (Our notation is that a * on the functional indicates that it is analogous to Glimm's, but includes all cumulative quadratic effects. The functional P is not starred as it is defined only slightly differently than in [17], where quadratic effects were counted but did not accumulate.) More specifically, these Glimm type functionals are defined at each time step of an approximate Glimm scheme solution as *sups* over all possible wave configurations that can be generated up to time T assuming that wave strengths are given exactly by the leading order linear and quadratic effects at interaction, these effects being determined by the Lie bracket structure constants at the state $\tilde{u}$ alone. In this way, V^* is defined so that the change in V^* between the incoming and the outgoing waves of an interaction diamond is non-positive except for third order errors; while the functional Q^* decreases by order D, the sum of the products of approaching waves entering the diamond, when the supnorm is small. The method of reorderings thus allows us to use analogous definitions of the functionals to Glimm's, only now they are exact at the quadratic level, whereas Glimm's were exact only at the linear level. In this way, the errors in our quadratic model (compared to the fully nonlinear problem) are *third* order. We demonstrate that the error between the change in the functionals recorded at the quadratic level and the change recorded in the full nonlinear problem is bounded by a third order error equal to the supnorm of the solution times D. Specifically, (we use the same notation as [17]),

$$P(J_+) - P(J_-) \leqq K_P SD , \tag{1.14}$$

$$V^*(J_+) - V^*(J_-) \leqq K_V SD , \tag{1.15}$$

and

$$Q^*(J_+) - Q^*(J_-) \leqq -D + K_Q SD\,, \tag{1.16}$$

where K_P, K_V and K_Q are positive constants that can be fixed apriori. We reiterate that in the large variation problem we must incorporate all future quadratic effects of interactions up to time T into the definitions of the functionals V^* and Q^*; and moreover, the quadratic effects represent the dominant contribution to the magnitude of these functionals. The functionals are constructed to satisfy the estimates (1.15)–(1.16) so that we can apply the proof strategy set out in [14]. (See also Glimm's original paper.)

If we let the "quadratic model" refer to the model for wave interactions in which the scattered waves are determined exactly by the Lie bracket structure constants at the state $\tilde{u}$, then the value of the functionals V^*, P and Q^* on a given sequence of waves is determined *entirely* at the quadratic level; and the problem of estimating the "third order errors" between the quadratic model and the full nonlinear problem, is equivalent to the problem of the *continuity* of the functionals defined at the quadratic level alone. The d-norm works perfectly for this method of analysis because it is stable under perturbation by a set of waves which sum to an arbitrarily small total variation. The third order errors introduce such waves into the scheme that are not accounted for at the quadratic level. In fact, the generation of wave strengths in the full nonlinear problem is a small perturbation of the strengths computed at the quadratic level alone, so that the *quadratic* system plays the same role in our argument that the *linear* system plays in Glimm's original argument.

Although we address the problem of 3×3 systems that have the same Lie bracket structure as Euler, our method of reducing the nonlinear problem to the problem of obtaining estimates up to the quadratic level, applies in principle to the large total variation small oscillation problem for systems of conservation laws in general. In order to state the general results that our methods can establish, we now discuss the mathematical issues in more detail.

The key to getting bounds in Glimm's method lies with the quadratic error potential Q, which is the global extension of the local interaction error. The potential Q is constructed by *a priori* inserting an error term for each pair of approaching waves, so that this term may be removed from Q when those waves interact. To motivate the definition of Q^*, note that the errors due to two consecutive groups of interactions (reorderings of waves) should be the sum of the errors due to the reorderings taken separately. We then need a continuity property to account for higher-order errors. This is the statement that the potential is stable under small perturbations of wave strengths. We extend the method of reorderings to include in Q^* all of the additional approaching waves generated by cumulative quadratic effects, so that Q^* is exact at the quadratic level. The new feature in our potential is that the interaction terms are given by wave strengths *at the time of interactions* (as estimated by projecting forward the quadratic effects of earlier interactions), rather than the initial wave strengths. We then show that continuity of the functional Q^*, uniform in mesh length Δx, is sufficient for the existence of solutions to general systems. In general, our bounds will depend on time and other factors, and restrictions used in bounding Q^* represent restrictions on the solution. Our methods are sufficient to prove the following theorem that applies to any system of conservation laws (1.2) that is strictly hyperbolic, and either genuinely nonlinear or linearly degenerate in each characteristic field: (We let $\|u_0\|_\infty \equiv \mathrm{Sup}|u_0(\cdot) - \tilde{u}|$, which is equivalent to the oscillation since $\tilde{u} = u_0(-\infty)$.)

Theorem 2. *Suppose that the nonlinear functional Q^* is bounded and continuous with time-dependent bounds. Then given any T and V_0, there is an $\varepsilon = \varepsilon(V_0, T) > 0$ such that if the initial data u_0 of the Cauchy problem* (1.2) *satisfies*

$$\|u_0\|_\infty \leqq \varepsilon \quad \text{and} \quad TV(u_0) \leqq V_0 ,$$

then the conservation law admits a weak solution up to time T with bounded oscillation and total variation. Moreover, the total variation of the solution is bounded by $V^*(t = 0) + O(\varepsilon)$.

Note that for resonant systems, Theorem 2 is vacuous after the blowup time, because Q^* will not be bounded.

We now discuss the class of systems considered in this paper. Motivated by gas dynamics, we restrict ourselves to the class of 3×3 conservation laws possessing a Riemann coordinate w. We show in the next section that after normalization, such systems have the same (unsigned) Lie bracket structure constants as does (1.3). The existence of the Riemann coordinate means that for smooth solutions, we can derive a transport equation

$$w_t + \lambda_k(u) w_x = 0$$

for the Riemann coordinate. This represents a weak decoupling of that family from the system, in that w is constant along k-characteristics

$$\frac{dx}{dt} = \lambda_k(u) ,$$

and so no growth in w occurs in smooth regions. Although the formation and interaction of weak shocks generates "entropy" w, this is a higher order effect which does not change the qualitative behavior of solutions. Indeed, the change in entropy upon shock formation and interaction is cubic in incident wave strength, so is neglected in our quadratic model. This allows us to treat the degenerate field as a static background source for the generation of sound waves. Moreover, the assumption means that *all sound waves interact linearly up to cubic errors*, which can again be ignored.

With these assumptions, we can describe the scattering of sound waves in our quadratic model, as follows. A single interaction of a sound wave with a contact causes a sound wave of the opposite family to be reflected, whose strength is the (weighted) product of the incident wave strengths. This reflected sound wave then interacts with other contacts, so that a pattern of multiply reflected and transmitted waves emerges. Our assumption implies that we can treat each scattered sound wave separately, and combine these linearly. We refer to a single scattered wave and its trajectory as a path.

A path is thus given by an initial wave, together with a list of interactions, where after interaction the path follows the *reflected* wave. We need only consider interactions between sound waves and entropy jumps. The strength of the wave contributed by an interaction is then the product of the incident sound wave strength with the corresponding entropy jump, and this entropy jump is determined by the initial configuration. The contribution due to a single path, which is a series of interactions, can then be calculated, and adding up contributions due to each path yields the following *path integral formula*. The solution is represented in the quadratic model by a sequence of individual waves $\gamma = (\gamma_1, \ldots, \gamma_n)$, and the interactions up to a certain time are represented by a reordering τ.

Theorem 3. *Suppose that the initial approximation is represented by γ and the approximation after reordering is $\delta = \tau\gamma$. Then the reordered sound wave δ_i at position i, (which represents the relative position of a wave in an approximate Glimm scheme solution), is given by*

$$\delta_i = \sum_j \gamma_j \sum_r \Lambda(i,r) \sum_{\Pi_r(j,i)} \gamma_{k_1} \cdots \gamma_{k_r} ,$$

where $\Pi_r(j,i)$ is the collection of paths starting at j and ending at i and consisting of r interactions. Here γ_j is the sound wave initially at position j, and the γ_k's are the strengths of the contacts which determine a particular path (the indices j and k refer to **relative** *wave positions); Λ is a weight which is independent of wave strengths.*

Once we have found this formula, we look for bounds for the functionals. Our approach is as follows: first suppose that each entropy jump is bounded by β. The contribution of a single path with r interactions is then $(\beta\Lambda)^r$. We are then left with the task of counting the number of paths which connect j to i through r interactions. To count the paths, we observe that a path starting at a fixed point can be determined by a sequence of interaction times $t_1 < \cdots < t_r$. Thus the number of paths is bounded by the number of choices of times $t_q = q\Delta t$, that is $\binom{T/\Delta t}{r}$. Combining these and using the binomial theorem then gives an exponential bound for the total variation generated by a single sound wave. Although this bound depends on the mesh length Δx, a similar argument is used to get bounds independent of Δx. For this, instead of considering each entropy jump separately, we group paths into "blocks" contained in x-intervals of length d, each of which generates only a small amount of growth. We overestimate the contribution due to each block by adjusting the length scale d downward, and count the number of decompositions into blocks as above. In this way, we get bounds for the Glimm approximation, uniform in the mesh length Δx.

Theorem 4. *If the sequence γ is such that its d-norm δ satisfies*

$$6\Lambda\delta < 1/2, \quad \text{for some fixed } d > 0 ,$$

then the total variation of the reordered sequence $\tau\gamma$ is bounded, and satisfies

$$V(\tau\gamma) \leqq V(\gamma) + V(\gamma)\exp(8\Lambda\delta T\lambda/d) .$$

In particular, this bound is uniform as the mesh size $\Delta x \to 0$. Here λ and Λ are constants determined by the flux.

As a corollary, we get the following (precise) re-statement of Theorem 1 which gives a large time large variation existence theorem for the Euler equations. This includes the important case of space-periodic solutions. In particular, we have shown that the Euler equations are non-resonant, so that solutions do not blow up in finite time.

Corollary. *Suppose that system (1.2) possesses a Riemann coordinate w. Suppose also that we are given large numbers V_0 and T, and positive d. Then there exists an $\varepsilon > 0$, such that if the initial data satisfies*

$$TV(u_0) < V_0, \quad \|u_0\|_\infty < \varepsilon \quad \text{and} \quad \|u_0\|_d = \delta < 1/12\Lambda ,$$

then a weak solution $u(x,t) \in \mathcal{U}$ exists up to time T. Moreover, this solution has exponentially bounded total variation,

$$TV(u(\cdot,t)) < TV(u_0)\exp(8\lambda\Lambda\delta t/d) + O(\varepsilon),$$

with similar bounds for the other norms. The finite speed of propagation then gives the same result for periodic initial data u_0, in which case the initial bounds apply on an interval of dependence up to time T.

Note that V_0 and T can be arbitrarily specified, as long as the sup-norm of the data is small enough. The constants λ and Λ are bounds for the wave-speed and interaction coefficients, respectively, and are determined by the flux. The requirement that $\delta < 1/12\Lambda$ serves to identify the appropriate length scale d.

Our theorem includes the case of initial data having Lipschitz continuous entropy. If the entropy $w(u_0)$ is Lipschitz, the ratio δ/d is bounded by the Lipschitz constant, and the theorem holds, where now the length-scale d is determined by the Lipschitz constant K, namely $d < 1/12\Lambda K$, and the corresponding growth rate is $\exp(8\lambda\Lambda K t)$. The Lipschitz norm has a length scale built in, but unlike the d-norm, is not stable under perturbation by weak waves, and so is not suitable for our analysis.

As a final comment, we note that one does not expect solutions to the equations of gas dynamics to grow exponentially, since the thermo-dynamic entropy is convex. In this case, the extra requirement that the total entropy $\int S$ be non-increasing seems to proscribe growth of the L_2-norm of solutions. This convexity of entropy is an extra symmetry, forcing the non-zero interaction coefficients to have opposite sign, which in turn leads to cancellation among newly generated waves, see [16]. In contrast, the system (1.11) is constructed to have constant wave-speeds and to possess a 2-Riemann coordinate, so that the nonlinearity is *entirely* manifested in the geometric coupling of different wave families. Moreover, although system (1.11) has the same nonzero Lie bracket components as the Euler system (1.3), the interaction coefficients in this example have the same sign, not alternating signs like Euler, and this is essential for the argument in [18] demonstrating that periodic solutions of (1.11) exhibit exponential growth, with growth rate given exactly by $O(KT/d)$, as in the theorem above. This example indicates that the length-scale d and corresponding d-norm are the correct quantities needed to describe the mechanism for growth of solutions identified here and it means that, *in the absence of extra assumptions, our theorem is sharp.*[3]

The paper is organized as follows. In Sect. 2, we recall Glimm's method and show that norms can be measured in terms of wave strengths. In Sect. 3, we recall and extend the method of reorderings and interaction maps, and describe the effects of the interactions represented by a reordering. In Sect. 4 we define the functionals and describe their properties. In Sect. 5 we carry out the induction for Glimm's method and prove Theorem 2. These results can be extended to general systems at the cost of clumsy notation. We then restrict ourselves to systems with a Riemann coordinate, and in Sect. 6 we define paths and derive the path integral formula. In Sect. 7, we count paths and show that the functionals are bounded and continuous.

[3] See [1] for an interesting new regularity result for solutions generated by Glimm's method that is related to the problem of the stability of solutions in the d-norm.

2. Preliminaries

2.1. The Glimm Scheme. The proof of Theorem 1 is based on an analysis of approximate solutions generated by the Glimm scheme [3, 13]. Our procedure is to obtain bounds for Glimm-type functionals that take account of interactions up to the quadratic level, and then to apply a general method of analysis that estimates norms for solutions in the fully nonlinear problem by these functionals. The general method for reducing the problem to the quadratic level applies in principle to an arbitrary $N \times N$ system – and we really need only restrict to 3×3 systems possessing a Riemann coordinate in order to demonstrate that our Glimm-type functionals are *bounded* and *continuous*. (The presence of resonance demonstrates that boundedness and continuity of the functionals will fail in general without special assumptions that restrict the Lie algebra of the eigenvector fields r^i.) Thus we now develop the theory for $N \times N$ systems and we do not restrict to 3×3 systems until Sect. 6. However, to keep notation as simple as possible, in Sect. 3.2 we describe the interaction maps in detail only for 3×3 systems.

We begin by recalling Glimm's approximation $\bar{u}$ of the Cauchy problem (1.2), and establish notation. We also describe Glimm's space-like I-curves, and define the functionals which will be used to show that the oscillation and total variation of the approximation are bounded for all times t. The bound on the oscillation allows us to define the approximation for all times, and the bound on total variation is used to extract a subsequence of approximations which converges to an exact (weak) solution of (1.2) as the grid size tends to zero.

We partition $\mathbf{R} \times \mathbf{R}^+$ by setting $x_j = j\Delta x$ and $t_k = k\Delta t$, where j and $k > 0$ are integers. The Glimm approximation consists of Riemann solutions pieced together in such a way that $\bar{u}$ is an exact weak solution in each of the strips $t_k < t < t_{k+1}$. Let $a = (a_1, a_2, \ldots)$ be an infinite equidistributed sequence, with each a_k taking values in the interval $(-1, +1)$; see [3, 8, 13]. This is the random sampling sequence, which is used to choose the approximation $\bar{u}$ on the line $t = t_k$. We use a *staggered grid* with sampling points $\theta_j^k \in \mathbf{R} \times \mathbf{R}^+$, where $j + k$ is odd, defined by

$$\theta_j^k = (x_j + a_k \Delta x, t_k)\,. \tag{2.1}$$

We define $\bar{u}$ as follows: for each j an odd integer, define constants $u_j^0 \in \mathcal{U}$ by $u_j^0 = u_0(x_j)$. Supposing that we have defined $\bar{u}(x,t)$ for all $x \in \mathrm{R}$ and $t < t_k$, together with constant states u_j^i, for all integers j and $i < k$ with $j + i$ odd, we show how to define the approximation $\bar{u}(x,t)$ for times $t < t_{k+1}$. For each integer j such that $j + k$ is odd, define the constant u_j^k by

$$u_j^k = \bar{u}(\theta_j^k -) = \lim_{t \to t_k^-} \bar{u}(x_j + a_k \Delta x, t)\,, \tag{2.2}$$

and define $\bar{u}$ on the line $t = t_k$ by

$$\bar{u}(x, t_k) = u_j^k, \quad \text{for } x_{j-1} < x < x_{j+1}\,. \tag{2.3}$$

Now define the approximation $\bar{u}$ in the strip $t_k < t < t_{k+1}$ by solving a Riemann problem at each grid point (x_{j+1}, t_k). The Riemann problem is the initial value problem for a jump discontinuity between constant states. We let $\langle u_L \| u_R \rangle$ denote the solution of the Riemann problem for left state u_L and right state u_R, which consists of a sequence of admissible elementary waves (shock waves, rarefaction

waves, or contact discontinuities), one from each eigenfamily, as first constructed by Lax, cf. [2, 13]. Thus, in Glimm's scheme, for each i $(= j + 1)$ with $i + k$ even, we solve the Riemann problem $\langle u_{i-1}^k \| u_{i+1}^k \rangle$, centered at the point (x_i, t_k). Therefore, we have N waves leaving each point (x_i, t_k), with the constant states u_{i-1}^k and u_{i+1}^k to the left and right, respectively. These "local" Riemann solutions can be pieced together to give an exact solution along the vertical segments $x = x_j, t_k < t < t_{k+1}$, for those j for which $j + k$ is odd, since the adjacent Riemann solutions share the states u_j^k.

This construction is valid as long as the individual waves do not meet: we thus impose a Courant–Friedrichs–Lewy (CFL) condition. This asserts that the mesh sizes in the grid are chosen so that

$$\Delta x > \Delta t \cdot \sup_i \{|\lambda_i(u)| : u \in \mathscr{U}\} . \tag{2.4}$$

We now describe Glimm's I-curves and establish notation for the sequel. We wish to bound the sup- and T.V.- norms of the approximation $\bar{u}(\cdot, t)$ as a function of x for each time t. Notice, that for all times t between the lines $t = t_k$ and $t = t_{k+1}$, the approximation $\bar{u}(\cdot, t)$ consists of exactly the same constant states and intermediate waves. It is therefore necessary only to keep track of the waves occurring in the scheme. We define the functionals on the class of I-curves in terms of those waves which cross each I-curve.

An I-curve is a continuous space-like curve made up of line segments joining sampling points θ_j^k and *either* θ_{j+1}^{k-1} *or* θ_{j+1}^{k+1} (but not both). In other words, an I-curve is given by the assigning to each integer j a positive $l(j)$, with $j + l(j)$ odd, satisfying $l(j + 1) = l(j) \pm 1$, and the I-curve consists of the segments connecting all points $\theta_j^{l(j)}$ and $\theta_{j+1}^{l(j+1)}$. For each integer $k \geqq 0$, there is a unique I-curve J_k connecting all points θ_j^k and $\theta_{j\pm1}^{k+1}$. Thus all waves appearing in the scheme between times t_k and t_{k+1} cross the curve J_k. In particular, J_0 is the unique I-curve which meets all mesh points on the line $t = 0$.

The I-curves admit a partial ordering given as follows: curve J' precedes curve J'' if the curve J'' lies towards later time; that is, if $\theta_j^{l(j)}$ and $\theta_j^{l'(j)}$ lie on J' and J'', respectively, then $l(j) \leqq l'(j)$ for each j. Clearly J_0 is the minimal curve. If J_+ is an immediate successor of J_-, then there is a single integer j', such that $l_-(j) = l_+(j)$ for $j \neq j'$, and $l_+(j') = l_-(j') + 2$.

The difference between these curves J_+ and J_- forms the *diamond* Δ_j^k centered at the point (x_j, t_k), where $k = l_-(j') + 1$. That is, the diamond Δ_j^k consists of the segments joining the points θ_{j-1}^k to $\theta_j^{k\pm1}$ and $\theta_j^{k\pm1}$ to θ_{j+1}^k. Thus for each j and $k > 0$ with $j + k$ even, there is unique diamond Δ_j^k enclosing the mesh point (x_j, t_k). We can compare diamonds to I-curves, by saying that a diamond Δ precedes J if it lies below J, or simply if the point (x_j, t_k) enclosed in the diamond lies below the I-curve J. Figure 1 is a schematic of the Glimm approximations together with I-curves and a single diamond.

We shall be considering those waves which cross a particular I-curve J, and we shall write $\gamma_j \in J$ if the wave γ_j crosses J. By a *sequence of waves* we mean a collection $\gamma = (\gamma_1, \ldots, \gamma_n)$ of consecutive single waves γ_j separated by constant states, which are usually suppressed. Thus there is a single constant state u_j between the waves γ_j and γ_{j+1}, and u_j is connected to u_{j-1} by a wave strength γ_j. We write $\gamma = (\gamma_1, \ldots, \gamma_n) \subset J$ for a sequence of waves crossing J. Similarly, we refer to waves and sequences thereof entering or leaving a diamond Δ. If the single wave γ_j is a k-wave, we shall say that it is in the k^{th} *family*.

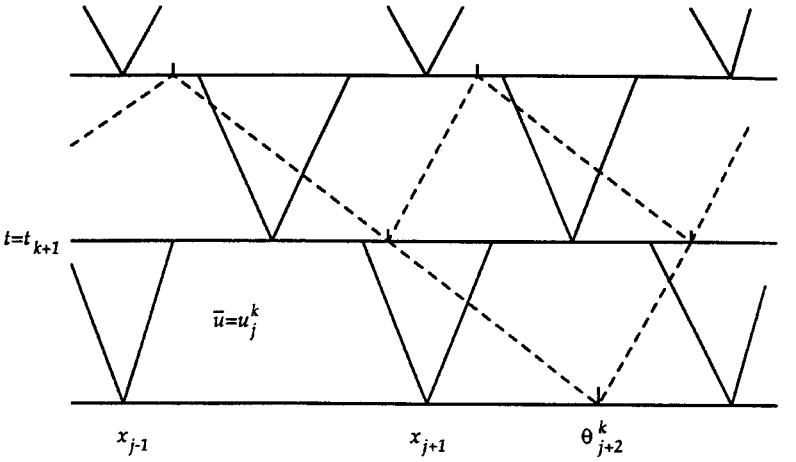

Fig. 1. The difference scheme

2.2. Norms. Glimm's original analysis uses the oscillation and total variation norms, given for a function $u(x)$ by

$$\|u\| = \sup |u(x) - u(x')| , \tag{2.5}$$

and

$$TV(u) = \sup \sum |u(x_{i+1}) - u(x_i)| , \tag{2.6}$$

respectively, the sum being taken over increasing finite partitions $\{x_i\}$. Note that the oscillation is equivalent to the sup-norm if an origin $\tilde{u}$ is specified, but the former is more convenient for our purposes, and we shall use both terms to mean the oscillation. We always assume that solutions take values in a small neighborhood $\mathcal{U}$ of $\tilde{u}$.

In this paper, we shall also require an estimate of the local variation of the solution over intervals of length d. We thus define the d-norm

$$\|u\|_d = \sup_a TV_{[a,a+d]} (u(\cdot))$$

to be the maximum variation of a function $u(x)$ over intervals of length d. Note that the d-norm increases at most linearly with d, $\|u\|_{d+d'} \leqq \|u\|_d + \|u\|_{d'}$. In the sequel, we shall require that the d-norm be small for some positive d, and the largest d for which this is small enough will determine the lengthscale on which wave interactions produce no significant growth effect. Any growth in solutions is driven by cumulative wave interactions on the larger lengthscale of the support of the solution. Our requirement that the d-norm be small represents no real restriction, according to the following lemma.

Lemma 2.1. *Suppose that the total variation of $u(\cdot)$ is finite. Then, given any $\varepsilon > 0$, there is a positive d (depending on the function u) such that*

$$\|u\|_d \leqq \|u\| + \varepsilon , \tag{2.7}$$

where $\|u\|$ is the oscillation.

Since we are restricting ourselves to solutions with small oscillation, we can always arrange that some d-norm is small by choosing ε small enough in the lemma. Our lengthscale is then determined by finding the *largest* d such that the d-norm of the solution is smaller than some constant which depends on the conservation law. In the lemma, our choice of d depends on the function u, so that the lengthscale is determined by the initial data. We shall see that this lengthscale in turn affects the rate of growth of solutions.

Proof. Suppose that the total variation of $u(\cdot)$ is given by V, and let η be given. Choose some partition $x_0 < x_1 < \cdots < x_n$ such that the total variation calculated by summing over all indices but one is near V, that is

$$V < \sum_{i \neq j-1, j} |u(x_{i+1}) - u(x_i)| + |u(x_{j+1}) - u(x_{j-1})| + \varepsilon ,$$

for each j. Now let

$$d = \min_i |x_i - x_{i-1}| ,$$

and let any interval $[a, b]$ with $b - a \leqq d$ be given. Then there is a j such that $x_{j-1} \leqq a \leqq b \leqq x_{j+1}$, and so we have

$$\sum_{i \neq j-1, j} |u(x_{i+1}) - u(x_i)| + TV_{[a,b]}(u) \leqq V .$$

Thus we obtain

$$TV_{[a,b]}(u) < |u(x_{j+1}) - u(x_{j-1})| + \varepsilon ,$$

and the lemma is proved. □

Our approximations are built up from Riemann solutions, so that they are piecewise constant with interpolations between (some of the) constant states, and can be treated as piecewise constant when considering oscillation and variation norms. It is convenient to consider functionals which are equivalent to these norms, but which are defined in terms of wave strengths without explicit reference to the intermediate states. Our notation is as follows. We consider the sequence of constant states $\{u_0, \ldots, u_n\}$ separated by the wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$ of index $(c_1, \ldots, c_n)$, cf. [17]. This means that the wave separating constant states u_{i-1} and u_i is a wave of (signed) strength γ_i from the c_i^{th} family, where c_i lies between 1 and N. (Although we can consider any wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$, we have in mind sequences of consecutive waves along an I-curve of the Glimm scheme.) We define the strength of a wave γ_i of the c_i^{th} family with left state u_L and right state u_R by

$$\gamma_i = l_{c_i}(\tilde{u}) \cdot (u_R - u_L) , \tag{2.8}$$

where $l_{c_i}(\tilde{u})$ is the c_i^{th} left eigenvector of Df evaluated at the origin $\tilde{u}$. (We let γ_i refer to both the wave as well as the signed strength of the wave.) Note that a given normalization of the eigenvectors $\{r^j\}$ determines the wave strength measure (2.8) through the normalization

$$l_i(u) \cdot r^j(u) = \delta_{ij} .$$

For genuinely nonlinear fields, a positive strength means that the wave is a rarefaction, while a negative strength indicates a shock. For any wave sequence

$\gamma = (\gamma_1, \ldots, \gamma_n)$, we define

$$V(\gamma) = \sum_{i=1}^{n} |\gamma_i| \quad \text{and} \quad S(\gamma) = \max_{j_1 < j_2} \left| \sum_{i=j_1, c_i = r}^{j_2} \gamma_i \right|, \tag{2.9}$$

where the latter denotes the largest sum of signed strengths of consecutive waves in the same family that occurs among $\gamma_1, \ldots, \gamma_n$. This measure of the sup-norm accounts for cancellation between shock and rarefaction waves, cf. [17]. Similarly, we define a functional for the local variation of the entropy,

$$H(\gamma) = \sup_i \sum_{k=i}^{l} |\gamma_k| ,$$

where the sum is over 2-waves ($c_k = 2$), and the distance between the 2-waves γ_i and γ_l is less than d. Note that the limits of the sum in the definition of H will depend on the mesh size $\Delta x \ll d$. It is convenient to define the strength of a 2-wave by the difference in the Riemann coordinate (entropy jump) across the wave.

These functionals are exact forms of the corresponding norms for linear equations, and the following lemma shows that they can be used for the nonlinear case as long as the oscillation is small enough.

Lemma 2.2. *If the total variation of the approximation $\bar{u}$ is bounded and the oscillation is small enough, these quantities are equivalent as norms to the functionals V and S defined in* (2.9), *respectively. For each k, we have*

$$u_k = u_0 + \sum_{j \leqq k} \gamma_j r^{c_j} + O(1) S , \tag{2.10}$$

where $O(1)$ is a constant depending on f and V. Here r^i denotes the i^{th} right eigenvector of the matrix Df, evaluated at the extreme left state u_0. In the sequel we assume that $u_0 = u_0(-\infty) = \tilde{u}$ is fixed.

The advantage gained by using the functionals S and V comes in being able to evaluate these quantities in terms of the wave strengths only, without having to know the intermediate constant states u_i explicitly.

Proof. Here and later, unless explicitly shown, all vector quantities are evaluated at the state $u_0 = \tilde{u}$, which we treat as the origin. Let O_1 denote a constant depending only on f. According to Lax's solution of the Riemann problem [2], we have

$$u_k = u_{k-1} + \gamma_k r^{c_k}(u_{k-1}) + O_1(|\gamma_k|^2) , \tag{2.11}$$

which we rewrite as

$$u_k - u_{k-1} = \gamma_k r^{c_k} + O_1 |\gamma_k| \, \|\bar{u}\| .$$

We immediately obtain

$$\begin{aligned} S(\gamma) &= \sup \left| \sum_{c_j = r} \gamma_j \right| = \sup \left| \sum_j \ell_r \cdot (\gamma_j r^{c_j}) \right| \\ &\leqq \sup |\textstyle\sum \ell_r \cdot (u_j - u_{j-1})| + O_1 \textstyle\sum |\gamma_j| \, \|\bar{u}\| \\ &\leqq O_1 \|\bar{u}\| (1 + V) . \end{aligned}$$

Similarly,

$$V(\gamma) = \sum |\gamma_i| = O_1 TV(\bar{u})(1 + \|\bar{u}\|)\,, \tag{2.12}$$

so that $V(\gamma)$ and $TV(\bar{u})$ are equivalent. We now show that some multiple of S bounds the oscillation $\|\bar{u}\|$.

Define

$$V_k = \sum_{i \leqq k} |\gamma_i| \leqq V\,,$$

$$E_k = u_k - \left(u_0 + \sum_{j \leqq k} \gamma_j r^{c_j}\right),$$

and

$$A_k = \max_{j \leqq k} |E_j|\,,$$

so that we must find a bound for each A_k. From (2.11) we get

$$u_{k+1} = u_k + \gamma_{k+1} r^{c_{k+1}} + |\gamma_{k+1}| O_1(S + |u_k - u_0|)\,,$$

where $r^{c_{k+1}}$ is evaluated at u_0. Then

$$\begin{aligned} |u_k - u_0| &= \left| E_k - \sum_{j \leqq k} \gamma_j r^{c_j} \right| \\ &\leqq |E_k| + \sum_{q=1}^{N} \left| \sum_{j \leqq k,\ c_j = q} \gamma_j r^q \right| \leqq A_k + O_1(S)\,, \end{aligned}$$

so that

$$\begin{aligned} A_{k+1} - A_k &\leqq |E_{k+1} - E_k| \\ &= |u_{k+1} - u_k - \gamma_{k+1} r^{c_{k+1}}| \\ &\leqq |\gamma_{k+1}|(O_1 S + O_1 A_k)\,. \end{aligned}$$

Taking $A_0 = 0$ and summing, we get

$$\begin{aligned} A_{k+1} &\leqq \sum_{j=1}^{k+1} |\gamma_j|(O_1 S + O_1 A_{j-1}) \\ &= O_1 V_{k+1} S + O_1 \sum_{j=1}^{k+1} |\gamma_j| A_{j-1}\,. \end{aligned}$$

We now write $|\gamma_{k+1}| = V_{k+1} - V_k$, and view A_k as a function of V_k, so that we have a discrete Gronwall inequality,

$$A_{k+1} \leqq O_1 SV + O_1 \sum_{j=1}^{k+1} A_{j-1}(V_j - V_{j-1})\,.$$

Thus,

$$A_{k+1} \leqq O_1 SV \exp(O_1 V_{k+1}) \leqq O_1 SV \exp(O_1 V) = SO(1)\,,$$

where $O(1)$ depends on f and V, and the lemma is proved. □

We note that by extending this expansion to include exact second order terms, a similar result holds, namely

$$u_k = u_0 + \sum_{i \leqq k} \gamma_i r^{c_i} + \frac{1}{2} \sum_{i \leqq k} \gamma_i^2 r^{c_i} \cdot \nabla r^{c_i} + \sum_{i<j} \gamma_i \gamma_j r^{c_i} \cdot \nabla r^{c_j} + O(S^2)\,.$$

Note that for large V, the quadratic terms in this expansion are of size $O(SV)$, and so are comparable to the linear terms.

2.3. Riemann Coordinates. We recall the definition of a Riemann coordinate, and note the difference between Riemann coordinates and Riemann invariants as defined by Lax. In this paper, we shall exclusively deal with Riemann coordinates.

Definition 2.3. *A* k**-Riemann coordinate** *for system* (1.2) *is a function* $w : \mathcal{U} \to \mathrm{R}$ *whose gradient* $\nabla_u w = l_k(u)$ *is a* k^{th} *left eigenvector of the flux matrix* Df, *so that*

$$\nabla_u w \cdot Df = \lambda_k(u) \nabla_u w$$

for the corresponding eigenvalue $\lambda_k(u)$.

The definition implies that for smooth solutions, we can derive a transport equation

$$w_t + \lambda_k(u) w_x = 0$$

for the Riemann coordinate, although we cannot in general rewrite this in conservation form. This represents a weak decoupling of the Riemann coordinate from the system, in that w is constant along k-characteristics

$$\frac{dx}{dt} = \lambda_k(u) ,$$

and so no growth in w occurs in smooth regions. Although the formation and interaction of weak shocks generates "entropy" w, we shall see that this is a higher order effect which does not change the qualitative behavior of solutions.

We remark that this definition is more restrictive than Lax's definition of a Riemann invariant, namely a function v satisfying $r^k \cdot \nabla v = 0$. Riemann invariants always exist for each family, but the existence of a Riemann coordinate implies the presence of certain symmetries. Indeed, a k-Riemann coordinate is a j-Riemann invariant for each $j \neq k$, and these are equivalent for 2×2 systems, but Riemann coordinates do not exist for general systems of three or more equations.

In the equations of gas dynamics, a consequence of the Law of Thermodynamics is that the entropy S is a Riemann coordinate. We interpret the entropy equation (for smooth solutions) as saying that changes in entropy are advected, but as there is no source term in the equation, no entropy is generated to leading order. Indeed, we shall see that all changes in entropy due to interactions of weak shocks are cubic in wave strength.

In the sequel we restrict our attention to 3×3 systems that possess a 2-Riemann coordinate, and we will refer to this as the entropy. For systems with a Riemann coordinate, the corresponding Lie algebra structure constant, or "interaction coefficient," vanishes. The following two lemmas provide a canonical form for the structure constants, cf. [19]. We include the proofs for completeness.

Lemma 2.4. *Let the Lie algebra structure constants* Λ_i^{jk} *be defined by*

$$[r^j, r^k] = \sum_i \Lambda_i^{jk} r^i , \tag{2.13}$$

and assume there exists a p-Riemann coordinate. Then we have

$$\Lambda_p^{jk} = 0 \quad \text{for all } j, k . \tag{2.14}$$

Moreover, for general hyperbolic systems (1.2), a judicious choice of normalization of the eigenvector fields $\{r^i\}$ can further simplify the structure constants at the state $\tilde{u}$:

Lemma 2.5. *Assume* (1.2) *is hyperbolic. There exists a normalization of* r^i *such that*

$$\Lambda_i^{jk}(\tilde{u}) = 0 \quad \text{for } i = j \text{ or } i = k\,. \tag{2.15}$$

Proof. To establish (2.14), assume first that

$$\nabla w = l_i(u)\,,$$

so that

$$(r^j \cdot \nabla)w = \delta_{ij}\,. \tag{2.16}$$

Now differentiate (2.16) along r^k to obtain

$$(r^k \cdot \nabla r^j) \cdot \nabla w + D^2 w(r^k, r^j) = 0\,. \tag{2.17}$$

But since $D^2 w(r^k, r^j)$ is symmetric in (r^k, r^j), we conclude from (2.17) that

$$l_i \cdot (r^k \cdot \nabla r^j - r^j \cdot \nabla r^k) = l_i \cdot [r^k, r^j] = 0\,,$$

from which (2.14) follows directly.

To establish (2.15), we must normalize the vector fields in such a way that all the coefficients Λ_i^{jk} vanish at the point $\tilde{u} \in \mathcal{U}$, unless i, j, and k are distinct.

To this end, suppose that for each p, we are given a function z_p, with

$$z_p(\tilde{u}) = 0 \quad \text{and} \quad \nabla_u z_p = \sum_j a_{pj}\, \ell_j$$

for some functions a_{pj}. We form the vectors

$$\bar{r}^p = e^{z_p}\, r^p \quad \text{and} \quad \bar{\ell}_p = e^{-z_p}\, \ell_p\,,$$

so that $\bar{r}^p|_{\tilde{u}} = r^p|_{\tilde{u}}$, and calculate directly that

$$r^j \cdot \nabla_u \bar{r}^k = e^{z_k}(a_{kj} r^k + r^j \cdot \nabla r^k)\,. \tag{2.18}$$

Now, if $\bar{\Lambda}_i^{jk}$ are the interaction coefficients corresponding to the $\bar{r}^p$'s,

$$\bar{\Lambda}_i^{jk} = \bar{\ell}_i \cdot [\bar{r}^j, \bar{r}^k]\,,$$

we have

$$\bar{\Lambda}_i^{jk} = e^{z_j + z_k - z_i}(a_{kj}\delta_{ik} + \Lambda_i^{jk} - a_{jk}\delta_{ji})\,. \tag{2.19}$$

We now choose the functions z_k, which determine a_{kj}, in a convenient way. We would like to set $a_{kj} = \Lambda_k^{kj}$, so that $\bar{\Lambda}_k^{kj} = 0$, for each $k \neq j$. To do this throughout the neighborhood $\mathcal{U}$ is not generally possible, but we can do it at the origin $\tilde{u}$.

For example, we may choose

$$a_{pp} = -\ell_p \cdot (r^p \cdot \nabla r^p)\,,$$

so that throughout $\mathcal{U}$,

$$\bar{\ell}_p \cdot (\bar{r}^p \cdot \nabla \bar{r}^p) = e^{z_p}(a_{pp} + \ell_p \cdot (r^p \cdot \nabla r^p)) = 0\,.$$

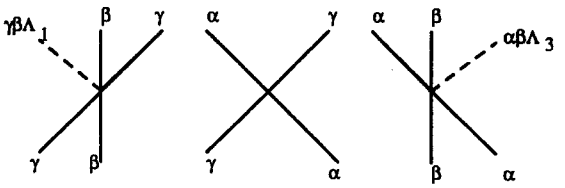

Fig. 2. Interactions

This choice of the function a_{pp} amounts to finding a solution z_p of the first order equation

$$r^p \cdot \nabla z_p = -\ell_p \cdot (r^p \cdot \nabla r^p)$$

in $\mathcal{U}$, where the right-hand side is known. This equation can be solved for prescribed Cauchy data on any non-characteristic hypersurface. We choose this surface and data so that the interaction coefficients vanish at the point $\tilde{u}$. For example, we choose a hyperplane whose tangent space at $\tilde{u}$ is the span of the vectors ℓ_j, where $j \neq p$, and take Cauchy data z_p which satisfies

$$\nabla z_p|_{\tilde{u}} = \sum_{j \neq p} \Lambda_p^{pj}|_{\tilde{u}} \ell_j, \quad \text{so that } r^j \cdot \nabla z_p|_{\tilde{u}} = \Lambda_p^{pj}(\tilde{u}), \quad \text{for } j \neq p\,.$$

Now, since $a_{pj} = r^j \cdot \nabla z_p$, we conclude from (2.19) that

$$\bar{\Lambda}_i^{jk}(\tilde{u}) = \Lambda_i^{jk}(\tilde{u}) \quad \text{for } i, j \text{ and } k \text{ distinct, and}$$

$$\bar{\Lambda}_i^{jk}(\tilde{u}) = 0 \quad \text{otherwise}\,. \qquad \square$$

In Sect. 6 we restrict to 3×3 systems in which the eigenvector fields $\{r^i\}$ are normalized so that (2.14) (with $p = 2$) and (2.15) hold, in which case the only nonzero coefficients Λ_i^{jk} at the state $\tilde{u}$ are

$$\Lambda_1 \equiv \Lambda_1^{32}(\tilde{u})\,, \tag{2.20}$$

and

$$\Lambda_3 \equiv \Lambda_3^{21}(\tilde{u})\,. \tag{2.21}$$

2.4. Interactions. The Riemann problem is the initial value problem in which the data consists of two constant states, whose solution consists of $N+1$ constant states separated by elementary waves. Suppose we are given two adjacent Riemann solutions, α separating u_L from u_M, and β separating u_M from u_R. The *Riemann interaction problem* is to resolve the Riemann problem with constant states u_L and u_R, and to express the resulting waves in terms of the strengths α and β. If the resulting waves are given by ε_k, then we have

$$\varepsilon_i = \alpha_i + \beta_i + \sum_{j>k} \alpha_j \beta_k \Lambda_i^{jk} + O(\|u - u_0\|) D(\alpha, \beta), \tag{2.22}$$

where $\Lambda_i^{jk} = \Lambda_i^{jk}|_{\tilde{u}}$ are the interaction coefficients, given by the Lie brackets of eigenvector fields (2.13). Estimate (2.22) is a refinement of Glimm's original estimate, cf. [17]. Here D is a quadratic functional measuring the amount of wave interaction, and is defined as follows: we say two waves are approaching if they would eventually interact in the absence of all other waves; this is the case when

the wave starting on the left is in the faster family, or both waves are in the same family and one of them is a shock. The amount of interaction D is then defined by

$$D(\alpha, \beta) = \sum_{\text{App}} |\alpha_p \beta_q| , \tag{2.23}$$

the sum being over all approaching pairs of waves. Note that all second order terms in the estimate (2.22) appear in D, so that

$$\varepsilon_i = \alpha_i + \beta_i + O(D) , \tag{2.24}$$

and moreover the coefficients A_i^{jk} in (2.22) can be evaluated at the origin $\tilde{u}$ rather than at the specific states u_L, etc.

Restricting to 3×3 systems possessing a Riemann coordinate, we assume that A_1 and A_3 given by (2.20) and (2.21) are the only nonzero interaction coefficients at the state $\tilde{u}$. With these simplifying assumptions, we can represent the interactions which occur schematically, with all strengths of outgoing waves given to within a third order error, as in Fig. 2. We again mention the fact that no new entropy waves are generated to leading order. This observation will be crucial in Sects. 6 and 7, and indeed it is known that without some assumption being made, solutions can explode in amplitude in finite time, [5, 18].

3. The Method of Reorderings

We recall the method of reorderings and define the nonlinear functionals which will be used to bound the oscillation and total variation of the Glimm approximants. The main idea in Glimm's analysis is to build a decreasing potential which measures the errors produced from future interactions. For initial data having small total variation, this potential was constructed by Glimm, and was shown to be bounded by the square of the total variation. Similarly, in case there is a coordinate system of Riemann invariants, a quadratic functional is sufficient.

However, in general systems of three or more equations, when the initial total variation is large, the quadratic effects of interactions become important and lead to a variety of destabilizing phenomena. In this case we must update the potential for interaction effects as new waves are scattered by earlier interactions.

It is enough to consider only quadratic effects, as these control higher order effects. It is important to note, however, that for general systems, it is not possible to bound the quadratic effects, and the solution may grow unboundedly unless restrictions are placed on the Lie algebra of eigenvector fields. Our discussion is general, but the notation for the interaction maps is developed in Sect. 3.2 only for 3×3 systems possessing a Riemann coordinate.

3.1. Reorderings. We briefly recall the theory of reorderings introduced in [17], and use these to build the nonlinear functionals. Our notation is as follows: we suppose that the mesh-size Δx is fixed, and consider the Glimm approximation corresponding to this mesh-size on a given spacelike I-curve. By finite speed of propagation we assume without loss of generality that the approximation is compactly supported. The approximation then consists of a sequence of constant states $\{u_0, \ldots, u_n\}$, each

pair of states being separated by a single wave. We shall refer to the wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$, where the wave γ_i separates the states u_{i-1} and u_i. If γ_i is a c_i-wave, we say that $(c_1, \ldots, c_n)$ is the index of the wave sequence. The index c_i identifies the family of γ_i, so that the wave strength determines the wave completely; for this reason we do not distinguish between the wave and its strength.

We shall model the evolution of the Glimm approximation through changes in this wave sequence, as follows. The propagation and interaction of waves correspond to changes in the order of the wave sequence, and consequent changes in the strengths of the individual waves. We shall model these phenomena in an abstract setting, thus reducing the problem to an algebraic one. The symmetric group S_n acts on the waves by permuting them, and we shall also allow waves to merge. This allows us to more accurately capture interaction errors and nonlinear decay, while requiring more care in definitions. We shall first consider a large class of maps, and then restrict to those which are admissible, both from a physical point of view, and as determined by Glimm's scheme.

A surjective map $\tau : \{1, \ldots, n\} \to \{1, \ldots, m\}$ defines an action on sequences of n waves through permutation of subscripts. That is, the map τ acts on the wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$ by shifting the wave γ_i from position i to position $\tau(i)$. If the new sequence is $\delta = \tau(\gamma)$, we have $\delta_{\tau(i)} \approx \gamma_i$, these being equal at the linear level. We say the map τ *reorders*, or is a reordering of, the sequence γ. We similarly define the action of τ on the index $c = c(\gamma)$ by permutation of subscripts. In order for the reordered index to be single-valued, we shall require that if $\tau(i) = \tau(j)$, then $c_i = c_j$. This corresponds to the collapse of two waves of the same family into a single wave. Note that an abstract reordering keeps track of the possible future relative positions of the original waves, but does not tell us actual positions, nor how the strengths of these waves change.

We can define a composition of maps as long as we take care that domains and ranges match up correctly, for example, if τ maps n numbers to m and σ maps m numbers to p, then the composition $\sigma\tau$ is a well-defined map of n numbers to p. In the sequel we shall implicitly regard all compositions as consistent, without further regard for the domains of maps.

The permutation τ identifies the positions of the reordered waves, but also carries information about the interactions that must have occurred in reordering the waves. In order to get at this structure, we define the *crossing set* C_τ by

$$C_\tau = \{(j,k) \mid j < k \text{ and } \tau(j) > \tau(k)\},$$

so that C_τ identifies the initial positions of those pairs of waves which will cross under the action of τ. Similarly, since our maps are not necessarily injective, we identify the *merge set*

$$M_\tau = \{(j,k) \mid j < k \text{ and } \tau(j) = \tau(k)\},$$

consisting of those pairs of waves which are joined under the action of τ. It is convenient to define the interaction set I_τ as the union of crossing and merge sets. We extend the action of σ, and refer to the pre-image under σ of C_τ as

$$\sigma' C_\tau = \{(j,k) \mid (\sigma j, \sigma k) \in C_\tau\},$$

with similar notation for other sets.

We now isolate those maps which give rise to "physical" reorderings of waves. Due to hyperbolicity, in order for a pair of waves to cross, that is to pass through

each other, the wave on the left should travel faster than that on the right. Physically, waves in the same family may merge to form one wave, but never cross. Thus for a reordering τ of the wave sequence γ, we require that in order for the waves γ_j and γ_k to cross (where $j < k$), the wave γ_j should be in the faster family, while waves which merge should be in the same family. We express this symbolically as

Condition R: *For every pair* $(j,k) \in C_\tau$, *we must have* $c_j > c_k$, (R) *while if* $(j,k) \in M_\tau$, *then* $c_j = c_k$.

We say that the reordering τ of γ is *admissible* if it satisfies (R), and we write $\tau \in A(\gamma)$, where $A(\gamma)$ is the set of all admissible reorderings of the sequence γ. We note that the set $A(\gamma)$ is determined only by the index $c = c(\gamma)$ of the sequence, and does not depend on the strengths (or actual speeds) of the individual waves, so that any reordering admissible for the sequence γ is admissible for all sequences with the same index as γ. We now enumerate some of the properties of admissible reorderings, which can be derived directly from Condition (R) cf. [17].

The product of admissible reorderings is admissible. Physically, once a pair of waves has crossed, these waves will diverge from each other, and any pair of waves can cross or merge at most once. Given a composition of two admissible reorderings, we express the composite crossing and merge sets as

$$C_{\sigma\tau} = C_\tau \cup \tau' C_\sigma, \tag{3.1}$$

and

$$M_{\sigma\tau} = M_\tau \cup \tau' M_\sigma\,, \tag{3.2}$$

where $\cup$ denotes a *disjoint* union.

According to Condition (R), even though the inverse $\tau'(j)$ of j is not well-defined, its index $c_{\tau'(j)}$ is. This enables us to check admissibility of a map σ of the *reordered* sequence $\tau(\gamma)$, by verifying that if $(j,k) \in I_\sigma$, then either $c_{\tau'(j)} > c_{\tau'(k)}$ or $c_{\tau'(j)} = c_{\tau'(k)}$, as appropriate.

In order for a pair of waves to cross or merge, all waves between them initially must cross or merge with one of them before they can interact. That is, only adjacent pairs of waves may cross or merge, and any admissible reordering consists of a series of pairwise interactions of adjacent pairs of single waves. This observation allows us to factor reorderings into admissible pairwise interactions.

We have two types of pairwise interactions, namely transpositions and joins. The transposition $\kappa = (k : k+1)$ flips k and $k+1$ while leaving all other places fixed, while the join ϕ^p maps positions p and $p+1$ back to p, and adjusts the other positions accordingly. Thus $\phi^p(i) = i - 1$ for $i > p$, and other places are left fixed. In order for these pairwise interactions to be admissible, we require $c_k > c_{k+1}$ or $c_p = c_{p+1}$, respectively. Note that each join contributes to the merge set only, while each transposition contributes to the crossing set only, that is

$$I_\phi = M_\phi \quad \text{and} \quad I_\kappa = C_\kappa\,,$$

while $C_\phi = M_\kappa = \emptyset$. In particular, joins do not contribute to crossing sets, although they may cause the labels of crossing waves to change.

We shall factor our reorderings into products of these pairwise interactions, and will make heavy use of induction when considering arbitrary reorderings. Note that this factorization is not unique, and indeed we shall consider maps with different (admissible) factorizations to be distinct as reorderings. Thus each reordering has an implicitly given factorization.

Each factorization of the reordering $\sigma = \lambda_s \cdots \lambda_1$, where each λ_r is a join or transposition, determines a time-like ordering on the interaction set I_σ, as follows. We say the pair (i,j) interacts before the pair (k,l), and write $(i,j) \lhd_\sigma (k,l)$, if there is a reordering μ, defined by $\mu = \lambda_m \cdots \lambda_1$ for some m, under which the pair (i,j) interacts, but under which the pair (k,l) does not interact. That is, $(i,j) \lhd_\sigma (k,l)$ if $(i,j) \in I_\mu$, while $(k,l) \notin I_\mu$. The factorization may be recovered by specifying an order relation on the set I_σ. We cannot, however, arbitrarily order the crossing set I_σ, for Condition (R) implies that certain wave pairs will always cross before others.

Corresponding to each crossing pair $(i,j) \in I_\sigma$ of the factored reordering $\sigma = \lambda_s \cdots \lambda_1$, we associate a "subordering" μ of σ, as follows. We set μ to be the *largest* reordering $\mu = \lambda_m \cdots \lambda_1$, under which the pair (i,j) does *not* interact. Thus if μ is associated to (i,j), and we set $\lambda = \lambda_{m+1}$, then the pair (i,j) interacts under the product $\lambda\mu$, and in fact we must have $\mu(i) + 1 = \mu(j)$, and $I_\lambda = \{\mu(i,j)\}$. With this notation, we see that $(k,l) \lhd_\sigma (i,j)$ if and only if $(k,l) \in I_\mu$. For the product $\pi = \sigma\tau$ of factored reorderings, it is clear that $(i,j) \lhd_\pi \tau'(k,l)$, for each $(i,j) \in I_\tau$ and $(k,l) \in I_\sigma$, and the relation $\lhd_\pi$ extends the order relations defined by the factorizations of σ and τ, respectively. We note that not all interacting pairs are comparable: indeed, if $(j,k) \in M_\tau$, and $(\tau i, \tau j) \in C_\sigma$, then both (i,j) and $(i,k) \in C_{\sigma\tau}$, but these cannot be compared under $\lhd_{\sigma\tau}$.

As was noted in [17], reorderings respect consecutive subsequences, so that there is a space-like ordering principle for waves generated by interactions with a fixed wave, namely the order is preserved or reversed. Other properties of reorderings mentioned in [17] are also valid here. Although the present definition of a reordering is slightly more general than was used in [17], our claims are straightforward applications of Condition (R), and more details can be found there.

3.2. Interaction Maps. We now recall the definition of interaction maps, making some changes for our particular assumptions. When two waves of different families interact, they generate waves of other families. This effect is accounted for by changing the strength of a nearby wave in the corresponding family. Interaction maps keep track of the waves whose strengths change as a result of the interaction. Thus given a factored reordering σ, an interaction map $\imath^\sigma$ acts on the crossing set C_σ as follows: the integer $\imath^\sigma(j,k)$ refers to the wave whose strength is changed due to the interaction of the waves γ_j and γ_k. We are using reorderings and interaction maps to model quadratic effects only, and so at this stage will ignore all cubic and higher-order effects, and treat quadratic effects as exact.

To simplify the notation, we describe the interaction maps for the class of 3×3 systems possessing a Riemann coordinate. We refer to this field as the entropy field, and to the others as acoustic fields. With this assumption, the only quadratic effects of interactions that arise come from the interaction of sound waves (i.e. 1- and 3-waves) with entropy changes (2-waves): in this case the affected wave is simply the nearest sound wave of the third family. We remark that although this assumption simplifies our notation and definitions somewhat, it is not necessary for the construction: see [17] for the general construction for systems of N equations.

Before defining interaction maps, we consider in detail what happens inside a diamond. Here we assume that our sampling is random in time, but uniform in space. Suppose for definiteness that a 3-wave enters the diamond from the left, and this is about to interact with a 2-wave entering the same diamond, see Fig. 3. Then the 2-wave must enter the diamond from the right, and moreover there must be a

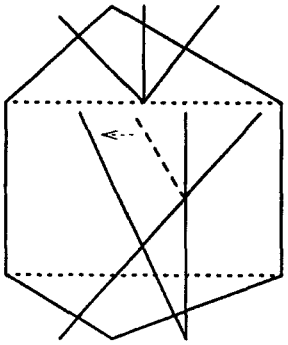

Fig. 3. A Single Diamond

1-wave (possibly of zero strength) entering the diamond between these two waves, also from the right. Thus the 3-wave first crosses the 1-wave, with no quadratic effects, and then interacts with the 2-wave, reflecting a quadratic 1-wave to the right of the original 1-wave. Up to third order, the effect of random sampling is to combine these waves linearly into a single Riemann solution emanating from the center of the diamond. For our purposes, we note that the newly generated 1-wave is combined with the first 1-wave to the left of the interacting pair at the time of interaction. Similar remarks hold for other interactions.

Using the intuition from the argument above, we now define the interaction maps. Suppose that we are given a factored reordering σ, which in turn determines an ordering $\lhd_\sigma$ on the interaction set I_σ. Corresponding to each pair $(i,j) \in I_\sigma$, we associate a subordering μ. We note that quadratic effects are generated only by pairs of waves that cross, and moreover, interactions between pairs of sound waves (i.e. 3-waves with 1-waves) generate only cubic effects, and can thus be ignored. We therefore define the interaction map $\imath^\sigma$ action on those wave pairs in the crossing set C_σ which include a 2-wave only.

In contrast to [17], we shall define interaction maps explicitly, and check that they satisfy the desired conditions. Thus, given a crossing pair $(j,k) \in C_\tau$, we let μ be its associated reordering, and define

$$\imath^\tau(j,k) = \begin{cases} \max\{i \mid c_i = \hat{c}, \mu(i) < \mu(j)\}, & \text{if } \hat{c} = 1\,, \\ \min\{i \mid c_i = \hat{c}, \mu(k) < \mu(i)\}, & \text{if } \hat{c} = 3\,, \end{cases} \tag{3.3}$$

where $\hat{c} \in \{1,3\}$ is the index of the third (acoustic) family, which is affected by but is not part of the interaction. This defines the interaction map for those pairs of interacting waves which generate quadratic effects. We note that since our definition is explicit, a factored reordering has a uniquely defined interaction map, in contrast to the general case described in [17], where several interaction maps may be associated to a single factorization.

We now check that interaction maps as defined here satisfy the conditions stated in [17]. It is evident from our definition that the interaction maps defined here satisfy conditions (i) to (iv) stated in [17], when appropriately interpreted. These conditions imply that interaction maps respect the order of waves, so that those waves generated by the interaction of consecutive waves of one family with a fixed wave are also consecutive. The other condition concerned interaction maps for those

reorderings which are compositions: in this case the condition is not identical, since our reorderings are not invertible. The correct condition in this context is as follows. Recall that $C_{\sigma\tau} = C_\tau \cup \tau' C_\sigma$, and suppose that interaction maps $\imath^\sigma, \imath^\tau$ and $\imath^{\sigma\tau}$ have been defined according to the above definition. Then we have

$$\imath^{\sigma\tau} = \imath^\tau \quad \text{on } C_\tau, \tag{3.4}$$

and

$$\tau\imath^{\sigma\tau} = \imath^\sigma\tau \quad \text{on } \tau' C_\sigma\,, \tag{3.5}$$

which holds because our definition of $\imath^{\sigma\tau}(j,k)$ relies only on the associated reordering μ, which is well-defined for each interacting pair. Of course, these statements make sense only for those wave pairs on which the interaction maps are defined.

We remark that instead of defining the interaction maps explicitly as we have done, we could have defined them by (backwards) induction, as in [17]. An inductive definition requires a choice, because reorderings are no longer invertible. However, if we require (3.3) to hold for transpositions, then Condition (R) and (3.5) fully determine the composite interaction map, leading to the above definition.

3.3. Changes in Wave Strengths. Now that we have defined interaction maps, which identify those waves whose strengths are affected by the interaction of nearby waves, it remains to give a detailed description of the changes in wave strengths due to specific reorderings. We shall describe these changes inductively, and later derive a compact formula for general reorderings. We are interested only in quadratic effects, and so shall ignore anything that is cubic or higher order in wave strengths.

Our notation is as follows: we start with a sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$ of waves, with index $c = (c_1, \ldots, c_n)$. Then the class $A(\gamma)$ of admissible (factored) reorderings is determined by c, and we suppose we are given some admissible reordering, say τ. This has the interaction map $\imath^\tau$ associated to it, as described above. We shall now describe the reordered sequence $\delta = \tau(\gamma)$ consisting of the waves in new positions and with strengths adjusted appropriately. We remark that the index $d = \tau(c)$ is already known, namely $d_k = c_{\tau' k}$, which is well-defined by (R).

Recall that we are implicitly assuming that each reordering is factored, and indeed, different factorizations lead to different results. In order to proceed with the induction, we first describe the effects of a single pairwise interaction.

Consider the single join $\phi = \phi^p$. As described earlier, this models the merging of two waves from the same family into one wave. *The effect of two nonlinear waves from the same family interacting is to add their strengths linearly, with errors of third order.* Thus, since we are ignoring cubic effects, we shall simply add the wave strengths. The two waves to be merged are γ_p and γ_{p+1}. If $\delta = \phi(\gamma)$ is the reordered sequence, we have

$$\delta_p = \gamma_p + \gamma_{p+1}, \tag{3.6}$$

and

$$\delta_{\phi(j)} = \gamma_j, \quad \text{for } j \neq p, p+1\,. \tag{3.7}$$

We now consider a transposition, which models the interaction of two waves from different families. According to our normalization, *the incident waves themselves are not affected (except for cubic errors). The effect of the interaction is to generate a new wave in the third family*, whose strength is the product of incident wave strengths, appropriately weighted. Since we are assuming that the second family has a Riemann coordinate, there are no new 2-waves generated. This was anticipated in our definition of interaction maps, and simplifies the notation. For

the definitions in full generality, see [17]. For other families, we do not allow new waves to be created, but instead adjust the strength of the nearest wave of the correct family, as described by the interaction map. This is exactly the effect of sampling in Glimm's scheme, as can be seen in Fig. 3, and corresponds to "conservation of the number of waves."

We again start with the sequence γ, and let the transposition be given by $\kappa = (k : k+1)$, and we shall describe the reordered sequence $\delta = \kappa(\gamma)$. Let $i = i^{\kappa}(k, k+1)$ refer to the wave whose strength is affected by the interaction, noting that c_i is that family distinct from those of the interacting waves, namely c_k and c_{k+1}. Then, since we are modelling only quadratic terms, we have

$$\delta_{\kappa(i)} = \gamma_i + \Lambda_{c_i} \gamma_k \gamma_{k+1}, \tag{3.8}$$

and

$$\delta_{\kappa(j)} = \gamma_j, \quad \text{for } j \neq i\,. \tag{3.9}$$

Note that apart from adjusting the strength of γ_i by the Lie bracket term, we have also switched the relative positions of γ_k and γ_{k+1}, as required. In particular, if we assume the existence of a 2-Riemann coordinate, we have $\Lambda_2 = 0$, so the transposition of a 3- and 1-wave pair has no effect on wave strengths.

We now inductively define the reordered wave sequence $\sigma(\gamma)$ for an arbitrary reordering σ. Since σ is a factored reordering, we may write $\sigma = \lambda_s \cdots \lambda_1$, where each λ_j is either a join or transposition. We simply define

$$\sigma(\gamma) = \lambda_s(\cdots(\lambda_1(\gamma)))\,,$$

where each successive step has been described above.

Although this definition is unwieldly for reorderings in general systems, we shall see that in our case it leads to a manageable expression in terms of path integrals. The advantage in this definition is that it yields an exact expression at the quadratic level, in the sense that if local interactions were exact at second order, then this formula would be exact. By contrast, in [17], by considering second order terms in initial waves only, we get an approximation after two interactions (dropping brackets of brackets), and the errors so generated become significant when the total variation is large.

Again, we emphasize that although we have defined the reordered wave sequence for special systems, our definitions hold for general systems of N equations. However, without some assumptions on the interaction coefficients Λ_i, the reordered sequences become difficult to work with. Knowing that some solutions blow up, we must make some assumptions on the interaction coefficients to obtain bounds.

4. The Functionals

In this section we define the nonlinear Glimm functionals and describe the problems of continuity and boundedness of these functionals. Our aim is to describe the total variation and sup-norm of the Glimm approximation at large times. Recall that these approximations are represented at any time step, or, more generally, on any space-like I-curve, as a sequence of waves. We shall define the nonlinear functionals in terms of these wave sequences and their reorderings. To obtain bounds on norms at large times, we must bound the functionals. The functionals include quadratic effects, and in order to get estimates which include all other nonlinear effects, we will need a continuity property for the functionals.

4.1. The Classical Induction. Before describing the functionals and their properties in detail, we briefly recall Glimm's original argument for bounding the total variation of the approximate solutions. Given a spacelike curve J, we define the total variation and error potential by

$$V(J) = \sum |\alpha_j|, \quad \text{and} \quad Q(J) = \sum_{\text{App}} |\alpha_j||\alpha_k| ,$$

respectively. Here the α_j's range over waves crossing the curve J, and the second sum is over approaching wave pairs, that is those which will interact at some later time. The local error $Q(\Delta) = D(\Delta)$ for waves entering a diamond is described similarly.

Since wave pairs interact once only, for any space-like curve J, we have

$$Q(J) \leqq V(J)^2 ,$$

which we refer to as *boundedness* of the functional Q (in terms of V).

For the induction step, we argue as follows, referring to Fig. 1 for notation. The waves entering the diamond Δ combine linearly in each family, up to a quadratic error which is bounded by $D(\Delta)$. Thus we have

$$V(J_+) \leqq V(J_-) + c_1 D(\Delta) , \tag{4.1}$$

where c is some generic constant depending only on the conservation law. We view this as a statement of the *continuity* of V. The corresponding statement for Q is more interesting. First, there are fewer approaching waves crossing J_+, as some waves have crossed inside the diamond. This means that there are fewer terms in $Q(J_+)$, and indeed the missing terms appear in $Q(\Delta)$. Thus we would like to write

$$\text{“}Q(J_+) + Q(\Delta) = Q(J_-) .\text{”} \tag{4.2}$$

However, this relation is not exact, as some of the waves have changed due to the interactions in the diamond, so that $Q(J_+)$ is defined using waves of different strengths. We are thus led to a question of continuity: namely, if we change wave strengths locally (at a single diamond), what is the change in the global functional Q? Glimm argued that since each wave approaches a subset of all other waves, and those waves leaving a single diamond may not interact, the total error in Q is the difference in wave strength multiplied by V, and so obtained the inequality

$$Q(J_+) \leqq Q(J_-) - Q(\Delta) + c_2 D(\Delta) V(J_-) , \tag{4.3}$$

where $D(\Delta) = Q(\Delta)$ bounds the difference in wave strength. We combine (4.1) and (4.3) by induction: if $V(J_-)$ is small enough, say $V(J_-) < 1/2c_2$, then

$$V(J_+) + 2c_1 Q(J_+) \leqq V(J_-) + 2c_1 G(J_-) ,$$

so this functional is decreasing. Thus we have the bound

$$V(J_+) < V(J_+) + 2c_1 Q(J_+) \leqq V(J_0) + 2c_1 Q(J_0) \leqq V_0(1 + 2c_1 V_0) ,$$

and if V_0 is small enough, then the inductive hypothesis is satisfied and we get total variation bounds, as required. In the sequel we will define Q^* to be an extension of Q which includes cumulative quadratic effects in such a way that (4.2) holds

exactly at the quadratic level, and the higher order errors due to interaction will be estimated by the *continuity* properties of the functional Q^*.

4.2. Definition of the Functionals. The above argument required V_0 to be small to get bounds on the total variation for large times. To get results for large total variation, we must more accurately describe the effects of interactions by taking quadratic effects into account. We will use the method of reorderings to do so. In [17], we constructed a potential P for the sup-norm, which included (some) quadratic effects. In that paper, the boundedness and continuity of this functional P were discussed, without identifying them as such. Here we shall describe analogous functionals V^* and Q^* for the total variation and the quadratic error potential, respectively. Our notation will be that the * indicates that the functional is the usual functional with cumulative quadratic effects included. For convenience we will assume the existence of a 2-Riemann coordinate, although the functionals can be defined for general systems.

In Sect. 2.2 we described the total variation, sup-norm and local d-variation of a sequence of constant states separated by waves. We shall similarly define all other functionals in terms of these wave sequences, and the extension to space-like curves is then clear, namely taking the supremum over all (consecutive) wave sequences crossing that curve. Thus we suppose that we are given a wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$, together with the associated machinery of reorderings described earlier.

Since we know that time-independent bounds are not available, we shall implicitly fix a time T of existence, and restrict the class of reorderings appropriately. That is, we admit only those reorderings corresponding to interactions taking place up to time T. Thus, because the system is hyperbolic, certain pairs of waves (initially very far apart) will not cross in time T, even though they are approaching. Note that the number of interacting pairs still depends on the mesh size. Henceforth the class of reorderings will be restricted to interactions taking place before time T, and we shall make the dependence on T explicit only when convenient.

In [17], we defined the action of a reordering on a wave sequence, and then defined the potential P as the supremum of the sup-norm of reordered sequences, over all reorderings. That is, we defined a new wave sequence $\imath^\tau\gamma$, and defined

$$P(\gamma) = \sup_\tau S(\imath^\tau\gamma)\,,$$

the sup being taken over all admissible reorderings and interaction maps. Note that the sequence $\imath^\tau\gamma$ is not the same as the sequence $\tau\gamma$ described in this paper, as the present treatment includes *cumulative* quadratic effects (brackets of brackets).

Our potential for the total variation is analogously defined as the supremum over reorderings of the variation of the reordered sequence,

$$V^*(\gamma) = \sup_\tau V(\tau\gamma)\,,$$

Note that our construction included quadratic effects, and so the functional V^* should describe the total variation at large times, except for (accumulated) cubic errors. We know that local cubic errors are controlled by quadratic terms, and continuity of the functional V^* will allow us to control the cumulative cubic errors.

We will use the local variation of the *entropy* of the solution in our estimates, and so we need a functional that measures this quantity. Since entropy jumps

are represented by 2-waves, we define a functional H which measures the local d-variation by

$$H(\gamma) = \sup_{i_0, i_d} \sum_{c_i=2} |\gamma_i| , \tag{4.4}$$

where the sum is over 2-waves between i_0 and i_d, and the supremum is over indices chosen such that γ_{i_0} and γ_{i_d} are at most distance d apart in the approximation. This quantity is bounded by $\|w\|_d$, if we define the strengths of each 2-wave as the entropy jump across that wave. Moreover, the fact that entropy is a Riemann coordinate implies that 2-waves undergo only cubic errors in wave interactions. In terms of reorderings, this means that the only elementary interactions resulting in changes in 2-wave strengths are merges of 2-waves, which do not increase the local 2-variation. Thus for any reordering τ, we have $H(\tau\gamma) \leqq H(\gamma)$. We emphasize that this functional has this special property only for systems possessing a Riemann coordinate.

Finally, we must define the quadratic error potential. As in Glimm's original construction, we wish to define a functional that *decreases* across a diamond, so that this decrease balances the increase in norms due to interaction error. The local interaction error is given by $O(1)D$ at quadratic level and $O(1)SD$ when quadratic terms are included (see (2.22)). Moreover, our error potential Q^* should reduce to D for local interactions, as does Q. We know that when considering the Glimm error potential Q, the error term involving V (and forcing V to be small) appeared when we were considering continuity of Q. We thus look for a functional Q^* satisfying a similar functional equation (4.2), but with better continuity properties.

The contribution to the error potential due to a pair of waves interacting is the product of their strengths. This product of strengths appears as a term in Q. However, the error appears only when the pair actually interacts, and at the time of interaction, these strengths will have been changed by earlier interactions. We thus build the functional Q^* by again taking products of wave strengths, but measuring them at the time of interaction, rather than initially. This is accomplished via the method of reorderings, as follows.

Suppose we are given a reordering τ of the sequence γ. Then the set I_τ determines which wave pairs interact, and thus which terms should appear in Q^*. We now find the strength of the waves at the time of interaction. To each interacting pair $(j,k) \in I_\tau$, we associated the "sub-ordering" μ of τ. The strengths of the waves γ_j and γ_k when this pair interacts, are just the corresponding strengths in the partially reordered sequence $\mu(\gamma)$. Note that this is consistent with our inductive definition of the reordered wave sequence.

We are now in a position to define the functional Q^*. With notation as above, we denote by γ_j^μ and γ_k^μ the reordered waves γ_j and γ_k before they interact, that is

$$\gamma_j^\mu = \eta_{j'} \quad \text{and} \quad \gamma_k^\mu = \eta_{k'} ,$$

where $j' = \mu(j)$, $k' = \mu(k)$ and $\eta = \mu(\gamma)$ is the partially reordered sequence. Note that the partial reordering μ is defined for all interacting wave pairs, and is different for each interacting pair. We now define

$$Q^*(\gamma, \tau) = \sum_{\text{App}} |\gamma_j^\mu||\gamma_k^\mu| = \sum_{C_\tau} |\gamma_j^\mu||\gamma_k^\mu| + \sum_{M_\tau}{}' |\gamma_j^\mu||\gamma_k^\mu| ,$$

where the $'$ on the second sum indicates that the sum is over approaching waves only, that is those for which one of γ_j^μ and γ_k^μ (with $c_j = c_k$) is negative. This is the usual property that rarefaction waves of the same family produce no error. We also

impose the restriction that each interacting pair appear only once in the sum: thus if $(k,l) \in M_\mu$, so that both $(j,k) \in I_\tau$ and $(j,l) \in I_\tau$, and $\gamma_k^\mu = \gamma_l^\mu$, the corresponding term $|\gamma_j^\mu||\gamma_k^\mu|$ should not appear twice. Finally, define the quadratic error potential $Q^*(J)$ by

$$Q^*(J) = \sup Q^*(\gamma, \tau)\,,$$

where the sup is over all finite consecutive wave sequences γ in J, and admissible factored reorderings $\tau \in A(\gamma)$.

4.3. Properties of the Functionals. We now describe the algebraic relations for V^* and Q^* which will be used in the sequel, and describe in detail what we mean by continuity and boundedness. We wish to bound the functionals at later times by the initial functionals, if possible. In order to do this, we shall describe the change in functionals under a reordering, which corresponds to one or more diamonds. In describing the change in functionals, we shall see that our inductive definition of reordered wave sequences leads to exact functional relations, thus explicitly separating the continuity properties as described above.

To describe the change in functionals, we consider the composition of reorderings. Thus suppose that $\tau \in A(\gamma)$, and let $\delta = \tau(\gamma)$ be the reordered sequence. Now if $\sigma \in A(\delta)$, the composition $\sigma\tau$ is admissible for γ. According to our inductive description of reordered wave sequences, we have exactly

$$\sigma(\delta) = \sigma\tau(\gamma)\,,$$

and we use this to compare the functionals at different times. Since this relation holds for any $\sigma \in A(\delta)$, we immediately find

$$V^*(\delta) = V^*(\tau\gamma) \leqq V^*(\gamma)\,,$$

which says that the functional for the total variation does not increase with time. Note that here we are considering only quadratic terms, and so cubic errors have not yet been taken into account. We will need continuity of the functionals in order to deal with these cubic effects.

We now write down an algebraic relation for the quadratic error potential Q^*. Using the same notation, we wish to compare $Q^*(\delta,\sigma)$ and $Q^*(\gamma,\sigma\tau)$. Note that if μ is associated to the interacting pair $(\tau j, \tau k) \in I_\sigma$, then the sub-ordering associated to the corresponding pair $(j,k) \in C_{\sigma\tau}$ is $\mu\tau$, and indeed we have $\delta_{\tau j}^\mu = \gamma_j^{\mu\tau}$, etc. Thus each term in $Q^*(\delta,\sigma)$ appears in $Q^*(\gamma,\sigma\tau)$. Conversely, those terms in $Q^*(\gamma,\sigma\tau)$ which do not appear in $Q^*(\delta,\sigma)$ are exactly those which interact under τ. Also, since the reordering associated to an interacting pair does not depend on later interactions, these terms are exactly those in $Q^*(\gamma,\tau)$, and we have

$$Q^*(\gamma,\sigma\tau) = Q^*(\tau\gamma,\sigma) + Q^*(\gamma,\tau)\,. \tag{4.5}$$

If τ is the reordering associated with a single diamond Δ between J_- and J_+, and σ is the reordering associated with changes between J_+ and some later time T, we interpret this relation as $Q^*(J_-) \approx Q^*(J_+) + Q^*(\Delta)$, so that the quadratic error potential does indeed decrease by the right amount. A more precise estimate can be stated once we have established continuity of the functionals.

Now that we have found the appropriate functional relations, which are *exact at the quadratic level*, we shall investigate what is needed to complete our induction argument. Since we will be bounding the functionals by their initial values, it is

clear that they must be bounded initially. Since the functional V^* is a potential for the total variation, we require that it be bounded initially in terms of the variation V, whereas the functional P which bounds the sup-norm at later times, must be bounded by S.

Our induction is on diamonds (or reorderings) in the Glimm scheme, that is on series of interactions, each of which produces quadratic effects plus errors which are cubic. As we have noted, our functional relations are exact only at the quadratic level. Thus when applying our formalism to the Glimm approximations, we must account for these errors. We first informally describe how this is accomplished.

We model the interaction of waves inside a diamond by a reordering which adjusts the wave strengths to account for the quadratic effects. From our interaction estimate (2.22), we know that the cubic errors are bounded by $O(SD)$, where the sup-norm S is small and we have control over $D(\Delta) \approx Q^*(\Delta)$. Thus suppose that γ represents the waves entering a diamond, and δ represents those leaving. We have that for some reordering ρ,

$$\delta = \rho(\gamma) + O(1)S(\gamma)D(\Delta)\,.$$

We have seen that we can compare the functionals defined on the sequences γ and $\rho(\gamma)$, and would like to compare these to those defined on δ. Since $\rho(\gamma)$ and δ are close, this is a problem of continuity of the functionals.

The continuity property of functionals can be described as follows: suppose that the wave sequences δ and δ' have the same index, but the wave strengths differ in a single position, say p. We shall say that some functional B is continuous, if

$$|B(\delta) - B(\delta')| < K_B|\delta_p - \delta'_p|\,.$$

We require that K_B be independent of the number of waves and mesh-size, while it may depend on norms of the solution or time. The requirement is that our estimates be uniform in mesh parameters, so that we can pass to the limit of Glimm approximations. Supposing that the sequences differ in one wave only is convenient and represents no restriction.

A similar remark holds for boundedness: we would like to know that the functionals, which represent norms at future times, can be bounded in terms of initial quantities. Thus we would like to know that

$$B(\delta) < B_0(V_0, S_0, T, \ldots)\,,$$

where the arguments of B_0 are known initially, $V_0 = V(\delta)$, etc. Of course, these bounds should be appropriate, so a bound for the sup-norm potential P in terms of V_0 alone would not be useful. We remark that in most circumstances, boundedness and continuity of functionals are equivalent, although they are conceptually different. In particular, if the bounds are smooth functions of norms, then the functionals are also continuous. This is just the statement that differentiable functions are Lipschitz continuous.

5. Existence of Solutions

We are now in a position to fill in the details of the existence theorem. We will make assumptions on the boundedness and continuity of the functionals, and then

derive our result. This means that the problem is essentially reduced to an algebraic one, namely finding bounds for the functionals which are described by finite sequences of waves. In the following sections, we find time-dependent bounds for the functionals, so that we get existence for arbitrarily large finite times. In general, any assumptions made in bounding the functionals translate to restrictions on the solution to the Cauchy problem. Throughout we assume implicitly that the sup-norm is small enough that all our estimates are uniform.

The bounds for the nonlinear functionals V^* and Q^* will depend on the time T of existence of solutions, as well as the total variation of the solution, and the local d-variation of the entropy. This local variation is measured in terms of the functional H defined earlier. For completeness, we shall make the time dependence explicit in the induction.

We use the notation developed earlier and state our assumptions in terms of the functionals. First, we assume the nonlinear functionals are bounded,

$$V^*(\gamma) \leqq V_0^*(V,T) \quad \text{and} \quad Q^*(\gamma) \leqq Q_0^*(V,T)\,, \tag{5.1}$$

where V_0^* and Q_0^* are increasing functions of their arguments, and $V = V(\gamma)$ is the total variation of the sequence γ. We note that in general these bounds also depend on other quantities, although we will not make them explicit here. Indeed, our bounds will depend on the lengthscale d, and corresponding sup- and d-norms. Next, we assume that these functionals are continuous,

$$V^*(\gamma) - V^*(\gamma') \leqq K_V(V,T)|\gamma - \gamma'| \tag{5.2}$$

and

$$Q^*(\gamma) - Q^*(\gamma') \leqq K_Q(V,T)|\gamma - \gamma'|\,, \tag{5.3}$$

for sequences γ and γ' differing in one wave in each family, and having the same index. Again K_V and K_Q are increasing functions of their arguments, and may depend on other quantities. We remark that these assumptions are equivalent when the bounds are smooth functions.

In [17] it was shown that the functional P for the supnorm is both continuous and bounded with explicit bounds, namely

$$P(\gamma) \leqq S(\gamma) + \Lambda S(\gamma)V(\gamma), \quad \text{and} \quad P(\gamma) - P(\gamma') \leqq K_P|\gamma - \gamma'|\,,$$

where $K_P = O(1)V$, and $O(1)$ and Λ depend only on the conservation law. We shall also make use of Glimm's estimate

$$V(J_+) \leqq V(J_-) + \Lambda D(\Delta)\,. \tag{5.4}$$

Theorem 1. *Suppose that the nonlinear functionals satisfy boundedness and continuity properties* (5.1)–(5.3). *Then given any T and V_0, there is an $\varepsilon > 0$ such that if the initial data u_0 of the Cauchy problem* (1.2) *satisfies*

$$\|u_0\|_\infty \leqq \varepsilon \quad \textit{and} \quad TV(u_0) \leqq V_0\,,$$

then the conservation law admits a weak solution with bounded oscillation and total variation. Moreover, the total variation of the solution is bounded up to time T, with bound $V_0^* + O(\varepsilon)$, and analogous bounds are available for other norms.

We remark that ε is defined in terms of V_0 and K_Q, so that $\varepsilon \to 0$ as $V_0 \to \infty$. In our analysis, K_Q depends on time, so that we get solutions for arbitrarily large but finite times. That is, given any time T, we have $\varepsilon = \varepsilon(V_0, T)$ and we have

existence up to time T. Similarly, any other assumptions which are made in finding K_Q must be imposed on the initial data. In our case, we require that the entropy of the initial data has local variation.

$$\|w(u_0)\|_d \leqq \eta < 1/3\Lambda, \quad \text{for some } d > 0 .$$

Here $\Lambda = \max\{\Lambda_i^{jk}\}$ is the maximum of the nonzero interaction coefficients, and is determined by the flux. For clarity in the proof, we shall make the dependence on the local variation $H \approx \|w(u)\|_d$ of the entropy explicit.

Although our main application is to 3×3 systems having a Riemann coordinate, this theorem holds for general $N \times N$ systems, with appropriate modifications. The functionals V^* and Q^* are again defined inductively on reorderings, and any restrictions made in finding bounds for V^* and Q^* must hold for the initial data.

Corollary 5.1. *Boundedness and continuity of the functionals V^* and Q^* are sufficient conditions for existence of solutions to general $N \times N$ systems. That is, we can choose $\varepsilon > 0$, depending on the bounds, such that the conclusions of the theorem follow.*

In general, such bounds are not available for large times when the total variation is large, and solutions are known to grow without bound, [5, 18]. Any restrictions used in finding the bounds for V^* and Q^* must again be imposed on the initial data and the solution.

Proof. As usual, we proceed by induction on space-like I-curves. There are two steps, first comparing the functionals evaluated at successive I-curves, and then carrying out the induction. We use the notation of Sect. 2.1, so that curves J_- and J_+ differ by a single diamond Δ, and J_0 is the initial I-curve. Again, since we are anticipating time-dependent bounds, we restrict the class of admissible reorderings to those spanning the finite time T.

Referring to Fig. 1, we suppose that the waves crossing curve J_- form the sequence γ, and those crossing J_+ are given by the sequence δ. Also, let the reordering τ represent the interactions inside the diamond Δ. According to (2.22), the local error is $O(1)SD(\Delta)$, and we have

$$\delta = \tau(\gamma) + O(1)S(\gamma)D(\Delta) .$$

By this we mean that the sequences δ and $\tau(\gamma)$ have the same index, and the wave strengths differ only in those waves leaving Δ, and in particular no pair of these waves approach, see [17].

Since the reordering τ models the diamond Δ, we expect that the quadratic errors $Q^*(\gamma, \tau)$ and $D(\Delta)$ are comparable. In fact, the only difference between these is that in Q^* we measure each quadratic term at the time of interaction. Since τ is a local reordering modeling only a single diamond, there are a limited number of interactions, so that the difference between these is cubic, and we have

$$Q^*(\gamma, \tau) = D(\Delta) + O(1)S(J_-)D(\Delta) .$$

We wish to compare the functionals for the sequences δ and γ, respectively. This in turn reduces to the continuity of the functionals, as follows. Consider the functional V^*, which satisfies $V^*(\tau\gamma) \leqq V^*(\gamma)$. We know that the sequences $\tau(\gamma)$

and δ have the same index, and that a few of the individual wave strengths differ by $O(SD)$. Then by continuity of V^* in (5.2), we have

$$\begin{aligned} V^*(\delta) - V^*(\tau\gamma) &< K_V \sum |\delta_p - (\tau\gamma)_p| \\ &\leqq K_V O(1) S(J_-) D(\Delta) , \end{aligned}$$

where $O(1)$ is uniform in a neighborhood and depends only on the flux. Now $V^*(J_+) = V^*(\delta)$ for some sequence δ, and since the corresponding sequence γ lies on J_-, we have

$$\begin{aligned} V^*(J_+) &\leqq V^*(\gamma) + K_V O(1) SD \\ &\leqq V^*(J_-) + K_V O(1) SD , \end{aligned} \tag{5.5}$$

which is the desired comparison of $V^*(J_+)$ and $V^*(J_-)$. As similar relation holds for the sup-norm functional P, namely

$$\begin{aligned} P(J_+) &\leqq P(\gamma) + K_P O(1) SD \\ &< P(J_-)(1 + K_P O(1) D(\Delta)) , \end{aligned} \tag{5.6}$$

for some sequence δ and corresponding $\gamma \subset J_-$, see [17]. We obtain a similar estimate for the change in the local variation H of the entropy. Since H is defined in terms of the 2-waves only, and wave interactions have no quadratic effects on 2-waves, we have

$$H(J_+) \leqq H(J_-) + K_H SD , \tag{5.7}$$

where K_H depends only on the conservation law. We have implicitly used algebraic relations for V^*, P and H, saying that these decrease at quadratic level after interaction.

We treat the quadratic error potential similarly: the error potential for J_+ is given by $Q^*(\delta, \sigma)$, where σ is an arbitrary reordering of δ, representing wave interactions up to some later time. Given the corresponding sequence $\gamma \subset J_-$, we shall compare this to $Q^*(\tau\gamma, \sigma)$ by invoking the continuity (5.3) of Q^*. We then use the functional relation (4.5) for Q^* to compare that with $Q^*(\gamma, \sigma\tau)$, which is in turn smaller than $Q^*(J_-)$. Carrying out the details, we have:

$$\begin{aligned} Q^*(\delta, \sigma) &\leqq Q^*(\tau\gamma, \sigma) + K_Q O(1) S(\gamma) D(\Delta) \\ &\leqq Q^*(\gamma, \sigma\tau) - Q^*(\gamma, \tau) + K_Q O(1) S(\gamma) D(\Delta) , \end{aligned}$$

which, when maximizing over σ, leads to

$$Q^*(J_+) \leqq Q^*(J_-) - D(\Delta)(1 - O(1) K_Q S) , \tag{5.8}$$

where we have used the local estimate $Q^*(\gamma, \tau) \approx D(\Delta)$. Thus we see that as long as S is small enough, the functional Q^* decreases by some fraction of D.

We now proceed with the induction. The new feature of the induction is that the constants K depend on the solution, so we must choose the initial sup-norm ε small enough that we get a uniform decrease in Q^*. Our separation of the sup- and total variation norms, together with the small error in Q^*, allows us to carry out the induction for large variation, as in [14]. In the local estimates above, we evaluated the K's at J_-: as these are increasing functions of their arguments, we shall overestimate them in the induction. In order to simplify notation, we include all $O(1)$ terms depending on the flux into the K's.

Given an initial total variation V_0, we set

$$Q_0^* = Q_0^*(V_0, 2\eta; T),$$

where 2η is an upper bound for the local variation H, and define

$$W = V_0 + 2\Lambda Q_0^* \tag{5.9}$$

and

$$M = (1 + \Lambda V_0)e^{2K_P Q_0^*}, \tag{5.10}$$

where we have set

$$K_P = K_P(W, 2\eta; T) \quad \text{and} \quad K_Q = K_Q(W, 2\eta; T).$$

We will show that

$$S(J) \leqq MS_0 \quad \text{and} \quad V(J) \leqq W$$

and $H(J) \leqq 2\eta$, as long as S_0 is small enough and $H_0 \leqq \eta$.

Choose the bound $\varepsilon > 0$ for S_0 small enough that

$$\varepsilon M K_Q < 1/2 \quad \text{and} \quad 2\varepsilon M K_H Q_0^* < \eta,$$

and suppose for the induction that

$$V(J_-) + 2\Lambda Q^*(J_-) \leqq W, \tag{5.11}$$

$$P(J_-) \leqq P(J_0)e^{2K_P(Q_0^* - Q^*(J_-))} \tag{5.12}$$

and

$$H(J_-) + 2S_0 M K_H Q^*(J_-) \leqq 2\eta. \tag{5.13}$$

These assumptions imply that $S(J_-) < P(J_0)e^{2K_P Q_0^*} \leqq MS_0$. By (5.8) we have

$$Q^*(J_+) - Q^*(J_-) \leqq -D(\Delta)(1 - K_Q M S_0) \leqq -D/2, \tag{5.14}$$

and so also $D \leqq 2(Q^*(J_-) - Q^*(J_+))$. Then by (5.4), we have

$$V(J_+) + 2\Lambda Q^*(J_+) \leqq V(J_-) + 2\Lambda Q^*(J_-) \leqq W,$$

and by (5.7),

$$H(J_+) + 2S_0 M K_H Q^*(J_+) \leqq H(J_-) + 2S_0 M K_H Q^*(J_-) \leqq 2\eta.$$

Similarly, by (5.6) and (5.14),

$$\begin{aligned} P(J_+) &\leqq P(J_-)(1 + K_P D) \\ &< P(J_0)e^{2K_P(Q_0^* - Q^*(J_-))}e^{2K_P(Q^*(J_-) - Q^*(J_+))}, \end{aligned}$$

where we have used $1 + x < e^x$. This completes the inductive step, and we have the bounds

$$V(J_+) < W \quad \text{and} \quad S(J_+) < MS_0,$$

and $H(J_+) < 2\eta$.

We have not made use of the functional V^* in proving existence of solutions. Instead, V^* gives us an estimate of the growth of the total variation of the solution. In particular, the dependence on time T of V^* gives an upper bound for the growth of total variation, which in turn drives the growth in the other norms. To see this,

observe that we are overestimating V by adding the factor $2\Lambda Q^*$, which controls the local error ΛD in V. As we have seen, however, the local error in V^* is $K_V SD$, which is much smaller. Thus we can control the growth of V^* by adding a smaller multiple of Q^*. Indeed, letting $c = 2K_V(W, 2\eta; T)MS_0$ and using the bounds found above, we have

$$\begin{aligned} V^*(J_+) + cQ^*(J_+) &\leqq V^*(J_-) + K_V SD + c(Q^*(J_-) - D/2) \\ &< V^*(J_-) + cQ^*(J_-)\,, \end{aligned}$$

so that $V(J_+) < V_0^* + cQ_0^*$, where c contains the small factor S_0. Now we see that for small values of ε, the total variation at later times is approximately given by the functional V^*. Furthermore, a knowledge of the initial wave configuration and the (non-local) reordering corresponding to the union of all diamonds below J_+ should give an accurate approximation for the actual variation along J_+. It appears that one should be able describe the pointwise large-time behaviour of the solution in this way. □

6. The Path Integral

In order to apply the general theorems of Sect. 5 to a given system, one must obtain the boundedness and continuity properties (5.1)–(5.3) for the functionals in this case. In particular, these conditions must hold for special reasons because they fail when T is larger than the blowup time in a resonant system of type (1.2). We now obtain the estimates (5.1)–(5.3) for 3×3 systems possessing a Riemann coordinate. In this case, our growth rate bounds will be determined by the d-norm, supnorm and total variation norms of the initial data. Thus we assume the existence of a Riemann coordinate, and use this to express reordered sequences in terms of a path integral formula. Given a wave sequence and set of reorderings, we have seen how the strengths of these waves change under the interactions represented by that reordering. However, our description is thus far inductive, and the resulting quantities are difficult to analyze. Our assumption means that the entropy field decouples, and can be treated as static at the quadratic level. This simplification allows us to describe reordered sequences in terms of a path integral formula.

6.1. Paths in Reorderings. Our first task is to describe paths in the framework of reorderings. The cause of instability in our systems is the generation of new waves due to interactions, which is compounded when these new waves themselves interact and generate more waves. In our construction, the generation of new waves is avoided by adjusting the strengths of nearby waves, thus conserving the number of waves. We now address the question of how these changes of strength accumulate.

When a pair (j,k) of waves interact, we identify the wave γ_i whose strength is to be adjusted via an interaction map, namely $i = \imath^\tau(j,k)$. The new (scattered) wave may now go on and interact with other waves, each time generating yet another wave. In our model we are merely adding each scattered wave strength to an existing wave: this means that those waves with which the scattered wave will interact are already known, namely they are those waves which interact with $i = \imath^\tau(j,k)$ *after* the pair (j,k) has interacted. We can now inductively describe a series of interactions, keeping track of the accumulated strength due to quadratic

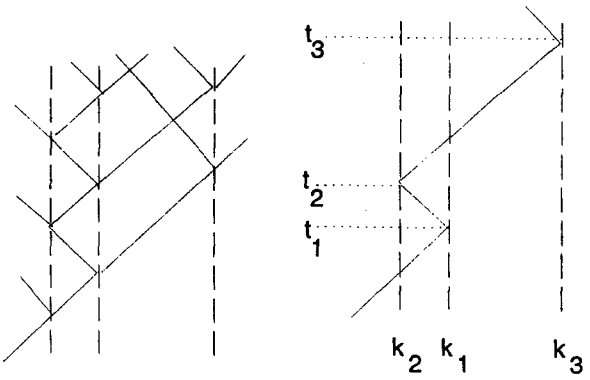

Fig. 4. Scattering pattern with a single path

effects of these interactions. Figure 4 illustrates the scattering of a single sound wave, and a single path which is part of that pattern.

Given a wave sequence and (factored) reordering τ, which determines the interaction map $\imath^\tau$, we make the following definition of a scattering path, which describes the changing position of a sound wave or a (multiply) scattered sound wave as it interacts with contact discontinuities.

Definition 6.1. *An* **r-path** *in* τ *is a set* $\{(j_p,k_p)\,|\,1 \leqq p \leqq r\} \subset C_\tau$, *satisfying*

$$(j_{p-1},k_{p-1}) \lhd_\tau (j_p,k_p) \quad \textit{and} \quad \imath^\tau(j_{p-1},k_{p-1}) = j_p\,,$$

for each $p < r$. *We shall say that the path begins at* j_1 *and ends at* $\tau(j_{r+1}) \equiv \tau\imath^\tau(j_r,k_r)$. *We analogously define a* 0*-path to be any index* i, *and say that it starts at* i *and ends at* $\tau(i)$.

A path is thus a means of tracking the *position* at which an accumulation of interaction effects is concentrated. In this definition, the k_p's represent the contact discontinuities (2-waves) which scatter the sound waves, while j_p represents the sound wave which has been generated as the result of the previous $p-1$ interactions. The k_p's can be viewed as the *relative* positions of the contact waves which scatter the sound waves, although the actual position of the contact changes due to sampling shifts. We shall see that we can also extract the accumulated wave strength associated to the interactions represented by that path.

As usual, we have defined paths independent of the actual strengths of waves appearing in the sequence, so that only the index of the wave sequence is needed to determine the possible paths in any reordering. We remark that paths feature only members of the crossing set C_τ, as these are the only interactions which generate quadratic effects by generating reflected waves.

According to our assumptions and our definition of interaction maps, images under $\imath^\tau$ are 1- and 3-waves only. Thus in the above definition it is implicit that each k_p (for $p > 1$) refers to a 2-wave, while j_p refers to a 1- or 3-wave, alternating between these as p changes. We extend this convention to (j_1,k_1), so that all nontrivial paths start at the 1- or 3-waves γ_{j_1}. We denote the set of all r-paths in τ beginning at $p = j_1$ and ending at $q = \tau(j_{r+1})$ by $\Pi_r(p,q,\tau)$.

As we have done previously, we now consider compositions of reorderings, and the paths associated to these. We would like to know that paths can be constructed

inductively, and that all paths can be obtained by a suitable local construction. The compound structure of paths is described in the following lemma. We shall use the symbol π_τ to refer to an arbitrary path in τ, so $\pi_\tau \subset C_\tau$. Our notation is that used earlier, so τ reorders the sequence γ, and σ reorders $\tau(\gamma)$, and the composition $\sigma\tau$ makes sense.

Lemma 6.2. *A path $\pi_{\sigma\tau}$ in $\sigma\tau$ determines unique paths π_τ and π_σ in τ and σ, respectively, via the definitions*

$$\pi_\tau = \pi_{\sigma\tau} \cap C_\tau \,, \tag{6.1}$$

and

$$\pi_\sigma = \tau\pi_{\sigma\tau} \cap C_\sigma \,, \tag{6.2}$$

respectively. We say that $\pi_{\sigma\tau}$ projects to π_τ and π_σ. Conversely, given paths π_τ and π_σ in τ and σ, respectively, such that π_σ starts at the end of π_τ, we can always find a path $\pi_{\sigma\tau}$ in $\sigma\tau$ which projects to π_τ and π_σ.

We note that different paths in $\sigma\tau$ may project to the same paths in σ and τ, in case τ is not one-to-one. When convenient, we shall denote a path which projects to π_τ and π_σ by $\pi_\sigma \circ \pi_\tau$.

Proof. Suppose that the path $\pi_{\sigma\tau}$ consists of $\{(j_p, k_p) | 1 \leqq p \leqq s\}$. Since these pairs are ordered by $\lhd_{\sigma\tau}$, and all pairs in C_τ cross before those in $\tau' C_\sigma$, there is a unique r such that $\{(j_p, k_p) | 1 \leqq p \leqq r\} \subset C_\tau$, and this is the path π_τ. To get a path in σ, we must advance the other crossing pairs by τ: thus $\{(\tau j_p, \tau k_p) | r+1 \leqq p \leqq s\}$ is a subset of C_σ, and that it forms a path follows immediately from the fact that $\tau\iota^{\sigma\tau} = \iota^\sigma\tau$ on $\tau' C_\sigma$. That these paths can be described in the above set notation is clear.

For the converse, we must be more careful, as the path $\pi_{\sigma\tau}$ is not uniquely defined in general. Suppose that π_τ is given by $\{(j_p, k_p) | 1 \leqq p \leqq r\}$, and $\pi_\sigma \subset C_\sigma$ is $\{(l_q, m_q) | 1 \leqq q \leqq s\}$, with $\tau\iota^\tau(j_r, k_r) = l_1$. We wish to construct a path $\pi_{\sigma\tau}$ which projects to these paths. It is clear that we should just copy π_τ initially, so it remains to extend $\pi_{\sigma\tau}$ in a way that projects to π_σ. Since τ is not necessarily one-to-one, there will be some choice involved, although the choice is restricted by the requirements for a path. For each q, choose $k_{r+q} \in \tau'(m_q)$. These are the only choices we have, as each j_{p+1} is determined inductively by $j_{p+1} = \iota^{\sigma\tau}(j_p, k_p)$. It remains to check that the path $\pi_{\sigma\tau} = \{(j_p, k_p) | 1 \leqq p \leqq r+s\}$ projects back to π_σ. To see this, note that $\tau(k_{r+q}) = m_q$ and, by induction,

$$\tau(j_{p+1}) = \tau\iota^{\sigma\tau}(j_p, k_p) = \iota^\sigma(\tau j_p, \tau k_p) = \iota^\sigma(l_{p-r}, k_{p-r}) = l_{p-r+1} \,.$$

This completes the proof of the lemma. □

In view of this lemma, we can describe the paths in a particular reordering by induction, after describing the paths for the elementary interactions, namely joins and transpositions. For a join $\phi = \phi_p$, there are no non-trivial paths: however, there are two 0-paths which end at the same point, namely $\{p\}$ and $\{p+1\}$. There is a single 1-path for the transposition $\kappa = (k : k+1)$, namely the path $\{(k, k+1)\}$ which coincides with the entire crossing set.

We emphasize the fact used in the above proof that any path is uniquely determined by the starting sound wave j_1 and the contact waves k_p, for a given interaction map. Also note that the only choice in the path $\pi_{\sigma\tau}$ came in the choice of the k_{r+q}'s. In particular, if τ is one-to-one on the k's (which are contacts), the

path $\pi_{\sigma\tau} = \pi_\sigma \circ \pi_\tau$ is uniquely determined, and a knowledge of all paths in τ and σ, respectively, gives a full description of the set of paths in $\sigma\tau$. In fact, by considering the gas dynamics equations in the Lagrangian formalism, (so $\lambda_2 \equiv 0$), and sampling randomly in time only, we can ensure that this is always the case.

The results of the lemma can be interpreted in terms of the sets Π_r as

$$\Pi_r(j,i,\sigma\tau) \subseteq \bigcup_{s,k} \Pi_{r-s}(k,i,\sigma) \circ \Pi_s(j,k,\tau)\,, \tag{6.3}$$

with equality when τ is one-to-one on contacts (2-waves). This abstract statement says that all paths of length r in $\sigma\tau$ are built up of paths of length s in τ concatenated with paths of length $r-s$ in σ. Note that the only restriction on the shorter paths is that the end of the first be the beginning of the second, but this could be in a number of different positions. Comparing (6.3) to the arithmetical relation

$$\binom{M+N}{r} = \sum \binom{M}{s}\binom{N}{r-s}$$

leads to a bound on the number of paths.

Corollary 6.3. *If the reordering τ is one-to-one on contacts and spans N time-steps, then*

$$\#\Pi_r(j,\cdot,\tau) \leqq \binom{N}{r}\,,$$

where $\#\Pi_r(j,\cdot,\tau)$ denotes the number of r-paths in τ, starting at any given j.

Proof. According to our construction, there is only one 2-wave entering or leaving each diamond. Also, each sound wave passes through a single diamond at each time step, so that it interacts with at most one contact at that time step. This means that any contact with which a particular sound wave interacts can be uniquely identified by the time of interaction. We extend this idea to count the total number of paths starting at a fixed point. Figure 4 illustrates this "projection" of interactions of a single sound wave and a path onto interaction times.

We wish to count the number of paths in τ which begin at position j. Thus suppose $\{(j_q,k_q)\}$ is a path, with $j_1 = j$. We know that the path is uniquely determined by the sound wave j_1, which is given, and the positions k_q of the 2-waves appearing in the path. Now, for each j_q, we can uniquely identify a time step t_q at which the crossing pair (j_q,k_q) interacts. We thus have a projection of the path onto an increasing sequence of time-steps $t_1 < t_2 < \cdots < t_r$. Moreover, since only one contact enters each diamond, different paths lead to different projections. This means that the number of paths starting at j is bounded by the number of increasing interaction times $0 < t_1 < t_2 < \cdots < t_r$. Since each $t_q = q\Delta t$ corresponds to an integer (counting time-steps), and there are at most N time-steps, we have

$$\#\Pi_r(j,\cdot,\tau) \leqq \binom{N}{r}\,. \qquad \square$$

We will later use a similar path-counting technique in obtaining bounds for the nonlinear functionals. Although we require that 2-waves do not merge in this proof, in finding bounds we will not require a one-to-one map from paths to sequences of interaction times.

6.2. Path Integral Formula. We now use the abstract paths to get a useful formulation of the wave strengths in the reordered sequence $\tau(\gamma)$. Recall that we are assuming the existence of a 2-Riemann coordinate, so that all images under interaction maps are either 1- or 3-waves. We now make the further simplifying assumption that 2-waves do not merge. This is true for gas dynamics as long as we use the Lagrangian formalism, for which the entropy has zero wave-speed. This assumption for 2-waves allows us to treat perturbations in that family (i.e. entropy fluctuations) as a static background source for changes in sound waves, which is fixed in time.

We can make the statement that no 2-waves merge more precise by adding an extra requirement to Condition (R), namely that if the pair $(j,k) \in M_\tau$, then we must have $c_j = c_k \mp 2$. This extra condition further restricts the set of admissible reorderings, but does not change any of our previous assertions in any way. Henceforth we shall assume that this extra condition is satisfied implicitly. We remark THAT THIS assumption is merely technical, as allowing two 2-waves to remain adjacent in the scheme instead of merging merely increases the set of reorderings and paths, but this increase does not affect the functional bounds at all, as the variation of the sequence never increases after a merge. A more formal analysis of this idea can be found in [17].

Since our definition uses only the index of a wave sequence, a path tracks only the position of the scattered wave we are tracing. In order to fully describe interaction effects, we must also take the actual wave strengths into account. Before writing down the path integral formula, we consider the strength associated to a single path. Thus suppose we are given a path $\pi_\tau = \{(j_p, k_p)\}$. This represents the wave generated as a result of the multiple interactions occurring in π_τ. We wish to find the strength contributed by this scattering path. As before, we assume that each $c_{k_p} = 2$, and take $c_{j_1} = 3$, say. We shall trace the strength of the scattered wave as the path is traversed, as in Fig. 4. Thus the original strength is γ_{j_1}, and when this wave interacts with γ_{k_1}, the reflected wave has strength $\varepsilon^{(1)} = \gamma_{j_1}\gamma_{k_1}\Lambda_1$. This reflected wave then interacts with wave γ_{k_2}, to produce $\varepsilon^{(2)} = \varepsilon^{(1)}\gamma_{k_2}\Lambda_3$. This process continues inductively, and after r interactions, say r even, the wave generated by this series of interactions has strength

$$\varepsilon^{(r)} = \varepsilon^{(r-1)}\gamma_{k_r}\Lambda_3 = \gamma_{j_1}\gamma_{k_1}\cdots\gamma_{k_r}(\Lambda_1\Lambda_3)^{r/2}\,.$$

This then is the amount of wave strength generated as a result of the interactions represented by the path π_τ. In this calculation, we have used the fact that 2-waves γ_{k_p} do not change at all across interactions. This observation allows us to use the initial 2-wave strength at later times. Moreover, because there is a 2-Riemann invariant, interactions between 3- and 1-waves have no (quadratic) effect, which means that different scattering patterns combine linearly, leading to the path integral formula, stated in the following theorem.

Theorem 2. *Given the wave sequence $\gamma = (\gamma_1, \ldots, \gamma_n)$ and factored reordering τ, let $\delta = \tau(\gamma)$ be the reordered wave sequence. Then for any i, the reordered wave δ_i is given by the* **path integral formula**

$$\delta_i = \sum_j \gamma_j \sum_{r \geqq 0} \Lambda(i,r) \sum_{\Pi_r(j,i,\tau)} \gamma_{k_1}\cdots\gamma_{k_r}\,.$$

Here $\Pi_r(j,i,\tau)$ is the collection of all r-paths in τ starting $j_1 = j$ and ending at $i = \tau\iota^\tau(j_r,k_r)$, and the weight $\Lambda(i,r)$ is given by

$$\Lambda(i,r) = \Lambda_c^{[(r+1)/2]}\Lambda_{c'}^{[r/2]}, \quad \text{where } c = c_{\tau' i} \quad \text{and} \quad c' = 4 - c\,.$$

We remark that this formula is applicable for all waves, although it is highly redundant for 2-waves. Indeed, no non-trivial paths end at 2-waves, and these do not merge, so that $\delta_{\tau i} = \gamma_i$ for 2-waves. Merges of sound waves are also implicit in the formula, as paths corresponding to merged waves end at the same place. The weight $\Lambda(i,r)$ consists of $r = [(r+1)/2] + [r/2]$ factors Λ, each corresponding to one reflection, and does not depend on the particular choice of path.

Proof. The proof is by induction on the factors of the reordering $\tau = \lambda_t \cdots \lambda_1$. If τ is a single join $\phi = \phi^q$, then there are no non-trivial paths, and the 0-paths $\{q\}$ and $\{q+1\}$ end at q, while other 0-paths $\{j\}$ end at $\phi(j)$. The formula then reduces to the definition of the sequence $\phi(\gamma)$. Similarly, if τ is a transposition $\kappa = (k : k+1)$, all 0-paths $\{j\}$ end at $\kappa(j)$, and the 1-path $\{(k,k+1)\}$ ends at $\kappa i = i = \iota^\kappa(k,k+1)$. Thus the formula reduces to $\delta_{k(j)} = \gamma_j$ for $j \neq i$, and $\delta_i = \gamma_{\kappa' i} + \gamma_k\gamma_{k+1}\Lambda_{c_i}$, which is again the definition of the sequence $\kappa(\gamma)$.

Now suppose that the path integral formula holds for τ. We consider the compositions $\phi^q\tau$ and $\kappa\tau$ separately. First consider the join $\phi = \phi^q$. Recall that ϕ^q maps q and $q+1$ to q, and changes the positions of all other waves appropriately. If $\sigma = \tau(\gamma)$ is given by the path integral formula, we must describe the new sequence $\delta' = \phi(\delta) = \phi\tau(\gamma)$. First, for $j \neq q$, we have simply $\delta'_j = \delta_{\phi' j}$, which is the sum over the sets $\Pi(\cdot, \phi' j, \tau)$. For the merged wave, we have $\delta'_q = \delta_q + \delta_{q+1}$, which is the sum over both sets $\Pi(\cdot, q, \tau)$ and $\Pi(\cdot, q+1, \tau)$. In the set notation of (6.3), we can express the set of paths in $\phi\tau$ as

$$\Pi_r(p,j,\phi\tau) = \Pi_r(p,\phi' j,\tau), \quad \text{for } j \neq q, \tag{6.4}$$

and

$$\Pi_r(p,j,\phi\tau) = \Pi_r(p,q,\tau) \cup \Pi_r(p,q+1,\tau)\,, \tag{6.5}$$

for each starting position p. These are exactly the sets over which we sum to get the δ'_j's. Moreover, the terms which are being summed are not changed, so that the path integral formula holds for $\phi\tau$.

We now consider the reordering $\kappa\tau$, where $\kappa = (k : k+1)$. Again we suppose that $\delta = \tau(\gamma)$ is given by the path integral formula, and we wish to show that the same is true for $\delta' = \kappa\delta = \kappa\tau(\gamma)$. For $j \neq \iota^\kappa(k,k+1)$, we have $\delta'_j = \delta_{\kappa' j}$, while by (6.3), r-paths in $\kappa\tau$ ending at j are exactly those in τ which end at $\kappa' j$, that is $\Pi_r(p,j,\kappa\tau) = \Pi_r(p,\kappa' j,\tau)$. Again we sum the same terms, so the formula holds for those $j \neq \iota^\kappa(k,k+1)$.

It remains to check the path integral formula for the wave δ'_i, where $\kappa' i = i = \iota^\kappa(k,k+1)$. According to the definition, we have

$$\delta'_i = \delta_i + \delta_k\delta_{k+1}\Lambda_{\bar{c}}\,,$$

where $\bar{c} = c_{\tau' i}$, and we must express this in terms of the γ's. According to our assumptions, either δ_k or δ_{k+1} must be a 2-wave, and the other a sound wave from the $\bar{c}'$ family, where $\bar{c}' = 4 - \bar{c}$. For definiteness, we suppose that δ_k is the 2-wave, the other case being similar. Then $\delta_k = \gamma_{\tau' k}$ by induction. Also, the waves δ_i and

δ_{k+1} are given inductively by the path integral formula. Thus we have

$$\delta_i' = \delta_i + \delta_{k+1}\gamma_{\tau' k}\Lambda_{\bar{c}} \tag{6.6}$$

$$= \sum_{r\geqq 0}\sum_{\Pi_r(\,\cdot\,,i,\tau)} \gamma_{j_1}\gamma_{k_1}\cdots\gamma_{k_r}\Lambda(i,r) \tag{6.7}$$

$$+ \gamma_{\tau' k}\Lambda_{\bar{c}}\sum_{r}\sum_{\Pi_r(\,\cdot\,,k+1,\tau)} \gamma_{j_1}\gamma_{k_1}\cdots\gamma_{k_r}\Lambda(k+1,r)\,. \tag{6.8}$$

By the remark following Lemma 6.2, any r-path in τ which ends at position $k+1$ extends to a unique $(r+1)$-path in $\kappa\tau$. Since τ is one-to-one on 2-waves, the $(r+1)^{\text{st}}$ pair of this path must be the pair $(\iota^{\tau}(j_r,k_r),\tau' k)\in C_{\kappa\tau}$, and the extended path ends at i. Moreover, since δ_{k+1} is a $\bar{c}'$-wave and δ_i is a $\bar{c}$-wave, we have

$$\Lambda_{\bar{c}}\Lambda(k+1,r) = \Lambda_{\bar{c}}\Lambda_{\bar{c}'}^{[(r+1)/2]}\Lambda_{\bar{c}}^{[r/2]} = \Lambda(i,r+1)\,.$$

Thus the second sum above is exactly the sum over paths in $\kappa\tau$ which include the pair $\tau'(k,k+1)$ and end at i, while the first sum is over paths ending at i but not including the pair $\tau'(k,k+1)$. Since these are exactly all paths in $\kappa\tau$ which ends at i, the path integral formula holds for δ_i', and the theorem is proved. □

7. The Functional Bounds

We now find time-dependent bounds for the nonlinear functionals for systems with a Riemann coordinate. By the results of Sect. 5, these bounds are all that are needed to deduce large-time existence of solutions. We shall derive the bounds in steps, first bounding the functional V^*, and obtaining the bound for Q^* and continuity of the functional as corollaries.

Before proceeding, we recall our assumptions in detail. We are assuming that the second (entropy) family possesses a Riemann coordinate, so that the second interaction coefficient Λ_2^{31} vanishes. Denoting the Riemann coordinate by $w:\mathscr{U}\to \mathrm{R}$, the strength of a 2-wave (which is a contact discontinuity) is defined by

$$\gamma = w(u_R) - w(u_L)\,, \tag{7.1}$$

where u_L and u_R are the states on either side of the wave. In particular, given a sequence of constant states separated by waves, the strength of all 2-waves can be found through the Riemann coordinate w. For convenience, we also assume that the second family is linearly degenerate, so that no 2-waves merge. We shall see that our bounds, and hence conditions for existence, depend on the entropy of the initial data. This is not surprising, because the (degenerate) entropy field does not grow or decay to leading order, so that entropy jumps in the solution persist.

Our estimates will depend on the local variation of the entropy of the data, quantified by the d-norm defined earlier,

$$\|w(u_0)\|_d = \sup_x TV_{[x,x+d]}(u_0)\,,$$

which we require to be small for some $d>0$. This is a weak assumption which simply says that the variation of the entropy is spread out over intervals of length d,

and so does not accumulate at any point. Since we are dealing with wave sequences only, we shall use the functional for the local d-variation of the sequence,

$$H(\gamma) = \sup \sum_{c_i=2} |\gamma_i| ,$$

where the sum (and sup) are over 2-waves positioned in an interval of length d. It is clear that H is bounded by $\|w(u)\|_d$. Throughout this section, we assume that d is a fixed positive number, and the local d-variation H is smaller than a constant determined only by the conservation law.

We begin by considering the functional V^* for the total variation. Recall that V^* is defined as the supremum of quantities $V(\tau\gamma)$, where γ is a wave sequence and τ is an admissible reordering. We shall find a bound for V^* by expressing the reordered sequence $\tau\gamma$ in terms of the path integral formula, and bounding the total variation of this reordered sequence. Since our bounds are time-dependent, we further restrict the class of reorderings to those which correspond to wave configurations up to time T, which we fix for the rest of this section.

The idea of the proof is as follows: according to the path integral formula, the amount of scattered wave strength generated by a single wave is

$$\sum_{\Pi_r} \gamma_{k_1} \cdots \gamma_{k_r} \Lambda(\cdot, r) ,$$

where we sum over all r-paths emanating from that wave. Now if each γ_{k_r} were bounded by β, then this is bounded by

$$\sum_r (\beta\Lambda)^r \#\Pi_r \leqq \sum \binom{N_T}{r} (\beta\Lambda)^r ,$$

according to Corollary 6.3. Now using the binomial theorem, we get the bound $(1 + \beta\Lambda)^{N_T} \leqq \exp(\beta\Lambda N_T)$. Although this bound depends on the mesh size through the number N_T of time-steps, we modify the method to obtain uniform bounds by accounting for the d-norm of the initial data. We know that when the variation is small, it does not grow. In particular, on the local lengthscale d, the local variation is small and growth will not occur. We therefore overestimate the amplification due to interactions on the small lengthscale d, and then combine these into the full lengthscale of the support of the solution, to bound the growth of the solution as above. Thus, instead of considering each 2-wave separately, we group them into "blocks" of size d, and use β as an estimate for the amount of wave strength generated by each block. The number of blocks is then $N_T = O(T/d)$, and we obtain uniform bounds by the above argument.

We proceed with the details. Fix the total variation $V = V(\gamma)$, time of existence T, mesh size $\Delta x = \lambda\Delta t$, as well as the local total d-variation of entropy $\delta = H(\gamma)$. We restrict the class of reorderings to those corresponding to a set of diamonds lying below the curve $t = T$. We prove the following theorem which gives the growth bound for V^*, and thus establishes the growth rate bound in Theorem 1 at the quadratic level.

Theorem 3. *If the sequence γ is such that $\eta = H(\gamma)$ satisfies*

$$6\Lambda\eta < 1/2, \quad \textit{for some fixed } d > 0 ,$$

then the functional V^* is bounded,

$$V^*(\gamma) \leqq V(\gamma) + V(\gamma)\exp(8\Lambda\eta T\lambda/d)\,.$$

In particular, this bound is uniform as the mesh size $\Delta x \to 0$.

Proof. Since V^* is defined by $V^* = \sup_\tau V(\tau\gamma)$, we choose a reordering τ and look for a bound for the variation of the reordered sequence.

According to the path integral formula in Theorem 2, the reordered sequence $\delta = \tau\gamma$ is given by

$$\delta_i = \sum_j \gamma_j \sum_{r\geqq 0} \Lambda(i,r) \sum_{\Pi_r(j,i,\tau)} \gamma_{k_1}\cdots\gamma_{k_r}\,, \tag{7.2}$$

where $\Pi_r(j,i,\tau)$ is the collection of r-paths in τ from $j = j_1$ to $i = \tau\iota^\tau(j_r,k_r)$, and $\Lambda(i,r) = \Lambda_c^{[(r+1)/2]}\Lambda_{c'}^{[r/2]}$. We have $V(\tau\gamma) = V(\delta) = \sum_i|\delta_i|$, and rearranging the sum gives

$$V(\tau\gamma) \leqq \sum_j |\gamma_j| \sum_{r\geqq 0} \sum_{\Pi_r(j,\cdot,\tau)} |\gamma_{k_1}\cdot\gamma_{k_r}|\Lambda^r\,, \tag{7.3}$$

where now the sum is over all paths in τ beginning at j, and we have set $\Lambda = \max\{|\Lambda_1|,|\Lambda_3|\}$. We shall find bound for each *amplification factor*

$$A_j = \sum_{r\geqq 1} \sum_{\Pi_r(j,\cdot,\tau)} |\gamma_{k_1}\cdot\gamma_{k_r}|\Lambda^r\,, \tag{7.4}$$

which measures the amount of total variation generated by multiple scattering of the sound wave γ_j. The functional V^* is then clearly bounded,

$$V^*(\gamma) \leqq V(\gamma)(1+A)\,,$$

where A is the bound for the amplification factor.

We first estimate the amplification factor due to a group of 2-waves having small total variation: this will be used to estimate the contribution due to blocks of size $3d$.

Claim. *If the total variation of 2-waves appearing in A_j is small, that is $\sum_k|\gamma_k|\Lambda < \varepsilon < 1$, then so is A_j,*

$$A_j \leqq \frac{\varepsilon}{1-\varepsilon}\,.$$

To see this, we simply write

$$\begin{aligned} A_j &\leqq \sum_{r\geqq 1} (\textstyle\sum|\gamma_{k_1}|\Lambda)\cdots(\sum|\gamma_{k_r}|\Lambda) \\ &\leqq \sum_{r\geqq 1} \varepsilon^r \leqq \frac{\varepsilon}{1-\varepsilon}\,. \end{aligned}$$

We will use this to over-estimate the contribution due to each block.

We now partition paths into blocks, as follows. For a fixed sound wave γ_j located at some point x_j of space, we partition the real line into (non-overlapping) intervals

$$I_m = (x_j + md, x_j + (m+1)d]$$

of length d. We also define the extended intervals

$$\hat{I}_m = (x_j + (m-1)d, x_j + (m+2)d]\,, \tag{7.5}$$

$$= I_{m-1} \cup I_m \cup I_{m+1} \tag{7.6}$$

of length $3d$. We shall say that γ_{k_i} (or simply k_i) lies in the interval I_m, if the spatial position of the wave γ_{k_i} is in I_m initially (i.e. before any wave interact).

Suppose we are given a single path starting at j, and determined by the set $\{k_1,\ldots,k_r\}$, each k_i referring to the corresponding 2-wave γ_{k_i}. We partition the path into blocks as follows. Let m_1 be such that k_1 lies in the interval I_{m_1}. Define the first block B_1 to be the maximal set of k_i's such that $k_1,\ldots,k_{b_1}$ lie in the extended interval $\hat{I}_{m_1}$. Now let m_2 refer to the interval I_{m_2} containing the first wave *not* in $\hat{I}_{m_1}$, that is k_{b_1+1}, see Fig. 5. Now define b_2 and the block B_2 to be the largest set of k_i's such that $k_{b_1+1},\ldots,k_{b_1+b_2}$ lie in $\hat{I}_{m_2}$.

Continuing this process, we partition the path into blocks $B_1,\ldots,B_\ell$, with the following useful properties. All 2-waves in a single block lie in an interval of length $3d$, and therefore have maximum total variation 3δ. We will thus be able to apply the above claim to each block, regardless of the number of waves in that block. Moreover, according to our construction, the distance between consecutive blocks is at least d. This will allow us to bound the number of blocks.

With these definitions in hand, we now reconsider the amplification factor A_j. By partitioning each path into blocks, we can rewrite

$$A_j \leqq \sum_{\ell} \left(\sum_{b_1} |\gamma_{k_1} \cdots \gamma_{k_{b_1}} A^{b_1}| \right) \cdots \left(\sum_{b_\ell} |\gamma_k \cdots \gamma_{k_r} A^{b_\ell}| \right)\,, \tag{7.7}$$

where $\sum b_i = r$ and each group of waves comes from block B_i. According to our construction, each block B_i comes from a set of waves occupying a spatial interval

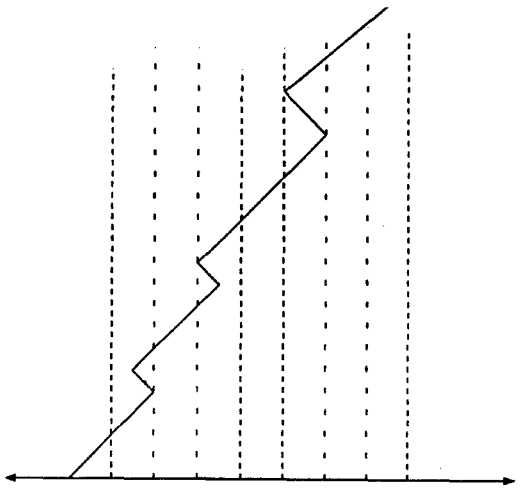

Fig. 5. Decomposition into blocks

of length $3d$, and so having local $3d$-variation given by 3δ. We can thus apply the claim above and bound each of these terms,

$$\sum_{b_1} |\gamma_{k_1} \cdots \gamma_{k_{b_1}} \Lambda^{b_1}| \leqq \frac{3\Lambda\delta}{1-3\Lambda\delta} \leqq 4\Lambda\delta\,,$$

etc., since $3\Lambda\delta < 1/4$. Using this bound in (7.7) then gives

$$A_j \leqq \sum_{\ell} (4\Lambda\delta)^{\ell} \#\mathscr{B}_{\ell}\,,$$

where $\#\mathscr{B}_{\ell}$ is the number of ways of choosing ℓ blocks $B_1, \ldots, B_{\ell}$.

Finally, we count the number of ways of choosing ℓ blocks, using the method of Corollary 6.3. Since block are separated in space by distance d, and wave-speeds are finite, it takes at least d/λ units of time to move between blocks. We again track the block by projecting to corresponding times of first interaction inside each block. Thus we must choose times $t_1 < \cdots < t_{\ell} < T$, where each $t_{i+1} - t_i > d/\lambda$. In Corollary 6.3, because each scattered wave had a known direction, the interaction times were enough to uniquely determine the path. In the present case, however, an even number of interactions in one block will not effect a direction change, so that there is a corresponding choice of direction for each interaction time (i.e. the next block could be either to the left or right of the present block). Thus the choices of ℓ times together with ℓ directions determine all possible ways of choosing ℓ blocks, and we have

$$\#\mathscr{B}_{\ell} \leqq \binom{T\lambda/d}{\ell} 2^{\ell}\,.$$

Thus we finally obtain a bound for the amplification factor,

$$A_j \leqq \sum_{\ell} \binom{T\lambda/d}{\ell} (8\Lambda\delta)^{\ell} \tag{7.8}$$

$$\leqq (1 + 8\Lambda\delta)^{T\lambda/d} \tag{7.9}$$

$$\leqq \exp(8\Lambda\delta T\lambda/d)\,, \tag{7.10}$$

where we have again used the binomial theorem. We note that this amplification factor is uniform in the mesh size. Finally, inserting this bound for the amplification factor into (7.3) yields the total variation bound, thus completing the proof. □

We remark that in bounding the amplification by partitioning paths into blocks, we have made heavy use of Lemma 6.2, which says that complex paths project to and are built up from smaller paths.

It is remarkable that the bound reduces to finding the amplification due to the passage of a single weak sound wave through a fluctuating entropy field, and that the amplification is essentially independent of the strength of the sound wave. The theorem shows that we have a bound for the variational potential V^*. We shall use this to show that the quadratic error potential Q^* is bounded, and the functionals V^* and Q^* are continuous.

Corollary 7.1. *Under the assumptions of Theorem* 3, *the quadratic error potential* Q^* *is also bounded, namely*

$$Q^*(\gamma, \tau) \leqq V^2 \exp(16\Lambda\delta T\lambda/d)\,.$$

Note that this can be viewed as analogous to the bound for Glimm's quadratic error potential, $Q \leqq V^2$, where we must now replace V by the potential V^* for the variation.

Proof. Given γ and τ, by Eq. (4.2) we have

$$Q^*(\gamma,\tau) = \sum_{C_\tau} |\gamma_j^\mu||\gamma_k^\mu| + \sum_{M_\tau}{}' |\gamma_j^\mu||\gamma_k^\mu| ,$$

where the sub-ordering μ is associated to the each pair $(j,k) \in C_\tau$. We now expand each γ_j^μ by the path integral formula, and rearrange the sum to get the bound. The idea is that since each wave pair interacts only once, and we can bound the amount of total variation generated by each wave, the total is bounded by the square of the variation bound. Using the path integral formula, we have

$$\gamma_k^\mu = \sum_l \gamma_l \sum_r \Lambda(k,r) \sum_{\Pi_r(l,k,\mu)} \gamma_{m_1} \cdots \gamma_{m_r} ,$$

with the analogous formula for γ_j^μ. Now, since paths in μ are paths in τ, we have $\Pi_r(l,k,\mu) \subseteq \Pi_r(l,k,\tau)$, so that we may use the amplification factor of the theorem to bound the total variation, even though the partial reordering μ changes for different interacting pairs. Substituting in the bounds found above, we get

$$Q^*(\gamma,\tau) \leqq \sum_{I_r} |\gamma_j^\mu||\gamma_k^\mu| \tag{7.11}$$

$$\leqq \sum_{j,l} |\gamma_j| A_j |\gamma_l| A_l \tag{7.12}$$

$$\leqq \left(\sum_j |\gamma_j| A_j \right)^2 , \tag{7.13}$$

and the result follows. □

We now address the problem of continuity of the functionals. Starting with two sequences which differ in one wave only, and a fixed reordering, we must show that the resulting sequences differ by an appropriately small amount.

As above, we start with a sequence γ and reordering τ. We have seen that the reordered sequence $\delta = \tau\gamma$ has bounded variation. We now wish to describe the effect of perturbing a single wave in the initial sequence. Thus, suppose that γ' is a sequence which coincides with γ except in one position, say p. We wish to bound the differences $V^*(\gamma) - V^*(\gamma')$ and $Q^*(\gamma,\tau) - Q^*(\gamma',\tau)$ by a multiple of $|\gamma_p - \gamma'_p|$.

Corollary 7.2. *The functionals V^* and Q^* are continuous: that is, there are functions $K_V(V,H)$ and $K_Q(V,H)$, also depending on time T, such that*

$$|V^*(\gamma) - V^*(\gamma')| \leqq K_V(V,H)|\gamma_p - \gamma'_p|, \tag{7.14}$$

and

$$|Q^*(\gamma) - Q^*(\gamma')| \leqq K_Q(V,H)|\gamma_p - \gamma'_p| . \tag{7.15}$$

Proof. Since the bound for the functional V^* depends smoothly on the quantities $H(\gamma)$ and $V(\gamma)$, and these norms change by at most $|\gamma_p - \gamma'_p|$, we have

$$|V^*(\gamma) - V^*(\gamma')| \leqq \frac{\partial V^*}{\partial V}|V(\gamma) - V(\gamma')| + \frac{\partial V^*}{\partial H}|H(\gamma) - H(\gamma')| \tag{7.16}$$

$$\leqq K_F(V,H)|\gamma_p - \gamma'_p| , \tag{7.17}$$

where K_V is an increasing function of its arguments, and we have set

$$V = \sup\{V(\gamma), V(\gamma')\} \quad \text{and} \quad H = \sup\{H(\gamma), H(\gamma')\} .$$

It is clear that Q^* is also continuous. □

Theorems 1 and 4 now follow from the general theorems of Sect. 5, using the bounds established in Theorem 3 and Corollaries 7.1 and 7.2, which are applicable to systems possessing a Riemann coordinate, which include the systems (1.1) and (1.11).

References

1. Cheverry, C.: Optique geometrique faiblement non linearire: Le cas general. Prepublication 95-11, Institut de Recherche Mathematique de Rennes, Mars 1995
2. Lax, P.D.: Hyperbolic systems of conservation laws, II. Comm. Pure Appl. Math. **10**, 537–566 (1957)
3. Glimm, J.: Solutions in the large for nonlinear hyperbolic systems of equations. Comm. Pure Appl. Math. **18**, 697–715 (1965)
4. Glimm, J., Lax, P.: Decay of solutions of systems of nonlinear hyperbolic conservation laws. Memoirs Amer. Math. Soc. **101** (1970)
5. Hunter, J.: Strongly nonlinear hyperbolic waves. Nonlinear Hyperbolic Equations-Theory, Computational Methods, and Applications, J. Ballman, R. Jeltsch (eds.) Viewig, 1989, pp. 257–268
6. Metivier, G., Joly, J.L., Rauch, J.: Resonant one dimensional non-linear geometrical optics. (to appear), 1992
7. Metivier, G.: A nonlinear instability for 3×3 systems of conservation laws. Preprint, 1993.
8. Liu, T.-P.: The deterministic version of the Glimm scheme. Commun. Math. Phys. **57**, 135–148 (1977)
9. Majda, A., Rosales, R.: Resonantly interacting weakly nonlinear hyperbolic waves in a single space variable. Stud. Appl. Math. **71**, 149–179 (1984)
10. Schochet, S.: Glimm's Scheme for Systems with Almost-Planar Interactions. Commun. PDE, **16**, 1423–1440 (1991)
11. Serre, D.: Richness and the classification of quasilinear hyperbolic system. IMA volumes #**29**, Berlin, Heidelberg, New York: Springer, 1991, 315–333
12. Sévennec, B.: Geometrie des systemes de lois de conservation, Memoire SMF #**56** (1994)
13. Smoller, J.: Shock-Waves and Reaction-Diffusion Equations. Berlin, Heidelberg, New York: Springer 1983
14. Temple, J.B.: Supnorm estimates in Glimm's method. J. Diff. Eqn. **83**, no. 1, 79–84 (1990)
15. Temple, B., Young, R.: The large time existence of periodic solutions for the compressible Euler equations. To appear in Proceedings of the IV International Workshop on Shock Waves and Conservation Laws, IMPA Brazil, July 1995
16. Tong, Y.: A functional integral approach to shock wave solutions of Euler equations with spherical symmetry. Commun. Math. Phys., to appear
17. Young, R.: Sup-norm Stability for Glimm's Scheme. Comm. Pure Appl. Math. **46**, 903–948 (1993)
18. Young, R.: Exact solutions to degenerate conservation laws. Preprint 1994
19. Young, R.: On elementary interactions for hyperbolic conservation laws. Preprint 1994
20. Zumbrun, K.: In preparation

Communicated by S.-T. Yau