Nonlinear diffusive–dispersive limits for scalar multidimensional conservation laws

Seok Hwang

Department of Mathematics, LaGrange College, LaGrange, GA 30240, USA
Received 11 March 2004; revised 21 October 2005
Available online 1 December 2005

Abstract

We consider a class of multidimensional conservation laws with vanishing nonlinear diffusion and dispersion terms. Under a condition on the relative size of the diffusion and dispersion coefficients, we show that the approximate solutions converge in a strong topology to the entropy solution of a scalar conservation law. Our proof is based on methodology developed in [S. Hwang, A.E. Tzavaras, Kinetic decomposition of approximate solutions to conservation laws: Applications to relaxation and diffusion–dispersion approximations, Comm. Partial Differential Equations 27 (2002) 1229–1254] which uses the averaging lemma.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Hyperbolic conservation laws; Diffusion–dispersion approximation; Kinetic formulation; Averaging lemmas

1. Introduction

Nonlinear hyperbolic conservation laws arise in the modeling of many problems from continuum mechanics, physics, chemistry, etc. The equation becomes parabolic when additional small scale dissipation mechanisms are taken into account: diffusion, heat conduction, capillarity in fluids, Hall effect in magnetohydrodynamics, etc. From a general standpoint, hyperbolic equations admit discontinuous solutions while parabolic equations have smooth solutions. Discontinuous solutions, understood in the generalized sense of
the distribution theory, are usually non-unique. It is therefore fundamental to understand which solutions are selected by a specific zero diffusion–dispersion limit. In this paper, we address this issue for a scalar multidimensional conservation law.

Consider a scalar multidimensional conservation law

$$\partial_t u + \text{div} F(u) = 0, \quad x \in \mathbb{R}^d, \quad t \in \mathbb{R}^+, \quad u(x, 0) = u_0(x), \quad x \in \mathbb{R}^d, \quad (1.1)$$

where the unknown function $u = u(x, t)$ is scalar-valued and the flux $F : \mathbb{R} \to \mathbb{R}^d$ is a given function. Then there are available two equivalent notions for weak solutions: the Kruzhkov entropy solution [9], stating that $u$ satisfies the entropy inequalities

$$\partial_t \eta(u) + \text{div} q(u) \leq 0 \quad \text{in} \ D',$$

for any entropy pair $\eta - q$ with $\eta$ convex, and the kinetic formulation of Lions et al. [13]. Both concepts lead to uniqueness, stability theorems and error estimates for approximate entropy solutions [9,15].

Starting with Tartar [21], entropy pairs are used to determine compactness for approximate solutions to (1.1). The compactness of a given family $\{u^\varepsilon\}$ of approximate solutions bounded in some $L^p$-norm ($p > 1$) appears to be determined by compactness of the entropy dissipation measure in the sense

$$\partial_t \eta(u^\varepsilon) + \text{div} q(u^\varepsilon) \quad \text{is precompact in} \ H^{-1}_{\text{loc},x,t}. \quad (1.3)$$

This has been proved in one-space dimension in both the $L^\infty$ and $L^p$ stability settings by Tartar [21] and Schonbek [19] (see [18] for a simplified proof using singular entropies and [22] for an analysis of the compensated compactness bracket in multidimension). Tartar’s framework is quite versatile and applies even to approximations that do not yield entropy solutions. In this paper, we will see how the kinetic formulation compactness framework of Lions et al. [13] can be easily adapted to analyze the structure (1.3) in multidimension.

For multidimensional conservation laws, convergence of approximate solutions is usually deduced by using a framework of DiPerna [3] and Szepessy [20]. The argument hinges on showing that a Young-measure solution (with certain regularity in time) that satisfies (1.2) for all convex $\eta$ and is a Dirac mass at $t = 0$ is in fact a regular weak solution. It yields compactness for bounded families of approximate entropy solutions, i.e., approximate solutions $\{u^\varepsilon\}$ that satisfy the dissipation structure

$$\partial_t \eta(u^\varepsilon) + \text{div} q(u^\varepsilon) \leq P^\varepsilon(u^\varepsilon) \quad (1.4)$$

with $P^\varepsilon(u^\varepsilon) \to 0$ in $D'$ as $\varepsilon \to 0$. An alternative compactness framework is proposed in [13] by means of the kinetic formulation and averaging lemmas (e.g., [16,17]). The framework in [13] is developed for approximations that still satisfy (1.4). Nevertheless, as we will see, it can be easily adapted to apply to the structure (1.3). This is achieved by first transforming the entropy dissipation structure of the problem at hand to an approximate transport equation. It results to an approximate transport equation where the right-hand side consists of “lower order
terms”; then the averaging lemma of Perthame and Souganidis [17] yields compactness of moments in $L^p$ spaces.

This idea was used to show the convergence of the relaxation approximation of Jin-Xin type, diffusion–dispersion approximation in [7] and kinetic models of BGK type in [6]. In this paper, we pursue this approach to get the convergence to a scalar conservation law by a nonlinear diffusive–dispersive approximation.

Consider the following approximation of (1.1) obtained by adding a nonlinear diffusion, $b : \mathbb{R}^d \to \mathbb{R}^d$, and a linear dispersion to the right-hand side of (1.1):

$$
\partial_t u + \text{div} F(u) = \varepsilon \sum_{j=1}^d \partial_{x_j} b_j(\nabla u) + \delta \sum_{j=1}^d \partial_{x_j x_j} u, \quad x \in \mathbb{R}^d, \ t \geq 0,
$$

$$
u(x, 0) = u_{0}^{\varepsilon, \delta}(x), \quad x \in \mathbb{R}^d.
$$

Let $u^{\varepsilon, \delta} : \mathbb{R}^d \times [0, T] \to \mathbb{R}$ be smooth solutions defined on an interval $[0, T]$ with a uniform $T$ independent of $\varepsilon, \delta$. Also, in (1.5), $u_{0}^{\varepsilon, \delta}$ is an approximation of the initial condition $u_0$ in (1.1).

Our main objective is to establish the condition under which, as $\varepsilon, \delta$ tend to 0, the approximate solutions $u^{\varepsilon, \delta}$ converge in a strong topology to the entropy solution of (1.1). When $\varepsilon = 0$, Eq. (1.5) is a generalized version of the well-known Korteweg–de Vries (KdV) equation, and the solutions become more and more oscillatory as $\delta \to 0$: the approximate solutions do not converge, see [10]. When $\delta = 0$, Eq. (1.5) reduces to a nonlinear parabolic equation resembling the pseudo-viscosity approximation of Von Neumann and Richtmyer [14]; in that regime, the solutions converge strongly to the entropy solution. Therefore, to ensure the convergence of the zero diffusion–dispersion approximation (1.5), it is necessary that diffusion dominate dispersion. Indeed, we show that the solutions of (1.5) tend to the entropy solution of (1.1) when $\varepsilon, \delta \to 0$ with $|\delta| \ll \varepsilon$.

For clarity, the main assumptions made in this paper are collected here.

(H1) For some constant $C_1, C_1' > 0$ and $m \geq 1$,

$$
|F'(u)| \leq C_1 + C_1'|u|^{m-1} \quad \text{for all } u \in \mathbb{R}.
$$

For the diffusion term, we fix $r \geq 1$ and let $b(\lambda) = (b_1(\lambda), \ldots, b_d(\lambda))$.

(H2) For some constant $C_2, C_3 > 0$,

$$
C_2|\lambda|^{r+1} \leq \lambda \cdot b(\lambda) \leq C_3|\lambda|^{r+1} \quad \text{for all } \lambda \in \mathbb{R}^d.
$$

(H3) The gradient matrix $Db(\lambda)$ is a positive definite matrix uniformly in $\lambda \in \mathbb{R}^d$.

We remark that the linear diffusion $b_j(\nabla u) = \partial_{x_j} u$ satisfies (H3).

Convergence of (1.5) has been established in the 1-dimensional case by LeFloch and Natalini [12], and in the multidimensional case by Correia and LeFloch [2] (see also [7,8]). In particular, the analysis of [2] uses the framework of DiPerna and Szepessy [3,20] and is based on the dissipative structure (1.4) which is valid only on the range
\[ \delta = o(\varepsilon^{(r+3)/(r+1)}) \]. We consider here a family of smooth solutions \( \{u^\varepsilon\} \) to (1.5) in the range \( \delta = O(\varepsilon^{(r+3)/(r+1)}) \) and show in Section 4 that \( \chi^\varepsilon = 1(u^\varepsilon, \xi) \) satisfies the transport equation

\[ \partial_t \chi^\varepsilon + F'(\xi) \cdot \nabla \chi^\varepsilon = \sum_{j=1}^{d} \partial_{x_j} \left( \bar{g}_j^\varepsilon + \partial_\xi g_j^\varepsilon + \bar{h}_j^\varepsilon + \partial_\xi h_j^\varepsilon \right) + \partial_\xi (m^\varepsilon + k^\varepsilon) \quad \text{in} \quad D'_{x,t,\xi}, \tag{1.6} \]

where

\[ 1(u, \xi) = \begin{cases} 1 & \text{if } u > 0, \\ 0 & \text{if } u = 0, \\ -1 & \text{if } u < 0 \end{cases} \tag{1.7} \]

is the usual Maxwellian associated with the kinetic formulation of scalar conservation laws, \( \bar{g}_j^\varepsilon, g_j^\varepsilon, \bar{h}_j^\varepsilon, h_j^\varepsilon \to 0 \) in \( L^r_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), \( r^* = \frac{r+1}{2} \), while \( m^\varepsilon, k^\varepsilon \) are bounded in measures (\( k^\varepsilon \) is not necessarily positive). Convergence is then obtained via the averaging lemma in [17]. In the limit \( \varepsilon \to 0 \), \( \chi^\varepsilon \to 1(u, \xi) =: \chi \) which satisfies

\[ \partial_t \chi + F'(\xi) \cdot \nabla \chi = \partial_\xi (m + k) \quad \text{in} \quad D'_{x,t,\xi}, \tag{1.8} \]

with \( m \) and \( k \) bounded measures. It turns out that if \( \delta = o(\varepsilon^{(r+3)/(r+1)}) \) the bounded measure \( m + k \) is positive and \( u \) is an entropy solution. By contrast, if \( \delta = O(\varepsilon^{(r+3)/(r+1)}) \), it is not clear whether the limit measure \( m + k \) is positive (see Remark 4.3 for a discussion).

2. Main results

Throughout it is assumed \( u_0 \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \) and the initial data in (1.5) are smooth function with compact support and uniformly bounded in \( L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \). Restricting attention to the diffusion-dominant regime we regard \( \delta = \delta(\varepsilon) \) and we suppose that \( u_{0,\delta}^\varepsilon \) approaches the initial condition \( u_0 \) of (1.1) in the sense that

\[ \lim_{\varepsilon \to 0} u_{0,\delta}^\varepsilon = u_0 \quad \text{in} \quad L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d), \quad \| u_{0,\delta}^\varepsilon \|_{L^2(\mathbb{R}^d)} \leq \| u_0 \|_{L^2(\mathbb{R}^d)}. \tag{2.1} \]

In this paper, we establish the following:

**Theorem 2.1.** Suppose that (H1)–(H3) hold with \( m \) and \( r \) such that \( 1 \leq m \leq \frac{2r}{r+1} \) and \( r \geq 1 \). In addition, \( F(u) \) satisfies the nondegeneracy condition (4.4) (or (4.3)).

(i) If \( \delta = O(\varepsilon^{(r+3)/(r+1)}) \), then solutions \( u^\varepsilon \) of (1.5) converge along a subsequence to a function \( u \) in \( L^s((0,T); L^p_{\text{loc}}(\mathbb{R}^d)) \), for all \( s < \infty \) and \( p < 2 \); the limiting \( u \) is a weak solution of (1.1).

(ii) If \( \delta = o(\varepsilon^{(r+3)/(r+1)}) \), then \( u \) is the unique Kruzhkov entropy solution of (1.1).
From this we obtain a result for diffusion with linear growth \((r = 1)\):

**Theorem 2.2.** Suppose \(F(u)\) satisfies \((H1)\) with \(m = 1\) and the nondegeneracy condition \((4.4)\) (or \((4.3))\). And \(b\) satisfies \((H2), (H3)\) with \(r = 1\).

(i) If \(\delta = O(\varepsilon^2)\), then solutions \(u^\varepsilon\) of \((1.5)\) converge along a subsequence to a function \(u\) in \(L^s((0, T); L^p_{\text{loc}}(\mathbb{R}^d))\), for all \(s < \infty\) and \(p < 2\); the limiting \(u\) is a weak solution of \((1.1)\).

(ii) If \(\delta = o(\varepsilon^2)\), then \(u\) is the unique Kružkov entropy solution of \((1.1)\).

3. A priori estimates

In this section we consider a sequence \(\{u^{\varepsilon, \delta}\}\) of smooth solutions of \((1.5)\) vanishing at infinity. The superscripts \(\varepsilon\) and \(\delta\) are omitted in this section except when the emphasis is necessary. We have the following estimates on smooth solutions \(u^{\varepsilon, \delta}\) of \((1.5)\) (for more estimates, see [2]).

**Lemma 3.1.** Suppose that \((H1)–(H3)\) hold with \(m\) and \(r\) such that \(1 \leq m \leq \frac{2r}{r+1}\) and \(r \geq 1\). For any solution of \((1.5)\) and \(t \in [0, T]\), we have

\[
\int_{\mathbb{R}^d} u(t)^2 \, dx + 2\varepsilon \int_0^t \int_{\mathbb{R}^d} \nabla u \cdot b(\nabla u) \, dx \, ds = \int_{\mathbb{R}^d} u_0^2 \, dx, \tag{3.1}
\]

\[
\varepsilon \int_0^t \int_{\mathbb{R}^d} |\nabla u|^{r+1} \, dx \, ds \leq C \int_{\mathbb{R}^d} u_0^2 \, dx, \tag{3.2}
\]

\[
\varepsilon^{r+3} \int_{\mathbb{R}^d} |\nabla u(t)|^{r+1} \, dx + \varepsilon^{\frac{2(r+2)}{r+1}} \int_0^t \int_{\mathbb{R}^d} |D^2 u|^2 \, dx \, ds \leq C. \tag{3.3}
\]

**Proof.** To derive the estimates (3.1) and (3.2), we multiply \((1.5)\) by \(\eta'(u)\) where \(\eta : \mathbb{R} \to \mathbb{R}\) is a sufficiently smooth function and define \(q : \mathbb{R} \to \mathbb{R}^d\) by \(q_j = \eta' F_j'\). Then we obtain

\[
\partial_t \eta(u) + \text{div} q(u) = \varepsilon \sum_{j=1}^d \partial_{x_j} \left( \eta'(u) b_j(\nabla u) \right) - \varepsilon \eta''(u) \sum_{j=1}^d (\partial_{x_j} u) b_j(\nabla u) \\
+ \delta \sum_{j=1}^d \partial_{x_j} \left( \eta'(u) \partial_{x_j} x_j u \right) - \frac{\delta}{2} \eta''(u) \sum_{j=1}^d \partial_{x_j} (\partial_{x_j} u)^2. \tag{3.4}
\]

If we integrate (3.4) over the whole of \(\mathbb{R}^d\) with \(\eta(u) = \frac{|u|^{\alpha+1}}{\alpha+1}\), we get:
\[
\frac{d}{dt} \int_{\mathbb{R}^d} |u|^{\alpha+1} \alpha + 1 \, dx = - \alpha \epsilon \int_{\mathbb{R}^d} |u|^\alpha \nabla u \cdot b(\nabla u) \, dx - \frac{\alpha \delta}{2} \int_{\mathbb{R}^d} d \sum_{j=1}^d |u|^{\alpha-1} \partial x_j (\partial x_j u)^2. \tag{3.5}
\]

Now if we integrate over \([0, t]\) and use (H2), with \(\alpha = 1\), we obtain (3.1) and (3.2).

To obtain (3.3), we differentiate (1.5) with respect to the variable \(x_k\) and then multiply by \(\nabla u\). If we integrate in \(\mathbb{R}^d\) and use integration by parts, we obtain

\[
\frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |\nabla u(t)|^2 \, dx - \int_{\mathbb{R}^d} \Delta u F'(u) \cdot \nabla u \, dx
\]

\[
= - \epsilon \int_{\mathbb{R}^d} \sum_{k=1}^d \nabla \partial x_k u \cdot D_b(\nabla u) \cdot \nabla \partial x_k u \, dx - \frac{\delta}{2} \sum_{j=1}^d \partial x_j \left( \sum_{k=1}^d d(\partial x_k x_j u)^2 \right) dx. \tag{3.6}
\]

After integrating over \([0, t]\) using (H1), we get

\[
\int_{\mathbb{R}^d} |\nabla u|^2 \, dx + 2 \epsilon \sum_{k=1}^d \int_0^t \int_{\mathbb{R}^d} \nabla \partial x_k u \cdot D_b(\nabla u) \cdot \nabla \partial x_k u \, dx
\]

\[
\leq \int_{\mathbb{R}^d} |\nabla u_0|^2 \, dx + 2C_1 \int_0^t \int_{\mathbb{R}^d} |D^2 u| |u|^{m-1} |\nabla u| \, dx \, dt
\]

\[
\leq \int_{\mathbb{R}^d} |\nabla u_0|^2 \, dx + \frac{C}{\epsilon} \int_0^t \int_{\mathbb{R}^d} |u|^{2m-2} |\nabla u|^2 \, dx \, dt + C_4 \epsilon \int_0^t \int_{\mathbb{R}^d} |D^2 u|^2 \, dx \, dt,
\]

and so, using (H3),

\[
\int_{\mathbb{R}^d} |\nabla u|^2 \, dx + C_5 \epsilon \int_0^t \int_{\mathbb{R}^d} |D^2 u|^2 \, dx \, dt \leq \int_{\mathbb{R}^d} |\nabla u_0|^2 \, dx + \frac{C}{\epsilon} \int_0^t \int_{\mathbb{R}^d} |u|^{2m-2} |\nabla u|^2 \, dx \, dt.
\]

By Hölder inequality and for \(m \leq \frac{r}{r+1}\),

\[
\int_{\mathbb{R}^d} |\nabla u|^2 \, dx + C_5 \epsilon \int_0^t \int_{\mathbb{R}^d} |D^2 u|^2 \, dx \, dt
\]

\[
\leq \int_{\mathbb{R}^d} |\nabla u_0|^2 \, dx + \frac{C}{\epsilon} \left[ \int_0^t \int_{\mathbb{R}^d} |\nabla u|^{r+1} \, dx \, dt \right]^{2/r+1} \left[ \int_0^t \int_{\mathbb{R}^d} |u|^2 \, dx \, dt \right]^{2/r+1}
\]

and now (3.3) follows from (3.1), (3.2). \(\square\)
4. Convergence proof

In preparation, recall that \( \eta-q \) with \( q = (q_j (u))_{j=1,\ldots,d} \) is an entropy–entropy flux pair if \( q_j' = a_j n' \). Such pairs describe the nonlinear structure of (1.1) and are represented in terms of the kernel \( \mathbb{1}(u, \xi) \) by the formulas

\[
\eta(u) - \eta(0) = \int_{\xi} \mathbb{1}(u, \xi) n'(\xi) d\xi, \quad q_j(u) - q_j(0) = \int_{\xi} \mathbb{1}(u, \xi) F_j'(\xi) n'(\xi) d\xi. \tag{4.1}
\]

Also, we use the limiting case of the averaging lemma proved in [17], see also [16]:

**Theorem 4.1.** Let \( \{f_n\}, \{g_{i,n}\} \) be two sequences of solutions to the transport equation

\[
\partial_t f_n + a(\xi) \cdot \nabla_x f_n = \partial_t \partial_{\xi} g_{0,n} + \sum_{i=1}^{d} \partial_{x_i} \partial_{\xi} g_{i,n}, \tag{4.2}
\]

where \( k \in \mathbb{N} \). Assume that \( a(\xi) \in C^\infty(\mathbb{R}) \) satisfies the nondegeneracy condition: for \( R > 0 \)

\[
\omega(\beta) = \sup_{\alpha \in \mathbb{R}, \omega \in S^{d-1}} \int_{\{||\xi|| \leq R\}} \left( |\alpha + \frac{a(\xi) \cdot \omega}{\beta}|^2 + 1 \right)^{-1} d\xi \to 0, \quad \text{as } \beta \to 0. \tag{4.3}
\]

If \( \{f_n\} \) is bounded in \( L^q(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), for some \( 1 < q < \infty \), and \( \{g_{i,n}\} \) is precompact in \( L^q(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), then the average

\[
\int_{\mathbb{R}} \psi(\xi) f_n(t, x, \xi) d\xi \quad \text{is precompact in } L^q(\mathbb{R}^d \times \mathbb{R}^+),
\]

for any \( \psi \in C^\infty_c(\mathbb{R}) \).

**Remark 4.2.** 1. The nondegeneracy condition (4.3) is equivalent to for all \( R > 0 \)

\[
\text{meas}\{\xi \in B_R | \alpha + a(\xi) \cdot \omega = 0\} = 0, \quad \forall \alpha \in \mathbb{R}, \, \omega \in S^{d-1}, \tag{4.4}
\]

where \( B_R = \{||\xi|| \leq R\} \) (for example, see [1,17]). Condition (4.4) can be interpreted geometrically that the curve \( \xi \mapsto a(\xi) \cdot \omega + \alpha \) is not locally contained in any hyperplane.

2. Assumption (4.3) (or (4.4)) on the behavior of \( a(\xi) \) is necessary; for example, there would no improvement of regularity in the linear case \( a(\xi) \equiv a \) which corresponds to \( F(\xi) \equiv a_\xi \) in the conservation law (in this case, if we simply choose some \( \omega \) such that \( \omega \perp a \) and \( \alpha = 0 \), then condition (4.3) (or (4.4)) is not satisfied).

3. By using cut-off functions, it is easy to show a variant of Theorem 4.1 stating that under the same hypotheses if \( \{f_n\} \) is bounded in \( L^q_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \) and \( \{g_{i,n}\} \) are precompact in \( L^q_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \) then the averages are precompact in \( L^q_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^+) \) for any \( \psi \in C^\infty_c(\mathbb{R}) \).
**Proof of Theorem 2.1.** Let $u^{\varepsilon, \delta}$ be a family of solutions to (1.5). The proof proceeds in three steps.

**Step 1.** Let $\eta - q$ be an entropy–entropy flux pair and denote by $u^{\varepsilon} = u^{\varepsilon, \delta}$. We multiply (1.5) by $\eta'(u^{\varepsilon})$ and obtain

$$
\partial_t \eta(u^{\varepsilon}) + \text{div } q(u^{\varepsilon}) = \varepsilon \sum_{j=1}^{d} \partial x_j \left( \eta'(u^{\varepsilon}) b_j(\nabla u^{\varepsilon}) \right) - \varepsilon \eta''(u^{\varepsilon}) \sum_{j=1}^{d} \partial x_j u^{\varepsilon} \cdot b_j(\nabla u^{\varepsilon})
$$

$$
+ \delta \sum_{j=1}^{d} \partial x_j \left( \eta'(u^{\varepsilon}) \partial x_j x_j u^{\varepsilon} \right) - \delta \eta''(u^{\varepsilon}) \sum_{j=1}^{d} \left( \partial x_j u^{\varepsilon} \right) \left( \partial x_j x_j u^{\varepsilon} \right).$$

(4.5)

Let $\varphi(x, t) \in C_{c}^\infty(\mathbb{R}^d \times \mathbb{R}^+)$ and let $\eta \in C_{c}^\infty(\mathbb{R})$ be viewed as a test function. By introducing the indicator function $\mathbbm{1}(u^{\varepsilon}, \xi)$, we have

$$
- \int_{x, t, \xi} \mathbbm{1}(u^{\varepsilon}, \xi) \partial_t \varphi(x, t) + \sum_{j=1}^{d} F_j'(\xi) \mathbbm{1}(u^{\varepsilon}, \xi) \partial x_j \varphi(x, t) \eta'(\xi) \, d\xi \, dx \, dt
$$

$$
= - \int_{x, t} \sum_{j=1}^{d} \left( \varepsilon b_j(\nabla u^{\varepsilon}) + \delta \partial x_j x_j u^{\varepsilon} \right) \eta'(u^{\varepsilon}) \partial x_j \varphi(x, t) \, dx \, dt
$$

$$
- \int_{x, t} \eta''(u^{\varepsilon}) \left( \varepsilon \sum_{j=1}^{d} \left( \partial x_j u^{\varepsilon} \right) \cdot b_j(\nabla u^{\varepsilon}) + \delta \sum_{j=1}^{d} \left( \partial x_j u^{\varepsilon} \right) \left( \partial x_j x_j u^{\varepsilon} \right) \right) \varphi(x, t) \, dx \, dt \quad (4.6)
$$

which is viewed as describing the action on tensor products $\varphi \otimes \eta'$.

We proceed to interpret (4.6) as an equation in $D'_{x, t, \xi}$. Let

$$
\chi^{\varepsilon} = \mathbbm{1}(u^{\varepsilon}, \xi),
$$

$$
H_j^{\varepsilon}(x, t) + \tilde{H}_j^{\varepsilon}(x, t) = \varepsilon b_j(\nabla u^{\varepsilon}) + \delta \partial x_j x_j u^{\varepsilon},
$$

$$
G^{\varepsilon}(x, t) + \tilde{G}^{\varepsilon}(x, t) = \varepsilon \sum_{j=1}^{d} \left( \partial x_j u^{\varepsilon} \right) \cdot b_j(\nabla u^{\varepsilon}) + \delta \sum_{j=1}^{d} \left( \partial x_j u^{\varepsilon} \right) \left( \partial x_j x_j u^{\varepsilon} \right).
$$

Note that $H_j^{\varepsilon}, \tilde{H}_j^{\varepsilon}, G^{\varepsilon}, \tilde{G}^{\varepsilon}$ are uniformly bounded in $L_{10c}^1(\mathbb{R}^d \times \mathbb{R}^+)$ from Lemma 3.1.

Let $\delta(u - \xi)$ be a Dirac delta function defined by $\langle \delta(u - \xi), \eta(\xi) \rangle = \eta(u)$. Then we wish to define $\delta(u^{\varepsilon} - \xi)G^{\varepsilon}(x, t)$ as a distribution in $D'_{x, t, \xi}$ by its action on tensor products

$$
\langle \delta(u^{\varepsilon} - \xi)G^{\varepsilon}, \varphi \otimes \eta' \rangle = \int_{x, t} G^{\varepsilon}(x, t) \varphi(x, t) \eta'(u^{\varepsilon}(x, t)) \, dx \, dt. \quad (4.7)
$$
This follows from the Schwartz kernel theorem (e.g., [5, Section 5.2]) as follows: define the linear map

\[ \mathcal{K} : C_c^\infty(\mathbb{R}) \to D'(\mathbb{R}^d \times \mathbb{R}^+) \] by \( \mathcal{K}\psi = G^\varepsilon(x,t)\psi(u^\varepsilon(x,t)) \).

If \( \psi_j \to 0 \) in \( C_c^\infty(\mathbb{R}) \) then \( \mathcal{K}\psi_j \to 0 \) in \( D'_{x,t,\xi} \). The kernel theorem implies that \( \delta(u^\varepsilon - \xi)G^\varepsilon \) is well defined as a distribution in \( D'_{x,t,\xi} \) and acts on tensor products via (4.7). Moreover,

\[ \langle \partial_\xi \delta(u^\varepsilon - \xi)G^\varepsilon, \varphi \otimes \eta' \rangle = - \int G^\varepsilon(x,t)\varphi(x,t)\eta''(u^\varepsilon(x,t)) \, dx \, dt. \quad (4.8) \]

Thus (4.6) is written as

\[
\partial_t \chi^\varepsilon + F'(\xi) \cdot \nabla \chi^\varepsilon \\
= \sum_{j=1}^d \partial_{x_j} \left( (H_j^\varepsilon(x,t) + \tilde{H}_j^\varepsilon(x,t)) \delta(u^\varepsilon - \xi) \right) + \partial_\xi \left( (G^\varepsilon(x,t) + \tilde{G}^\varepsilon(x,t)) \delta(u^\varepsilon - \xi) \right) \\
=: \sum_{j=1}^d \partial_{x_j} (\pi_j^\varepsilon + \tilde{\pi}_j^\varepsilon) + \partial_\xi (m^\varepsilon + k^\varepsilon) \quad \text{in} \ D'_{x,t,\xi}. \quad (4.9)
\]

**Step 2.** The next objective is to estimate the terms \( \pi_j^\varepsilon, \tilde{\pi}_j^\varepsilon, m^\varepsilon, \) and \( k^\varepsilon \).

Let \( \theta(x,t,\xi) \subset C_c^\infty(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \). We estimate first the terms \( \pi_j^\varepsilon \). Using the estimates in Lemma 3.1 and the hypothesis (H1), we see that

\[
\left| \left[H_j^\varepsilon \delta(u^\varepsilon - \xi), \theta(x,t,\xi) \right] \right| \\
= \left| \int_{x,t} \left( \epsilon b_j(\nabla u^\varepsilon) \theta(x,t,u^\varepsilon(x,t)) \right) \, dx \, dt \right| \leq \epsilon \int_{x,t} \left| b_j(\nabla u^\varepsilon) \right| \left| \theta(x,t,u^\varepsilon(x,t)) \right| \, dx \, dt \\
\leq \epsilon \left( \int_{x,t} \left| \nabla u^\varepsilon \right|^{r+1} \, dx \, dt \right)^{\frac{r}{r+1}} \cdot \left\| \theta(x,t,u^\varepsilon) \right\|_{L^{r+1}_{x,t}} \leq C \epsilon^{\frac{1}{r+1}} \left\| \theta(x,t,u^\varepsilon) \right\|^{\frac{r}{r+1}}_{L^{r+1}_{x,t}} \leq C \epsilon^{\frac{1}{r+1}} \left\| \theta \right\|^{r+1}_{L^{r+1}_{x,t}(W^{r+1}_{x,\xi})}.
\]

Here we used the following:

\[
\int_{x,t} \theta^{r+1}(x,t,u^\varepsilon) \, dx \, dt = \int_{x,t} \int_{-\infty}^{u^\varepsilon(x,t)} (r+1) \theta^r \theta_\xi \, d\xi \, dx \, dt \\
\leq (r+1) \int_{x,t} \left( \int_{-\infty}^{u^\varepsilon(x,t)} \theta^{r+1} \, d\xi \right)^{\frac{r}{r+1}} \cdot \left( \int_{-\infty}^{u^\varepsilon(x,t)} (\theta_\xi)^{r+1} \, d\xi \right)^{\frac{1}{r+1}} \, dx \, dt.
\]
\[ I \leq C \left[ \int_{x,t,\xi} \theta^{r+1} d\xi \, dx \, dt + \int_{x,t,\xi} |\partial_{\xi} \theta|^{r+1} d\xi \, dx \, dt \right] \]

\[ \leq C \|\theta\|_{L^{r+1}_{x,t}(W^{1,r+1}_{\xi})}. \]

This shows that \( \pi^\varepsilon_j \to 0 \) in \( L^{r^*}_{x,t}(W^{-1,q^*}_{\xi}) \), \( r^* = \frac{r+1}{r} \) as \( \varepsilon \to 0 \), or

\[ \pi^\varepsilon_j = \bar{g}^\varepsilon_j + \partial_{\xi} g^\varepsilon_j \quad \text{with} \quad \bar{g}^\varepsilon_j, g^\varepsilon_j \to 0 \quad \text{in} \quad L^q_{x,t,\xi}. \]

Next, we estimate the terms \( \bar{\pi}^\varepsilon_j \). Using the estimates in Lemma 3.1, we see that for \( \delta = O(\varepsilon^{(r+3)/(r+1)}) \)

\[ \left| \langle \bar{H}^\varepsilon \delta(u^\varepsilon - \xi), \theta(x, t, \xi) \rangle \right| \]

\[ \leq \delta \int_{x,t} \left| \partial_{x_j} x_j u^\varepsilon \right| \left| \theta(x, t, u^\varepsilon(x, t)) \right| \, dx \, dt \]

\[ \leq \delta \left( \int_{x,t} \left| \partial_{x_j} x_j u^\varepsilon \right|^2 \, dx \, dt \right)^{1/2} \left\| \theta(x, t, u^\varepsilon(x, t)) \right\|_{L^2_{x,t}} \]

\[ \leq C \delta \frac{r+2}{r+1} \left( \frac{2(r+2)}{r+1} \int_{x,t} \left| \partial_{x_j} x_j u^\varepsilon \right|^2 \, dx \, dt \right)^{1/2} \left\| \theta \right\|_{L^2_{x,t}(H^1_\xi)} \]

\[ \leq C \delta \frac{1}{r+1} \left\| \theta \right\|_{L^2_{x,t}(H^1_\xi)}. \]

This shows that \( \bar{\pi}^\varepsilon_j \to 0 \) in \( L^2_{x,t}(H^{-1}_{\xi}) \) as \( \varepsilon \to 0 \), or

\[ \bar{\pi}^\varepsilon_j = \bar{h}^\varepsilon_j + \partial_{\xi} h^\varepsilon_j \quad \text{with} \quad \bar{h}^\varepsilon_j, h^\varepsilon_j \to 0 \quad \text{in} \quad L^2_{x,t,\xi}. \]

Thirdly, consider the term \( m^\varepsilon = G^\varepsilon \delta(u^\varepsilon - \xi) \). Using the estimates (3.1) and (H2), we have

\[ \left| \langle m^\varepsilon, \theta \rangle \right| = \left| \langle \delta(u^\varepsilon - \xi) G^\varepsilon, \theta \rangle \right| \leq \varepsilon \int_{x,t} |\nabla u^\varepsilon| |b(\nabla u^\varepsilon)| \left| \theta(x, t, u^\varepsilon) \right| \, dx \, dt \]

\[ \leq \varepsilon \int_{x,t} |\nabla u^\varepsilon|^{r+1} \, dx \, dt \cdot \sup_{x,t,\xi} \left| \theta(x, t, \xi) \right| \leq C \sup_{x,t,\xi} \left| \theta(x, t, \xi) \right|. \]

So, \( m^\varepsilon \) lies in a bounded set of the space of bounded measures \( \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \) (the dual of \( C_0(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), the continuous functions that vanish at infinity). The Sobolev embedding theorem implies \( \mathcal{M} \) is embedded in \( W^{-1,p} \), \( 1 \leq p < \frac{d+2}{d+1} \), and thus \( m^\varepsilon \) is precompact in \( W^{-1,p}_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), for \( 1 \leq p < \frac{d+2}{d+1} \).
Finally, we estimate the term \( k^\varepsilon = \bar{G}^\varepsilon \delta (u^\varepsilon - \xi) \). Observe that

\[
|\langle k^\varepsilon, \theta \rangle| = |(\delta (u^\varepsilon - \xi) \bar{G}^\varepsilon, \theta)| \leq \delta \int_{x,t} |\nabla u^\varepsilon| |D^2 u^\varepsilon| |\theta(x, t, u^\varepsilon)| \, dx \, dt
\]

\[
\leq \delta \int_{x,t} \left[ \mu |D^2 u^\varepsilon|^2 \cdot |\theta(x, t, \xi)| + \frac{1}{\mu} |\nabla u^\varepsilon|^2 \cdot |\theta(x, t, \xi)| \right] \, dx \, dt
\]

\[
\leq \delta \left[ \int_{x,t} \mu |D^2 u^\varepsilon|^2 \cdot |\theta(x, t, \xi)| \, dx \, dt \\
+ \frac{1}{\mu} \left( \int_{x,t} |\theta|^{\frac{r+1}{r}} \, dx \, dt \right)^{\frac{r}{r+1}} \cdot \left( \int_{x,t} |\nabla u^\varepsilon|^{r+1} \, dx \, dt \right)^{\frac{2}{r+1}} \right]
\]

\[
\leq C\delta \left[ \mu \varepsilon^{-2} + \frac{1}{\mu} \varepsilon^{-2} \right] \sup_{x,t,\xi} |\theta(x, t, \xi)|.
\]

If we take \( \mu = \varepsilon \) and \( \delta = O(\varepsilon^{(r+3)/(r+1)}) \), we obtain

\[
|\langle k^\varepsilon, \theta \rangle| \leq C\|\theta\|_{C^0},
\]

and \( k^\varepsilon \) lies in a bounded set of the space of bounded measures.

On the other hand, if \( \delta = o(\varepsilon^{(r+3)/(r+1)}) \), we have

\[
k^\varepsilon = \bar{G}^\varepsilon \delta (u^\varepsilon - \xi) \rightharpoonup 0 \quad \text{weak* in } M_{x,t,\xi}.
\]

\[ (4.10) \]

**Step 3.** In summary, the function \( \chi^\varepsilon = 1(u^\varepsilon, \xi) \) satisfies the (approximate) transport equation

\[
\partial_t \chi^\varepsilon + F'(\xi) \cdot \nabla \chi^\varepsilon = \sum_{j=1}^{d} \partial_{x_j} \left( \bar{g}^\varepsilon_j + \partial_\xi g^\varepsilon_j + \bar{h}^\varepsilon_j + \partial_\xi h^\varepsilon_j \right) + \partial_\xi (m^\varepsilon + k^\varepsilon) \quad \text{in } D'_{x,t,\xi},
\]

\[ (4.11) \]

where \( \bar{g}^\varepsilon_i, g^\varepsilon_i, \bar{h}^\varepsilon_i, h^\varepsilon_i \to 0 \) in \( L^r_{loc}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), \( r^* = \frac{r+1}{r} \), while \( m^\varepsilon \) and \( k^\varepsilon \) are bounded in measures (\( k^\varepsilon \) is not necessarily positive) and precompact in \( W^{-1,p}_{loc}(\mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}) \), for \( 1 \leq p < \frac{d+2}{2} \). By the averaging lemma (Theorem 4.1),

\[
\int_{\xi} 1(u^\varepsilon, \xi) \psi(\xi) \, d\xi \quad \text{is precompact in } L^p_{loc}, \ 1 < p < q^*,
\]

for \( \psi(\xi) \in C^\infty_c(\mathbb{R}) \) and \( q^* = \min\left\{ \frac{d+2}{d+1}, \frac{r+1}{r} \right\} \).
Let $R$ be a large positive number and consider $\psi \in C_c^\infty(\mathbb{R})$ such that $\psi = 1$ on $(-R, R)$ and $0 \leq \psi \leq 1$. Then

$$
\left| u^\varepsilon - \int_{-R}^R \mathbb{1}(u^\varepsilon, \xi) \psi(\xi) \, d\xi \right| = \left| \int_{-R}^R \mathbb{1}(u^\varepsilon, \xi)(1 - \psi(\xi)) \, d\xi \right|
\leq \int_{-R}^R \mathbb{1}(u^\varepsilon, \xi) \, d\xi + \int_{-\infty}^{-R} \mathbb{1}(u^\varepsilon, \xi) \, d\xi
= (u^\varepsilon - R)^+ + (u^\varepsilon + R)^-.
$$

Moreover,

$$
\int (u^\varepsilon - R)^+ + (u^\varepsilon + R)^- \, dx \, dt \leq \int_{|u^\varepsilon| > R} |u^\varepsilon| \, dx \, dt \leq \frac{1}{R} \int_0^T \int |u^\varepsilon|^2 \, dx \, dt \leq \frac{C}{R}.
$$

We conclude that $\{u^\varepsilon\}$ is Cauchy in $L^1_{\text{loc}, x, t}$. Since $u^\varepsilon \in b L^\infty(L^2)$, it follows that (along subsequences) $u^\varepsilon \to u$ in $L^p_{\text{loc}}, p < 2$, and almost everywhere and that $u \in L^\infty(L^2)$.

Next, we pass to the limit $\varepsilon \to 0$ in (4.11). Let $\delta = O(\varepsilon (r^3 + 1))$. Then

$$
\chi^\varepsilon = \mathbb{1}(u^\varepsilon, \xi) \to \chi = \mathbb{1}(u, \xi) \quad \text{a.e. and in } L^p_{\text{loc}, x, t}(L^1_{\xi}), 1 \leq p < 2,
$$

$$
m^\varepsilon = G^\varepsilon \delta(u^\varepsilon - \xi) \rightharpoonup m \quad \text{weak* in } \mathcal{M}_{x, t, \xi},
$$

$$
k^\varepsilon = \bar{G}^\varepsilon \delta(u^\varepsilon - \xi) \rightharpoonup k \quad \text{weak* in } \mathcal{M}_{x, t, \xi},
$$

(4.12)

and $\chi$ satisfies

$$
\partial_t \chi + F'(\xi) \cdot \nabla \chi = \partial_\xi (m + k) \quad \text{in } D'_{x, t, \xi}.
$$

(4.13)

In this case the bounded measure $m + k$ may, in general, be nonpositive. By contrast, for $\delta = o(\varepsilon (r^3 + 1))$, by (4.10), the function $\chi = \mathbb{1}(u, \xi)$ satisfies the kinetic formulation of Lions, Perthame and Tadmor

$$
\partial_t \chi + F'(\xi) \cdot \nabla \chi = \partial_\xi m,
$$

with $m$ a positive, bounded measure, and thus $u$ is the unique entropy solution of (1.1) (see [16]).

**Remark 4.3.** The limit $m + k$ is a bounded measure but we do not know whether $m + k$ is a positive measure. The sign of $m + k$ depends on the sign of

$$
G^\varepsilon(x, t) + \bar{G}^\varepsilon(x, t) = \varepsilon \sum_{j=1}^d (\partial_{x_j} u^\varepsilon) \cdot b_j(\nabla u^\varepsilon) + \delta \sum_{j=1}^d (\partial_{x_j} u^\varepsilon)(\partial_{x_j} u^\varepsilon),
$$
and it could in principle happen that the terms $\varepsilon(\partial_x u^\varepsilon) \cdot b_j(\nabla u^\varepsilon)$ and $\delta(\partial_x u^\varepsilon)(\partial_x x^j u^\varepsilon)$ are comparable near shocks.

Note that for 1-dimensional scalar conservation law, when the flux is not genuinely nonlinear in the sense of Lax, diffusive–dispersive approximation can produce in the range $\delta \sim \varepsilon^2$ nonclassical shocks that dissipate the energy but do not dissipate all the convex entropies (see [4,11] for a survey of this subject).

References