UC Berkeley
UC Berkeley Previously Published Works

Title
Adjoint and Compensated Compactness Methods for Hamilton–Jacobi PDE

Permalink
https://escholarship.org/uc/item/0d94n6m3

Journal
Archive for Rational Mechanics and Analysis, 197(3)

ISSN
1432-0673

Author
Evans, Lawrence C.

Publication Date
2010-09-01

DOI
10.1007/s00205-010-0307-9

Peer reviewed

Adjoint and Compensated Compactness Methods for Hamilton–Jacobi PDE

Lawrence C. Evans

Communicated by C. Dafermos

Abstract

We investigate the vanishing viscosity limit for Hamilton–Jacobi PDE with nonconvex Hamiltonians, and present a new method to augment the standard viscosity solution approach. The main idea is to introduce a solution σ^ε of the adjoint of the formal linearization, and then to integrate by parts with respect to the density σ^ε. This procedure leads to a natural phase space kinetic formulation and also to a new compensated compactness technique.

1. Introduction

The Crandall–Lions theory of viscosity solutions for Hamilton–Jacobi partial differential equations provides fundamental existence, uniqueness and stability theorems, even for Hamiltonians H that are nonconvex. To date, however, there has been little progress in the understanding of the precise nature of the vanishing viscosity limiting process and of the gradient shock structure for viscosity solutions, except when H is convex (or concave). This paper introduces some new tools for these problems, most importantly a nonlinear adjoint technique.

1.1. Basic equations

Given a smooth Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ and smooth initial data $g : \mathbb{R}^n \to \mathbb{R}$, we consider the corresponding initial-value problem for the regularized Hamilton–Jacobi equation:

$$\begin{cases} u_t^\varepsilon + H(Du^\varepsilon) = \varepsilon \Delta u^\varepsilon & \text{in } \mathbb{R}^n \times (0, \infty) \\ u^\varepsilon = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{1.1}$$

Here we write $Du^\varepsilon = D_x u^\varepsilon$ for the gradient in the variable $x = (x_1, \ldots, x_n)$.

Given next a time $t_1 > 0$ and a Radon probability measure α on $\mathbb{R}^n$, we introduce this terminal value problem, the adjoint of the linearization of (1.1):

$$\begin{cases} -\sigma_t^\varepsilon - \operatorname{div}(\sigma^\varepsilon DH(Du^\varepsilon)) = \varepsilon \Delta \sigma^\varepsilon & \text{in } \mathbb{R}^n \times [0, t_1) \\ \sigma^\varepsilon = \alpha & \text{on } \mathbb{R}^n \times \{t = t_1\}. \end{cases} \tag{1.2}$$

Our plan is to use the solutions $\sigma^\varepsilon = \sigma^{\varepsilon,\alpha}$ of the adjoint problem (1.2), for various choices of the terminal conditions α, to extract information about the limiting behavior of u^ε as $\varepsilon \to 0$.

Our introducing the PDE (1.2) is directly inspired by weak KAM theory (see [9]), but there is in fact a rich history of similar ideas arising in the optimal control of ODE (in the Pontryagin maximum principle) and of PDE. LIONS's book [15] provides many PDE examples. In these contexts the solutions of certain differential equations involving adjoints of linearizations can be interpreted as Lagrange multipliers. Our approach, however, is more directly motivated by a desire to go beyond the usual maximum principle, sup-norm techniques embodied in the theory of viscosity solutions: see for instance BARDI-CAPUZZO DOLCETTA [1]. We want, rather, to develop integration by parts methods, which turn out to be available, provided we integrate against the density σ^ε.

Generalizations of the methods in this paper to static Hamilton–Jacobi type PDE on bounded domains will appear in TRAN [21]. Let me call attention also to the interesting paper [13] by Glimm et al. which is one of the few in the literature devoted to understanding the gradient shock structure of solutions for nonconvex H in many space dimensions.

Notation. We will denote a typical point of $\mathbb{R}^{n+1}$ as $q = (p, p_{n+1})$ for $p \in \mathbb{R}^n$, $p = (p_1, \ldots, p_n)$, and $p_{n+1} \in \mathbb{R}$. We will also write $\nabla w = (Dw, w_t) = (w_{x_1}, \ldots, w_{x_n}, w_t)$ for the full gradient of a function $w = w(x, t)$. More generally, we use "D" to denote the gradient of a function of the n variables x, and "∇" to denote the gradient of a function of the $n + 1$ variables (x, t). We write $H = H(p)$ and $g = g(x)$.

1.2. Standard estimates, convergence to viscosity solution

For simplicity, we assume throughout that

$$g : \mathbb{R}^n \to \mathbb{R} \quad \text{is a smooth function with compact support.}$$

Let us first record the elementary estimates that for each time $t_1 > 0$:

$$\sup_{\mathbb{R}^n \times [0, t_1]} |u^\varepsilon|, |Du^\varepsilon|, |u_t^\varepsilon| \leq C \quad \text{and} \quad \sigma^\varepsilon \geqq 0, \int_{\mathbb{R}^n} \sigma^\varepsilon \, dx = 1 \ (0 \leqq t < t_1). \tag{1.3}$$

(Our observing that $\sigma^\varepsilon \geq 0$ is the only use of the maximum principle in this paper.) Furthermore, $u^\varepsilon \to u$ locally uniformly, where u is the unique viscosity solution

of the Hamilton–Jacobi equation

$$\begin{cases} u_t + H(Du) = 0 & \text{in } \mathbb{R}^n \times (0, \infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{1.4}$$

Since u is Lipschitz continuous, u is differentiable almost everywhere with respect to $(n + 1)$-dimensional Lebesgue measure; and in addition

$$u_t(x, t) + H(Du(x, t)) = 0 \tag{1.5}$$

at each point (x, t) where u is differentiable.

We note that the viscosity solution approach, although amazingly successful, leaves open many questions, both as to the structure of possible singularities of u and also as to the fine details of the convergence $u^\varepsilon \to u$. For instance, although the PDE (1.5) holds almost everywhere, we do not know in general whether $Du^\varepsilon \to Du$ almost everywhere. The techniques of this paper provide some new understanding about these issues.

2. A second derivative estimate

The following second derivative estimate will be very useful later:

Theorem 2.1. *There exists a constant C, independent of $\varepsilon > 0$, such that*

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \left(|D^2 u^\varepsilon|^2 + |Du_t^\varepsilon|^2 \right) \sigma^\varepsilon \, dx \, dt \leqq C. \tag{2.1}$$

We are using the notation $|D^2 u^\varepsilon|^2 := u_{x_i x_j}^\varepsilon u_{x_i x_j}^\varepsilon$ and $|Du_t^\varepsilon|^2 := u_{x_i t}^\varepsilon u_{x_i t}^\varepsilon$, the implicit summation for $i, j = 1, \ldots, n$. In the estimate (2.1), $\sigma^\varepsilon = \sigma^{\varepsilon, \alpha}$ is the solution of (1.2) corresponding to any given Borel probability measure α on $\mathbb{R}^n$.

Proof. Let $w^\varepsilon = \frac{1}{2}(|Du^\varepsilon|^2 + |u_t^\varepsilon|^2)$. Then

$$w_t^\varepsilon + DH(Du^\varepsilon) \cdot Dw^\varepsilon = \varepsilon \Delta w^\varepsilon - \varepsilon \left(|D^2 u^\varepsilon|^2 + |Du_t^\varepsilon|^2 \right).$$

We multiply by σ^ε and integrate by parts, using (1.2) to deduce

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \left(|D^2 u^\varepsilon|^2 + |Du_t^\varepsilon|^2 \right) \sigma^\varepsilon \, dx \, dt + \int_{\mathbb{R}^n} w^\varepsilon(x, t_1) \, d\alpha$$
$$= \int_{\mathbb{R}^n} w^\varepsilon(x, 0) \sigma^\varepsilon(x, 0) \, dx.$$

Since $w^\varepsilon(x, 0)$ is bounded independently of ε, this and (1.3) imply (2.1). $\quad\square$

To illustrate the usefulness of (2.1), we present a quick new proof of an estimate for the rate of convergence of u^ε to u.

Theorem 2.2. *We have the estimate*

$$\sup_{\mathbb{R}^n \times [0, t_1]} \left| \frac{\partial u^\varepsilon}{\partial \varepsilon} \right| \leqq C \left(\frac{t_1}{\varepsilon} \right)^{\frac{1}{2}}. \tag{2.2}$$

Consequently,

$$\sup_{\mathbb{R}^n \times [0, t_1]} |u^\varepsilon - u| \leqq C (t_1 \varepsilon)^{\frac{1}{2}}. \tag{2.3}$$

Proof. 1. According to standard parabolic estimates, the function u^ε is smooth in the parameter ε away from $\varepsilon = 0$. We therefore can differentiate (1.1) with respect to ε:

$$u_{\varepsilon t}^\varepsilon + DH(Du^\varepsilon) \cdot Du_\varepsilon^\varepsilon = \varepsilon \Delta u_\varepsilon^\varepsilon + \Delta u^\varepsilon,$$

the subscript ε denoting the partial derivative.

Select any time $t_1 > 0$ and any point $x_1 \in \mathbb{R}^n$, and then let σ^ε solve the adjoint PDE (1.2), with the terminal condition that $\alpha = \delta_{x_1}$. We multiply the foregoing PDE by σ^ε and integrate, to discover that

$$u_\varepsilon^\varepsilon(x_1, t_1) = \int_0^{t_1} \int_{\mathbb{R}^n} \Delta u^\varepsilon \sigma^\varepsilon \, \mathrm{d}x \, \mathrm{d}t.$$

There is no term corresponding to time $t = 0$, since $u_\varepsilon^\varepsilon = g_\varepsilon \equiv 0$ there. Consequently our inequalities (2.1) and (1.3) and Cauchy–Schwarz imply

$$|u_\varepsilon^\varepsilon(x_1, t_1)| \leqq \int_0^{t_1} \int_{\mathbb{R}^n} |\Delta u^\varepsilon| \sigma^\varepsilon \, \mathrm{d}x \, \mathrm{d}t \leqq C \left(\frac{t_1}{\varepsilon} \right)^{\frac{1}{2}}.$$

2. The foregoing estimate is useful since $\varepsilon^{-\frac{1}{2}}$ is integrable near $\varepsilon = 0$. Indeed, we deduce for all $x \in \mathbb{R}^n$, $t > 0$ that

$$|u^{\varepsilon_1}(x, t) - u^{\varepsilon_2}(x, t)| \leqq C (t \varepsilon_1)^{\frac{1}{2}}$$

provided $0 < \varepsilon_2 < \varepsilon_1$ and $0 \leqq t \leqq t_1$. Let $\varepsilon_2 \to 0$ to derive (2.3). □

Estimate (2.3), which shows that the full sequence $\{u^\varepsilon\}_{0 \leqq \varepsilon \leqq 1}$ converges uniformly, is originally due to Fleming, whose clever papers [11, 12] employ stochastic game theory. Crandall and Lions used viscosity solution techniques to provide a simpler proof in [6]. (We can also derive an estimate similar to (2.2) by applying the conventional maximum principle to the function $v^\varepsilon := \varepsilon^{\frac{1}{2}} u_\varepsilon^\varepsilon + |Du^\varepsilon|^2 - \lambda t$.)

3. Measures on phase space

In this section we employ the densities σ^ε to introduce a natural phase space "kinetic" formulation for the Hamilton–Jacobi equation (1.4). This undertaking is inspired by the kinetic formulations for conservation laws developed by Lions et al. [17, 18], although the technical differences are many. See also Perthame's book [19], and refer to Section 7 for some discussion contrasting our approach with theirs.

3.1. Existence of the measures μ, σ

First, we introduce a measure μ on "phase space" that records the dynamics in both the "physical space" variables x and t and the "momentum" variables q. The measure σ will then be the projection of μ onto physical space.

Approximating a point mass by the average over a cylinder. Hereafter we focus attention upon some given point $(x_1, t_1) \in \mathbb{R}^n \times (0, \infty)$. We will be particularly interested in the adjoint PDE (1.2) with $\alpha = \delta_{x_1}$ at time t_1, and will sometimes (as in Section 2) use this as the terminal condition.

However, since we also want to study the limiting behavior of ∇u^ε near this point, we will also need to approximate the point mass at (x_1, t_1) by averages over the space-time cylinders

$$C(x_1, t_1, r) := B(x_1, r) \times [t_1, t_1 + r] \tag{3.1}$$

for small $r > 0$. So fix $r > 0$ and let $\sigma^\varepsilon_{s,r}$ denote the solution of (1.2) with terminal data, the function

$$\alpha_r := \frac{1}{|B(x_1, r)|} \chi_{B(x_1, r)}. \tag{3.2}$$

at time s, where $t_1 \leqq s \leqq t_1 + r$. Next average with respect to s:

$$\sigma^\varepsilon_r(x, t) := \frac{1}{r} \int_{t_1}^{t_1+r} \sigma^\varepsilon_{s,r}(x, t) \, ds \tag{3.3}$$

for $(x, t) \in \mathbb{R}^n \times [0, t_1)$, and observe that σ^ε_r solves the adjoint PDE (1.2). Set

$$r_j := \frac{1}{j}. \tag{3.4}$$

Theorem 3.1.   (i) *There exists a subsequence $\varepsilon_m \to 0$ and for each $j = 1, \ldots$ a nonnegative Radon measure μ_{r_j} on $\mathbb{R}^n \times [0, t_1] \times \mathbb{R}^{n+1}$ satisfying*

$$\lim_{\varepsilon_m \to \infty} \int_0^{t_1} \int_{\mathbb{R}^n} \Phi(x, t, \nabla u^{\varepsilon_m}) \sigma^{\varepsilon_m}_{r_j} \, dx \, dt = \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\mu_{r_j} \tag{3.5}$$

for $j = 1, \ldots$ and all continuous, bounded functions Φ.
   (ii) *There exists a subsequence $\{r_m\}$ of $\{r_j\}$ and a nonnegative measure $\mu = \mu_{x_1}$ on $\mathbb{R}^n \times [0, t_1] \times \mathbb{R}^{n+1}$ such that*

$$\lim_{r_m \to 0} \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\mu_{r_m} = \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\mu \tag{3.6}$$

for all continuous, bounded functions Φ.

Proof. Fix r_1. Let $\{\Phi_l\}_{l=1}^{\infty}$ be a countable dense subset of $C([0, t_1] \times \mathbb{R}^n \times \mathbb{R}^{n+1})$. Using a standard diagonal argument, we extract a subsequence $\varepsilon_k \to 0$ such that the limit

$$\phi_l := \lim_{\varepsilon_k \to 0} \int_0^{t_1} \int_{\mathbb{R}} \Phi_l(x, t, \nabla u^{\varepsilon_k}) \sigma_{r_1}^{\varepsilon_k} \, dx \, dt$$

exists for $l = 1, \ldots$. The mapping $\Phi_l \mapsto \phi_l$ is linear, and we have the estimate

$$|\phi_l| \leqq C \sup |\Phi_l| \quad (l = 1, \ldots).$$

The Riesz Representation Theorem provides the existence of a Radon measure μ_{r_1} satisfying (3.6) for the given radius r_1. Continuing, and again using the diagonal argument, we extract a further subsequence $\{\varepsilon_m\}$ for which (3.6) holds for all r_j.

The second assertion is a consequence of the weak compactness of measures with uniformly bounded mass. □

Notation. We hereafter always take

$$\mu := \mu_{x_1}, \quad \sigma := \mathrm{proj}_{x,t} \mu$$

unless otherwise stated, where $\mathrm{proj}_{x,t} \mu$ means the projection of μ onto $\mathbb{R}^n \times [0, t_1]$.

We now show that we can "slice" the measure σ at each time.

Lemma 3.2. *For all bounded continuous $\phi = \phi(x, t)$ we can write*

$$\int_0^{t_1} \int_{\mathbb{R}^n} \phi(x, t) \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \phi(x, t) \, d\gamma_t \, dt \tag{3.7}$$

where $\gamma_t(\cdot)$ is a Radon probability measure on $\mathbb{R}^n$.

Proof. According to the basic estimates (1.3), the projection of σ onto the time interval $[0, t_1]$ is one-dimensional Lebesgue measure $\mathcal{L}^1$. Then, for example, Theorem 10 in Chapter 1 of [8] implies the decomposition (3.7). □

In view of the Lemma, we write

$$d\sigma = d\gamma_t \, dt. \tag{3.8}$$

Similarly, for each $r = r_j$ we have the decomposition

$$d\sigma_r = d\gamma_{r,t} \, dt \tag{3.9}$$

for $\sigma_r = \mathrm{proj}_{x,t} \mu_r$.

Next we show that the measure γ_t is defined for every time $0 \leqq t \leqq t_1$:

Theorem 3.3.   (i) *The mapping $t \mapsto \gamma_t$ is continuous into the space $\mathcal{M}$ of probability measures on $\mathbb{R}^n$, taken with the weak topology. In particular, γ_t is defined for each time $0 \leqq t \leqq t_1$. A similar statement holds for $\gamma_{r,t}$ for each $r = r_j$.*
  (ii) *For each time $0 \leqq t \leqq t_1$ and $r = r_j$, we have*

$$\sigma_r^{\varepsilon}(\cdot, t) \rightharpoonup \gamma_{r,t} \quad \text{weakly in } \mathcal{M} \tag{3.10}$$

as $\varepsilon = \varepsilon_m \to 0$.

(iii) *Furthermore, for each time $0 \leq t \leq t_1$*

$$\gamma_{r,t} \rightharpoonup \gamma_t \quad weakly \ in \ \mathcal{M} \tag{3.11}$$

as $r = r_m \to 0$.

Proof. 1. The basic estimates (1.3) show for each smooth $\phi = \phi(x)$ that

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^n} \phi \sigma_r^\varepsilon \, \mathrm{d}x \right| \leq C \|D\phi\|_{L^\infty} + \varepsilon C \|D^2\phi\|_{L^\infty}. \tag{3.12}$$

Fix $r = r_j$. Selecting, then, a smooth function ϕ, a time $0 \leq t < t_1$ and a small $h > 0$, we have

$$\left| \int_{\mathbb{R}^n} \phi \sigma_r^\varepsilon (x, t) \, \mathrm{d}x - \frac{1}{h} \int_t^{t+h} \int_{\mathbb{R}^n} \phi \sigma_r^\varepsilon (x, s) \, \mathrm{d}x \, \mathrm{d}s \right| \leq Ch,$$

for $C = C(\phi)$. Consequently the decomposition (3.9) gives

$$\limsup_{\varepsilon_m \to 0} \left| \int_{\mathbb{R}^n} \phi \sigma_r^{\varepsilon_m} (x, t) \, \mathrm{d}x - \frac{1}{h} \int_t^{t+h} \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_{r,s} \, \mathrm{d}s \right| \leq Ch \tag{3.13}$$

for all $h > 0$. Therefore

$$\int_{\mathbb{R}^n} \phi \sigma_r^{\varepsilon_m} (x, t) \, \mathrm{d}x \to \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_{r,t} \tag{3.14}$$

for $\mathcal{L}^1$-almost everywhere time t.

This limit holds for a given smooth function ϕ. We now select a countable collection of smooth functions $\{\phi_l\}_{l=1}^\infty$, dense in the space of bounded continuous functions, so that (3.14) holds for each $\phi = \phi_l$ and almost every time. Since the functions $\sigma_r^{\varepsilon_m} (\cdot, t)$ are nonnegative and bounded in L^1, owing to (1.3), the limit (3.14) then holds as well for each continuous ϕ. Thus $\sigma_r^{\varepsilon_m} (\cdot, t)$ converges weakly in $\mathcal{M}$ to $\gamma_{r,t}$ for almost every time $0 \leq t \leq t_1$.

Furthermore, the estimate (3.12) implies that the mappings $t \mapsto \sigma_r^\varepsilon$ are uniformly Lipschitz continuous into $(C^2)^*$, the dual space of C^2. Consequently we can, if necessary, redefine $\gamma_{r,t}$ on a subset of the time interval $[0, t_1]$ of Lebesgue measure zero, to ensure that $t \mapsto \gamma_{r,t}$ is Lipschitz continuous into $(C^2)^*$, and consequently continuous into $\mathcal{M}$, taken with the weak topology. Then sending $h \to 0$ in (3.13) shows that $\sigma_r^{\varepsilon_m} (\cdot, t)$ converges to $\gamma_{r,t}$ in $(C^2)^*$, and so, also, $\mathcal{M}$, for each time $0 \leq t \leq t_1$. This proves assertion (ii) and the second part of (i).

2. Now (3.13) and (3.14) imply

$$\left| \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_{r,t} - \frac{1}{h} \int_t^{t+h} \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_{r,s} \, \mathrm{d}s \right| \leq Ch.$$

Therefore (3.8) gives

$$\limsup_{r_m \to 0} \left| \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_{r_m,t} - \frac{1}{h} \int_t^{t+h} \int_{\mathbb{R}^n} \phi \, \mathrm{d}\gamma_s \, \mathrm{d}s \right| \leq Ch$$

Since the mappings $t \mapsto \gamma_{r_m,t}$ are uniformly Lipschitz continuous into $(C^2)^*$, we can argue as above to deduce assertions (i) and (iii). $\quad \square$

Next we slice μ:

Theorem 3.4. *For all continuous* $\Phi = \Phi(x, t, q)$ *we can write*

$$\int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\mu = \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\nu_{x,t}(q) \, d\sigma \quad (3.15)$$

where $\nu_{x,t}(\cdot)$ *is a Radon probability measure on* $\mathbb{R}^{n+1}$.

This follows again from Theorem 10 in Chapter 1 of [8], the proof of which invokes the theory of derivates from Section 2.9 in FEDERER [10]. The probability measures $\nu_{x,t}$ are analogs of the Young measures introduced by L. Tartar in compensated compactness theory, with the difference that now the background measure is σ and not Lebesgue measure.

We will later use the notation

$$\overline{\Phi}(x, t) := \int_{\mathbb{R}^{n+1}} \Phi(q, x, t) \, d\nu_{x,t},$$

to denote the average with respect to $\nu_{x,t}$. Note also that in view of (1.3) the support of each $\nu_{x,t}$ lies within some fixed ball in $\mathbb{R}^{n+1}$.

3.2. The support of the measure μ and σ

In this and the next section we begin our investigation as to the structure of the measure $\mu = \mu_{x_1}$. We start by examining the support of μ.
Notation. Let us hereafter write

$$\Sigma := \{q \in \mathbb{R}^{n+1} \mid p_{n+1} + H(p) = 0\}$$

for the hypersurface in $\mathbb{R}^{n+1}$ that is the graph of $-H$.

Theorem 3.5.   (i) *For σ-almost all points* (x, t), *we have*

$$\mathrm{spt} \, \nu_{x,t} \subseteq \Sigma. \quad (3.16)$$

(ii) *Furthermore,*

$$\mathrm{spt} \, \sigma \subseteq \{|x - x_1| \leqq M(t_1 - t)\}. \quad (3.17)$$

for times $0 \leqq t \leqq t_1$, *where* $M := \sup\{|DH(Du^\varepsilon)|\}$.
(iii) *In particular,*

$$\lim_{t \to t_1-} \gamma_t = \delta_{x_1} \quad \text{as measures on } \mathbb{R}^n. \quad (3.18)$$

The second assertion is that the support of σ lies within a cone with vertex (x_1, t_1): this illustrates finite speed of propagation.

Proof. 1. The estimate (2.1) holds for each $\sigma^\varepsilon_{s,r}$ and hence also for the average σ^ε_r defined by (3.3). Therefore (1.1) implies

$$\int_0^{t_1} \int_{\mathbb{R}^n} (u^\varepsilon_t + H(Du^\varepsilon))^2 \sigma^\varepsilon_r \, dx \, dt = \varepsilon^2 \int_0^{t_1} \int_{\mathbb{R}^n} (\Delta u^\varepsilon)^2 \sigma^\varepsilon_r \, dx \, dt \leqq C\varepsilon.$$

Select $r = r_j$, and let $\varepsilon = \varepsilon_m \to 0$. Then let $r_m \to 0$:

$$\int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} (p_{n+1} + H(p))^2 \, d\mu = 0.$$

This gives (3.16).

2. Let $\phi : \mathbb{R} \to \mathbb{R}$ be smooth and satisfy $\phi(z) = 0$ for $z \leqq 0$, $\phi(z) > 0$ for $z > 0$, $\phi' \geqq 0$. Put

$$\zeta(x, t) := \phi(|x - x_1| + M(t - t_1) - \delta),$$

for a small number $\delta > 0$. Then

$$\begin{aligned}
\frac{d}{dt} \int_{\mathbb{R}^n} \zeta \sigma^\varepsilon_r \, dx &= \int_{\mathbb{R}^n} (\zeta_t + D\zeta \cdot DH(Du^\varepsilon) - \varepsilon \Delta \zeta) \sigma^\varepsilon_r \, dx \\
&= \int_{\mathbb{R}^n} (M + \frac{x - x_1}{|x - x_1|} \cdot DH(Du^\varepsilon)) \phi' \sigma^\varepsilon_r \, dx + O(\varepsilon) \\
&\geqq O(\varepsilon).
\end{aligned}$$

Integrating and let $r = r_j$. Sending $\varepsilon = \varepsilon_m \to 0$ and then $r = r_m \to 0$, we deduce that for each time $0 \leqq t \leqq t_1$

$$\int_{\mathbb{R}^n} \phi(|x - x_1| + M(t - t_1) - \delta) \, d\gamma_t = 0.$$

That this holds for each $\delta > 0$ implies (3.17). The assertion (3.18) follows, since each γ_t is a probability measure. $\square$

3.3. Evolution equations for μ and σ

Next we derive the dynamics for the measure μ and for its projection σ.

Theorem 3.6. *There exists a symmetric, nonnegative definite $(n + 1) \times (n + 1)$ matrix of Borel measures $M = ((m^{kl}))$ on $\mathbb{R}^n \times [0, t_1] \times \mathbb{R}^{n+1}$ such that*

$$\mu_t + DH(p) \cdot D_x \mu = -(m^{kl})_{p_k p_l}. \tag{3.19}$$

The phase space transport equation (3.19) is our *kinetic formulation* of the Hamilton–Jacobi PDE (1.4). We call M the *matrix of dissipation measures*. Note that M depends upon the point x_1 and the terminal time t_1. The implicit summation in (3.19) is for $k, l = 1, \ldots, n + 1$.

Proof. 1. First we observe that on account of the estimate (2.1), the expressions

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Psi(x, t, \nabla u^\varepsilon) u^\varepsilon_{x_i x_k} u^\varepsilon_{x_j x_k} \sigma^\varepsilon_r \, dx \, dt$$

and

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Psi(x, t, \nabla u^\varepsilon) u^\varepsilon_{t x_k} u^\varepsilon_{t x_k} \sigma^\varepsilon_r \, dx \, dt$$

for each $i, j = 1, \dots, n$, are bounded if Ψ is bounded. As in the proof of Theorem 3.1, we may assume, upon passing if necessary to a further subsequences of $\varepsilon = \varepsilon_m \to 0, r = r_m \to 0$ and reindexing, that

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Psi(x, t, \nabla u^\varepsilon) u^\varepsilon_{x_i x_k} u^\varepsilon_{x_i x_l} \sigma^\varepsilon_r \, dx \, dt \to \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Psi \, dm^{kl} \quad (3.20)$$

and

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Psi(x, t, \nabla u^\varepsilon) u^\varepsilon_{t x_i} u^\varepsilon_{t x_i} \sigma^\varepsilon_r \, dx \, dt \to \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Psi \, dm^{n+1,n+1}$$

$$(3.21)$$

for all bounded, continuous Ψ and for appropriate measures m^{kl} $(k, l = 1, \dots, n)$ and $m^{n+1,n+1}$. Similarly, for $k = 1, \dots, n$ we have

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Psi(x, t, \nabla u^\varepsilon) u^\varepsilon_{x_k x_i} u^\varepsilon_{t x_i} \sigma^\varepsilon_r \, dx \, dt \to \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Psi \, dm^{k,n+1}$$

$$(3.22)$$

The matrix of measures $M = ((m^{kl}))$ is nonnegative and symmetric, owing to the form of the left-hand sides of the previous three equalities.

2. Now let $\Phi = \Phi(x, t, q)$ be smooth and have compact support in $\mathbb{R}^n \times (0, t_1) \times \mathbb{R}^{n+1}$. Put

$$w^\varepsilon(x, t) := \Phi(x, t, \nabla u^\varepsilon).$$

Then

$$\begin{aligned}
w^\varepsilon_t + DH(Du^\varepsilon) \cdot Dw^\varepsilon - \varepsilon \Delta w^\varepsilon = {}& \Phi_t + DH(Du^\varepsilon) \cdot D_x \Phi \\
& - \varepsilon(\Phi_{p_k p_l} u^\varepsilon_{x_k x_i} u^\varepsilon_{x_l x_i} + 2\Phi_{p_k p_{n+1}} u^\varepsilon_{x_k x_i} u^\varepsilon_{t x_i} \\
& + \Phi_{p_{n+1} p_{n+1}} u^\varepsilon_{t x_i} u^\varepsilon_{t x_i}) + E^\varepsilon. \quad (3.23)
\end{aligned}$$

where

$$E^\varepsilon := \varepsilon(\Delta_x \Phi + 2\Phi_{x_i p_k} u^\varepsilon_{x_i x_k} + 2\Phi_{x_i p_{n+1}} u^\varepsilon_{x_i t}). \quad (3.24)$$

In these formulas Φ is evaluated at $(x, t, \nabla u^\varepsilon)$.

3. Next we multiply (3.23) by σ_r^ε and integrate by parts, using (1.2) to simplify:

$$\int_0^{t_1} \int_{\mathbb{R}^n} \left(-\Phi_t - D_x\Phi \cdot DH(Du^\varepsilon)\right) \sigma_r^\varepsilon \, dx \, dt$$

$$+\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \left(\Phi_{p_k p_l} u^\varepsilon_{x_k x_i} u^\varepsilon_{x_l x_i} + 2\Phi_{p_k p_{n+1}} u^\varepsilon_{x_k x_i} u^\varepsilon_{t x_k}\right.$$

$$\left. +\Phi_{p_{n+1} p_{n+1}} u^\varepsilon_{t x_k} u^\varepsilon_{t x_k}\right) \sigma_r^\varepsilon \, dx \, dt = \int_0^{t_1} \int_{\mathbb{R}^n} E^\varepsilon \sigma_r^\varepsilon \, dx \, dt.$$

We estimate that

$$\left| \int_0^{t_1} \int_{\mathbb{R}^n} E^\varepsilon \sigma_r^\varepsilon \, dx \, dt \right| \leqq C\varepsilon + C\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \left(|D^2 u^\varepsilon| + |Du_t^\varepsilon|\right) \sigma_r^\varepsilon \, dx \, dt$$

$$\leqq C\varepsilon + C\varepsilon^{\frac{1}{2}}.$$

according to (2.1). Letting $\varepsilon = \varepsilon_m \to 0$, $r = r_m \to 0$ and recalling (3.20)–(3.22), we discover

$$\int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \left(-\Phi_t - D_x\Phi \cdot DH(p)\right) d\mu + \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi_{p_k p_l} \, dm^{kl} = 0, \tag{3.25}$$

where now Φ is evaluated at (x, t, q) and the implicit summation is for $k, l = 1, \ldots, n+1$. As Φ is arbitrary, we see that μ is a weak solution of the evolution equation (3.19). □

An alternate way to end the proof would be to observe that the left-hand side of (3.19) is a distribution with bounded support that is nonpositive when acting on functions convex in q. According to Lions [16], then, it has the form given on the right-hand side of (3.19). M. Christ [4] has shown me a different, elegant proof of this.

Deriving dynamics for σ is much simpler:

Theorem 3.7. *We have*

$$\sigma_t + \text{div}(\sigma \overline{DH}) = 0 \quad \text{in } \mathbb{R}^n \times [0, t_1). \tag{3.26}$$

Proof. Let $\zeta = \zeta(x, t)$ be smooth, with compact support. According to (1.2),

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta_t \sigma_r^\varepsilon + \sigma_r^\varepsilon DH(Du^\varepsilon) \cdot D\zeta \, dx \, dt = \varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} \Delta\zeta \sigma_r^\varepsilon \, dx \, dt.$$

Let $\varepsilon = \varepsilon_m \to 0$, $r = r_m \to 0$:

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta_t d\sigma + \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} DH(p) \cdot D\zeta \, d\mu = 0.$$

This implies (3.26). □

3.4. Limit from below for the measure μ

It is not at all clear where the mapping $(x, t) \mapsto \nu_{x,t}$ is continuous. But we can show that a certain limit from below exists at the terminal point (x_1, t_1).

Theorem 3.8. (i) *There exists a Radon probability measure* $\nu^- = \nu^-_{x_1, t_1}$ *on* $\mathbb{R}^{n+1}$ *such that the limit*

$$\lim_{t \to t_1^-} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi(x, t, q) \, d\nu_{x,t} \, d\gamma_t = \int_{\mathbb{R}^{n+1}} \Phi(x_1, t_1, q) \, d\nu^- \quad (3.27)$$

exists for all continuous Φ.
(ii) *Furthermore,*

$$\operatorname{spt} \nu^- \subseteq \Sigma. \quad (3.28)$$

Strictly speaking, the expression on the left side of (3.27) is defined only for $\mathcal{L}^1$-almost everywhere time t. So we really mean that the limit exists as $t \to t_1^-$, possibly omitting a set of measure zero.

Proof. 1. Suppose first that $\Phi = \Phi(q)$ is smooth and convex in the variable q. Let $\zeta = \zeta(t)$ be smooth and nonnegative, with compact support in $(0, t_1)$. Then according to (3.25),

$$\int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \zeta' \Phi = \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \zeta \Phi_{p_k p_l} \, dm^{kl} \geqq 0.$$

This is true for all ζ as above, and therefore for almost all $0 < s < r < t_1$

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi \, d\nu_{r,x} \, d\gamma_r \leqq \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi \, d\nu_{s,x} \, d\gamma_s.$$

Thus $t \mapsto \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi \, d\nu_{t,x} d\gamma_t$ is nonincreasing, for times t restricted to a subset of $[0, t_1)$ of full measure, and consequently has a limit as $t \to t_1^-$. If $\Phi = \Phi(q)$ is not convex in q, we write $\Phi = \Phi_1 - \Phi_2$, where Φ_1 and Φ_2 are convex, to deduce that the limit on the left of (3.27) exists in this case also.
2. By approximation, the limit on the left of (3.27) exists for all continuous $\Phi = \Phi(x, t, q)$ and is a bounded linear functional of Φ. Therefore the limit is represented by integration against a probability measure. In view of (3.18), this measure is a Dirac mass at (x_1, t_1) times a probability measure ν^- on $\mathbb{R}^{n+1}$.
3. The assertion (3.28) follows from (3.27) and (3.16). □

3.5. A formula for ∇u

Our next goal is deriving a representation formula for the gradient $\nabla u(x_1, t_1)$, provided it exists, in terms of the measure ν^-. We have introduced the approximations introduced by (3.2) and (3.3) to make this possible.

Theorem 3.9. *Suppose that u is differentiable at (x_1, t_1) and that (x_1, t_1) is a Lebesgue point for ∇u in $\mathbb{R}^n \times (0, \infty)$. Then*

$$\nabla u(x_1, t_1) = \int_{\mathbb{R}^{n+1}} q \, dv^-. \tag{3.29}$$

So $\nabla u(x_1, t_1)$ is the center of mass of the measure v^- in $\mathbb{R}^{n+1}$.

Proof. 1. We first claim that

$$Du(x_1, t_1) = \int_{\mathbb{R}^n} Dg(x) \, d\gamma_0 \tag{3.30}$$

and

$$u_t(x_1, t_1) = - \int_{\mathbb{R}^n} H(Dg(x)) \, d\gamma_0. \tag{3.31}$$

where we recall that $d\sigma = d\gamma_t \, dt$. In other words, $\nabla u(x_1, t_1)$ is the average of the gradient of the initial data with respect to the probability measure γ_0. To prove (3.30), put $w^\varepsilon := u^\varepsilon_{x_k}$ and note that

$$w^\varepsilon_t + DH(Du^\varepsilon) \cdot Dw^\varepsilon = \varepsilon \Delta w^\varepsilon.$$

Multiply by $\sigma^\varepsilon_{s,r}$ and integrate by parts, to find

$$\int_{\mathbb{R}^n} u^\varepsilon_{x_k}(x, s)\alpha_r \, dx = \fint_{B(x_1, r)} u^\varepsilon_{x_k}(x, s) \, dx = \int_{\mathbb{R}^n} g_{x_k} \sigma^\varepsilon_{s,r}(x, 0) \, dx,$$

the slash through the integral sign denoting an average. Next, average with respect to $t_1 \leq s \leq t_1 + r$:

$$\fint_{C(x_1, t_1, r)} u^\varepsilon_{x_k}(x, s) \, dx \, ds = \int_{\mathbb{R}^n} g_{x_k} \sigma^\varepsilon_r(x, 0) \, dx.$$

Put $r = r_j$, let $\varepsilon = \varepsilon_m \to 0$ and note that $u^\varepsilon_{x_k} \rightharpoonup u_{x_k}$ weakly in $L^2_{loc}(\mathbb{R}^n \times (0, \infty))$, since $u^\varepsilon \to u$ locally uniformly. Thus (3.10) implies

$$\fint_{C(x_1, t_1, r_j)} u_{x_k}(x, s) \, dx \, ds = \int_{\mathbb{R}^n} g_{x_k} \, d\gamma_{r_j, 0}.$$

Now let $r_m \to 0$, remembering that (x_1, t_1) is a Lebesgue point for Du and recalling from (3.11) that $\gamma_{r_m, 0}$ converges weakly to γ_0. This proves (3.30).

2. Similarly, we have

$$\fint_{C(x_1, t_1, r)} u_t(x, s) \, dx \, ds = \int_{\mathbb{R}^n} (-H(Dg(x)) + \varepsilon \Delta g)\sigma^\varepsilon_r(x, 0) \, dx;$$

and this formula in the limit $\varepsilon = \varepsilon_m \to 0$, $r = r_m \to 0$ implies (3.31).

3. Now select a smooth function $\zeta = \zeta(t)$ such that $\zeta(0) = 1$, $\zeta(t_1) = 0$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^n} \zeta \nabla u^\varepsilon \sigma_r^\varepsilon \, \mathrm{d}x = \int_{\mathbb{R}^n} \zeta' \nabla u^\varepsilon \sigma_r^\varepsilon \, \mathrm{d}x.$$

We integrate, and then take a sequence of smooth functions ζ_k approximating the function ζ satisfying $\zeta \equiv 1$ on $[0, t_1 - h]$, $\zeta(t_1) = 0$, with ζ linear on $[t_1 - h, t_1]$. We obtain the identity

$$\int_{\mathbb{R}^n} \nabla u^\varepsilon(x, 0) \sigma_r^\varepsilon(x, 0) \, \mathrm{d}x = \frac{1}{h} \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \nabla u^\varepsilon \sigma_r^\varepsilon \, \mathrm{d}x \, \mathrm{d}t.$$

Let $\varepsilon = \varepsilon_m \to 0$ and then $r = r_m \to 0$:

$$\int_{\mathbb{R}^n} \nabla u(x, 0) \, \mathrm{d}\gamma_0 = \frac{1}{h} \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} q \, \mathrm{d}\mu.$$

According to the formulas (3.30) and (3.31), the term on the left equals $\nabla u(x_1, t_1)$. Hence when we now send $h \to 0$ and recall Theorem 3.8, the identity (3.29) follows. □

4. Special cases

In this section we discuss various special cases.

4.1. When σ is a point mass at $t = 0$

A particularly easy situation occurs when the measure σ is a Dirac mass at time 0.

Theorem 4.1. *Assume for some point $x_0 \in \mathbb{R}^n$ that*

$$\gamma_0 = \delta_{x_0}. \tag{4.1}$$

Then

(i) $q = (Dg(x_0), -H(Dg(x_0)))$ *μ-almost everywhere,*
(ii) $M \equiv 0$,
(iii) $x_1 = x_0 + t_1 DH(Dg(x_0))$, *and*
(iv) *the support of σ is the line segment*

$$\{x_0 + t DH(Dg(x_0)) \mid 0 \leqq t \leqq t_1\}. \tag{4.2}$$

We interpret this as saying that the characteristic for the Hamilton–Jacobi PDE (1.4) starting at x_0 at $t = 0$ does not hit a shock before time t_1. This characteristic is the straight line (4.2).

Proof. Let $\Phi(q) := \frac{1}{2}(|p - Dg(x_0)|^2 + |p_{n+1} + H(Dg(x_0))|^2)$. Since Φ is convex, for each time $0 \leqq s \leqq t_1$ we have

$$
\varepsilon \int_0^s \int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 + |Du_t^\varepsilon|^2 \, \sigma_r^\varepsilon \, dx \, dt
$$

$$
+ \frac{1}{2} \int_{\mathbb{R}^n} |Du^\varepsilon(x, s) - Dg(x_0)|^2 + |u_t^\varepsilon(x, s) + H(Dg(x_0))|^2 \sigma_r^\varepsilon(x, s) \, dx
$$

$$
\leqq \frac{1}{2} \int_{\mathbb{R}^n} |Du^\varepsilon(x, 0) - Dg(x_0)|^2 + |u_t^\varepsilon(x, 0) + H(Dg(x_0))|^2 \sigma_r^\varepsilon(x, 0) \, dx
$$

$$
= \frac{1}{2} \int_{\mathbb{R}^n} |Dg(x) - Dg(x_0)|^2 + |\varepsilon \Delta g(x) - H(Dg(x))
$$

$$
+ H(Dg(x_0))|^2 \sigma_r^\varepsilon(x, 0) \, dx.
$$

Sending $\varepsilon = \varepsilon_m \to 0, r = r_m \to 0$, and recalling (4.1), we see that the term on the right goes to zero. It follows that

$$
\int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} |p - Dg(x_0)|^2 + |p_{n+1} + H(Dg(x_0))|^2 \, d\mu = 0. \qquad (4.3)
$$

Since also $\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} |D^2 u^\varepsilon|^2 + |Du_t^\varepsilon|^2 \sigma_r^\varepsilon \, dx \, dt \to 0$, (3.20)–(3.22) imply $M = M^\gamma \equiv 0$.

In view of (4.3)

$$
\overline{DH} = \int_{\mathbb{R}^{n+1}} DH(p) \, d\nu_{x,t} = DH(Dg(x_0)).
$$

Thus the evolution equation (3.26) for σ gives

$$
\sigma_t + \operatorname{div}(\sigma \, DH(Dg(x_0))) = 0 \quad \text{in } \mathbb{R}^n \times [0, t_1);
$$

and so $x_1 = x_0 + t_1 DH(Dg(x_0))$. $\quad\square$

4.2. H uniformly convex

Assume, next, that H is uniformly convex:

$$
D^2 H \geqq \theta I \qquad (4.4)
$$

for some constant $\theta > 0$.

As a further illustration of our methods, we present first a simple new proof of a standard one-sided second derivative estimate:

Theorem 4.2. *Assume that H satisfies* (4.4). *Then there exists a constant C such that*

$$
u_{\xi\xi}^\varepsilon(x, t) \leqq \frac{C}{t} \qquad (4.5)
$$

for all $x \in \mathbb{R}^n, t > 0$ and all $\varepsilon > 0$ and $|\xi| = 1$.

We likewise have the simpler bound

$$u^{\varepsilon}_{\xi\xi} \leqq C \tag{4.6}$$

if g is C^2.

Proof. Differentiate the PDE (1.1) twice with respect to ξ, to find

$$(u^{\varepsilon}_{\xi\xi})_t + H_{p_k}(Du^{\varepsilon})(u^{\varepsilon}_{\xi\xi})_{x_k} + H_{p_k p_l}(Du^{\varepsilon})u^{\varepsilon}_{x_k\xi}u^{\varepsilon}_{x_l\xi} = \varepsilon\Delta u^{\varepsilon}_{\xi\xi}.$$

Let σ^{ε} solve (1.2) with the terminal condition $\alpha = \delta_{x_1}$ at time t_1. We multiply the foregoing PDE by $t^2\sigma^{\varepsilon}$ and integrate:

$$t_1^2 u^{\varepsilon}_{\xi\xi}(x_1, t_1) + \int_0^{t_1}\int_{\mathbb{R}^n} t^2 H_{p_k p_l}(Du^{\varepsilon})u^{\varepsilon}_{x_k\xi}u^{\varepsilon}_{x_l\xi}\sigma^{\varepsilon}\,\mathrm{d}x\,\mathrm{d}t = 2\int_0^{t_1}\int_{\mathbb{R}^n} tu^{\varepsilon}_{\xi\xi}\sigma^{\varepsilon}\,\mathrm{d}x\,\mathrm{d}t.$$

We use (4.4) to further calculate that

$$\begin{aligned}
t_1^2 u^{\varepsilon}_{\xi\xi}&(x_1, t_1) + \theta\int_0^{t_1}\int_{\mathbb{R}^n} t^2|D^2u^{\varepsilon}|^2\sigma^{\varepsilon}\,\mathrm{d}x\,dt\\
&\leqq C\int_0^{t_1}\int_{\mathbb{R}^n} t|D^2u^{\varepsilon}|\sigma^{\varepsilon}\,\mathrm{d}x\,dt\\
&\leqq \theta\int_0^{t_1}\int_{\mathbb{R}^n} t^2|D^2u^{\varepsilon}|^2\sigma^{\varepsilon}\,\mathrm{d}x\,dt + \frac{C}{\theta}\int_0^{t_1}\int_{\mathbb{R}^n}\sigma^{\varepsilon}\,\mathrm{d}x\mathrm{d}t\\
&\leqq \theta\int_0^{t_1}\int_{\mathbb{R}^n} t^2|D^2u^{\varepsilon}|^2\sigma^{\varepsilon}\,\mathrm{d}x\,dt + \frac{C}{\theta}t_1.
\end{aligned}$$

$\square$

Theorem 4.3. *Assume that H satisfies the uniform convexity condition* (4.4). *Suppose the terminal measure α has a smooth density. Then we have the estimate*

$$\int_0^{t_1}\int_{\mathbb{R}^n}(|D^2u^{\varepsilon}|^2 + |Du^{\varepsilon}_t|^2)\sigma^{\varepsilon}\,\mathrm{d}x\,\mathrm{d}t \leqq C, \tag{4.7}$$

for a constant C independent of ε. Consequently

$$M \equiv 0.$$

In particular, our program of introducing the adjoint PDE (1.2) and the dissipation measures M is not especially interesting if H is convex. But we will see later that for nonconvex H, M can be nontrivial even for times $0 \leqq t < t_1$.

Proof. 1. Differentiate the PDE (1.1) twice with respect to x_j and sum on j:

$$(\Delta u^{\varepsilon})_t + H_{p_k}(Du^{\varepsilon})(\Delta u^{\varepsilon})_{x_k} + H_{p_k p_l}(Du^{\varepsilon})u^{\varepsilon}_{x_k x_j}u^{\varepsilon}_{x_l x_j} = \varepsilon\Delta(\Delta u^{\varepsilon}).$$

Multiply by σ^{ε} and integrate:

$$\int_0^{t_1}\int_{\mathbb{R}^n} H_{p_k p_l}(Du^{\varepsilon})u^{\varepsilon}_{x_k x_j}u^{\varepsilon}_{x_l x_j}\sigma^{\varepsilon}\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^n}\Delta u^{\varepsilon}\,\mathrm{d}\alpha = \int_{\mathbb{R}^n}\Delta g\sigma^{\varepsilon}(x, 0)\,\mathrm{d}x.$$

In view of (4.4) and our hypotheses on g and α, we derive the inequality

$$\int_0^{t_1}\int_{\mathbb{R}^n}|D^2u^{\varepsilon}|^2\sigma^{\varepsilon}\,\mathrm{d}x\,\mathrm{d}t \leqq C.$$

2. We next differentiate (1.1) twice with respect to t, and as above derive the identity

$$\int_0^{t_1} \int_{\mathbb{R}^n} H_{p_k p_l}(Du^\varepsilon) u^\varepsilon_{x_k t} u^\varepsilon_{x_l t} \sigma^\varepsilon \, dx \, dt + \int_{\mathbb{R}^n} u^\varepsilon_{tt}(x, t_1) \, d\gamma$$
$$= \int_{\mathbb{R}^n} u^\varepsilon_{tt}(x, 0) \sigma^\varepsilon(x, 0) \, dx. \tag{4.8}$$

Now $u^\varepsilon_{tt} = -DH(Du^\varepsilon)Du^\varepsilon_t + \varepsilon \Delta u^\varepsilon_t$. Since α has a smooth density, we calculate at time t_1 that

$$\int_{\mathbb{R}^n} u^\varepsilon_{tt} \, d\alpha = -\int_{\mathbb{R}^n} DH(Du^\varepsilon)Du^\varepsilon_t \, d\alpha + O(\varepsilon)$$
$$= -\int_{\mathbb{R}^n} DH(Du^\varepsilon)D\left(u^\varepsilon_t + \lambda\right) \, d\alpha + O(\varepsilon)$$
$$= \int_{\mathbb{R}^n} H_{p_k p_l} u^\varepsilon_{x_k x_l} \left(u^\varepsilon_t + \lambda\right) \, d\alpha + O(1)$$

Now, according to (4.6) and the convexity of H, $H_{p_k p_l} u^\varepsilon_{x_k x_l} \leq C$. Thus, if we select the constant λ above so that $u^\varepsilon_t + \lambda \leq 0$, we have

$$\int_{\mathbb{R}^n} H_{p_k p_l} u^\varepsilon_{x_k x_l} \left(u^\varepsilon_t + \lambda\right) \, d\alpha \geq -C.$$

Then (4.8) and (4.4) provide the estimate

$$\int_0^{t_1} \int_{\mathbb{R}^n} \left|Du^\varepsilon_t\right|^2 \sigma^\varepsilon \, dx \, dt \leq C.$$

$\square$

4.3. H homogeneous

For this section this we assume that H is positively homogeneous of degree one:

$$H(\lambda p) = \lambda H(p) \quad (p \in \mathbb{R}^n, \lambda > 0). \tag{4.9}$$

This implies

$$DH(p) \cdot p = H(p). \tag{4.10}$$

Theorem 4.4. *Suppose that H satisfies the homogeneity condition* (4.9) *and that* $\alpha = \delta_{x_1}$.
Then

$$\text{spt} \, \sigma \subseteq \{(x, t) \mid u(x, t) = u(x_1, t_1)\}, \tag{4.11}$$

and so σ is supported within a level set of u.

Proof. According to (4.10) we have

$$u_t^\varepsilon + DH(Du^\varepsilon) \cdot Du^\varepsilon = \varepsilon \Delta u^\varepsilon,$$

and therefore

$$v_t^\varepsilon + DH(Du^\varepsilon) \cdot Dv^\varepsilon = \varepsilon \Delta v^\varepsilon - \varepsilon \psi''(u^\varepsilon)|Du^\varepsilon|^2 \qquad (4.12)$$

for $v^\varepsilon := \Psi(u^\varepsilon)$. Multiplying (4.12) by σ_r^ε, integrating and as usual sending $\varepsilon = \varepsilon_m \to 0, r = r_m \to 0$ gives

$$\Psi(u(x_1, t_1)) = \int_{\mathbb{R}^n} \Psi(u(x, t)) \, d\gamma_t$$

for each time $0 \leqq t \leqq t_1$. Let $\alpha \neq u(x_1, t_1)$. Replacing Ψ by a sequence Ψ_k such that

$$\Psi_k(z) \to \begin{cases} 1 & z = \alpha \\ 0 & z \neq \alpha, \end{cases}$$

we deduce

$$\int_{\mathbb{R}^n \cap \{u(\cdot, t) = \alpha\}} d\gamma_t = 0.$$

for each $0 \leqq t \leqq t_1$. □

5. Compensated compactness

We next modify the compensated compactness technique of MURAT and TARTAR to the case at hand, to derive some integral formulas that will turn out to contain information about the structure of the measure μ.

Suppose that $\Phi = \Phi(q)$ is smooth. Put

$$\rho_r^\varepsilon := \varepsilon \left(\Phi_{p_k p_l}(\nabla u^\varepsilon) u_{x_k x_i}^\varepsilon u_{x_l x_i}^\varepsilon + 2\Phi_{p_k p_{n+1}}(\nabla u^\varepsilon) u_{x_k x_i}^\varepsilon u_{t x_i}^\varepsilon \right.$$
$$\left. + \Phi_{p_{n+1} p_{n+1}}(\nabla u^\varepsilon) u_{t x_i}^\varepsilon u_{t x_i}^\varepsilon \right) \sigma_r^\varepsilon.$$

Then (2.1) implies

$$\int_0^{t_1} \int_{\mathbb{R}^n} |\rho_r^\varepsilon| \, dx \, dt \leqq C,$$

the constant C independent of ε and r. We may consequently assume, passing as necessary to further subsequences, that as $\varepsilon = \varepsilon_m \to 0$ and then $r = r_m \to 0$

$$\rho_r^\varepsilon \rightharpoonup \rho \text{ weakly in the sense of measures,} \qquad (5.1)$$

for some signed Borel measure ρ on $\mathbb{R}^n \times [0, t_1]$.

5.1. Integral formulas

Theorem 5.1. (i) *We have the identity*

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta(\overline{p_{n+1}\Phi} + \overline{p_i H_{p_i}\Phi}) \, d\sigma$$

$$= -\int_0^{t_1} \int_{\mathbb{R}^n} u(\zeta_t \overline{\Phi} + \zeta_{x_i} \overline{H_{p_i}\Phi}) \, d\sigma + \int_0^{t_1} \int_{\mathbb{R}^n} \zeta u \, d\rho, \qquad (5.2)$$

for all smooth functions $\zeta = \zeta(x, t)$ with compact support in $\mathbb{R}^n \times (0, t_1)$.
(ii) *For all smooth functions $\eta = \eta(x, t)$ with compact support, we have the further formula*

$$\int_0^{t_1} \int_{\mathbb{R}^n} \eta_t \overline{\Phi} + \eta_{x_i} \overline{H_{p_i}\Phi} \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \eta \, d\rho. \qquad (5.3)$$

Proof. 1. Put

$$w^\varepsilon := \Phi(\nabla u^\varepsilon).$$

Then

$$w_t^\varepsilon + DH \cdot Dw^\varepsilon = \varepsilon \Delta w^\varepsilon - \varepsilon \left(\Phi_{p_k p_l} u_{x_k x_i}^\varepsilon u_{x_l x_i}^\varepsilon \right.$$
$$\left. + 2\Phi_{p_k p_{n+1}} u_{x_k x_i}^\varepsilon u_{t x_i}^\varepsilon + \Phi_{p_{n+1} p_{n+1}} u_{t x_i}^\varepsilon u_{t x_i}^\varepsilon \right);$$

and the PDE (1.2) implies

$$\left(w^\varepsilon \sigma_r^\varepsilon \right)_t + \left(H_{p_i} w^\varepsilon \sigma_r^\varepsilon \right)_{x_i} = \varepsilon \left(\sigma_r^\varepsilon \Delta w^\varepsilon - w^\varepsilon \Delta \sigma_r^\varepsilon \right) - \rho_r^\varepsilon$$
$$= \varepsilon \left(\sigma_r^\varepsilon w_{x_i}^\varepsilon - w^\varepsilon \sigma_{r,x_i}^\varepsilon \right)_{x_i} - \rho_r^\varepsilon. \qquad (5.4)$$

2. We multiply the previous identity by $u^\varepsilon \zeta$ and integrate:

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta(u_t^\varepsilon + u_{x_i}^\varepsilon H_{p_i}) w^\varepsilon \sigma_r^\varepsilon \, dx \, dt = -\int_0^{t_1} \int_{\mathbb{R}^n} u^\varepsilon (\zeta_t + \zeta_{x_i} H_{p_i}) w^\varepsilon \sigma_r^\varepsilon \, dx \, dt$$

$$+ \varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} (\zeta u^\varepsilon)_{x_i} \left(\sigma_r^\varepsilon w_{x_i}^\varepsilon - w^\varepsilon \sigma_{r,x_i}^\varepsilon \right) \, dx \, dt + \int_0^{t_1} \int_{\mathbb{R}^n} \zeta u^\varepsilon \rho_r^\varepsilon \, dx \, dt.$$

Denote the term on the left by A_r^ε and the three terms on the right by $B_r^\varepsilon, C_r^\varepsilon, D_r^\varepsilon$:

$$A_r^\varepsilon = B_r^\varepsilon + C_r^\varepsilon + D_r^\varepsilon.$$

Recalling Theorem 3.1, we see that as $\varepsilon = \varepsilon_m \to 0$, and then $r = r_m \to 0$:

$$A_r^\varepsilon \to \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \zeta(p_{n+1} + p_i H_{p_i}) \Phi(q) \, d\mu$$

$$= \int_0^{t_1} \int_{\mathbb{R}^n} \zeta(\overline{p_{n+1}\Phi} + \overline{p_i H_{p_i}\Phi}) \, d\sigma.$$

Likewise,

$$B_r^\varepsilon \to - \int_0^{t_1} \int_{\mathbb{R}^n} u(\zeta_t \overline{\Phi} + \zeta_{x_i} \overline{H_{p_i} \Phi}) \, d\sigma,$$

since $u^\varepsilon \to u$ locally uniformly. Furthermore (5.1) implies

$$D_r^\varepsilon \to \int_0^{t_1} \int_{\mathbb{R}^n} \zeta u \, d\rho.$$

3. Next, we claim that

$$C_r^\varepsilon \to 0. \tag{5.5}$$

To confirm this, note first that

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} |(\zeta u^\varepsilon)_{x_i} \sigma_r^\varepsilon w_{x_i}^\varepsilon| \, dx \, dt \leqq \varepsilon C \int_0^{t_1} \int_{\mathbb{R}^n} (|D^2 u^\varepsilon| + |Du_t^\varepsilon|) \sigma_r^\varepsilon \, dx \, dt \leqq C\varepsilon^{1/2}$$

according to the estimate (2.1). Also,

$$\varepsilon \left| \int_0^{t_1} \int_{\mathbb{R}^n} (\zeta u^\varepsilon)_{x_i} w^\varepsilon \sigma_{r,x_i}^\varepsilon \, dx \, dt \right| = \varepsilon \left| \int_0^{t_1} \int_{\mathbb{R}^n} ((\zeta u^\varepsilon)_{x_i} w^\varepsilon)_{x_i} \sigma_r^\varepsilon \, dx \, dt \right| \leqq C\varepsilon^{1/2}.$$

This proves (5.5).

4. To prove (5.3), multiply the identity (5.4) by η, integrate by parts, and let $\varepsilon = \varepsilon_m \to 0, r = r_m \to 0$. □

Discussion. At this point, if we knew that u were smooth on the support of σ, we could put $\eta = u\zeta$ in (5.3) and combine with (5.2), to deduce

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta (\overline{p_{n+1}\Phi} + \overline{p_i H_{p_i} \Phi}) \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \zeta (u_t \overline{\Phi} + u_{x_i} \overline{H_{p_i} \Phi}) \, d\sigma.$$

This identity is valid for all smooth ζ and Φ, and consequently

$$p_{n+1} + p_i H_{p_i}(p) = u_t(x, t) + u_{x_i}(x, t) H_{p_i}(p) \quad \mu\text{-almost everywhere.} \tag{5.6}$$

However, since u is only Lipschitz continuous, we must work harder to extract useful information from (5.2) and (5.3).

5.2. Compensated compactness at the point (x_1, t_1)

We can, in fact, deduce formula (5.6) at the terminal point if u is differentiable there:

Theorem 5.2. *Suppose u is differentiable at (x_1, t_1). Then*

$$p_{n+1} + p_i H_{p_i}(p) = u_t(x_1, t_1) + u_{x_i}(x_1, t_1) H_{p_i}(p) \quad \nu^-\text{-almost everywhere.}$$

$$\tag{5.7}$$

This formula says geometrically that for almost every point q in the support of ν^-, $\nabla u(x_1, t_1)$ lies in the plane tangent to the hypersurface $\Sigma = \{p_{n+1} + H(p) = 0\}$ at q. We discuss later how to use this information.

Proof. Let

$$\phi(x, t) := Du(x_1, t_1) \cdot (x - x_1) + u_t(x_1, t_1)(t - t_1) + u(x_1, t_1).$$

denote the linear approximation to u at (x_1, t_1). Put $\eta = \phi\zeta$ in (5.3):

$$\int_0^{t_1} \int_{\mathbb{R}^n} (\phi\zeta_t + \zeta u_t(x_1, t_1))\overline{\Phi} + (\phi\zeta_{x_i} + \zeta u_{x_i}(x_1, t_1))\overline{H_{p_i}\Phi} \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \phi\zeta \, d\rho.$$

Combine with (5.2) to find:

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta(\overline{p_{n+1}\Phi} + \overline{p_i H_{p_i}\Phi}) \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \zeta(u_t(x_1, t_1)\overline{\Phi} + u_{x_i}(x_1, t_1)\overline{H_{p_i}\Phi}) \, d\sigma$$

$$+ \int_0^{t_1} \int_{\mathbb{R}^n} u - \phi \, d\rho + \int_0^{t_1} \int_{\mathbb{R}^n} (\phi - u)(\zeta_t \overline{\Phi} + \zeta_{x_i} \overline{H_{p_i}\Phi}) \, d\sigma.$$

Next, select a small number $h > 0$ and put $\zeta(x, t) = a(t)b(x)$, where $a(t_1) = 0$, $a \equiv 0$ on $[0, t_1 - h]$ and $b \equiv 1$ on the support of σ:

$$\int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} a(\overline{p_{n+1}\Phi} + \overline{p_i H_{p_i}\Phi}) \, d\sigma = \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} a(u_t(x_1, t_1)\overline{\Phi}$$

$$+ u_{x_i}(x_1, t_1)\overline{H_{p_i}\Phi}) \, d\sigma + \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} a(u - \phi) \, d\rho + \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} (\phi - u)a'\overline{\Phi} \, d\sigma.$$

$$(5.8)$$

Now pick smooth $a(t) = a_h(t)$ so that

$$\frac{1}{h} \int_{t_1-h}^{t_1} a_h(t) \, dt = 1, \quad |a_h| \leq C, \quad |a_h'| \leq \frac{C}{h}$$

for some constant C. Then

$$\left| \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} a_h(u - \phi) \, d\rho \right| = o(h),$$

since u is differentiable at (x_1, t_1); and likewise

$$\left| \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} (\phi - u)a_h' \overline{\Phi} \, d\sigma \right| \leq \frac{o(h)}{h} \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} d\sigma = o(h).$$

We divide (5.8) by h and let $h \to 0$, recalling Theorem 3.8 to conclude that

$$\int_{\mathbb{R}^{n+1}} (p_{n+1} + p_i H_{p_i})\Phi \, d\nu^- = \int_{\mathbb{R}^{n+1}} (u_t(x_1, t_1) + u_{x_i}(x_1, t_1)H_{p_i})\Phi \, d\nu^-.$$

The validity of this identity for all Φ implies (5.7). $\quad\square$

Remark: compensated compactness for conservation laws versus Hamilton–Jacobi PDE. Theorems 5.1 and 5.2 together comprise a new application of the Div-Curl Lemma to Hamilton–Jacobi PDE. It is interesting to compare and contrast our approach with the more customary div-curl methods for conservation laws, due to Tartar, Murat, DiPerna, Chen, etc: see for instance Tartar [20] or [8]. The latter seem inherently to require that we work in $n = 1$ space dimension, since only then does the divergence identity $u_t + v_x = 0$ imply $u = w_x$, $v = -w_t$ for some potential w. (But see Bagnerini et al. [3].) By contrast, our techniques are rather more natural, since we have an obvious candidate for a curl–free vector field even in $n > 1$ dimensions, namely ∇u^ε.

But our conclusion (5.7) is not so strong as that for conservation laws, as it is only a single algebraic/geometric relation among variables in $n + 1$ dimensions.

5.3. Almost everywhere convergence of ∇u^ε

Conditions ensuring ν^- is a point mass. As noted above, the identity (5.7) implies that for almost every point q in the support of ν^-, $\nabla u(x_1, t_1)$ lies in the plane tangent to the surface $\Sigma = \{p_{n+1} + H(p) = 0\}$ at q. In addition, from (1.5) we know that

$$\nabla u(x_1, t_1) \in \Sigma; \tag{5.9}$$

and from (3.28) that

$$\mathrm{spt}\, \nu^- \subseteq \Sigma. \tag{5.10}$$

Also, according to Theorem 3.9,

$$\nabla u(x_1, t_1) = \int_{\mathbb{R}^{n+1}} q \, d\nu^-, \tag{5.11}$$

provided we assume also that (x_1, t_1) is also a Lebesgue point for ∇u.

Now, if the mapping $p \mapsto H(p)$ is either strictly convex or strictly concave, conditions (5.9)–(5.11) alone imply that $\nu^- = \delta_{\nabla u(x_1,t_1)}$. The point is that *adding in our new geometric condition (5.7) leads to the same conclusion for certain nonconvex Hamiltonians.*

Example. For instance, suppose the surface Σ has the shape illustrated in Fig. 1, for $n = 1$. If $\nabla u(x_1, t_1)$ lies on the part of the curve illustrated, our condition (5.7) implies that the only other possible points belonging to the support of ν^- are q_1 and q_2, as drawn. Thus $\mathrm{spt}\, \nu^- \subseteq \{\nabla u(x_1, t_1), q_1, q_2\}$. However $\nabla u(x_1, t_1)$ is then an extreme point of the closed convex hull of $\mathrm{spt}\, \nu^-$, and consequently (5.11) implies $\nu^- = \delta_{\nabla u(x_1,t_1)}$. The same conclusion follows, even more easily, if $\nabla u(x_1, t_1)$ lies on the left or right parts of Σ.

This argument applies also for certain nonconvex Hamiltonians in $n > 1$ variables, for example $H(p) = (|p|^2 - 1)^2$, but certainly fails for other examples.
Pointwise convergence of ∇u^ε. A fundamental open problem is to determine precise conditions on the Hamiltonian H, implying that ν^- is a point mass and,
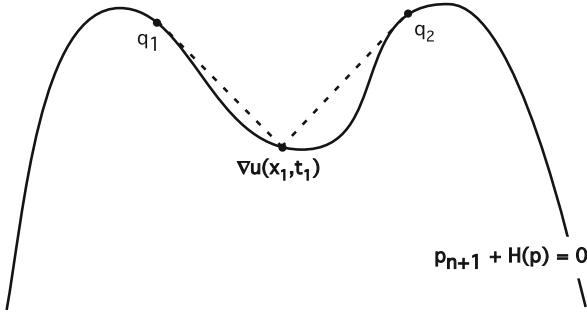
Fig. 1. Geometry of the support of ν^-

therefore, that $\nabla u^\varepsilon \to \nabla u$ almost everywhere. Inspired by the previous example, we provide a partial answer to this question.

Define

$$\Lambda \subseteq \Sigma$$

to consist of those points $q^* = (p_1^*, \ldots, p_{n+1}^*) \in \Sigma$ with the property that q^* is an extreme point of the set

$$\overline{\mathrm{co}}\{q = (p_1, \ldots, p_{n+1}) \in \Sigma \mid p_{n+1} + p_i H_{p_i}(p) = p_{n+1}^* + p_i^* H_{p_i}(p)\},$$

where "$\overline{\mathrm{co}}$" denotes the closed, convex hull.

For instance, $\Lambda = \Sigma$ if H is uniformly convex or concave, and also $\Lambda = \Sigma$ in the example illustrated in Fig. 1.

Theorem 5.3. *If $\Lambda = \Sigma$, then*

$$\nabla u^\varepsilon \to \nabla u \quad \text{almost everywhere.} \tag{5.12}$$

Proof. 1. Since $\{\nabla u^\varepsilon\}$ is bounded, there exists a subsequence $\varepsilon_j \to 0$ and for almost every point (x, t) a Radon probability measure $\rho_{x,t}$ on $\mathbb{R}^{n+1}$ such that

$$\int_{C(x_1, t_1, r)} \Phi(\nabla u^{\varepsilon_j}(x, t)) \, dx \, dt \to \int_{C(x_1, t_1, r)} \int_{\mathbb{R}^{n+1}} \Phi(q) \, d\rho_{x,t} \, dx \, dt \tag{5.13}$$

for all cylinders $C(x_1, t_1, r) \subset \mathbb{R}^n \times [0, \infty)$ and all continuous functions Φ. These are the usual Young measures for the weak convergence $\nabla u^\varepsilon \rightharpoonup \nabla u$. We may suppose that all the sequences $\{\varepsilon_m\}$ discussed earlier are subsequences of the given sequence $\{\varepsilon_j\}$ for which (5.13) holds.

2. Almost every point (x_1, t_1) is a Lebesgue point for the functions $(x, t) \mapsto \int_{\mathbb{R}^{n+1}} \Phi(q) \, d\rho_{x,t}$ for every Φ. At such a point $\nabla u(x_1, t_1)$ exists. Then if Φ is convex and $t_1 \leqq s \leqq t_1 + r$, we have

$$\fint_{B(x_1,r)} \Phi(\nabla u^\varepsilon(x, s)) \, dx = \int_{\mathbb{R}^n} \Phi(\nabla u^\varepsilon(x, s))\alpha_r(x) \, dx$$

$$\leqq \int_{\mathbb{R}^n} \Phi(\nabla u^\varepsilon(x, t))\sigma_{r,s}^\varepsilon \, dx$$

for each time $0 \leqq t < t_1$. Average with respect to s:

$$\fint_{C(x_1,t_1,r)} \Phi(\nabla u^\varepsilon)\, dx\, ds \leqq \int_{\mathbb{R}^n} \Phi(\nabla u^\varepsilon(x,t))\sigma_r^\varepsilon\, dx.$$

Consequently, if $\alpha = \alpha(t)$ is a smooth, nonnegative function with integral one on the interval $[t_1 - h, t_1]$, we have

$$\fint_{C(x_1,t_1,r_j)} \int_{\mathbb{R}^{n+1}} \Phi(q)\, d\rho_{x,s} dx\, ds = \lim_{\varepsilon_m \to 0} \fint_{C(x_1,t_1,r_j)} \Phi(\nabla u^{\varepsilon_m})\, dx\, ds$$

$$\leqq \lim_{\varepsilon_m \to 0} \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \alpha \Phi(\nabla u^{\varepsilon_m}(x,t))\, \sigma_{r_j}^{\varepsilon_m}\, dx\, dt$$

$$= \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \alpha \Phi\, d\mu_{r_j}.$$

Remember that (x_1, t_1) is a Lebesgue point for the function $(x,t) \mapsto \int_{\mathbb{R}^{n+1}} \Phi(q)\, d\rho_{x,t}$ and send $r = r_m \to 0$:

$$\int_{\mathbb{R}^{n+1}} \Phi(q)\, d\rho_{x_1,t_1} \leqq \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \alpha \Phi\, d\mu.$$

This inequality is valid for all functions α as above, and so

$$\int_{\mathbb{R}^{n+1}} \Phi(q)\, d\rho_{x_1,t_1} \leqq \frac{1}{h} \int_{t_1-h}^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi\, d\mu.$$

Finally, send $h \to 0$ and recall (3.27), to discover that

$$\int_{\mathbb{R}^{n+1}} \Phi(q)\, d\rho_{x_1,t_1} \leqq \int_{\mathbb{R}^{n+1}} \Phi(q)\, d\nu^-. \tag{5.14}$$

2. Now, Theorem 3.9 tells us that

$$q^* := \nabla u(x_1, t_1) = \int_{\mathbb{R}^{n+1}} q\, d\nu^-; \tag{5.15}$$

and furthermore $p_{n+1} + p_i H_{p_i}(p) = p_{n+1}^* + p_i^* H_{p_i}(p)$ for points $q \in \operatorname{spt} \nu^-$. Since $q^* \in \Lambda = \Sigma$, q^* is an extreme point of the closed convex hull of spt μ. But because (5.15) asserts that q^* is also the center of mass of ν^-, it follows that $\nu^- = \delta_{q^*}$.

Our taking $\Phi(q) = |q - q^*|^2$ in (5.14) lets us now deduce that $\rho_{x_1,t_1} = \delta_{q^*}$ also. This conclusion, valid for almost all points (x_1, t_1), implies $\nabla u^\varepsilon \to \nabla u$ almost everywhere. □

5.4. Compensated compactness for times $0 \leqq t < t_1$

Our integral identities also contain information for times $0 \leqq t < t_1$. Since u need not be smooth, we need to introduce some approximations. Assume, therefore, that $\{u^\delta\}_{0 < \delta < 1}$ is some collection of smooth functions satisfying

$$\max_{\mathbb{R}^n \times [0,t_1]} |Du^\delta, u_t^\delta| \leqq C, \quad \max_{\mathbb{R}^n \times [0,t_1]} |u^\delta - u| = o(1), \quad (5.16)$$

and

$$u_{x_i}^\delta \rightharpoonup \beta^i \ (i = 1, \ldots, n), \quad u_t^\delta \rightharpoonup \beta^{n+1} \text{ weakly in } L^2(d\sigma) \quad (5.17)$$

as $\delta \to 0$, for some vector field $\boldsymbol{\beta} = (\beta^1, \ldots, \beta^{n+1})$.

Theorem 5.4. *We have*

$$p_{n+1} + p_i H_{p_i}(p) = \beta^{n+1}(x, t) + \beta^i(x, t) H_{p_i}(p) \quad \mu\text{-almost everywhere.} \quad (5.18)$$

So, for almost every point q in the support of $\nu_{x,t}$, $\boldsymbol{\beta}(x, t)$ lies in the plane tangent to the surface Σ at q. Even if we know nothing about the location of $\boldsymbol{\beta}(x, t)$, we can still deduce that the tangent planes to Σ corresponding to points in spt $\nu_{x,t}$ must all intersect at at least one point.

Proof. Put $\eta = u^\delta \zeta$ in (5.3):

$$\int_0^{t_1} \int_{\mathbb{R}^n} \left(u^\delta \zeta_t + \zeta u_t^\delta \right) \overline{\Phi} + \left(u^\delta \zeta_{x_i} + \zeta u_{x_i}^\delta \right) \overline{H_{p_i} \Phi} \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} u^\delta \zeta \, d\rho.$$

Subtract from (5.2) and recall (5.16), to discover

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta (\overline{p_{n+1} \Phi} + \overline{p_i H_{p_i} \Phi}) \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \zeta \left(u_t^\delta \overline{\Phi} + u_{x_i}^\delta \overline{H_{p_i} \Phi} \right) d\sigma + o(1).$$

Sending $\delta \to 0$ and remembering the weak convergence (5.17), we find

$$\int_0^{t_1} \int_{\mathbb{R}^n} \zeta (\overline{p_{n+1} \Phi} + \overline{p_i H_{p_i} \Phi}) \, d\sigma = \int_0^{t_1} \int_{\mathbb{R}^n} \zeta (\beta^{n+1} \overline{\Phi} + \beta^i \overline{H_{p_i} \Phi}) \, d\sigma.$$

That this identity is valid for all ζ and Φ implies (5.18). $\quad\square$

One interesting possibility is that some part of the measure σ is singular and is supported on a smooth shock hypersurface Γ, across which ∇u is discontinuous. See Fig. 6. We suppose that u is smooth on each side of Γ, and let $\nabla u^\pm$ denote the limits on ∇u on each side of the shock surface

Theorem 5.5. *For σ-almost every point $(x, t) \in \Gamma$, we have*

$$p_{n+1} + p_i H_{p_i}(p) = u_t^\pm(x, t) + u_{x_i}^\pm(x, t) H_{p_i}(p) \quad \nu_{x,t}\text{-almost everywhere.} \quad (5.19)$$
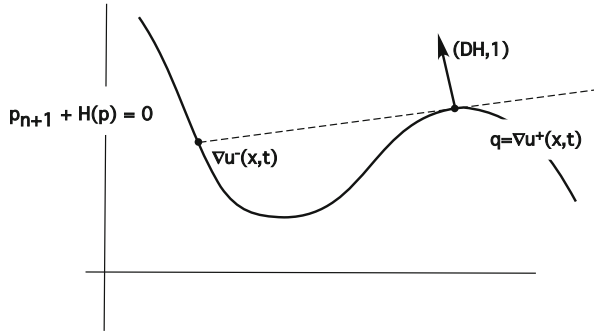
Fig. 2. $\nabla u^{\pm}$ at a right contact discontinuity

So, for almost every point q in the support of $\nu_{x,t}$, both the points $\nabla u^+(x,t)$ and $\nabla u^-(x,t)$ lie in the plane tangent to the surface $p_{n+1} + H(p) = 0$ at q. Note additionally that both points $\nabla u^{\pm}(x,t)$ also lie on this surface. A particular case is that spt $\nu_{x,t}$ is a point mass, at either $\nabla u^+(x,t)$ or $\nabla u^-(x,t)$. Then (5.19) implies that the former corresponds to a right contact discontinuity, corresponding to the geometry illustrated in Fig. 2. The latter case gives a left contact discontinuity. In this sense *our compensated compactness arguments "predict" the onset of contact discontinuities.*

Proof. Let ν denote a unit normal vector field to Γ in $\mathbb{R}^{n+1}$, and think of ν as pointing in the "right hand" direction: see Fig. 6 in Section 7. Smoothly extend ν off of Γ.

Put $u^\delta := u(x \pm \delta\nu, t)$. Then $u^\delta \to u$ locally uniformly and $\nabla u^\delta \to \nabla u^{\pm}$ on Γ. Apply formula (5.18). $\square$

6. Dissipation and contact discontinuities

6.1. *Contact discontinuities in $n = 1$ dimension*

In this section we work out an interpretation of the matrix of dissipation measures for the special case of a piecewise smooth solution u of (1.4) in $n = 1$ spatial dimension. We henceforth assume the geometry of the characteristics and shocks illustrated in Fig. 3. This picture corresponds to illustrations found in several papers about shock structure for scalar conservation laws with nonconvex flux functions in $n = 1$ space dimension: see the articles by DAFERMOS [7], BALLOU [2], MARSON [14], etc. and also ZHENG's book [22].

We assume that the shock wave $\Gamma = \{x = s(t)\}$ is a smooth right contact discontinuity, as drawn. We denote by $q^+ = q^+(t)$ and $q^- = q^-(t)$ the right- and left-hand values of the gradient $\nabla u = (u_x, u_t)$ along Γ at time t_1. The geometric locations of $q^{\pm}$ are illustrated in Fig. 4, the dashed line connecting q^+ and q^- being tangent to the curve Σ at q^+, since the latter is a contact discontinuity.
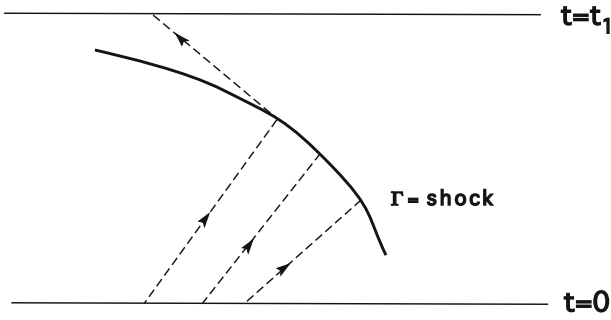
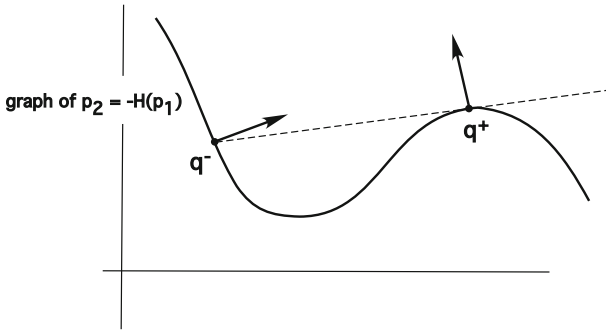**Fig. 3.** A right contact discontinuity curve for ∇u

Fig. 4. Left and right states for a right contact discontinuity

Next, let $\Phi : \mathbb{R}^2 \to \mathbb{R}$ be a smooth function and set

$$\phi(t) := \int_{\mathbb{R}} \Phi(\nabla u) \, d\sigma.$$

According to Theorem 3.19, we have

$$\dot{\phi} = -\int_{\mathbb{R}} \int_{\mathbb{R}^2} \Phi_{p_k p_l}(q) \, dm^{kl}. \tag{6.1}$$

Our intention is to express the abstract expression on the right explicitly in terms of the geometry of the contact discontinuity and the graph of H, from Figs. 3 and 4. (That the graph $\Sigma = \{p_{n+1} + H(p) = 0\}$ lies below the segment connecting $q^{\pm}$ follows from viscosity solution admissibility conditions: see [5].)

For definiteness, suppose $\gamma = \delta_{x_1}$ and the point (x_1, t_1) lies above the shock, as illustrated in Fig. 3. For the moment, we think of the measure σ as evolving backwards in time, and is thus a point mass along the characteristic until it intersects the shock Γ at some time $0 < t^* < t_1$. We further assume that for each time $t \leqq t^*$, σ decomposes into a point mass of magnitude $\sigma^+(t)$ at $x = s(t)$ along the shock Γ and an absolutely continuous part with smooth density σ^- to the left of the shock. Then

$$\phi(t) = \int_{-\infty}^{s(t)} \Phi(\nabla u) \sigma^- \, dx + \Phi(q^+(t)) \sigma^+(t). \tag{6.2}$$

We want to calculate $\dot{\phi}$ and also to understand how $\sigma^{\pm}$ and $q^{\pm}$ evolve for times $t \leqq t^*$.

Theorem 6.1. *Under the foregoing assumptions,*

$$\dot{\phi} = - \left(\Phi(q^+) + \nabla\Phi(q^+) \cdot (q^- - q^+) - \Phi(q^-) \right) (H'(p^+) - H'(p^-))\sigma^-.$$

(6.3)

Furthermore,

$$\dot{\sigma}^+ = (H'(p^-) - H'(p^+))\sigma^-,$$

(6.4)

and

$$\dot{q}^+\sigma^+ = (q^- - q^+)\dot{\sigma}^+.$$

(6.5)

In these formulas, $\sigma^- = \sigma^-(t)$ denotes the left hand limit of the density $\sigma^-(x, t)$ at $x = s(t)$, along the shock curve Γ. Observe that the right-hand side of (6.3) is not preserved under an interchange of q^- and q^+: this asymmetry occurs since we have a right contact discontinuity and not a left contact discontinuity. We will see also that (6.4) and (6.5) are infinitesimal versions of the conservation of the mass of σ and the conservation of integral of q with respect to σ.
Interpretations.

   (i) Since $H'(p^-) > H'(p^+)$ and $\sigma^- \geqq 0$, we see in particular that if Φ is convex, then $\dot{\phi} \leq 0$, as of course (6.1) predicts. Also (6.4) implies

$$\dot{\sigma}^+ \geqq 0.$$

We understand this to mean that as time goes forward, more and more of the mass of the measure σ coalesces onto the shock, brought there along the characteristics from the left that collide with the shock.
 (ii) Formula (6.3) also implies that *the rate of dissipation is cubic in the shock strength* $|q^+ - q^-|$.
 (iii) Comparing (6.3) with (6.1), we see that the matrix M of dissipation measures is supported along the line segment connecting q^+ and q^-.

Proof. 1. Differentiate (6.2) with respect to t_1:

$$\dot{\phi} = \Phi(q^-)\sigma^-\dot{s} + \int_{-\infty}^{s(t)} (\Phi(\nabla u)\sigma^-)_t \, dx + \nabla\Phi(q^+) \cdot \dot{q}^+\sigma^+ + \Phi(q^+)\dot{\sigma}^+$$

$$= \Phi(q^-)\sigma^-\dot{s} - \int_{-\infty}^{s(t)} (H'(u_x)\Phi(\nabla u)\sigma^-)_x \, dx + \nabla\Phi(q^+) \cdot \dot{q}^+\sigma^+ + \Phi(q^+)\dot{\sigma}^+$$

$$= \Phi(q^-)\sigma^-\dot{s} - H'(p^-)\Phi(q^-)\sigma^- + \nabla\Phi(q^+) \cdot \dot{q}^+\sigma^+ + \Phi(q^+)\dot{\sigma}^+,$$

where $q^{\pm} = (p_1^{\pm}, p_2^{\pm}) = (p^{\pm}, p_2^{\pm})$. Since $\dot{s} = H'(p^+)$ along the contact discontinuity, we therefore deduce that

$$\dot{\phi} = \Phi(q^-)(H'(p^+) - H'(p^-))\sigma^- + \nabla\Phi(q^+) \cdot \dot{q}^+\sigma^+ + \Phi(q^+)\dot{\sigma}^+. \quad (6.6)$$
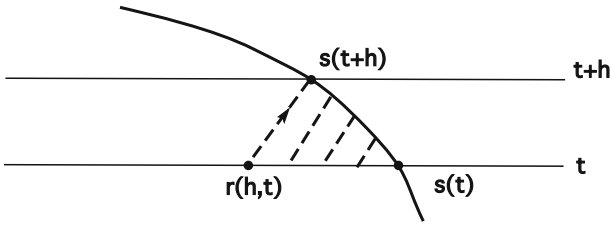
Fig. 5. Characteristics intersecting the shock during $[t, t + h]$

2. Next we prove (6.4). To see this, recall that $\int_{\mathbb{R}} d\sigma = 1$ for each time, and therefore

$$\int_{r(h,t)}^{s(t)} \sigma^- \, \mathrm{d}x = \sigma^+(t + h) - \sigma^+(t)$$

for small $h > 0$, for the point

$$r(h, t) := s(t + h) - h H'(p^-(t + h)).$$

See Fig. 5. Our differentiating with respect to h and putting $h = 0$ gives

$$(H'(p^-) - \dot{s})\sigma^- = \dot{\sigma}^+.$$

This implies (6.4), since $\dot{s} = H'(p^+)$.

3. Now insert (6.4) into (6.6):

$$\dot{\phi} = (\Phi(q^+) - \Phi(q^-))\dot{\sigma}^+ + \nabla\Phi(q^+) \cdot \dot{q}^+\sigma^+. \tag{6.7}$$

The final observation is that, according to (6.1), if we put $\Phi(q) = q$, the left-hand side of (6.7) is identically zero:

$$0 = (q^+ - q^-)\dot{\sigma}^+ + \dot{q}^+\sigma^+.$$

This is (6.5).

We plug this identity back into (6.7) and recall (6.4), ending up with the stated formula (6.3). $\quad\square$

6.2. Contact discontinuities in $n > 1$ dimensions

Next, we extend the previous calculations to the case of a smooth contact discontinuity Γ in higher dimensions, on both sides of which u is smooth.

We write $\Gamma(t) := \Gamma \cap (\mathbb{R}^n \times \{t\})$ for the shock at time $t \geqq 0$, and assume $\Gamma(t)$ to be a smooth hypersurface in $\mathbb{R}^n$. Let ν denote a unit normal vector field to Γ in $\mathbb{R}^{n+1}$; and for each time let n denote a corresponding unit normal vector field to $\Gamma(t)$ in $\mathbb{R}^n$. We regard ν as pointing in the "right hand" direction, along which u has a contact discontinuity. Write $S(t)$ for the region in $\mathbb{R}^n$ to the "left" of $\Gamma(t)$: see Fig. 6.
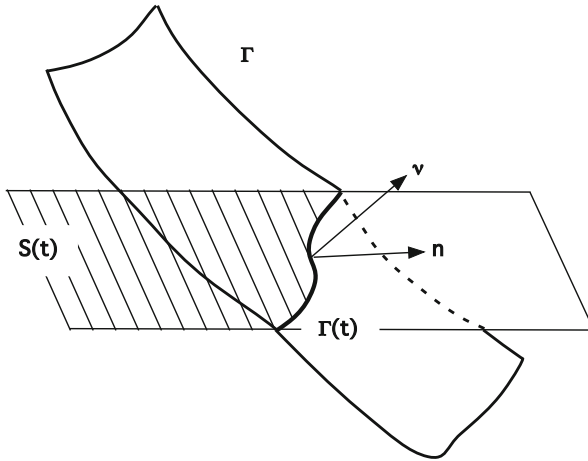
Fig. 6. A discontinuity surface for ∇u

We introduce, as well, the functions $q^{\pm} = (p^{\pm}, p_{n+1}^{\pm})$ to denote the left- and right-hand limits of ∇u along Γ. Since we have a contact discontinuity from the right, the normal velocity v of $\Gamma(t)$ with respect to the direction n is

$$v = DH(p^+) \cdot n. \tag{6.8}$$

Furthermore,

$$\nu = (\hat{\nu}, \nu_{n+1}) = \frac{1}{(1+v^2)^{1/2}}(n, -v). \tag{6.9}$$

Finally, we will suppose that the measure σ decomposes into one part supported along $\Gamma(t)$, with smooth density σ^+ with respect to $(n-1)$-dimensional Hausdorff measure $\mathcal{H}^{n-1}$, and another part supported in $S(t)$, with smooth density σ^- with respect to n-dimensional Lebesgue measure.

Given, then, a smooth function $\Phi : \mathbb{R}^{n+1} \to \mathbb{R}$, we put

$$\phi(t) := \int_{\mathbb{R}^n} \Phi(\nabla u) \, d\sigma = \int_{S(t)} \Phi(\nabla u)\sigma^- \, dx + \int_{\Gamma(t)} \Phi(q^+)\sigma^+ \, d\mathcal{H}^{n-1}, \tag{6.10}$$

and recall that

$$\dot{\phi} = -\int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \Phi_{p_k p_l}(q) \, dm^{kl}. \tag{6.11}$$

Once again we wish to understand the right-hand side of (6.11) explicitly in terms of the geometry of Γ and the graph of H.

Theorem 6.2. *We have*

$$\dot{\phi} = -\int_{\Gamma(t)} \left(\Phi(q^+) + \nabla\Phi(q^+) \cdot (q^- - q^+) - \Phi(q^-) \right)$$
$$\times (DH(p^+) - DH(p^-)) \cdot n\sigma^- \, d\mathcal{H}^{n-1}. \tag{6.12}$$

The formula (6.12) is a higher dimensional analog of (6.3). As before, we see that the rate of dissipation is cubic in the shock strength $|q^+ - q^-|$.

Proof. 1. We begin by recording for smooth functions $f = f(x, t)$ the kinematic formula

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Gamma(t)} f \, \mathrm{d}\mathcal{H}^{n-1} = \int_{\Gamma(t)} \frac{Df}{Dt} - f v \mathbb{H} \, \mathrm{d}\mathcal{H}^{n-1}, \qquad (6.13)$$

where $\mathbb{H}$ denotes the mean curvature of $\Gamma(t)$ and

$$\frac{Df}{Dt} := f_t + v Df \cdot n.$$

We employ (6.13) in differentiating (6.10) with respect to t_1:

$$\dot{\phi} = \int_{\Gamma(t)} \Phi(q^-)\sigma^- v \, \mathrm{d}\mathcal{H}^{n-1} + \int_{S(t)} (\Phi(\nabla u)\sigma^-)_t \, \mathrm{d}x$$

$$+ \int_{\Gamma(t)} \nabla \Phi(q^+) \cdot \frac{Dq^+}{Dt} \sigma^+ + \Phi(q^+) \frac{D\sigma^+}{Dt} - \Phi(q^+)\sigma^+ v \mathbb{H} \, \mathrm{d}\mathcal{H}^{n-1}.$$

$$(6.14)$$

Furthermore,

$$\int_{S(t)} (\Phi(\nabla u)\sigma^-)_t \, \mathrm{d}x = - \int_{S(t)} \mathrm{div}(DH(Du)\Phi(\nabla u)\sigma^-) \, \mathrm{d}x$$

$$= - \int_{\Gamma(t)} \Phi(q^-) DH(p^-) \cdot n\sigma^- \, \mathrm{d}\mathcal{H}^{n-1}.$$

Since $v = DH(p^+) \cdot n$ along the contact discontinuity, it follows that

$$\dot{\phi} = \int_{\Gamma(t)} \Phi(q^-)(DH(p^+) - DH(p^-)) \cdot n\sigma^- \, \mathrm{d}\mathcal{H}^{n-1}$$

$$+ \int_{\Gamma(t)} \nabla \Phi(q^+) \cdot \frac{Dq^+}{Dt} \sigma^+ + \Phi(q^+) \frac{D\sigma^+}{Dt} - \Phi(q^+)\sigma^+ v \mathbb{H} \, \mathrm{d}\mathcal{H}^{n-1}.$$

$$(6.15)$$

2. Next, we prove

$$\frac{D\sigma^+}{Dt} - v\mathbb{H}\sigma^+ = (DH(p^-) - DH(p^+)) \cdot n\sigma^-, \qquad (6.16)$$

a generalization of (6.4).

To do so, first fix a small number $h > 0$. Then select any small smooth region $\Delta(t) \subset \Gamma(t)$. We evolve $\Delta(t)$ into smooth surfaces $\Delta(s) \subset \Gamma(s)$ for times $t \leqq s \leqq t + h$, so that the velocity of each point along $\partial \Delta(s)$ is normal to $\Gamma(s)$. The surfaces $\{\Delta(s)\}_{t \leqq s \leqq t+h}$ sweep out the region $\Sigma(t, h) \subset \Gamma$ illustrated in Fig. 7.

Finally, from each point belonging to $\Sigma(t, h)$, we construct to the "left" the backwards characteristic with constant slope $(DH(p^-), 1)$. These, at time t_1, give us the region $R(t, h) \subset \mathbb{R}^n$ drawn in Fig. 7. Finally, we let $\Lambda(t, h)$ denote
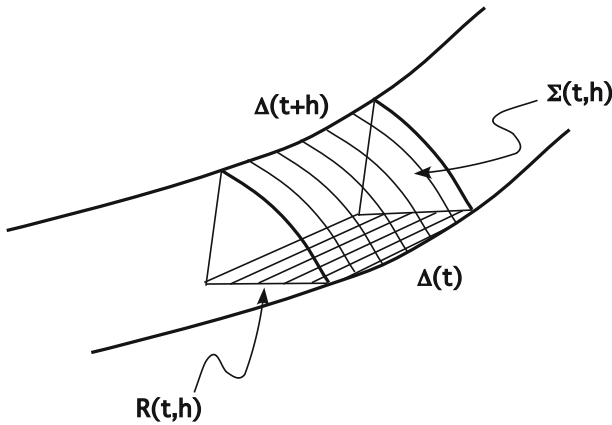
Fig. 7. The surfaces $R(t, h) \subset \mathbb{R}^n$ and $\Sigma(t, h) \subset \Gamma$

the $(n + 1)$-dimensional solid whose sides are $\Sigma(t, h)$, $R(t, h)$ and the unions of the backwards characteristics from $\partial \Sigma(t, h)$ over the time interval $[t, t + h]$. Since $\int_{\mathbb{R}^n} \, d\sigma = 1$ for each time,

$$\int_{R(t,h)} \sigma^- \, dx = \int_{\Delta(t+h)} \sigma^+ \, d\mathcal{H}^{n-1} - \int_{\Delta(t)} \sigma^+ \, d\mathcal{H}^{n-1}. \qquad (6.17)$$

We need, first, to rewrite the term on the left. To do so, note that $\sigma_t^- + \mathrm{div}(DH(Du)\sigma^-) = 0$ to the left of the shock, and therefore the Divergence Theorem implies

$$\int_{R(t,h)} \sigma^- \, dx = \int_{\Sigma(t,h)} (v^{n+1} + DH(p^-) \cdot \hat{v})\sigma^- \, d\mathcal{H}^n.$$

But since v is perpendicular to Γ and $(DH(p^+), 1)$ is tangent,

$$v^{n+1} + DH(p^+) \cdot \hat{v} = 0.$$

Consequently, owing to (6.9) we have

$$\int_{R(t,h)} \sigma^- \, dx = \int_{\Sigma(t,h)} (DH(p^-) - DH(p^+)) \cdot \hat{v}\sigma^- \, d\mathcal{H}^n$$

$$= \int_{\Sigma(t,h)} (DH(p^-) - DH(p^+)) \cdot n \frac{\sigma^-}{(1 + v^2)^{1/2}} \, d\mathcal{H}^n.$$

We differentiate with respect to h and then set $h = 0$:

$$\frac{d}{dh} \int_{R(t,h)} \sigma^- \, dx \bigg|_{h=0} = \int_{\Delta(t)} (DH(p^-) - DH(p^+)) \cdot n\sigma^- \, d\mathcal{H}^{n-1}. \quad (6.18)$$

Next, we apply the kinematic formula (6.13), with $\Delta(t)$ replacing $\Gamma(t)$, to deduce

$$\frac{d}{dt} \int_{\Delta(t)} \sigma^+ \, d\mathcal{H}^{n-1} = \int_{\Delta(t)} \frac{D\sigma^+}{Dt} - \sigma^+ v \mathbb{H} \, d\mathcal{H}^{n-1}. \qquad (6.19)$$

Note that there is no additional term involving an integral along $\partial\Delta(t)$, since the velocity there is normal to $\Gamma(t)$.

Using (6.18) and (6.19) in (6.17), we see that

$$\int_{\Delta(t)} (DH(p^-) - DH(p^+)) \cdot n\sigma^- \, d\mathcal{H}^{n-1} = \int_{\Delta(t)} \frac{D\sigma^+}{Dt} - \sigma^+ v\mathbb{H} \, d\mathcal{H}^{n-1}.$$

This implies (6.16), since the region $\Delta(t)$ is arbitrary.

3. Formula (6.16) lets us simplify (6.15):

$$\dot{\phi} = \int_{\Gamma(t)} (\Phi(q^-) - \Phi(q^+))(DH(p^+) - DH(p^-)) \cdot n\sigma^- \, d\mathcal{H}^{n-1}$$

$$+ \int_{\Gamma(t)} \nabla\Phi(q^+) \cdot \frac{Dq^+}{Dt}\sigma^+ \, d\mathcal{H}^{n-1}. \tag{6.20}$$

To go further, we need the final claim that

$$\frac{Dq^+}{Dt}\sigma^+ = (q^+ - q^-)(DH(p^+) - DH(p^-)) \cdot n\sigma^-, \tag{6.21}$$

which is a higher dimensional variant of (6.5).

We prove this by modifying as follows the calculations from Step 2 in the proof. Since $\int_{\mathbb{R}^n} \nabla u d\sigma$ is constant in time,

$$\int_{R(t,h)} \nabla u\sigma^- \, dx = \int_{\Delta(t+h)} q^+\sigma^+ \, d\mathcal{H}^{n-1} - \int_{\Delta(t)} q^+\sigma^+ \, d\mathcal{H}^{n-1}. \tag{6.22}$$

Because $(\sigma^- \nabla u)_t + \mathrm{div}(DH(Du)\nabla u\sigma^-) = 0$ in the region to the left of the shock, the Divergence Theorem implies

$$\int_{R(t,h)} \nabla u\sigma^- \, dx = \int_{\Sigma(t,h)} (v^{n+1} + DH(p^-) \cdot \hat{v})q^-\sigma^- \, d\mathcal{H}^n$$

$$= \int_{\Sigma(t,h)} (DH(p^-) - DH(p^+)) \cdot n\frac{q^-\sigma^-}{(1 + v^2)^{1/2}} \, d\mathcal{H}^n.$$

Hence

$$\frac{d}{dt} \int_{R(t,h)} \nabla u\sigma^- \, dx \bigg|_{h=0} = \int_{\Delta(t)} (DH(p^-) - DH(p^+)) \cdot nq^-\sigma^- \, d\mathcal{H}^{n-1}. \tag{6.23}$$

Also,

$$\frac{d}{dt} \int_{\Delta(t)} q^+\sigma^+ \, d\mathcal{H}^{n-1} = \int_{\Delta(t)} q^+ \left(\frac{D\sigma^+}{Dt} - \sigma^+ v\mathbb{H} \right) + \frac{Dq^+}{Dt}\sigma^+ \, d\mathcal{H}^{n-1}. \tag{6.24}$$

Using (6.23), (6.24) and (6.16) in (6.22), we deduce

$$\int_{\Delta(t)} (q^- - q^+)(DH(p^-) - DH(p^+)) \cdot n\sigma^- \, d\mathcal{H}^{n-1} = \int_{\Delta(t)} \frac{Dq^+}{Dt}\sigma^+ \, d\mathcal{H}^{n-1}.$$

Since the region $\Delta(t)$ is arbitrary, the identity (6.21) follows.

4. We employ (6.21) in (6.20), at last to conclude the proof of (6.12). $\quad\square$

7. Scalar conservation laws

Our kinetic formulation for Hamilton–Jacobi PDE is inspired by the corresponding approach to conservation laws due to LIONS, PERTHAME and TADMOR [17,18] (see also PERTHAME [19]), but our methods, when applied to the latter, differ. In this section we sketch out what our techniques give for a scalar conservation law in n space variables.

Given a smooth flux function $F : \mathbb{R} \to \mathbb{R}^n$, $F = F(z) = (F_1(z), \ldots, F_n(z))$, and $g : \mathbb{R}^n \to \mathbb{R}$, $g = g(x)$, we introduce the regularized scalar conservation law

$$\begin{cases} u_t^\varepsilon + (F_i(u^\varepsilon))_{x_i} = \varepsilon \Delta u^\varepsilon & \text{in } \mathbb{R}^n \times (0, \infty) \\ u^\varepsilon = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{7.1}$$

Given, next, a probability measure γ on $\mathbb{R}^n$ and a time $t_1 > 0$, we introduce the following terminal value problem

$$\begin{cases} -\sigma_t^\varepsilon - (F_i'(u^\varepsilon)\sigma^\varepsilon)_{x_i} = \varepsilon \Delta \sigma^\varepsilon & \text{in } \mathbb{R}^n \times [0, t_1) \\ \sigma^\varepsilon = \alpha & \text{on } \mathbb{R}^n \times \{t = t_1\}. \end{cases} \tag{7.2}$$

We have the estimates:

$$\sup_{\mathbb{R}^n \times [0,\infty)} |u^\varepsilon| \leqq C \quad \text{and} \quad \sigma^\varepsilon \geqq 0, \int_{\mathbb{R}^n} \sigma^\varepsilon \, dx = 1 \ (0 \leqq t < T). \tag{7.3}$$

Furthermore, $u^\varepsilon \to u$ strongly in L^p, where u is the unique entropy solution of the conservation law

$$\begin{cases} u_t + (F_i(u))_{x_i} = 0 & \text{in } \mathbb{R}^n \times (0, t_1] \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{7.4}$$

By analogy with Theorem 3.1, there exists a sequence $\varepsilon = \varepsilon_m \to 0$ and a non-negative measure μ on $\mathbb{R}^n \times [0, t_1] \times \mathbb{R}$ such that

$$\int_0^{t_1} \int_{\mathbb{R}^n} \Phi(x, t, u^\varepsilon)\sigma^\varepsilon \, dx \, dt \to \int_0^{t_1} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \Phi(x, t, z) \, d\mu \tag{7.5}$$

for all continuous, bounded functions Φ.

Theorem 7.1. (i) *We have the estimate*

$$\varepsilon \int_0^{t_1} \int_{\mathbb{R}^n} |Du^\varepsilon|^2 \sigma^\varepsilon \, dx \, dt \leqq C. \tag{7.6}$$

(ii) *There exists a nonnegative measure m on $\mathbb{R}^n \times [0, t_1] \times \mathbb{R}$ such that*

$$\mu_t + F_i'(z)\mu_{x_i} = -m_{zz}. \tag{7.7}$$

We call m the *dissipation measure*. Note that in our approach, m depends upon the measure α and the terminal time t_1. The implicit summation in (7.7) is for $i = 1, \ldots, n$. The transport equation (7.7) is an analog of (3.19), but is much simpler.

Proof. 1. Let $w^\varepsilon = \frac{1}{2}(u^\varepsilon)^2$. Then

$$w^\varepsilon_t + F'_i(u^\varepsilon)w^\varepsilon_{x_i} = \varepsilon\Delta w^\varepsilon - \varepsilon|Du^\varepsilon|^2.$$

We multiply by σ^ε and integrate by parts:

$$\varepsilon\int_0^{t_1}\int_{\mathbb{R}^n}|Du^\varepsilon|^2\sigma^\varepsilon\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^n}w^\varepsilon(x,t)\,\mathrm{d}\alpha = \int_{\mathbb{R}^n}w^\varepsilon(x,0)\sigma^\varepsilon(x,0)\,\mathrm{d}x.$$

This implies (7.6).

2. Let $\Phi = \Phi(x,t,z)$ be smooth and have compact support in $\mathbb{R}^n \times (0,t_1) \times \mathbb{R}$. Put

$$w^\varepsilon(x,t) := \Phi(x,t,u^\varepsilon).$$

Then

$$w^\varepsilon_t + F'_i(u^\varepsilon)w^\varepsilon_{x_i} = \varepsilon\Delta w^\varepsilon + \Phi_t + F'_i(u^\varepsilon)\Phi_{x_i} - \varepsilon\Phi_{zz}|Du^\varepsilon|^2 + E^\varepsilon$$

for

$$E^\varepsilon := \varepsilon\left(\Delta_x\Phi + 2\Phi_{x_iz}u^\varepsilon_{x_i}\right).$$

In these formulas ϕ is evaluated at (x,t,u^ε).

We multiply by σ^ε and integrate by parts, using (7.2) to simplify:

$$\int_0^{t_1}\int_{\mathbb{R}^n}(-\Phi_t - F'_i(u^\varepsilon)\Phi_{x_i})\sigma^\varepsilon\,\mathrm{d}x\,\mathrm{d}t + \varepsilon\int_0^{t_1}\int_{\mathbb{R}^n}\Phi_{zz}|Du^\varepsilon|^2$$
$$= \int_0^{t_1}\int_{\mathbb{R}^n}E^\varepsilon\sigma^\varepsilon\,\mathrm{d}x\,\mathrm{d}t.$$

Passing to limits as $\varepsilon = \varepsilon_m \to 0$ gives the weak formulation of (7.7). □

References

1. BARDI, M., CAPUZZO-DOLCETTA, I.: *Optimal Control and Viscosity Solutions of Hamilton–Jacobi-Bellman Equations*, with appendices by M. Falcone and P. Soravia, Birkhäuser, 1997
2. BALLOU, D.: Solutions to nonlinear hyperbolic Cauchy problems without convexity conditions. *Trans. Am. Math. Soc.* **152**, 441–460 (1970)
3. BAGNERINI, P., RASCLE, M., TADMOR, E.: Compensated compactness for 2D conservation laws. *J. Hyperbolic Differ. Equ.* **2**, 697–712 (2005)
4. CHRIST, M.: On linear functionals which are positive on convex functions (preprint)

5. Crandall, M.G., Evans, L.C., Lions, P.-L.: Some properties of viscosity solutions of Hamilton Jacobi equations. *Trans. Am. Math. Soc.* **282**, 487–502 (1984)
6. Crandall, M.G., Lions, P.-L.: Two approximations of solutions of Hamilton Jacobi equations. *Math. Comp.* **43**, 1–19 (1984)
7. Dafermos, C.: Regularity and large time behavior of solutions of a conservation law without convexity. *Proc. R. Soc. Edinb. Sect. A* **99**, 201–239 (1985)
8. Evans, L.C.: *Weak Convergence Methods for Nonlinear Partial Differential Equations*, CBMS 74, American Mathematical Society, 1990
9. Evans, L.C.: A survey of partial differential equations methods in weak KAM theory. *Comm. Pure Appl. Math.* **57**, 445–480 (2004)
10. Federer, H.: *Geometric Measure Theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer, 1969
11. Fleming, W.: The convergence problem for differential games II. *Advances in Game Theory*, Princeton University Press, Princeton, 195–210, 1964
12. Fleming, W.: The Cauchy problem for degenerate parabolic equations. *J. Math. Mech.* **13**, 987–1008 (1964)
13. Glimm, J., Kranzer, H.C., Tan, D., Tangerman, F.M.: Wave fronts for Hamilton–Jacobi equations: the general theory for Riemann solutions in $\mathbb{R}^n$. *Comm. Math. Phys.* **187**, 647–677 (1997)
14. Marson, A.: Nonconvex conservation laws and ordinary differential equations. *J. London Math. Soc.* (2) **69**, 428–440 (2004)
15. Lions, J.-L.: *Optimal Control of Systems Governed by Partial Differential Equations*, translated by S. K. Mitter, Springer, 1971
16. Lions, P.-L.: Identification du cone dual des fonctions convexes et applications. *C. R. Acad. Sci. Paris Sr. I Math.* **326**, 1385–1390 (1998)
17. Lions, P.-L., Perthame, B., Tadmor, E.: Kinetic formulation of the isentropic gas dynamics and p-systems. *Comm. Math. Phys.* **163**, 415–431 (1994)
18. Lions, P.-L., Perthame, B., Tadmor, E.: A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Amer. Math. Soc.* **7**, 169–191 (1994)
19. Perthame, B.: *Kinetic Formulation of Conservation Laws*. Oxford Lecture Series in Mathematics and its Applications, vol. 21. Oxford University Press, Oxford, 2002
20. Tartar, L.: Compensated compactness and applications to partial differential equations. *Nonlinear Analysis and Mechanics*, Heriot-Watt Symposium, Vol. IV, Research Notes in Math 39, Pitman, 136–212, 1979
21. Tran, H.V.: Paper to appear
22. Zheng, Y.: *Systems of Conservation Laws. Two-dimensional Riemann problems*. Progress in Nonlinear Differential Equations and their Applications, 38. Birkhäuser, 2001

Department of Mathematics,
University of California,
Berkeley, USA.
e-mail: evans@math.berkeley.edu