AN ENTROPY-BASED GLIMM-TYPE FUNCTIONAL

LAURA CARAVENNA

Abstract. We consider the Cauchy problem for a scalar conservation law in one space dimension

$$\begin{cases} u_t + f(u)_x = 0 \\ u(t=0) = \bar u \end{cases} \qquad \text{with } f \in \mathrm{C}^2(\mathbb{R},\mathbb{R}),\ \bar u \in \mathrm{BV}(\mathbb{R}),\ u(t,x) : \mathbb{R}^+ \times \mathbb{R} \mapsto \mathbb{R}.$$

We introduce, in this simple setting, a new Glimm-type interaction potential: the time marginal of the entropy dissipation measure of a uniformly convex entropy. We show that the Glimm estimates hold for this functional.

1. Introduction

Consider the Cauchy problem for the system of scalar conservation laws in one space dimension

$$u_t + f(u)_x = 0 \qquad f \in \mathrm{C}^2(\mathbb{R}^n, \mathbb{R}^n) \tag{1a}$$

$$u(t=0,\, x) = \bar u(x) \qquad u(t,x) : \mathbb{R}^+ \times \mathbb{R} \mapsto \mathbb{R}^n \tag{1b}$$

Even with C^∞ initial data, classical solutions generally break down in finite time. For finding global solutions, we need then to consider (1) in distributional sense. Nevertheless, while classical solutions lack of global existence, weak solutions lack of uniqueness: the problem is then to characterize the right one. Many conditions have been considered to single out the right solution; a particularly important one, inspired by the second law of thermodynamics, is the entropy condition.

A mathematical *entropy* is a function $\eta \in \mathrm{W}^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ for which there exist $q \in \mathrm{W}^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ such that

$$q' = \eta' f'. \tag{2}$$

Classical solutions satisfy $\eta(u)_t + q(u)_x = 0$ for free, but for weak solutions it is a too stringent requirement. A weak solution satisfies the *entropy condition* if, for all convex entropy η, we have

$$\eta(u)_t + q(u)_x \leq 0 \qquad \text{in the sense of distributions.} \tag{3}$$

We shall call *entropy dissipation measure* the opposite of the non-positive distribution in (3), which actually is a nonnegative Radon measure.

In the scalar case, when the flux is strictly convex several weaker conditions are sufficient, from the first one by Oleinik ([10]) up to (3) required for a single convex η instead of a whole family ([11], [6]). However, at the present, the one working also with non-convex fluxes and with systems is the entropy condition. In particular, existence and uniqueness have been proven in the scalar case by Kruzhkov, for L^∞ initial data ([9]). The generalization to system is not straightforward, since the flux

The author is grateful to professor Stefano Bianchini. She would like to thank him for having suggested the problem, and for his help.

$t \mapsto u(t)$ is no more contractive in the L^1 distance, and here the natural setting is the space BV. By means of a new functional, the *wave interaction functional*, Glimm was able to control the total strength of waves, and consequently achieve key estimates for compactness ([8]). This, in turn, allows to establish the well-posedness of the Cauchy problem for strictly hyperbolic systems in one space dimension, with small, BV initial data, when each characteristic field is either genuinely non-linear or linearly degenerate. In the subsequent research, other approximation schemes have been developed; in this paper, we will refer to the wave front-tracking algorithm ([2]). Each of them, nevertheless, requires a functional analogous to the Glimm's one as a basic mathematical tool for BV and stability estimates.

Aim of this paper is to suggest that the Glimm functional can actually be related to a physically meaningfull entity, namely the entropy. More precisely, we want to propose as a Glimm-type functional the time marginal of the entropy dissipation measure, w.r.t. a uniformly convex entropy. Although that functional is needed in the case of systems, here we focus only on the scalar case, since the constructions can possibly be generalized.

We introduce in Section 2 the functionals we are interested in. At first we briefly recall a classical wave interaction potential, $\mathcal{Q}$, from [4]. We propose then two new functionals, $\mathcal{D}$ and $\mathcal{E}$. $\mathcal{D}$ has a definition somehow similar to $\mathcal{Q}$'s one, but we point out that it can also be derived considering a uniformly convex entropy: it is the time marginal of the entropy dissipation measure. Unlikely, it has no semicontinuity property and it is neither monotone along entropy solutions, as $\mathcal{Q}$ was. Nevertheless, we can proceed as follows: inspired by the known functional $V + \kappa_0 \mathcal{Q}$, we define

$$\mathcal{E} = \mathrm{TV} - \kappa^{-1}\mathcal{D} \qquad\qquad \text{for a suitably small constant } \kappa^{-1}.$$

Section 3 deals mainly with $\mathcal{E}$. In spite of $\mathcal{D}$'s bad behavior, $\mathcal{E}$ satisfies the Glimm estimates presented in [4], as shown in Lemma 3.1. Moreover, it actually has the sought semicontinuity and monotonicity properties. More precisely, we give a proof of the following theorem.

Theorem. *The following statements hold.*

Semicontinuity*. Let* $\{u_\nu\}_\nu, u$ *be in* $\mathrm{BV}(\mathbb{R})$ *with total variation uniformly bounded. If* $u_\nu \to u$ *in* $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$*, then* $\mathcal{E}(u) \leq \liminf_\nu \mathcal{E}(u_\nu)$*.*

Monotonicity*. Consider the scalar, one dimensional conservation law*

$$u_t + f(u)_x = 0 \qquad\qquad \textit{with the flux } f \textit{ in } \mathrm{C}^2(\mathbb{R}).$$

The functional $\mathcal{E}$ *decreases along its entropy solutions with bounded variation.*

The guidelines for demonstrating the first point are from the similar proof, for $\mathcal{Q}$, in [3]. For the second point, we employ both semicontinuity and the front-tracking approximation. $\mathcal{E}$'s monotonicity is first shown, by direct computation, just on piecewise constant entropy solutions. After that, a suitable sequence approximating the initial data is chosen, for starting the front-tracking algorithm. Semicontinuity, finally, yields the thesis.

Notation 1.1. We will be dealing with real functions, of one or two variables. When they belong to $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$, they are always assumed to be right continuous. When they belong to $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^+ \times \mathbb{R})$ and solve (1a), their restriction at every time, which is $\mathrm{BV}_{\mathrm{loc}}$, is assumed to be right continuous.

2. The Functionals

In this section we recall a commonly used Glimm-type functional, $\mathcal{Q}$, referring to its expression in [4]. We define then two new functionals, $\mathcal{D}$ and $\mathcal{E} = \mathrm{TV} - \kappa^{-1}\mathcal{D}$. $\mathcal{D}$ has fewer good properties than $\mathcal{Q}$; in spite of that, it is intimately related to one strictly convex entropy and, however, it can play its role: the aim of the next section will be to show that, in the scalar case, $\mathcal{E}$ provides a Glimm-type functional. Namely, we will see that it decreases along the BV entropy solutions of scalar conservation laws in one space dimension. In order to prepare that, at the end of this section we state some elementary properties of the above functionals.

Notation 2.1. Let X be a disjoint union of countably many intervals, not necessarily open or bounded. Let $\mathrm{TV}(\cdot)$ denote the total variation and $\mathrm{BV}_1(X) := \{u \in \mathrm{BV}(X) : \mathrm{TV}(u) \leq 1\}$. The functionals $\mathcal{Q}$, $\mathcal{D}$, $\mathcal{E}$ will be defined over $\mathrm{BV}(X)$; nevertheless, when evaluating them on a family of $\mathrm{BV}(X)$ functions $\{u(t)\}_{t\in\mathbb{R}^+}$, we will rename them as functions from $\mathbb{R}^+$ to $\mathbb{R}$.

2.1. **The Glimm-type functional $\mathcal{Q}$.** Given $u \in \mathrm{BV}(X)$, we consider the variation of the product measure on X^2 given by $du(x)df(u)(y) - du(y)df(u)(x)$, denoted by $|\cdot|$. $\mathcal{Q}$ is defined as

$$\mathcal{Q}(u) := \frac{1}{4}|du(x)df(u)(y) - du(y)df(u)(x)|(X^2). \tag{4}$$

Remark 2.2. When one considers a piecewise constant function, $\mathcal{Q}$ takes an easy form: if $u = \sum_{i=1}^N u^i \chi_{(x_{i-1},x_i)}$, then

$$\mathcal{Q}(u) = \frac{1}{2}\sum_{1\leq i<j\leq N-1}\left|\begin{pmatrix} u^{i+1} - u^i \\ f(u^{i+1}) - f(u^i)\end{pmatrix} \wedge \begin{pmatrix} u^{j+1} - u^j \\ f(u^{j+1}) - f(u^j)\end{pmatrix}\right|.$$

This last one was its first definition, given in [4], and the one in (4) is just a l.s.c. extension in the weak topology.

Two fundamental properties of $\mathcal{Q}$ are recalled in the next theorems.

Theorem 2.3 (Semicontinuity)**.** *Let $\{u_\nu\}_\nu, u : X \mapsto \mathbb{R}$ be a sequence with uniformly bounded total variation. If $u_\nu \to u$ in $\mathrm{L}^1_{\mathrm{loc}}(X)$, then $\mathcal{Q}(u) \leq \liminf_\nu \mathcal{Q}(u_\nu)$.*

Proof. The proof is an adjustment of the one in [2], p. 203, for an analogous functional. □

Consider now the scalar conservation law

$$u_t + f(u)_x = 0 \qquad \text{with } f \in \mathrm{W}^{1,\infty}_{\mathrm{loc}}(\mathbb{R}), \tag{5}$$

and let $u(t,x)$ be an entropy solution. Given an initial data $u_0 \in \mathrm{BV}$, then $u(t)$ remains in BV for all the times ([2], Th. 6.1). In particular, we can consider $\mathcal{Q}(t) := \mathcal{Q}(u(t))$ and this function turns out to be monotone.

Theorem 2.4 (Monotonicity)**.** *Assume that $f : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz and let $u(t)$ be the unique entropy solution to* (5) *with initial data $u_0 \in \mathrm{BV}$. Then the map $t \mapsto \mathcal{Q}(u(t))$ is non-increasing in time.*

Proof. The proof follows from Th. 1 and 3 in [4]. One can moreover rearrange the proof in [2], p. 209, for an analogous functional. □

2.2. **The entropy–based functional $\mathcal{D}$.** Firstly, we now give a geometric definiton of the functional, similar to the one of $\mathcal{Q}$ and given by an easy formula (Def. 2.7). This is inspired by [5], Chap. 8.5. After an example, we will show that, when evaluated along the solutions of a scalar conservation law, this functional is related to a uniformly convex entropy. By that relation, we will arrive to an equivalent, more meaningful definition of $\mathcal{D}$ (Eq. 12).

Assume that $f : [-1, 1] \mapsto \mathbb{R}$ is a continuous function.

Definition 2.5. The *convex envelope* of f in a subinterval $[a, b]$ is defined as

$$f_*^{[a,b]} := \sup \{ \ g \mid g : [a, b] \mapsto \mathbb{R} \text{ convex}, \ g \leq f \ \};$$

Similarly, the *concave envelope* of f in a subinterval $[a, b]$ is defined as

$$f^*_{[a,b]} := \inf \{ \ g \mid g : [a, b] \mapsto \mathbb{R} \text{ concave}, \ g \geq f \ \}.$$

Definition 2.6. Define the positive function $\mathcal{A} : [-1, 1]^2 \mapsto \mathbb{R}$ as follows: (Fig. 1)

- if $a \leq b$, as the area between f and its convex envelope in $[a, b]$;
- if $a > b$, as the area between the concave envelope in $[b, a]$ and f itself:

Namely, $\mathcal{A}(a, b) := \chi_{\{a<b\}} \int_a^b \left(f - f_*^{[a,b]}\right) + \chi_{\{b<a\}} \int_b^a \left(f^*_{[b,a]} - f\right)$.

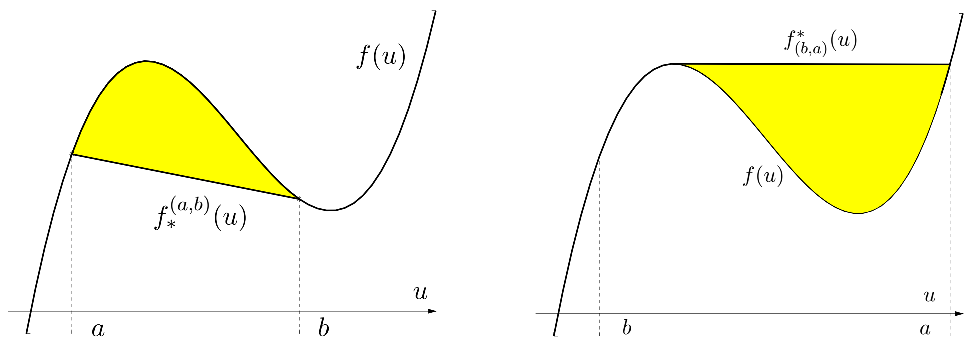

FIGURE 1. The area represents the value of $\mathcal{A}(a, b)$.

Definition 2.7. Given a function $u \in \mathrm{BV}(X)$, it has at most a countable points of discontinuity $\{x_i\}_i$; denote by $\{u_i^-, u_i^+\}_i$ the left and right limits at that points of jump. We define (Fig. 2)

$$\mathcal{D}(u) := \sum_i \mathcal{A}(u_i^-, u_i^+). \tag{6}$$

Consider now the scalar conservation law

$$u_t + f(u)_x = 0 \qquad \text{with } f \in \mathrm{W}^{1,\infty}_{\mathrm{loc}}(\mathbb{R}). \tag{7}$$

and let $u(t, x)$ be an entropy solution. As well as for $\mathcal{Q}$, we are interested in the behavior of the map

$$t \mapsto \mathcal{D}(u(t)).$$

Given an initial data $\bar{u} \in \mathrm{BV}$, then $u(t)$ remains in BV for all the times ([2], Th. 6.1). In particular, we are actually allowed to consider $\mathcal{D}(t) := \mathcal{D}(u(t))$. Before investigating its meaning, let's observe its behavior, and compare it with $\mathcal{Q}$'s one, with an example.

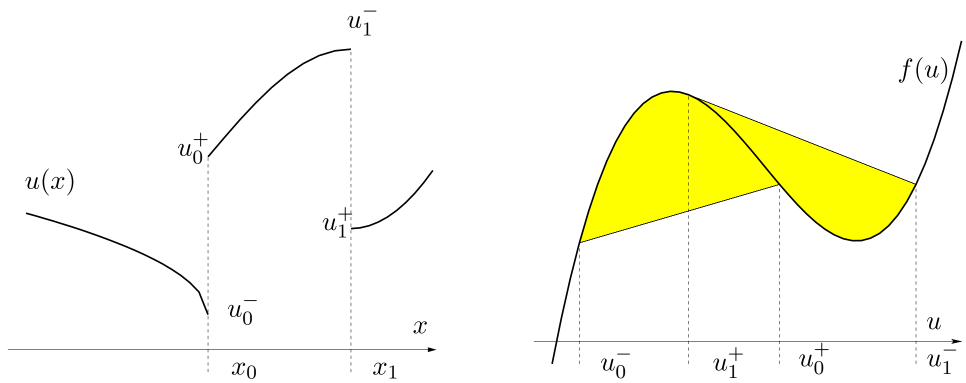

FIGURE 2. On the left we have drawn a function $u(x)$, on the right the shadowed area corresponds to the value $\mathcal{D}(u)$.

Example 2.8. Consider to have a piecewise constant function with three approaching jumps, say between u_0, u_1, u_2 and u_3, which simultaneously collide. At first $\mathcal{D}$'s value is $\mathcal{A}(u_0, u_1) + \mathcal{A}(u_1, u_2) + \mathcal{A}(u_2, u_3)$, while at the end is $\mathcal{A}(u_0, u_3)$. If the interactions are monotone, meaning that $(u_{i+1} - u_i)(u_{j+1} - u_j) > 0$, then it increases as much as $\mathcal{Q}$ decreases: of the shadowed area. Otherwise it is more difficult: $\mathcal{D}$ can decrease, and its variation is in general different from the one of $\mathcal{Q}$.

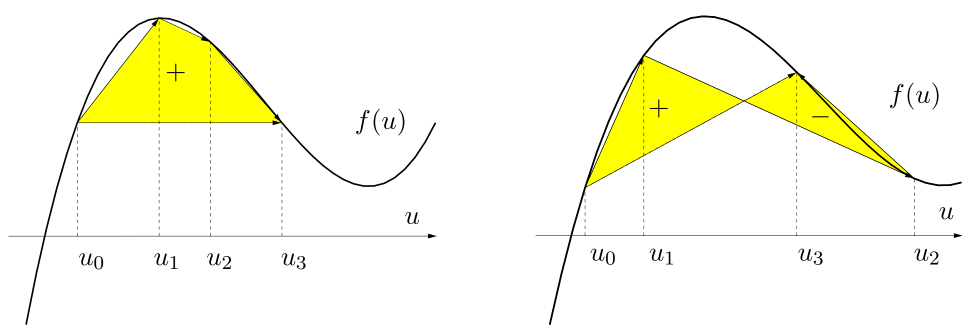

FIGURE 3. Variation of $\mathcal{D}$: on the left a monotone interaction, on the right an interaction with cancellation. $\mathcal{D}$ increases of the area with $+$, decreases with $-$.

Consider now the following definition, similar to the one in [5], p. 76.

Definition 2.9. The *entropy dissipation* is the nonnegative Radon measure

$$\mu := -(\eta(u)_t + q(u)_x)\,. \tag{8}$$

When the initial data $\bar{u}$ is in BV, then $u(t, x)$ is in $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^2)$ ([2], almost Th. 6.1). By the entropy inequality, the distribution $\eta(u)_t + q(u)_x$ is non-positive, thus in particular it is a Radon measure: the definition above thus is well posed. As a consequence of the BV structure, joined with the conservation law (7), we get the following, more concrete expression of μ.

Lemma 2.10. *Given an initial data $\bar{u} \in \mathrm{BV}$, let $u = u(t,x)$ be the unique entropy solution of* (7)*. Given any locally Lipschitz entropy pair (η, q) with $\eta(0) = q(0) = 0$, we can then write*

$$\mu = \sum_i \begin{pmatrix} \eta(u^+)-\eta(u^-) \\ q(u^+)-q(u^-) \end{pmatrix} \cdot \nu_i \, d\mathcal{H}^1 \restriction_{\Gamma_i}, \tag{9}$$

where

- *$\{\Gamma_i\}_{i\in\mathbb{N}}$, pairwise disjoint sets in the (x, t) plane, are the images of C^1, $||f'||_\infty$-Lipschitz curves $(\gamma_i(t), t)_{t\in I_i}$ parametrized with time;*
- *ν_i is the unit vector field orthogonal to Γ_i;*
- *$u^\pm = u^\pm(t, \gamma_i(t))$ are such that, for every $t \in I_i$,*

$$\lim_{\rho\to 0} \fint_{B_\rho^{\pm\nu_i}(\gamma_i(t),\, t)} |u - u^\pm| = 0\,, \tag{10}$$

  where $B_\rho^{\pm\nu}(x,t)$ are the half balls in which, respectively, $\pm\nu \cdot \binom{x'-x}{t'-t} \geq 0$.
- *$\mathcal{H}^1$-a.e. where $u^+ \neq u^-$ we have, in the plane (x, t),*

$$\nu = \frac{1}{\sqrt{1+\lambda^2}} \begin{pmatrix} 1 \\ -\lambda \end{pmatrix}, \qquad \lambda = \frac{f(u^+) - f(u^-)}{u^+ - u^-}.$$

Remark 2.11. The set $\cup_i \Gamma_i$ cover the jump set of $\eta(u)$, but it is generally larger. Nevertheless, as we are going to prove, the above expression still holds, due to the coefficients of $\mathcal{H}^1 \restriction_{\Gamma_i}$. In fact, they vanishes $\mathcal{H}^1$-a.e. outside the jump set.

Proof. We recall first of all the regularity due to u's membership in BV. In fact (see Th. 6.1 in [2]), $\forall t > 0$ we have that $u \in \mathrm{BV}((0,t) \times \mathbb{R}; \mathbb{R})$. By the general theory of BV functions (Federer-Vol'pert Th. 3.78 in [1]), we can then partition the domain into two Borel sets, $\mathcal{C}$ and $\mathcal{J}$, such that:

- u is approximately continuous at $\mathcal{H}^1$-a.e. every point (t, x) of $\mathcal{C}$, meaning that there exists u^0 s.t. $\lim_{\rho\to 0} \fint_{\{|t'-t|+|x'-x|<\rho\}} |u(t', x') - u^o| \,\mathrm{d}t' \,\mathrm{d}x' = 0$;
- every point of $\mathcal{J}$ is of approximate jump discontinuity, meaning that there exist $\tilde{\nu}(t, x), \tilde{u}^\pm(t, x)$ such that $\lim_{\rho\to 0} \fint_{B_\rho^{\pm\tilde{\nu}_i}(t,x)} |u(y) - \tilde{u}^\pm(t, x)| \,\mathrm{d}y = 0$, where $B_\rho^{\pm\tilde{\nu}}(x)$ are the semi-balls in which, respectively, $\pm\tilde{\nu} \cdot (y - x) > 0$. Moreover, $\mathcal{J}$ is countably $\mathcal{H}^1$-rectifiable.

By rectifiability, there exist a sequence of disjoint C^1 manifolds $\{\Gamma_i\}_i$ such that $\mathcal{H}^1(\mathcal{J} \setminus \cup_i \Gamma_i) = 0$ ([7], Th. 3.1.15 p. 227 and Th. 3.2.18 p. 255). Denote by $\nu(x) = \{\nu_i(x)\}_i$ be the normal fields to that manifolds and by $u^\pm$ the traces of u on it ([1], Th. 3.77). It should be $\nu(x) = \tilde{\nu}(x)$ and $\tilde{u}^\pm = u^\pm$, for $\mathcal{H}^1$-a.e. x, since the triple $(u^\pm, \nu)$ for which (10) holds is unique up to a permutation of $u^\pm$ and a change of sign of ν.

By Vol'pert chain rule ([1], Th. 3.99), η shares the same structure: $\eta(u)$, $q(u)$ are $\mathrm{BV}((0,t) \times \mathbb{R}; \mathbb{R})$ and

$$D\eta(u) = \eta'(u)\nabla u \mathcal{L}^2 + \eta'(u) d_c u + (\eta(u^+) - \eta(u^-))\nu d\mathcal{H}^1 \restriction_{\mathcal{J}}, \tag{11}$$

where the measure $d_c u$ vanishes on Hausdorff 1-dimensional sets; for q we have an analogous expression. In particular, the jump part of η's distributional derivative is supported on $\mathcal{J}$ and it is equal to $(\eta(u^+) - \eta(u^-))\, \nu \, d\mathcal{H}^{n-1}(x) \restriction_{\mathcal{J}}$. Notice now that you can see an approximate continuity point as a special jump point where $u^+ = u^-$, thus in those points the coefficient $\eta(u^+) - \eta(u^-)$ vanishes; since, moreover, u is approximately continuous at $\mathcal{H}^1$-a.e. point outside $\mathcal{J}$ and $\mathcal{H}^1(\mathcal{J} \setminus \cup_i \Gamma_i) =$

0, it is clear that its distributional derivative can be written also as $(\eta(u^+) - \eta(u^-))\nu(x)\, d\mathcal{H}^{n-1}(x) \restriction_{\cup_i \Gamma_i}$. For q it is similar.

The fact that u solves (7) carries additional properties. By the Rankine–Hugoniot condition (Eq (4.8) in [2]), when $u^+ \neq u^-$, then ν gives the same positive direction as $\binom{1}{-\lambda}$, with $\lambda = \frac{f(u^+)-f(u^-)}{u^+-u^-}$, $|\lambda| \leq ||f'||_\infty$. This implies that each Γ_i can be parametrized with time: it is the image of a C^1, $||f'||_\infty$-Lipschitz curve $(\gamma_i(t), t)_{t\in I_i}$ and, for $\mathcal{L}^1$-a.e. t, we have $\frac{d}{dt}(\gamma_i(t),\, t)$ orthogonal to $\nu((t, \gamma_i(t)))$.

Since u is a weak solution of $u_t + f(u)_x = 0$, we can see that the measure $[\partial_t u + \partial_x f(u)]\mathcal{L}^1 + [d_c^t u + d_c^x f(u)]$ vanishes. By the relation $q' = \eta' f'$ and others similar to (11), we get then

$$\begin{aligned}&[\partial_t \eta(u) + \partial_x q(u)]\mathcal{L}^2 + [d_c^t \eta(u) + d_c^x q(u)] = \\ &\quad \eta'(u)\big\{[\partial_t u + \partial_x f(u)]\mathcal{L}^1 + [d_c^t u + d_c^x f(u)]\big\} = 0.\end{aligned}$$

Because of this, in the entropy dissipation measure survives only the jump part:

$$\begin{aligned}\mu = \eta(u)_t + q(u)_x &\equiv \begin{pmatrix} \eta(u^+)-\eta(u^-) \\ q(u^+)-q(u^-)\end{pmatrix} \cdot \nu\, d\mathcal{H}^{n-1} \restriction_{\mathcal{J}} \\ &\equiv \sum_i \begin{pmatrix} \eta(u_i^+)-\eta(u_i^-) \\ q(u^+)-q(u^-)\end{pmatrix} \cdot \nu_i\, d\mathcal{H}^{n-1} \restriction_{\Gamma_i} . \qquad \square\end{aligned}$$

In particular, notice that μ's t-marginal is thus absolutely continuous w.r.t. the Lebesgue measure $\mathcal{L}^1$: let's investigate more its density.

Notation 2.12. We fix, from now on, the uniformly convex entropy η, and a corresponding entropy flux q, as

$$\eta(u) := \frac{u^2}{2}\,, \qquad q(u) := uf(u) - \int^u f(v)\, dv\,.$$

As noticed in [5], p. 221, a short computation gives the equality

$$\begin{aligned}\begin{pmatrix} \eta(u_i^+)-\eta(u_i^-) \\ q(u_i^+)-q(u_i^-)\end{pmatrix} \cdot \nu_i &= \dot\gamma_i\big[\eta(u_i^+) - \eta(u_i^-)\big] - \big[q(u_i^+) - q(u_i^-)\big] \\ &= \int_{u_i^-}^{u_i^+} f - \frac{f(u_i^+) + f(u_i^-)}{2}(u_i^+ - u_i^-)\,.\end{aligned}$$

By the chord condition (see (8.4.3) in [5]), f between u_i^- and u_i^+ has to stay on the right side of the chord with these two endpoints. The right hand side thus is exactly $\mathcal{A}(u_i^-, u_i^+)$, arriving at the expression

$$\sum_i \big[\mathcal{A}(u^+(\gamma_i(t), t),\, u^-(\gamma_i(t), t))\, d\mathcal{L}^1(t) \restriction_{I_i}\big].$$

This, by the first definition of $\mathcal{D}$, coincides with

$$\mathcal{D}(u(t))\, d\mathcal{L}^1(t) \restriction_{\mathbb{R}+} .$$

Actually, we see that, in this setting, one can define $\mathcal{D}$, as well, as the density of μ's t-marginal: we just found

$$\mu([r, s] \times \mathbb{R}) = \int_r^s \mathcal{D}(t)dt. \tag{12}$$

$\mathcal{D}$ will turn out to be a BV function and its first expression will give us its right continuous representative in the L^1 class (Corollary 3.7).

Remark 2.13. The functional $\mathcal{D}$ is neither lower, nor upper semicontinuous with respect to $\mathrm{L}^1_{\mathrm{loc}}(X)$ convergence. Take for example $X = [-1, 1]$, the flux $f(u) = 3u^2$ and consider the sequences:

- $u_k := \chi_{[-1,-\frac{k}{k+1})}$, or, as well, $u_k := \chi_{[0,\frac{1}{k})}$ and $u := 0$.
  It holds $\mathcal{D}(u_k) = 1/2 > 0 = \mathcal{D}(u)$, thus $\mathcal{D}(u) < \liminf_k \mathcal{D}(u_k)$.
- $u_k := \chi_{[-1,0)} + (1 - kx)\chi_{[0,\frac{1}{k})}$, $u := \chi_{[-1,0)}$.
  It holds $\mathcal{D}(u) = 1/2 > 0 = \mathcal{D}(u_k)$, thus this time $\mathcal{D}(u) > \liminf_k \mathcal{D}(u_k)$.

Notice that in the first case the total variation strictly decreases from 1 to 0: we'll see (in Theorem 3.3) that you can't have $\mathcal{D}(u) < \liminf_k \mathcal{D}(u_k)$ without loosing total variation.

2.3. **The functional $\mathcal{E}$.** In the previous subsection we introduced the functional $\mathcal{D}$. In Example 2.8, we noticed that $\mathcal{D}$ is not monotone along solutions of a scalar conservation law, while $\mathcal{Q}$ instead is. Moreover, it is not even semicontinuous with respect to $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R})$ convergence (Rem. 2.13), while $\mathcal{Q}$ is. Nevertheless, the above properties are shared by the following functional $\mathcal{E}$. This is what, in the next section, we're going to suggest as an entropy-based, Glimm-type functional.

Definition 2.14. We define the positive functional $\mathcal{E} : \mathrm{BV}_1 \mapsto \mathbb{R}$ as

$$\mathcal{E} := \mathrm{TV} - \kappa^{-1}\mathcal{D}, \qquad \text{with } \kappa^{-1} = \frac{3}{||f''||_\infty}.$$

Even though we could defined $\mathcal{E}$ on the whole BV, we are interested in considering it just in BV_1, since the above properties will hold in this space. Suitably changing the constant κ^{-1} with $T\kappa^{-1}$, it would be equivalent to consider BV_T.

Notation 2.15. Consider a piecewise constant function $u = \sum_{i=0}^q u^i \chi_{[x_i,x_{i+1})}$, with $\inf_X = x_0 < \cdots < x_{q+1} = \sup_X$. We define then $\mathcal{E}(u^0, \ldots, u^q) := \mathcal{E}(u)$. We recall, moreover, that BV_T denotes $\{u \in \mathrm{BV} : \mathrm{TV}(u) \le T\}$.

2.4. **Simple properties.** In the present subsection we are going to underline simple analytic properties of $\mathcal{D}$ (Lemma 2.19, Remark 2.21). They are mainly based on areas' estimates, which now are presented (2.16, 2.17, 2.18).

Remark 2.16. By the trapezoidal interpolation formula ([12], Page 375, Eq. 9.11), the area between the graph of f and a chord on it with projection of length p, is less or equal to $\frac{||f''||_\infty}{12}|p|^3$. Notice moreover that f_* (resp. f^*) can differ from f only on intervals and there it coincides with the chord on f between the ends of the interval. As a straightforward consequence of the inequality $|x|^3 + |y|^3 \le (|x| + |y|)^3$, we get thus that precisely the same estimate holds for the area between f and its convex (or concave) envelope, in an interval of length p: this is to say

$$\mathcal{A}(a, b) \le \frac{||f''||_\infty}{12}|b - a|^3\,.$$

From that, we get also that the area between the concave and the convex envelope of f in an interval of length p is dominated by $\frac{||f''||_\infty}{6}|p|^3$.

Moreover, we can easily estimate also the area of a triangle with vertices on the graph of f, as in the following lemma.

Lemma 2.17. *The area of a triangle with vertices on the graph of f is less or equal to*

$$\frac{||f''||_\infty}{4} p_1 p_2 (p_1 + p_2)\,,$$

where p_1, p_2 are the length of the shorter projections of the sides.

Proof. After a suitable translation, we can assume that the vertices are $(-p_1,\, f(-p_1))$, $(0,\, 0)$, $(p_2,\, f(p_2))$. The area then will be $\frac{1}{2}|p_2 f(-p_1) + p_1 f(p_2)|$; using Mac-Laurin's expansion we get, for twice the area,

$$
\begin{aligned}
&|p_2 f(-p_1) + p_1 f(p_2)| \\
&= \left| p_2 \left[-f'(0)p_1 + \frac{f''(\xi_1)}{2} p_1^2 \right] + p_1 \left[f'(0) p_2 + \frac{f''(\xi_2)}{2} p_2^2 \right] \right| \\
&= \frac{1}{2} |f''(\xi_1) p_2 p_1^2 + f''(\xi_2) p_1 p_2^2| \\
&\leq \frac{||f''||_\infty}{2} (p_1 + p_2) p_1 p_2 \,. \qquad \square
\end{aligned}
$$

Finally, we have an estimate of the difference between the areas corresponding to different intervals in terms of the difference between their extremes:

Lemma 2.18. *Suppose $|a - a'| \leq \varepsilon$. Then it holds*

$$
|\mathcal{A}(a,\, b) - \mathcal{A}(a',\, b)| \leq \frac{||f''||_\infty}{4} |b - a'|\, |b - a| \varepsilon + \frac{||f''||_\infty}{12} \varepsilon^3 \,.
$$

Proof. It suffices to consider the case $a < a' < b$. The difference between the two areas is now given on the one hand by $\mathcal{A}(a,\, a')$, on the other hand by the area between the convex envelope f_* of f in $(a,\, b)$ and the curve $\tilde{f}$ made of the two convex envelopes of f in the subintervals $(a,\, a')$ and $(a',\, b)$ (Fig. 4). The first area, by Remark 2.16, is $\mathcal{A}(a,\, a') \leq \frac{||f''||_\infty}{12} \varepsilon^3$. For the second area, just observe that can be contained in a triangle with one vertex in $(a',\, f(a'))$ and the others on f, as well, with projection respectively in $[a,\, a']$ and in $[a',\, b]$: by Lemma 2.17, the area is less or equal to $\frac{||f''||_\infty}{4} |b - a'|\, |b - a| \varepsilon$. $\square$

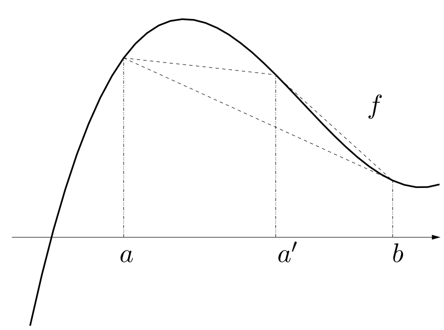

FIGURE 4. Picture of Lemma 2.18

Lemma 2.19. *Let $\kappa = \frac{||f''||_\infty}{3}$. The functional $\mathcal{A} : [-1,1]^2 \mapsto \mathbb{R}^+$ satisfies*

$i_1)$ ***Superadditivity:*** *for all $a, b, c \in [-1,1]$ it holds*

$$
|c - a| = |c - b| + |b - a| \implies \mathcal{A}(a,c) \geq \mathcal{A}(a,b) + \mathcal{A}(b,c). \tag{13}
$$

i_2) ***Lipschitz condition:*** *for all* $a, b, c \in [-1, 1]$ *it holds*

$$|\mathcal{A}(a, c) - \mathcal{A}(a, b)| \leq \kappa |b - c|. \tag{14}$$

Thus for all $\underline{w}$, $\underline{w}'$ *in* $[-1, 1]^2$ *it holds*

$$|\mathcal{A}(\underline{w}) - \mathcal{A}(\underline{w}')| \leq |w_1 - w_1'| + |w_2 - w_2'| \leq \sqrt{2}\kappa ||\underline{w}' - \underline{w}||. \tag{15}$$

i_3) ***Superlinearity:*** $\exists \sigma \in \mathrm{C}^0(\mathbb{R}^+; \mathbb{R}^+)$ *such that for all* $a, b \in [-1, 1]$ *it holds*

$$\mathcal{A}(a, b) \leq \sigma(|b - a|)|b - a| \quad \text{with} \quad \sigma(0) = 0, \tag{16}$$

$$\sigma(r) \quad \text{and} \quad (1 - \kappa^{-1}\sigma(r))r \quad \text{nondecreasing for } r \in [0, 1]. \tag{17}$$

Proof. By definition, whenever $a \leq b \leq c$ we have $f_*^{(a,b)} \geq f_*^{(a,c)}{}_{\restriction(a,b)}$ and, as well, $f^*_{(a,b)} \leq f^*_{(a,c)}{}_{\restriction(a,b)}$; from this i_1) follows, since the other possible configuration, $c \leq b \leq a$, is completely analogous. The proof of i_2) is similar to the one of Lemma 2.18. Consider finally $\sigma(r) = \frac{||f''||_\infty}{12} r^2 = \kappa \frac{r^2}{4}$, then Remark 2.16 gives exactly (16) in i_3); moreover, since for $r \in [0, 1]$ both $\kappa \frac{r^2}{4}$ and $r - \frac{r^3}{4}$ are monotone increasing, we obtain also (17). □

Remark 2.20. We have the bounds

$$\mathcal{A}(w, z) \overset{(13)}{\leq} \mathcal{A}(v, z) - \mathcal{A}(v, w) \overset{(14)}{\leq} \kappa |z - w|. \tag{18}$$

It follows immediately that, when restricted to $\{|z - w| \leq 1\}$, the functionals $\mathcal{D}$ and $\mathcal{A}$ are bounded: $0 \leq \mathcal{D} \leq \kappa$ and $0 \leq \mathcal{E} \leq 1$.

Remark 2.21. It is really important to notice that $\mathcal{D}$ *remains constant solving a Riemann problem*: $\mathcal{D}(t) \equiv \mathcal{D}(0) \equiv \mathcal{A}(u^-, u^+)$. In particular, when evaluated on a piecewise constant entropy solutions of a scalar conservation law in one space variable, it is a right continuous function of the time, as well as TV and $\mathcal{E}$.

3. $\mathcal{E}$ as a Glimm-Type Functional

The aim of this section is to show that $\mathcal{E}$ decreases along the exact solution u of a scalar conservation law, namely that $\mathcal{E}(u(t_2)) - \mathcal{E}(u(t_1)) \leq 0$ when $0 \leq t_1 \leq t_2$ (Theorem 3.6). For this purpose, at first it is useful to know its semicontinuity w.r.t. $\mathrm{L}^1_{\mathrm{loc}}$ convergence (Theorem 3.3). Secondly, for the lack of continuity, we need to approximate a given $\mathrm{L}^1_{\mathrm{loc}}$ function u with a sequence $\{u_k\}_k$ such that $\mathcal{E}(u_k)$ converges to $\mathcal{E}(u)$ (Lemma 3.4). At that point the thesis will rely on showing that the functional decreases on the piecewise constant front-tracking approximations of u, as shown in Lemma 3.5. In order to do this, we have to see that, collapsing all the jumps of a piecewise constant function in one single jump, $\mathcal{E}$ does not increase (Lemma 3.1).

Lemma 3.1. *Let J be an open real interval and $u : J \mapsto \mathbb{R}$ a piecewise constant function with* $\mathrm{TV}(u) \leq 1$. *Consider a function $\hat{u} : J \mapsto \mathbb{R}$ with only one jump from the first to the last value of u. Then we have that $\mathcal{E}(\hat{u}) \leq \mathcal{E}(u)$.*

Proof. Let $u_0, \ldots, u_q$ be the values of u, from $-\infty$ to $+\infty$, u_0, u_q will be the values of $\hat{u}$. Consider at first the three values u_0, u_{q-1}, u_q: let's prove $\mathcal{E}(u_0, u_q) \leq \mathcal{E}(u_0, u_{q-1}, u_q)$. All the possible cases, considering similar problems with suitably reflected flux and initial data, can be reduced to the following ones (Fig. 5):

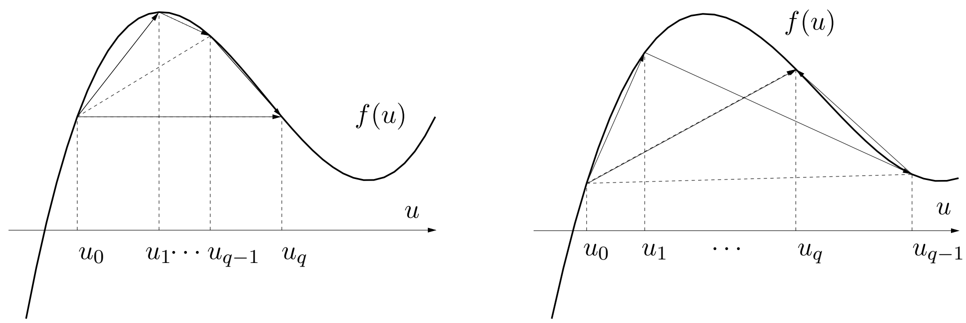

FIGURE 5. Collapsing the jumps $\mathcal{E}$ does not increase.

- $u_0 \leq u_{q-1} < u_q$. By jumps' monotonicity, the total variation remains constant. Moreover, by $\mathcal{A}$'s superadditivity,

$$\mathcal{D}(u_0, u_q) = \mathcal{A}(u_0, u_q) \geq \mathcal{A}(u_0, u_{q-1}) + \mathcal{A}(u_{q-1}, u_q) = \mathcal{D}(u_0, u_{q-1}, u_q). \tag{19}$$

We gain thus $\mathcal{E}(u_0, u_q) \leq \mathcal{E}(u_0, u_{q-1}, u_q)$.
- $u_0 \leq u_q < u_{q-1}$. Due to cancellation, we have that the total variation decreases of $2(u_{q-1} - u_q)$. Nevertheless, $\mathcal{D}$ can decrease, too. To see it, remember that $\mathcal{D}(u_0, u_q) \equiv \mathcal{A}(u_0, u_q)$, even if we have other jumps in between. Then, by the structure of $\mathcal{A}$, $\mathcal{D}$'s decrease is well controlled by the size of the jump cancelled:

$$\begin{aligned}\mathcal{D}(u_0, u_q) - \mathcal{D}(u_0, u_{q-1}, u_q) &\overset{R.2.21}{=} \mathcal{A}(u_0, u_q) - \mathcal{A}(u_0, u_{q-1}) - \mathcal{A}(u_{q-1}, u_q) \\ &\overset{(18)}{\geq} -\kappa(u_{q-1} - u_q) - \kappa(u_{q-1} - u_q) = -2\kappa(u_{q-1} - u_q).\end{aligned}$$

Adding the total variations, we reach again $\mathcal{E}(u_0, u_q) \leq \mathcal{E}(u_0, u_{q-1}, u_q)$.

If $q = 2$, we just proved the claim. If $q > 2$, using an induction hypothesis on the first $q - 1$ jumps we have $\mathcal{E}(u_0, u_{q-1}) \leq \mathcal{E}(u_0, u_1, \ldots, u_{q-1})$. Now,

$$\begin{aligned}\mathcal{E}(u_0, u_{q-1}, u_q) &= \mathcal{E}(u_0, u_{q-1}) + \mathcal{E}(u_{q-1}, u_q) \\ &\leq \mathcal{E}(u_0, u_1, \ldots, u_{q-1}) + \mathcal{E}(u_{q-1}, u_q) = \mathcal{E}(u).\end{aligned}$$

Since we proved $\mathcal{E}(u_0, u_q) \leq \mathcal{E}(u_0, u_{q-1}, u_q)$, we get $\mathcal{E}(\hat{u}) \leq \mathcal{E}(u)$. □

Remark 3.2. Compare (19) with the estimates in [4] and in [2]. We see that, for interactions with non-decreasing total variation, the decrease of $\kappa\mathcal{E}$ is the same as the decrease of $\mathcal{Q}$.

Let's show now the semicontinuity of $\mathcal{E}$.

Theorem 3.3 (lower semicontinuity)**.** *Let* $\{u_\nu\}_\nu, u$ *be in* $\mathrm{BV}_1(X)$*. If* $u_\nu \to u$ *in* $\mathrm{L}^1_{\mathrm{loc}}(X)$*, then* $\mathcal{E}(u) \leq \liminf_\nu \mathcal{E}(u_\nu)$*.*

Proof. The guideline of this proof is from [3]. Let's suppose first that $u_\nu \to u$ in $\mathrm{L}^1(X)$. Rename u_ν to be a subsequence $\{\tilde{u}_\nu\}_\nu$ such that $\liminf_\nu \mathcal{E}(u_\nu) = \lim_\nu \mathcal{E}(\tilde{u}_\nu)$. Up to a new subsequence, we can suppose to have pointwise convergence at every point of continuity of all $u, \{u_\nu\}_\nu$, thus outside the countable set of their jumps. Fix $\varepsilon > 0$. Let $J := \{x_1, \ldots, x_{N(\varepsilon)}\}$ be such that each jump of u outside J has size less than $\frac{\varepsilon}{2}$.

For every $\eta > 0$, you can choose $y_1^\ell < x_1 < y_1^r < \cdots < y_N^\ell < x_N < y_N^r$ such that, called $J_i := (y_i^\ell, y_i^r)$,

($\Diamond$) $u, \{u_\nu\}_\nu$ are continuous at y_i^ℓ, y_i^r for all i.
($\triangle$) $\mathrm{TV}(u(J_i \setminus \{x_i\})) < \eta$.

Let's write $u_{\nu,i}^\ell := u_\nu(y_i^\ell)$, $u_{\nu,i}^r := u_\nu(y_i^r)$, $u_i^\pm := u(x_i^\pm)$. We have the following estimates:

i_1) $\mathcal{E}(u \restriction_{J_i}) \leq \liminf_\nu \mathcal{E}(u_\nu \restriction_{J_i}) + 4\eta$, $\forall i = 1, \ldots, N$.
By condition ($\Diamond$), we have pointwise convergence of u_ν to u at y_i^ℓ, y_i^r: thus $\exists \bar{\nu}_i(\varepsilon, \eta)$ such that, for all $\nu > \bar{\nu}_i$, the ends of u's jumps at x_i are close to the values of u_ν at the ends of the subintervals:

$$|u_{\nu,i}^r - u_i^+| + |u_{\nu,i}^\ell - u_i^-| \leq \eta\,. \tag{20}$$

From that we have also

$$\big||u_{\nu,i}^r - u_{\nu,i}^\ell| - |u_i^+ - u_i^-|\big| \leq \eta\,. \tag{21}$$

Let's compare, in each J_i, our functions u, u_ν with intermediate terms given by auxiliary functions with only one jump from $u_{\nu,i}^\ell$ to $u_{\nu,i}^\ell$. Recall that, as shown in Lemma 3.1, collapsing jumps $\mathcal{E}$ does not increase: thus $\mathcal{E}(u_{\nu,i}^\ell, u_{\nu,i}^r) \leq \mathcal{E}(u_\nu \restriction_{J_i})$ for all ν. Let's consider now u, first look at the total variation. By the choice of y_j^ℓ, y_j^r

$$\begin{aligned} \mathrm{TV}(u \restriction_{J_i}) &\overset{(\triangle)}{\leq} |u_i^+ - u_i^-| + \eta \\ &\overset{(21)}{\leq} |u_{\nu,i}^r - u_{\nu,i}^\ell| + 2\eta. \end{aligned} \tag{22}$$

Let's look now $\mathcal{D}$. By (20) and by Lipschitz condition of $\mathcal{A}$ it holds

$$\begin{aligned} \mathcal{D}(u \restriction_{J_i}) &\geq \mathcal{A}(u_i^-,\, u_i^+) \\ &\overset{j_2)+(20)}{\geq} \mathcal{A}(u_{\nu,i}^\ell,\, u_{\nu,i}^r) - 2\kappa\eta \equiv \mathcal{D}(u_{\nu,i}^\ell,\, u_{\nu,i}^r) - 2\kappa\eta. \end{aligned} \tag{23}$$

Dividing (23) by κ and then subtracting it to (22), we find that

$$\mathcal{E}(u \restriction_{J_i}) \leq \mathcal{E}(u_{\nu,i}^\ell,\, u_{\nu,i}^r) + 4\eta. \tag{24}$$

From this we have

$$\mathcal{E}(u \restriction_{J_i}) \overset{(24)}{\leq} \mathcal{E}(u_{\nu,i}^\ell,\, u_{\nu,i}^r) + 4\eta, \overset{\mathrm{L.3.1}}{\leq} \mathcal{E}(u_\nu \restriction_{J_i}) + 4\eta.$$

Taking the liminf over ν, we get the claim.

i_2) $\mathcal{E}(u \restriction_{\mathbb{R}\setminus\cup_i J_i}) \leq \liminf_\nu \mathcal{E}(u_\nu \restriction_{\mathbb{R}\setminus\cup_i J_i}) + \kappa^{-1}\sigma(\varepsilon)$.
By construction, each jump of u in $\mathbb{R} \setminus \cup_i J_i$ is less than $\varepsilon/2$: we can thus partition this domain into subintervals I_k such that $\mathrm{TV}(u \restriction_{I_k}) < \varepsilon$ and

u, u_ν are continuous at their extremes, neglecting the extremes of X. Superlinearity[a] yields, if $\{\delta_j^{k,\nu}\}_j$ are the size of $u_\nu \restriction_{I_k}$'s jumps,

$$\begin{aligned}
&\mathcal{E}(u \restriction_{I_k}) - \mathcal{E}(u_\nu \restriction_{I_k}) \\
&\quad \leq \mathrm{TV}(u \restriction_{I_k}) - \mathrm{TV}(u_\nu \restriction_{I_k}) + \kappa^{-1}\mathcal{D}(u_\nu \restriction_{I_k}) \\
&\quad \overset{(16)}{\leq} \mathrm{TV}(u \restriction_{I_k}) - \mathrm{TV}(u_\nu \restriction_{I_k}) + \kappa^{-1}\sum_j \sigma(\delta_j^k)\delta_j^k \\
&\quad \overset{(17)}{\leq} \mathrm{TV}(u \restriction_{I_k}) - \mathrm{TV}(u_\nu \restriction_{I_k}) + \kappa^{-1}\sigma(\mathrm{TV}(u_\nu \restriction_{I_k}))\mathrm{TV}(u_\nu \restriction_{I_k})
\end{aligned}$$

Given a continuous, nondecreasing function $g : \mathbb{R} \to \mathbb{R}$, for every real sequence $\{z_k\}_k$ we have that $g(\liminf z_k) = \liminf g(z_k)$; applying this to $z \mapsto (1 - \sigma(z))z$, one can see that, continuing from above,

$$\begin{aligned}
&\limsup_\nu \big\{\mathcal{E}(u \restriction_{I_k}) - \mathcal{E}(u_\nu \restriction_{I_k})\big\} \\
&\quad \leq \mathrm{TV}(u \restriction_{I_k}) - \liminf_\nu \big\{\big(1 - \kappa^{-1}\sigma(\mathrm{TV}(u_\nu \restriction_{I_k}))\big)\mathrm{TV}(u_\nu \restriction_{I_k})\big\} \\
&\quad = \mathrm{TV}(u \restriction_{I_k}) - \big(1 - \kappa^{-1}\sigma\big(\liminf_\nu \mathrm{TV}(u_\nu \restriction_{I_k})\big)\big) \liminf_\nu \mathrm{TV}(u_\nu \restriction_{I_k}) \\
&\quad \overset{(17)}{\leq} \kappa^{-1}\sigma(\mathrm{TV}(u \restriction_{I_k}))(\mathrm{TV}(u \restriction_{I_k})),
\end{aligned}$$

where we used that the sequence $\{u_\nu \restriction_{I_k}\}_\nu$ converges to $u \restriction_{I_k}$ in L^1, thus, by semicontinuity, $\mathrm{TV}(u \restriction_{I_k}) \leq \liminf_\nu \mathrm{TV}(u_\nu \restriction_{I_k})$. By construction of I_k, moreover, we have $\mathrm{TV}(u \restriction_{I_k}) \leq \varepsilon$: this yields

$$\mathcal{E}(u \restriction_{I_k}) - \liminf_\nu \mathcal{E}(u_\nu \restriction_{I_k}) \overset{(17)}{\leq} \kappa^{-1}\sigma(\varepsilon)\mathrm{TV}(u \restriction_{I_k}). \tag{25}$$

By the continuity at the extremes of I_k, summing up we find:

$$\begin{aligned}
&\mathcal{E}(u \restriction_{\mathbb{R}\setminus\cup_i J_i}) - \liminf_\nu \mathcal{E}(u_\nu \restriction_{\mathbb{R}\setminus\cup_i J_i}) \\
&\quad = \limsup_\nu \left\{\sum_k \mathcal{E}(u \restriction_{I_k}) - \mathcal{E}(u_\nu \restriction_{I_k})\right\} \\
&\quad \leq \sum_k \left\{\mathcal{E}(u \restriction_{I_k}) - \liminf_\nu \mathcal{E}(u_\nu \restriction_{I_k})\right\} \\
&\quad \overset{(25)}{\leq} \sum_k \kappa^{-1}\sigma(\varepsilon)\mathrm{TV}(u \restriction_{I_k}) = \kappa^{-1}\sigma(\varepsilon)\mathrm{TV}(u \restriction_{\mathbb{R}\setminus\cup_i J_i}) \leq \kappa^{-1}\sigma(\varepsilon).
\end{aligned}$$

---

[a] Here it will be really a key point. Without superlinearity we can't have, generally speaking, semicontinuity. Consider, for example, $\mathcal{A}(a,b) := |b-a|$ and $X = (0,1)$; take

- $v_n(x) := \sum_{k=1}^n \frac{k}{n-1}\chi_{\left[\frac{k-1}{n},\frac{k}{n}\right)}(x)$, converging to $v(x) := x$ in L^1; $1 = \mathcal{E}(v) > \mathcal{E}(v_n) = 1-\kappa$.
- $v_n(x) := (1-nx)\chi_{\left(0,\frac{1}{n}\right)}(x)$, converging in L^1 to $v := 0$; $0 = \mathcal{E}(v) < \mathcal{E}(v_n) = 1$.

Finally, by the continuity at y_i^ℓ, y_i^r, collecting the results we have

$$\begin{aligned}
\mathcal{E}(u) &= \mathcal{E}(u \restriction_{\mathbb{R}\setminus\cup_i J_i}) + \sum_{i=1,\dots,N(\varepsilon)} \mathcal{E}(u \restriction_{J_i}) \\
&\overset{2.}{\leq} \liminf_\nu \mathcal{E}(u_\nu \restriction_{\mathbb{R}\setminus\cup_i J_i}) + \sum_{i=1,\dots,N} \mathcal{E}(u \restriction_{J_i}) + \kappa^{-1}\sigma(\varepsilon) \\
&\overset{1.}{\leq} \liminf_\nu \mathcal{E}(u_\nu \restriction_{\mathbb{R}\setminus\cup_i J_i}) + \sum_{i=1,\dots,N} \liminf_\nu \mathcal{E}(u_\nu \restriction_{J_i}) + 4\eta N(\varepsilon) + \kappa^{-1}\sigma(\varepsilon) \\
&\leq \liminf_\nu \left\{ \mathcal{E}(u_\nu \restriction_{\mathbb{R}\setminus\cup_i J_i}) + \sum_{i=1,\dots,N} \mathcal{E}(u_\nu \restriction_{J_i}) \right\} + 4\eta N(\varepsilon) + \kappa^{-1}\sigma(\varepsilon) \\
&= \liminf_\nu \mathcal{E}(u_\nu) + 4\eta N(\varepsilon) + \kappa^{-1}\sigma(\varepsilon).
\end{aligned}$$

Take now the limit over $\eta \to 0$, you get

$$\mathcal{E}(u) \leq \liminf_\nu \mathcal{E}(u_\nu) + \kappa^{-1}\sigma(\varepsilon).$$

Thanks to superlinearity, L. 2.19 j_3), taking the limit when $\varepsilon \to 0$ we are done.

Consider now that $\{u_\nu\}_\nu$ converges to u in $\mathrm{L}^1_{\mathrm{loc}}(X)$ but not in $\mathrm{L}^1(X)$. For every $K > 0$, we can consider $X' = X \cap (-K, K)$ and there, by the poof above, having $\mathrm{L}^1(X')$ convergence we guess

$$\mathcal{E}(u \restriction_{X'}) \leq \liminf_\nu \mathcal{E}(u_\nu \restriction_{X'}) \leq \liminf_\nu \mathcal{E}(u_\nu).$$

If we take the limit when $K \to \infty$, the left hand side goes to $\mathcal{E}(u)$, while the right hand side remains constantly $\liminf_\nu \mathcal{E}(u_\nu)$, thus the thesis. □

Lemma 3.4. *Assume to have a real interval J, a sequence $\{\varepsilon_\nu\}_\nu$ decreasing to 0 and a function $u \in \mathrm{BV}(J)$. There exists a sequence $\{u_\nu\}_\nu$ such that*

- *$u_\nu \in \mathrm{BV}(J)$ is piecewise constant and $||u - u_\nu||_\infty \leq \varepsilon_\nu$; moreover, if u's limits at the endpoints of J are in $\varepsilon_\nu\mathbb{Z}$, the ones of u_ν can be taken equal to them.*
- *$\mathrm{TV}(u_\nu) \leq \mathrm{TV}(u)$, $\mathrm{TV}(u_\nu) \to \mathrm{TV}(u)$ and $\mathcal{D}(u_\nu) \to \mathcal{D}(u)$.*

Proof. Let $u \in \mathrm{BV}(J)$. Construct the sequence as follows. Put $x_0 := \inf J$, which belongs to $\mathbb{R} \cup \{-\infty\}$, and take $u_\nu(x_0) \in \varepsilon_\nu\mathbb{Z}$ such that $|u(x_0^+) - u_\nu(x_0)| < \varepsilon_\nu$ — there are at most two possible choices. Take then recursively

$$x_{i+1} := \min\{x > x_i : |u(x) - u_\nu(x_i)| \geq \varepsilon_\nu\}$$

and define $u_\nu(x_{i+1}) \in \varepsilon_\nu\mathbb{Z}$ between $u_\nu(x_i)$ and $u(x_{i+1})$ such that $|u(x_{i+1}) - u_\nu(x_{i+1})| < \varepsilon_\nu$. The distance between $u(x_{i-1})$ and $u(x_{i+1})$ or between $u(x_i)$ and $u(x_{i+1})$ is more or equal to ε_ν: we have thus a finite number of points, say $N - 1$. If u's limits at the endpoints of J are in $\varepsilon_\nu\mathbb{Z}$, then you want the same limit for u_ν: add thus two other points, $x_{N+1} := \sup J$ and $x_N := \frac{x_{N-1}+x_{N+1}}{2}$, and define

$$u_\nu(x_N) := \begin{cases} u(x_{N+1}^-) & \text{if } u(x_{N+1}^-) \in \varepsilon_\nu\mathbb{Z}, \\ u_\nu(x_{N-1}) & \text{otherwise.} \end{cases}$$

Finally put

$$u_\nu(x) := \sum_{i=1}^{N} u_\nu(x_{i-1})\chi_{[x_{i-1}, x_i)} + u_\nu(x_N)\chi_{[x_N, x_{N+1}]}.$$

If $J \neq \overline{J}$, just forget u_ν's precise values at the extremes. Being $||u - u_\nu||_\infty \leq \varepsilon_\nu$, we have the $\mathrm{L}^1_{\mathrm{loc}}(J)$ convergence. Observe that the total variation of u_ν can be bigger than the one of u just because of the approximation at x_0: if this happens, restart from the beginning taking the other possible choice for $u_\nu(x_0)$. Now $\mathrm{TV}(u_\nu) \leq \mathrm{TV}(u)$, then, by semicontinuity, $\mathrm{TV}(u_\nu) \to \mathrm{TV}(u)$.

Let's show now that $\mathcal{D}(u_\nu) \to \mathcal{D}(u)$. Fix $\varepsilon > 0$. Let $y_1, \ldots, y_k$ be such that

$$\mathcal{D}(u) - \sum_{h=1}^{k} \mathcal{A}(u(y_h)^-, u(y_h)^+) < \varepsilon. \tag{26}$$

Fix now $0 < \eta < \frac{1}{3} \min \left\{ |u(y_h^+) - u(y_h^-)| \right\}_{h=1}^{k}$. For every $\varepsilon_\nu < \eta$, consider $u_\nu \in \mathrm{BV}(J)$ constructed above. Let $\{x_i\}_{i \in I_\nu}$ be the set of points where u_ν has a jump higher than η: by construction it will contain the $\{y_i\}_i$ and $|I_\nu| < \frac{\mathrm{TV}(u)}{\eta}$. Using $\mathcal{A}$'s superlinearity, we get

$$\begin{aligned}
&\mathcal{D}(u_\nu) - \sum_{i \in I_\nu} \mathcal{A}(u_\nu(x_i^-), u_\nu(x_i^+)) \\
&\quad = \sum_{i \notin I_\nu} \mathcal{A}(u_\nu(x_i^-), u_\nu(x_i^+)) \\
&\quad \overset{j_3)}{\leq} \sum_{i \notin I_\nu} \sigma(u_\nu(x_i^-) - u_\nu(x_i^+)) |u_\nu(x_i^-) - u_\nu(x_i^+)| \\
&\quad \overset{(17)}{\leq} \sigma(\eta) \sum_{i \notin I_\nu} |u_\nu(x_i^-) - u_\nu(x_i^+)| \leq \sigma(\eta) \mathrm{TV}(u_\nu) \leq \sigma(\eta) \mathrm{TV}(u).
\end{aligned}$$

Moreover, for the choice of $\{y_i\}_{i=1}^k$, contained in $\{x_i\}_{i \in I_\nu}$, from (26)

$$\mathcal{D}(u) - \sum_{i \in I_\nu} \mathcal{A}(u(x_i^-), u(x_i^+)) < \varepsilon.$$

Finally, by $\mathcal{A}$'s Lipschitz condition, being $||u - u_\nu||_\infty \leq \varepsilon_\nu$

$$\begin{aligned}
&\sum_{i \in I_\nu} |\mathcal{A}(u(x_i^-), u(x_i^+)) - \mathcal{A}(u_\nu(x_i^-), u_\nu(x_i^+))| \\
&\quad \overset{(2)}{\leq} \sum_{i \in I_\nu} 2\kappa \varepsilon_\nu \leq \frac{2\kappa \mathrm{TV}(u)}{\eta} \varepsilon_\nu.
\end{aligned}$$

Adding the three pieces above, we have

$$|\mathcal{D}(u) - \mathcal{D}(u_\nu)| \leq \varepsilon + \sigma(\eta) \mathrm{TV}(u) + \frac{2\kappa \mathrm{TV}(u)}{\eta} \varepsilon_\nu.$$

Taking first the limsup over ν, then the limit over $\eta \to 0$, finally over $\varepsilon \to 0$, we reach the thesis. □

Lemma 3.5. *$\mathcal{E}$ decreases along piecewise constant front-tracking solutions of*

$$u_t + f(u)_x = 0, \qquad\qquad f \in \mathrm{C}^2(\mathbb{R}).$$

Proof. There is no loose of generality in assuming $\mathrm{TV}(u(0)) \leq 1$, up to a rescaling of the flux f, or, equivalently, of the function $\mathcal{E}$. Fix a time t: $\mathcal{E}(t)$ is given by the sum, over the points of jump, of $\mathcal{E}(u_i^-, u_i^+)$. Depending only on the right and left limits at jump's points, $\mathcal{E}(t)$ can vary only when interactions take place. Just to fix the ideas, look at the first instant of interaction. By the additive form, with respect to jumps, it suffices moreover to show that it does not increase when all the discontinuities lines intersect in the (x,t) plane at one point $(\bar{x},\bar{t})$. Let's consider this case.

By the construction of the solution of a Riemann problem and the structure of $\mathcal{D}$ (Remark 2.21), we have that for $t > \bar{t}$ it holds $\mathcal{E}(t) = \mathcal{E}(\bar{t})$. Moreover, looking at $t < \bar{t}$ we see that $u(\bar{t})$ is built just collapsing all the jumps of $u(t)$ in one point. Thus, as shown in Lemma 3.1, $\mathcal{E}$ does not increase. □

Theorem 3.6. *Consider the scalar, one dimensional conservation law*

$$u_t + f(u)_x = 0, \qquad f \in \mathrm{C}^2(\mathbb{R}). \tag{27}$$

The functional $\mathcal{E}$ decreases along its entropy solutions with bounded variation.

Proof. Let $u \in \mathrm{BV}((0,t) \times \mathbb{R})$, $\forall t > 0$, be an entropy solution of (27). There is no loose of generality in assuming $\mathrm{TV}(u(0)) \leq 1$, up to a rescaling of the flux f, or, equivalently, of the function $\mathcal{E}$. The restriction of u at any time will be $\mathrm{BV}(\mathbb{R})$, with its total variation decreasing in time ([2], Th. 6.1): thus $\mathrm{TV}(u(t)) \leq 1$. Given $r > s \geq 0$, we want to show that

$$\mathcal{E}(u(r)) - \mathcal{E}(u(s)) \leq 0.$$

By the semigroup property, we can suppose $s = 0$. Use the wave front tracking algorithm to approximate u with a sequence $\{u_\nu\}_\nu$ of piecewise constant functions. For starting it, approximate $u(0)$ as in Lemma 3.4: we have thus $\mathcal{E}(u_\nu(0)) \to \mathcal{E}(u(0))$ and $\mathrm{TV}(u_\nu) \leq 1$. By Lemma 3.5, $\mathcal{E}$ decreases along the approximations:

$$\mathcal{E}(u_\nu(r)) - \mathcal{E}(u_\nu(0)) \leq 0.$$

From this and the semicontinuity (Lemma 3.3) we gain that

$$\begin{aligned}
&\mathcal{E}(u(r)) - \mathcal{E}(u(0)) \\
&\quad \leq \liminf_\nu \mathcal{E}(u_\nu(r)) - \lim_\nu \mathcal{E}(u_\nu(0)) \\
&\quad = \liminf_\nu \Big\{ \mathcal{E}(u_\nu(r)) - \mathcal{E}(u_\nu(0)) \Big\} \leq 0\,,
\end{aligned}$$

as we were looking for. □

Corollary 3.7. *$\mathcal{E}(t)$ and $\mathrm{TV}(t)$ are the right continuous representative of non increasing functions. Moreover, $\mathcal{D}(t)$ is the right continuous representative of a $\mathrm{BV}(\mathbb{R}^+)$ function.*

Proof. Recall that the map $t \mapsto u(t)$ is continuous with values in L^1 ([2], Th. 6.3). As a consequence, the maps $\mathrm{TV}(t)$ and $\mathcal{E}(t)$ are lower semicontinuous also with respect to time. TV and $\mathcal{E}$ are monotone, precisely not increasing, and thus, by the lower semicontinuity, right continuous w.r.t. time, for $t > 0$: for every $h > 0$

$$\begin{aligned}
\mathcal{E}(t+h) \leq \mathcal{E}(t) \leq \liminf_{k \to 0^+} \mathcal{E}(t+k), \\
\limsup_{h \to 0^+} \mathcal{E}(t+h) \leq \mathcal{E}(t) \leq \liminf_{k \to 0^+} \mathcal{E}(t+k).
\end{aligned}$$

Hence $\mathcal{D}$, which is their difference, is both right continuous and with bounded variation. □

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* Oxford Clarendon Press, 2000.

[2] A. Bressan. *Hyperbolic Systems of Conservation Laws. The One-Dimensional Cauchy Problem.* Oxford University Press, 2000.

[3] P. Baiti, A. Bressan. Lower semicontinuity of weighted path lenght in BV. *Geometrical Optics and Related Topics*, pages 31–58, 1997. F. Colombini e N. Lerner Eds., Birkhauser.

[4] S. Bianchini, A. Bressan. On a Lyapunov functional relating shortening curves and viscous conservation laws. *Nonlinear Anal.*, 51:649–662, 2002.

[5] C. M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics.* Berlin Springer-Verlag, 2005.

[6] C. De Lellis, F. Otto, and M. Westdickenberg. Minimal entropy condition for Burgers equation. *Quart. Appl. Math.*, 62(4):687–700, 2004.

[7] H. Federer. *Geometric Measure Theory.* Berlin Springer-Verlag, 1969.

[8] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.*, 18(4):697–715, 1965.

[9] S. N. Kruzhkov. First-order quasilinear equations with several space variables. *Math. USSR Sbornik*, 10:217–273, 1970.

[10] O. A. Oleinik. Discontinuous solutions of nonlinear differential equations. *AMS Transl.*, 26:95–172, 1963. Series 2.

[11] E. Y. Panov. Uniqueness of the Cauchy problem for a first order quasilinear equation with one admissible strictly convex entropy. *Math. Notes*, 55(5):517–525, 1994.

[12] A. Quarteroni, R. Sacco, F. Saleri. *Numerical Mathematics.* Springer, 2000.

Laura Caravenna, S.I.S.S.A., via Beirut 4, 34014 Trieste, Italy
E-mail address: l.caravenna@sissa.it