Regularizing effect of nonlinearity in multidimensional scalar conservation laws

Gianluca Crippa[1], Felix Otto[2], and Michael Westdickenberg[2]

[1] Scuola Normale Superiore
Piazza dei Cavalieri 7, I–56126 Pisa, Italy
g.crippa@sns.it

[2] Institut für Angewandte Mathematik, Universität Bonn
Wegelerstraße 10, D–53115 Bonn, Germany
otto@iam.uni-bonn.de
mwest@iam.uni-bonn.de

1 Introduction

In these notes we will deal with *scalar multidimensional conservation laws*, which are first order partial differential equations of the form

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{f}(u) = 0 \quad \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n. \tag{1}$$

Here $\mathbf{f}\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ is a smooth map which will be called the flux function. The solution $u\colon \mathbb{R}_+ \times \mathbb{R}^n \longrightarrow \mathbb{R}$ can be viewed as a conserved quantity: indeed, if we integrate the equation over a bounded set $\Omega \subset \mathbb{R}^n$ with smooth boundary and we apply the divergence theorem, we see that the total amount of u in Ω changes in time according to the flux of $\mathbf{f}(u)$ across the boundary $\partial\Omega$. This interpretation motivates the interest of this class of equations: many situations in nature are modeled on the general principle that physical quantities are neither created nor destroyed, and their variation in a domain is due to the flux across the boundary.

The mathematical treatment of these equations, however, is challenging. As we will show in the next section, there is a lack of existence of smooth solutions, even for onedimensional equations: starting from smooth initial data we can lose regularity in finite time, due to a blow-up of the space derivative. Hence we are forced to consider distributional solutions, but this implies a loss of uniqueness: we will indicate via some simple explicit examples how it is possibile to associate to the same initial data several (in fact, infinitely many) weak solutions.

To restore the uniqueness we are forced to add some condition to our notion of solution. This is encoded in the concept of *entropy solution*: we ask that some nonlinear functions of u are dissipated along the flow. This is inspired by the second principle of thermodynamics and is consistent with a widely used approximation procedure, the vanishing viscosity technique. A celebrated theorem by Kružkov ensures existence, uniqueness and stability for this class of solutions of scalar conservation laws.

A consequence of Kružkov theory is the propagation of BV-regularity (see Section 3 for details): if we start with initial data with bounded variation, then the entropy solution has bounded variation for all times. We remark in passing that the BV framework is natural in the context of conservation laws, since entropy solutions in general develop discontinuities (which will be called *shocks*) in finite time, even starting with smooth initial data.

If the initial data is just L^∞ we cannot expect BV regularity of the solution (see the examples in Section 4). It turns out, however, that under quite general assumptions the equation has a regularizing effect. Clearly, this requires nondegeneracy of the flux $\mathbf{f}$: remember that in the case of a linear flux function $\mathbf{f}$ the equation reduces to a transport equation with constant speed, hence no regularization will be possible. We will therefore require genuine nonlinearity of the flux function (see condition (15)). Under this assumption we will describe two different regularization results.

For a flux function which is the higher dimensional analogue of Burgers' flux we can obtain fractional regularity of the entropy solution; we will show that it belongs to a Besov space whose order depends on the space dimension. We will also show the sharpness of this result by constructing explicit examples. See Propositions 4.3 and 4.4. These results are original and we give a rather elementary proof of them in the Appendix. Our approach completely avoids the use of Fourier transform methods, which are otherwise typical in the context of velocity averaging lemmas, the usual tool in this framework.

Even if BV regularity of the solution is not expected, we can prove that entropy solutions nevertheless have a similar structure as BV functions. We can identify a *jump set* on which the solution has strong traces; outside this set the solution has a weak form of continuity (see Theorem 4.5 and [11] for the precise statement).

Our proof relies on the *kinetic formulation* of the conservation law (see Section 5), which encodes the entropy inequalities in form of a linear transport equation with measure-valued right hand side. The structure theorem then follows from blow-up techniques and a complete characterization of the states that are obtained via blow-ups.

2 Background Material

We are mainly interested in how the *nonlinearity* of $\mathbf{f}$ affects the regularity of entropy solutions of the scalar conservation law (1). It turns out that a nonlinear flux in general does not allow for smooth solutions. To see this, we first use the chain rule to rewrite (1) in the form

$$\frac{\partial u}{\partial t} + \mathbf{f}'(u) \cdot \nabla u = 0.$$

This identity implies that u is constant along the characteristic lines of (1), that is, along the trajectories of the ordinary differential equation

$$\dot{X}(t) = \mathbf{f}\big(u(t, X(t))\big) \quad \forall t \in \mathbb{R}_+.$$

In particular, the characteristics are straight lines. Assuming that the solution u of (1) attains initial data $u(0, \cdot) = u_0$, we obtain the implicit relation

$$u(t, x) = u_0\big(x - t\mathbf{f}'(u(t, x))\big) \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n. \tag{2}$$

If the flux function is nonlinear, however, and thus $\mathbf{f}'$ is nonconstant, then the characteristic lines typically intersect somewhere and so the identity (2) is no longer well-defined. This shows that smooth solutions of (1) in general cannot exist globally in time. Indeed, one can check that an upper bound for the lifespan of smooth solutions is given by $T_\infty = \max(-\kappa, 0)^{-1}$, where

$$\kappa := \operatorname*{ess\,inf}_{x \in \mathbb{R}^n} \mathbf{f}''(u_0(x)) \cdot \nabla u_0(x), \tag{3}$$

see Theorem 6.1.1 in [10]. As an example, notice that in the onedimensional case $n = 1$ with a convex flux $\mathbf{f}'' \geqslant 0$, formula (3) indicates global existence of smooth solutions, if the initial data is nondecreasing $u_0' \geqslant 0$. One can easily check that indeed in this situation characteristic lines have a fan-like shape and therefore never cross. In general, however, the nonlinearity of the flux $\mathbf{f}$ forces us to consider weak solutions of (1) instead of smooth ones.

In the following, we will be concerned with functions $u \in \mathrm{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ satisfying the scalar conservation law (1) in distributional sense only. These functions are called weak solutions. The price we have to pay for this broader solution concept is a lack of uniqueness. Let us consider Burgers' equation

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} = 0 \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R} \tag{4}$$

with initial data

$$u_0(x) := \left\{ \begin{array}{rl} -1 & \text{if } x < 0 \\ 1 & \text{if } x \geqslant 0 \end{array} \right\} \quad \forall x \in \mathbb{R}. \tag{5}$$

Initial value problems for onedimensional conservation laws with piecewise constant initial data are called *Riemann problems*. One can check that

$$u_r(t, x) := \left\{ \begin{array}{rl} -1 & \text{if } x/t < -1 \\ 1 & \text{if } x/t > 1 \\ x/t & \text{otherwise} \end{array} \right\} \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}$$

is a weak solution of (4) and (5). The function u_r is homogeneous of degree zero and locally Lipschitz continuous outside the origin. Notice that across the lines defined by $|x/t| = 1$ the x-derivative of u_r is discontinuous. The solution u_r is called a *rarefaction wave*. Another weak solution of (4) and (5) is

$$u_s(t,x) := u_0(x) \quad \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}.$$

This solution is called a *shock* solution because it is discontinuous along the timeaxis. More generally, if u is a solution of the scalar conservation law (1), then a shock is a discontinuity of u across a (sufficiently regular) hypersurface $\mathcal{J} \subset \mathbb{R}_+ \times \mathbb{R}^n$ that is timelike in the sense that in every point $(t,x) \in \mathcal{J}$, the normal vector $\mu(t,x) \in \mathbb{R}^{n+1}$ is not parallel to the timeaxis. This implies that the normal vector can always be written in the form

$$\mu = (-s, \nu)^T \quad \text{with } s \in \mathbb{R} \text{ and } \nu \in S^{n-1}.$$

The number s is called the *shock speed* because it determines how fast $\mathcal{J}$ is propagating in spatial direction ν. The fact that u is a weak solution of (1) implies that the *Rankine-Hugoniot condition* holds in every point $(t,x) \in \mathcal{J}$: If u^+ denotes the limit of u when approaching (t,x) from that side of the hyperplane the normal $\mu = (-s,\nu)^T$ is pointing to, and if u^- is the limit from the opposite side (recall that u is discontinuous along $\mathcal{J}$), then

$$-s\big[u^+ - u^-\big] + \big[\mathbf{f}(u^+) - \mathbf{f}(u^-)\big] \cdot \nu = 0.$$

The limits u^+ and u^- are called *traces* of u on the hypersurface $\mathcal{J}$.

In order to restore uniqueness, one typically imposes an *entropy condition* which in the case of scalar conservation laws can be written in the form of a family of differential inequalities. That is, one discards the majority of weak solutions, keeping only those that satisfy in distributional sense

$$\frac{\partial \eta(u)}{\partial t} + \nabla \cdot \mathbf{q}(u) \leqslant 0 \tag{6}$$

for all *convex entropy-entropy flux pairs* $(\eta, \mathbf{q})$ defined by $\mathbf{q}' = \eta' \mathbf{f}'$ and $\eta'' \geqslant 0$. Weak solutions satisfying the entropy condition are called *entropy solutions*. This entropy condition is inspired by the second law of thermodynamics. It is consistent with a widely used method for constructing weak solutions of conservation laws, called the *vanishing viscosity method*: One considers

$$\frac{\partial u_\varepsilon}{\partial t} + \nabla \cdot \mathbf{f}(u_\varepsilon) = \varepsilon \Delta u_\varepsilon \quad \text{with } \varepsilon > 0, \tag{7}$$

which is easily shown to have unique smooth solutions, and then sends ε to zero. Notice the after multiplying (7) by $\eta'(u_\varepsilon)$ and using $\mathbf{q}' = \eta' \mathbf{f}'$ we get

$$\frac{\partial \eta(u_\varepsilon)}{\partial t} + \nabla \cdot \mathbf{q}(u_\varepsilon) = \varepsilon \Delta \eta(u_\varepsilon) - \eta''(u_\varepsilon) |\nabla u_\varepsilon|^2,$$

which converges to (6) in the limit $\varepsilon \to 0$, by convexity of η. It turns out that the concept of entropy solutions restores exactly the right amount of rigidity to obtain well-posedness for the initial value problem: In his seminal paper [19], Kružkov proved existence and uniqueness of entropy solutions of (1) for initial data $u_0 \in \mathrm{L}^\infty(\mathbb{R})$. They coincide with the solutions obtained by the vanishing viscosity method. Kružkov's proof is based on the observation that entropy solutions generate a contractive semigroup in $\mathrm{L}^1(\mathbb{R}^n)$: If u and v are entropy solutions of (1) with initial data u_0 and v_0 respectively, then

$$\|u(t) - v(t)\|_{\mathrm{L}^1(\mathbb{R}^n)} \leqslant \|u_0 - v_0\|_{\mathrm{L}^1(\mathbb{R}^n)} \quad \forall t \geqslant 0. \tag{8}$$

As a consequence of this estimate one also obtains stability in the space $\mathrm{BV}(\mathbb{R}^n)$ of functions of bounded variation: If the data $u_0 \in \mathrm{BV}(\mathbb{R}^n)$, then also $u(t) \in \mathrm{BV}(\mathbb{R}^n)$ for all $t > 0$. Recall that a function is of bounded variation if the distributional first derivative is a measure (we refer the reader to Section 3 for further information). Notice also that this stability holds trivially for linear fluxes. As a consequence of the well-known *structure theorem*, entropy solutions of scalar conservation laws with BV-regularity automatically have the structure we expect: The solution is (approximately) continuous outside a set $\mathcal{J} \subset \mathbb{R}_+ \times \mathbb{R}^n$ which locally looks like a Lipschitz continuous hypersurface and across which the solution is discontinuous. On this shock set, strong traces can be defined and the Rankine-Hugoniot condition is satisfied. Moreover, the entropy condition (6) reduces to the shock admissibility condition

$$-s\big[\eta(u_+) - \eta(u_-)\big] + \big[\mathbf{q}(u_+) - \mathbf{q}(u_-)\big] \cdot \nu \leqslant 0 \tag{9}$$

for all points in $\mathcal{J}$. This last statement follows from a generalization of the *chain rule* for functions of bounded variation (see Section 3). In that sense, $\mathrm{BV}(\mathbb{R}^n)$ is a natural space for entropy solutions of conservation laws.

Recall, however, that Kružkov's result ensures well-posedness even for initial data $u_0 \in \mathrm{L}^\infty(\mathbb{R}^n)$. What can be said about the regularity of entropy solutions in this case? Clearly, if the flux is linear, then (1) is simply a linear transport equation with constant velocity, and therefore the solution u at any positive time cannot be more regular than the initial data. But for nonlinear fluxes more can be said: It turns out that while on the one hand the nonlinearity prevents global existence of smooth solutions of (1), on the other hand it also has a regularizing effect! The first result in that direction, due to Oleĭnik [22], is that for onedimensional scalar conservation laws with uniformly convex flux, entropy solutions satisfy the one-sided Lipschitz condition

$$\sup_{x\in\mathbb{R}} \frac{\partial u}{\partial x}(t,x) \leqslant \frac{1}{ct} \quad \forall t > 0, \tag{10}$$

where $c := \inf_{u\in\mathbb{R}} \mathbf{f}''(u)$. Notice that (10) only allows for decreasing jumps. Since u is bounded it follows that initial data in $\mathrm{L}^\infty(\mathbb{R})$ is instantaneously regularized to $\mathrm{BV}(\mathbb{R})$ locally. The slightly more precise estimate

$$\sup_{x\in\mathbb{R}} \frac{\partial}{\partial x}\mathbf{f}'\big(u(t,x)\big) \leqslant \frac{1}{t} \quad \forall t>0$$

was proved by Hoff [17]. We also refer to [27, 9, 5] for further generalizations. Oleĭnik's estimate (10) can be proved using the vanishing viscosity approximation. For simplicity we only consider the case of Burgers' equation (4): Let u_ε be the unique smooth solution of the parabolic approximation

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{1}{2}\frac{\partial u_\varepsilon^2}{\partial x} = \varepsilon \frac{\partial^2 u_\varepsilon}{\partial x^2} \tag{11}$$

and set $v_\varepsilon := \frac{\partial u_\varepsilon}{\partial x}$. Then v_ε satisfies the equation

$$\frac{\partial v_\varepsilon}{\partial t} + v_\varepsilon^2 + u_\varepsilon \frac{\partial v_\varepsilon}{\partial x} = \varepsilon \frac{\partial^2 v_\varepsilon}{\partial x^2}. \tag{12}$$

Now we use the comparison principle for parabolic equations and the fact that the function $V(t,x) := 1/t$ is a solution of (12) to obtain the estimate

$$v_\varepsilon(t,x) \leqslant V(t,x) = \frac{1}{t} \quad \forall (t,x) \in \mathbb{R}_+ \times \mathbb{R}.$$

Let $w_\varepsilon(t,x) := u_\varepsilon(t,x) - x/t$. Then we have for all $t>0$ and $x_0 \leqslant x_1$

$$w_\varepsilon(t,x_1) - w_\varepsilon(t,x_0) = u_\varepsilon(t,x_1) - u_\varepsilon(t,x_0) - \frac{1}{t}(x_1 - x_0) \leqslant 0, \tag{13}$$

so that $w_\varepsilon(t,\cdot)$ is a decreasing function, and thus of bounded variation locally. The same is true for $u_\varepsilon(t,\cdot)$. From (11) we now conclude that u_ε is of bounded variation both in space and time locally, which implies strong convergence. We obtain (13) for the limit $u = \lim_{\varepsilon\to 0} u_\varepsilon$ as well, and then (10) follows.

A different approach to regularity results for entropy solutions of scalar conservation laws uses the kinetic formulation introduced by Lions, Perthame and Tadmor in [20]. To motivate this approach, we recall that smooth solutions u of the scalar conservation law (1) are constant along characteristics, which are straight lines. This implies that the level sets of u, that is, the sets

$$E_v(t) := \{x \in \mathbb{R}^n \colon u(t,x) \geqslant v\} \quad \forall v \in \mathbb{R},$$

are moving with constant velocity given by $\mathbf{f}'(v)$. Notice that level sets are ordered: If $v \geqslant \hat{v}$, then $E_v \subseteq E_{\hat{v}}$. For nonlinear fluxes, level sets corresponding to different values of v are traveling with different speeds. Therefore the ordering usually breaks down at some time. This corresponds to the fact that smooth solutions of (1) typically do not exist globally. The ordering of the E_v can be restored by a projection step, which in fact is related to entropy being dissipated. We refer the reader to [6] for further details. This heuristics can be made rigorous and leads to the following result: A function $u \in \mathrm{L}^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$ is an entropy solution of (1) if and only if the function defined by

$$\chi(v,u(t,x)) := \begin{cases} +1 & \text{if } 0 < v \leqslant u(t,x) \\ -1 & \text{if } u(t,x) \leqslant v < 0 \\ 0 & \text{otherwise} \end{cases}$$

for a.e. $(v,t,x) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n$, satisfies the kinetic equation

$$\frac{\partial \chi(v,u)}{\partial t} + \mathbf{f}'(v) \cdot \nabla_x \chi(v,u) = \frac{\partial \mu}{\partial v} \quad \text{in } \mathscr{D}'(\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n), \tag{14}$$

where $\mu \in \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n)$ is some nonnegative measure. We refer the reader to [20] and to Section 5 below. Notice that μ is intrinsically defined by the entropy solution since it captures the entropy dissipation due to the nonsmoothness of u. A plethora of intriguing results on scalar conservation laws has been obtained from the kinetic formulation, in particular when combined with a technique called *velocity averaging*. This method was invented in the context of kinetic equations, and it played a central role in the global existence result for Boltzmann's equation by DiPerna and Lions, see [13]. The key observation is that moments of solutions to kinetic equations enjoy more regularity than one might expect a priori. For scalar conservation laws this method implies compactness and regularity of entropy solutions.

Consider the kinetic equation (14): For fixed $v \in \mathbb{R}$ this equation provides information about the derivative of $\chi(v,u)$ in direction of $(1, \mathbf{f}'(v)) \in \mathbb{R}^{n+1}$. On the other hand, the map $v \mapsto \chi(v,u(x))$ is of bounded variation uniformly in x. These observations can be combined to obtain regularity for the average $\int_{\mathbb{R}} \chi(v,u)\,dv = u$. Here regularity means either boundedness in some Sobolev space or strong $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$-precompactness for sequences of (approximate) entropy solutions. For these results an assumption on the nondegeneracy of the flux $\mathbf{f}$ is necessary: Notice that if the flux is linear, then the scalar conservation law (1) is simply a transport equation. That is, the initial data is uniformly transported into the direction $\mathbf{f}'(v)$, which is independent of $v \in \mathbb{R}$. We can therefore not expect any regularizing effect: At any positive time, the solution is not better behaved that the initial data. The most natural assumption on the nondegeneracy of the flux $\mathbf{f} \colon \mathbb{R} \longrightarrow \mathbb{R}^n$ is the following:

There is no open interval $I \subset \mathbb{R}$ such that $\mathbf{f}'(v)$ is contained in an $(n-1)$-dimensional affine space for all $v \in I$.

In the one-dimensional case this means that the flux f is not affine on any open set. A more formal restatement of the same assumption is that

$$\forall \xi \in \mathbb{S}^n \quad \text{the set } \big\{ v \in \mathbb{R} \colon \xi_0 + \mathbf{f}'(v) \cdot \xi' = 0 \big\} \text{ contains no open intervals,}$$

where $\xi = (\xi_0, \xi') \in \mathbb{R} \times \mathbb{R}^n$. In order to apply the velocity averaging argument we need a slightly stronger assumption, which takes the following form:

$$\forall \xi \in \mathbb{S}^n \quad \mathscr{L}^1\Big(\big\{ v \in \mathbb{R} \colon \xi_0 + \mathbf{f}'(v) \cdot \xi' = 0 \big\} \Big) = 0. \tag{15}$$

That is, the set of values $v \in \mathbb{R}$ for which the characteristic directions $\mathbf{f}'(v)$ stay inside the hyperplane determined by the normal vector $\xi \in \mathbb{S}^n$ has zero Lebesgue measure. A sufficient condition for (15) to hold is that

$$\forall v \in \mathbb{R} \quad \text{the vectors } \mathbf{f}''(v), \ldots, \mathbf{f}^{(n+1)}(v) \text{ are linearly independent.} \tag{16}$$

Indeed choose $\xi = (\xi_0, \xi') \in \mathbb{S}^n$ and let $h(v) := \xi_0 + \mathbf{f}'(v) \cdot \xi'$ for $v \in \mathbb{R}$. Assume that there exists an open interval $I \subset \mathbb{R}$ with $h(v) = 0$ for all $v \in I$. Then

$$\begin{aligned} I &\subset \big\{v \in \mathbb{R} \colon h'(v) = \ldots = h^{(n)} = 0\big\} \\ &= \big\{v \in \mathbb{R} \colon \mathbf{f}''(v) \cdot \xi' = \ldots = \mathbf{f}^{(n+1)}(v) \cdot \xi' = 0\big\} \\ &= \big\{v \in \mathbb{R} \colon \mathbf{f}^{(k)}(v) \perp \xi' \text{ for all } k \in \{2, \ldots, n+1\}\big\} \\ &\subset \big\{v \in \mathbb{R} \colon \mathbf{f}''(v), \ldots, \mathbf{f}^{(n+1)}(v) \text{ are linearly dependent}\big\} = \varnothing \end{aligned}$$

because n vectors in $\mathbb{R}^n$ all contained in a hyperplane cannot form a basis. Assumption (16) is satisfied for the generalized Burgers' flux

$$\mathbf{f}(v) := \Big(\tfrac{1}{2}v^2, \ldots, \tfrac{1}{n+1}v^{n+1}\Big) \quad \forall v \in \mathbb{R}, \tag{17}$$

which we will study in more detail later on. If (15) holds, then any sequence of entropy solutions contains a subsequence which converges strongly in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$, see [20]. A more quantitative version of (15) is that

$$\forall \xi \in \mathbb{S}^n \quad \mathscr{L}^1\Big(\big\{v \in \mathbb{R} \colon |\xi_0 + \mathbf{f}'(v) \cdot \xi'| \leqslant \delta\big\}\Big) \leqslant C\delta^\alpha \tag{18}$$

for all $\delta > 0$ and some $\alpha \in (0, 1]$. Assumption (18) yields Sobolev regularity for entropy solutions. The regularity one obtains depends on the nondegeneracy of $\mathbf{f}$, that is, on α. We refer the reader to [16, 14, 3, 20, 4, 18, 26, 24] for further information about the velocity averaging argument. Its proof typically relies on Littlewood-Paley type decompositions, interpolation arguments and a spectral decomposition adapted to the "velocity direction".

3 Entropy Solutions with BV-regularity

Let $\Omega \subset \mathbb{R}^{n+1}$ be an open subset. A function $u \in \mathrm{L}^1(\Omega)$ is called of bounded variation if its distributional derivative Du is an $\mathbb{R}^{n+1}$-valued measure with finite total variation in Ω. We denote by $\mathrm{BV}(\Omega)$ the space of functions of bounded variation, which is a Banach space with respect to the norm

$$\|u\|_{\mathrm{BV}(\Omega)} := \|u\|_{\mathrm{L}^1(\Omega)} + \|Du\|_{\mathrm{M}(\Omega)}.$$

The local space $\mathrm{BV}_{\mathrm{loc}}(\Omega)$ is then defined in the usual way. It turns out that

$$\|u\|_{\mathrm{BV}(\Omega)} \approx \|u\|_{\mathrm{L}^1(\Omega)} + \sup_{h \neq 0} |h|^{-1} \|u(\cdot + h) - u\|_{\mathrm{L}^1(\Omega)}$$

(see Remark 3.25 in [2]), which together with the L^1-contraction estimate (8) and (1) implies that entropy solutions of (1) are BV-stable: If the initial data $u_0 \in \mathrm{L}^\infty \cap \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}^n)$, then also the entropy solution $u \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R}^n)$. The importance of this observation for the theory of scalar conservation laws comes from the fact that BV-functions have a very particular structure. To explain this statement we need the following definitions.

Definition 3.1 (Rectifiable sets). *Let $\Omega \subset \mathbb{R}^{n+1}$ be open. A subset $\mathcal{J} \subset \Omega$ is called $\mathscr{H}^n$-rectifiable if $\mathcal{J} = E \cup \bigcup_{k\in\mathbb{N}} E_k$, where $\mathscr{H}^n(E) = 0$ and each E_k is contained in an n-dimensional Lipschitz continuous submanifold of Ω. Here $\mathscr{H}^n$ is the n-dimensional Hausdorff measure.*

Definition 3.2 (Orientation). *Let $\Omega \subset \mathbb{R}^{n+1}$ be open and let $\mathcal{J} \subset \Omega$ be an $\mathscr{H}^n$-rectifiable set. An orientation of $\mathcal{J}$ is a Borel vector field $\mu\colon \mathcal{J} \longrightarrow \mathbb{S}^n$ with the property that for $\mathscr{H}^n$-a.e. $\mathbf{x} \in \mathcal{J}$, the vector $\mu(\mathbf{x})$ is normal to $\mathcal{J}$.*

Definition 3.3 (Traces of u). *Let $\Omega \subset \mathbb{R}^{n+1}$ be open and let $\mathcal{J} \subset \Omega$ be an $\mathscr{H}^n$-rectifiable set oriented by a normal vector field μ. We say that two Borel functions $u^\pm\colon \mathcal{J} \longrightarrow \mathbb{R}$ are the traces of u on $\mathcal{J}$ if for $\mathscr{H}^n$-a.e. $\mathbf{y} \in \mathcal{J}$*

$$\lim_{r\to 0} \left(\fint_{B_r^+(\mathbf{y})} |u(\mathbf{x}) - u^+(\mathbf{y})|\, d\mathbf{x} + \fint_{B_r^-(\mathbf{y})} |u(\mathbf{x}) - u^-(\mathbf{y})|\, d\mathbf{x} \right) = 0,$$

where $B_r^\pm(\mathbf{y}) := B_r(\mathbf{y}) \cap \{\mathbf{x} \in \mathbb{R}^{n+1}\colon\ \pm\mu(\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) > 0\}$.

The Structure Theorem (see Section 3.9 in [2]) states that the derivative of a BV-function u can be decomposed into three parts: We have

$$Du = (Du)_a + (Du)_c + (Du)_j,$$

where the first component $(Du)_a$ is absolutely continuous with respect to the Lebesgue measure $\mathscr{L}^{n+1}$, and the second part $(Du)_c$ (called the Cantor part) is a singular measure that, however, is small in some sense. The third part $(Du)_j$ (called the jump part) can be written in the form

$$(Du)_j = (u^+ - u^-)\mu\, \mathscr{H}^n \lfloor \mathcal{J}. \tag{19}$$

Here $\mathcal{J} \subset \Omega$ is an $\mathscr{H}^n$-rectifiable set oriented by a unit normal vector field $\mu\colon \mathcal{J} \longrightarrow \mathbb{S}^n$, and the functions $u^\pm\colon \mathcal{J} \longrightarrow \mathbb{R}$ are the traces of u on $\mathcal{J}$. It is possible to generalize the classical chain rule to functions of bounded variation: If u is a BV-function and $g\colon \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz continuous, then also the composition $g(u)$ is of bounded variation. Therefore its derivative $Dg(u)$ can be decomposed into three terms as above, and in particular

$$\big(Dg(u)\big)_j = \big(g(u^+) - g(u^-)\big)\mu\, \mathscr{H}^n \lfloor \mathcal{J},$$

where the rectifiable set $\mathcal{J}$ and the functions μ and $u^\pm$ are the same as (19). We refer the reader to Theorem 3.101 in [2]. A function $u\colon \mathbb{R}^+ \times \mathbb{R}^n \longrightarrow \mathbb{R}$ of bounded variation is then a weak solution of (1) if the following holds:

- With the subscript a denoting the absolutely continuous parts, we have

$$\left(\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{f}(u)\right)_a = \left(\frac{\partial u}{\partial t}\right)_a + \mathbf{f}'(u) \cdot (\nabla u)_a = 0, \tag{20}$$

  and a similar statement is true for the Cantor parts.
- With the same notation as in (19), we have

$$-s\big(u^+ - u^-\big) + \big(\mathbf{f}(u^+) - \mathbf{f}(u^-)\big) \cdot \nu = 0 \quad \mathscr{H}^n\text{-a.e. in } \mathcal{J}, \tag{21}$$

  where (s, ν) is defined by $\mu\sqrt{1+s^2} = (-s, \nu)$ and $\nu \in \mathbb{S}^{n-1}$. The number s is the shock speed, and (21) is called the Rankine-Hugoniot condition.

Under the assumption of BV-regularity, also the entropy condition simplifies: Let $(\eta, \mathbf{q})$ be any convex entropy/entropy flux pair with $\mathbf{q}' = \eta'\mathbf{f}'$.

- With the subscript a denoting the absolutely continuous parts, we have

$$\left(\frac{\partial \eta(u)}{\partial t} + \nabla \cdot \mathbf{q}(u)\right)_a = \eta'(u)\left(\left(\frac{\partial u}{\partial t}\right)_a + \mathbf{f}'(u) \cdot (\nabla u)_a\right) = 0$$

  because of (20), and a similar statement is true for the Cantor parts.
- With the same notation as in (21), we have

$$-s\big(\eta(u^+) - \eta(u^-)\big) + \big(\mathbf{q}(u^+) - \mathbf{q}(u^-)\big) \cdot \nu \leqslant 0 \quad \mathscr{H}^n\text{-a.e. in } \mathcal{J}.$$

These considerations show that BV-regularity is a very desirable property for entropy solutions of scalar conservation laws. As we will discuss in the next section, however, entropy solutions are typically not of bounded variation. Let us introduce one more definition that will be used there.

Definition 3.4. *Let $\Omega \subset \mathbb{R}^{n+1}$ be an open set and $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. We say that the function u has vanishing mean oscillation (VMO) at a point $\mathbf{y} \in \Omega$ if*

$$\lim_{r\to 0} \fint_{B_r(\mathbf{y})} \left| u(\mathbf{x}) - \fint_{B_r(\mathbf{y})} u(\mathbf{z})\, d\mathbf{z} \right| d\mathbf{x} = 0.$$

We say that u is approximately continuous at $\mathbf{y} \in \Omega$ if u has VMO there and

$$\lim_{r\to 0} \fint_{B_r(\mathbf{y})} u(\mathbf{z})\, d\mathbf{z} = u(\mathbf{y}).$$

All Lebesgue-measurable functions are approximately continuous $\mathscr{L}^{n+1}$-a.e. For BV-functions this statement can be improved to approximate continuity $\mathscr{H}^n$-a.e. off the jump set $\mathcal{J}$. We refer the reader to Section 3.9 in [2].

4 Structure of Entropy Solutions

As explained in Section 2, unique entropy solutions do exist even for rough initial data that is not in $\mathrm{BV}(\mathbb{R}^n)$. What can be said about the structure of these solutions? Under appropriate assumptions on the nondegeneracy of the flux, entropy solutions do have some extra regularity (as follows from the velocity averaging arguments), but typically they are not of bounded variation. In this section we discuss in more detail the regularizing effects due to the interplay between the nonlinearity of the problem and the entropy condition. It turns out that the distinction between time and space variables is not essential. Therefore we consider a slightly more general situation:

Definition 4.1. *Let $\Omega \subset \mathbb{R}^n$ be an open set, and let $\mathbf{f}\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ be a smooth flux function. For any convex entropy $\eta\colon \mathbb{R} \longrightarrow \mathbb{R}$ let the corresponding entropy flux $\mathbf{q}_\eta\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ be defined (up to a constant) by the condition*

$$\mathbf{q}_\eta'(v) = \eta'(v)\mathbf{f}'(v) \quad \forall v \in \mathbb{R}. \tag{22}$$

The function $u\colon \Omega \longrightarrow (0,1)$ is called a generalized entropy solution if

$$\nabla \cdot \mathbf{f}(u) = 0 \quad \text{in } \mathscr{D}'(\Omega) \tag{23}$$

and if for all convex entropy/entropy flux pairs $(\eta, \mathbf{q}_\eta)$

$$\nabla \cdot \mathbf{q}_\eta(u) \in \mathrm{M}_{\mathrm{loc}}(\Omega). \tag{24}$$

Remark 4.2. Notice that in (24) we do not require that the entropy dissipation measure has a sign. Our definition therefore includes certain weak solutions of (23) that contain non-classical (entropy violating) shocks.

We first discuss the regularity of generalized entropy solutions. For definiteness, we consider only the higher-dimensional version of Burgers' flux:

$$\mathbf{f}(v) := \Big(v, \tfrac{1}{2}v^2, \dots, \tfrac{1}{n}v^n\Big) \quad \forall v \in \mathbb{R}. \tag{25}$$

Proposition 4.3. *Let $\Omega \subset \mathbb{R}^n$ be open. There exists a constant $C > 0$ with the following property: Let u be a generalized entropy solution u corresponding to the generalized Burgers' flux* (25)*. Assume that there is no entropy dissipation in Ω: for all convex entropy/entropy flux pairs $(\eta, \mathbf{q}_\eta)$ defined by* (22)

$$\nabla \cdot \mathbf{q}_\eta(u) = 0 \quad \text{in } \mathscr{D}'(\Omega).$$

For any compact subset $K \subset \Omega$ and $R := \operatorname{dist}(K, \mathbb{R}^n \setminus \Omega)$ we then have

$$\sup_{(x,y)\in K\times K} \frac{|u(x)-u(y)|}{|x-y|^{\frac{1}{n-1}}} \leqslant C R^{-\frac{1}{n-1}} \Big(1 + \|u\|_{\mathrm{L}^\infty(K)}\Big)^3.$$

The exponent $1/(n-1)$ is optimal.

For Burgers' equation with $n = 2$ we recover the well-known fact that entropy solutions are locally Lipschitz continuous in open sets that do not meet the shock set. We postpone the proof of the the first part of Proposition 4.3 to the Appendix, see page 37. To prove the optimality of the Hölder exponent $1/(n-1)$ let $\Omega := (0,1)^{n-1} \times \mathbb{R}$. Then we construct a solution of (23) that only depends on x_1 and x_n and is constant along characteristics. Consider

$$u_0(x_n) := \begin{cases} x_n^\alpha & \text{if } x_n > 0, \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

for some number $\alpha > 0$. Then the function u, implicitly defined by

$$u(x) = u_0\big(x_n - x_1 \mathbf{f}_n'(u(x))\big) \quad \forall x \in \Omega,$$

is a solution of (23) as follows from easy inspection. In view of (26), this gives

$$u(x) = \begin{cases} \big(x_n - x_1 u(x)^{n-1}\big)^\alpha & \text{if } x_n - x_1 u(x)^{n-1} \geqslant 0, \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

In particular, we have $u(x) \geqslant 0$, thus $u(x) = 0$ if $x_n < 0$. Indeed, for $x_n < 0$ the first case in (27) never applies since $x_1 u(x)^{n-1} \geqslant 0$. We rewrite (27) as

$$u(x)^{1/\alpha} + x_1 u(x)^{n-1} = x_n \quad \text{if } x_n \geqslant 0. \tag{28}$$

Solving this identity for $u(x)$ yields

$$u(x) \approx \begin{cases} x_n^\alpha & \text{if } \alpha > 1/(n-1), \\ (x_n/x_1)^{1/(n-1)} & \text{otherwise} \end{cases}$$

for small $x_n > 0$. The second case shows that rough data with α small is regularized to Hölder continuity with exponent $1/(n-1)$. On the other hand, if α is large and thus u_0 is smooth, then the solution u stays smooth.

We now consider the case with entropy dissipation.

Proposition 4.4. *Let $\Omega \subset \mathbb{R}^n$ be open and let $\mathbf{e}_m$ denote the mth unit basis vector of $\mathbb{R}^n$. Then there exists a constant $C > 0$ with the following property: Let u be a compactly supported, generalized entropy solution corresponding to the generalized Burgers' flux* (25)*. For all $m \in \{1, \dots, n\}$ we then have*

$$\begin{aligned} &\sup_{h \neq 0} |h|^{-\frac{1}{n+1}} \|u(\cdot + h\mathbf{e}_m) - u\|_{\mathrm{L}^1(\Omega)} \\ &\qquad \leqslant C\bigg(|\operatorname{spt} u|^m \, \|u\|_{\mathrm{L}^1(\Omega)}^{n-m} \, \|\mu\|_{\mathrm{M}(\mathbb{R}\times\Omega)} \bigg)^{\frac{1}{n+1}}, \end{aligned} \tag{29}$$

where μ is the kinetic entropy dissipation measure (introduced in Theorem 5.1 below), which is finite. The exponent $1/(n+1)$ is optimal.

Notice that (29) actually implies that u is in the Besov space $\mathrm{B}^{1/(n+1)}_{1,\infty}(\Omega)$. For Burgers' equation with $n = 2$ we obtain a differentiability of one third. The regularity $1/(n+1)$ was obtained independently in [24] for entropy solutions of (23) (for which the entropy dissipation in (24) is a nonnegative measure). Their proof uses a new Fourier multiplier estimate combined with the bootstrap argument already employed in [20]. In contrast to this, we give a rather elementary proof of Proposition 4.4 in the Appendix, which completely avoids the use of Fourier methods, see page 40. The optimality of the result again follows from an explicit example: Let $\Omega := (0,1)^{n-1} \times \mathbb{R}$ and consider nonincreasing sequences $\Delta_k \in \ell^1$ and $c_k \in \ell^{n+1}$ to be specified later. Let

$$a_k^- := \sum_{l=1}^{k} \Delta_l \quad \text{and} \quad a_k^+ := a_k^- - \Delta_k/2$$

for all $k \in \mathbb{N}$. Then we define the function

$$u(x) := \sum_{k=1}^{\infty} c_k \mathbf{1}_{I_k}(x_n + s_k x_1) \quad \forall x \in \Omega, \tag{30}$$

where $I_k := [a_k^+, a_k^-]$ and $s_k := \frac{1}{n} c_k^{n-1}$ for all $k \in \mathbb{N}$. We claim that (30) is a generalized entropy condition of (23), and prove first that the Rankine-Hugoniot condition is satisfied along each discontinuity. Indeed, consider

$$\mathcal{J}_k^{\pm} := \big\{ x \in \Omega \colon x_n + s_k x_1 = a_k^{\pm} \big\}$$

with unit normal vectors $\nu_k := (-s_k, 0, \ldots, 0, 1)/\sqrt{1+s_k^2}$. Then we have

$$\nabla \cdot \mathbf{f}(u) = \sum_{k=1}^{\infty} \frac{-s_k \mathbf{f}_1(c_k) + \mathbf{f}_n(c_k)}{\sqrt{1+s_k^2}} \left(\mathscr{H}^{n-1} \lfloor \mathcal{J}_k^+ - \mathscr{H}^{n-1} \lfloor \mathcal{J}_k^- \right) = 0,$$

using that $\mathbf{f}(0) = 0$. To check assumption (24), consider for any convex entropy $\eta \colon \mathbb{R} \longrightarrow \mathbb{R}$ the corresponding entropy flux $\mathbf{q}_\eta$ defined by

$$\mathbf{q}_\eta(u) := \int_0^u \mathbf{f}'(v) \eta'(v) \, dv \quad \forall u \in \mathbb{R}.$$

Then the entropy dissipation is given as

$$\nabla \cdot \mathbf{q}_\eta(u) = \sum_{k=1}^{\infty} \frac{-s_k \mathbf{q}_{\eta,1}(c_k) + \mathbf{q}_{\eta,n}(c_k)}{\sqrt{1+s_k^2}} \left(\mathscr{H}^{n-1} \lfloor \mathcal{J}_k^+ - \mathscr{H}^{n-1} \lfloor \mathcal{J}_k^- \right)$$

because $\mathbf{q}_\eta(0) = 0$. Notice that after an integration by parts

$$\begin{aligned} -s_k \mathbf{q}_{\eta,1}(c_k) + \mathbf{q}_{\eta,n}(c_k) &= \int_0^{c_k} \eta'(v) \big(-\tfrac{1}{n} c_k^{n-1} + v^{n-1} \big) \, dv \\ &= -\tfrac{1}{n} \int_0^{c_k} \eta''(v) \big(c_k^{n-1} v - v^n \big) \, dv, \end{aligned}$$

which implies that the total entropy dissipation satisfies

$$\|\nabla \cdot \mathbf{q}_\eta(u)\|_{\mathrm{M}(\Omega)} \leqslant C \sum_{k=1}^{\infty} |c_k|^{n+1} < \infty,$$

for some constant C which only depends on n and the sup of η'' on a compact set. This shows that u is indeed a generalized entropy solution of (23). On the other hand, we can estimate the finite difference

$$h^{-\alpha}\|u(\cdot + h\mathbf{e}_n) - u\|_{\mathrm{L}^1(\Omega)} \geqslant h^{-\alpha} \sum_{\Delta_k/2 \geqslant h} 2h|c_k| \quad \forall h > 0, \tag{31}$$

where $\mathbf{e}_n$ is the nth standard basis vector of $\mathbb{R}^n$ and $\alpha > 0$ is some number. Indeed, for all k with $\Delta_k/2 \geqslant h$ we have $(h\mathbf{e}_n + I_k) \cap I_{k+1} = \varnothing$. Let

$$\Delta_k := k^{-(1+\varepsilon)} \quad \text{and} \quad c_k := k^{-(1+\varepsilon)/(n+1)}$$

for all $k \in \mathbb{N}$ and some $\varepsilon > 0$. Then $\Delta_k \in \ell^1$ and $c_k \in \ell^{n+1}$ as required and

$$h^{-\alpha} \sum_{\Delta_k/2 \geqslant h} 2h|c_k| \geqslant C h^{-\alpha+1/(n+1)+\varepsilon/(1+\varepsilon)} \quad \forall h > 0, \tag{32}$$

with C some constant independent of h. Assume now that $\alpha > 1/(n+1)$. Then there exists $\varepsilon > 0$ small such that the exponent in (32) is negative. We conclude that the left-hand side of (31) blows up as $h \to 0$. This shows that the maximal regularity we can hope for is $\alpha = 1/(n+1) < 1$.

We conclude that typically a generalized entropy solution is not of bounded variation. One might wonder, however, whether a generalized entropy solution still has the same structure as a BV-function. This is indeed the case in the sense made precise in the following theorem.

Theorem 4.5 (De Lellis, Otto, Westdickenberg). *Let $\Omega \subset \mathbb{R}^n$ be open and assume that the flux $\mathbf{f}\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ satisfies the nondegeneracy condition*

$$\forall \xi \in \mathbb{S}^{n-1}: \quad \mathscr{L}^1\big(\{v \in \mathbb{R} : \mathbf{f}'(v) \cdot \xi = 0\}\big) = 0. \tag{33}$$

Let u be a generalized entropy solution in the sense of Definition 4.1. Then there exists an $\mathscr{H}^{n-1}$-rectifiable set $\mathcal{J} \subset \Omega$ such that:

- *For all $y \in \mathcal{J}$ the function u has strong traces on $\mathcal{J}$.*
- *For all $y \notin \mathcal{J}$ the function u has vanishing mean oscillation.*

Remark 4.6. To simplify the presentation, we will consider only the case of a classical entropy solution in the following, for which the entropy dissipation in (24) is a *nonnegative* measure. We refer the reader to [20] for the necessary modifications in the general case of measures which change sign.

Theorem 4.5 shows that generalized entropy solutions have a BV-like fine structure without actually being of bounded variation. It is an open problem whether VMO can be improved to approximate continuity of u outside $\mathcal{J}$ (see however the result by De Lellis and Rivière [12]). Another open problem is to prove that there is no entropy dissipation outside of $\mathcal{J}$. For entropy solutions of bounded variation this follows from the BV-chain rule and (23). Our Theorem 4.5 only ensures that the entropy dissipation restricted to $\mathcal{J}$ has the correct structure: If $\nu\colon \mathcal{J} \longrightarrow \mathbb{S}^{n-1}$ denotes a normal vector field along $\mathcal{J}$, then we have for all convex entropy/entropy flux pairs $(\eta, \mathbf{q}_\eta)$ that

$$\Big(\nabla\cdot\mathbf{q}_\eta(u)\Big)\lfloor\mathcal{J} = \Big(\mathbf{q}_\eta(u^+) - \mathbf{q}_\eta(u^-)\Big)\cdot\nu\,\mathscr{H}^{n-1}\lfloor\mathcal{J}.$$

To prove that $\nabla\cdot\mathbf{q}_\eta(u) = 0$ outside $\mathcal{J}$ would probably require some analogue of the BV-chain rule for generalized entropy solutions. This is a hard problem. We refer the reader to [1] for some results in that direction.

Notice that the fine structure result contains the existence of strong traces on codimension-one rectifiable subsets as a subproblem. This is relevant for understanding how entropy solutions of nondegenerate scalar conservation laws attain their initial or boundary data. The problem has first been studied by Vasseur [25] who uses techniques quite similar to ours. In particular, the idea of "blowing up" a neighborhood of a given point (see Definition 5.7) was introduced there. We also refer the reader to [7, 8, 23] for related results.

5 Kinetic formulation, Blow-ups and Split States

In this section we provide some tools we will need later. We start by proving a variant of the kinetic formulation for scalar conservation laws, introduced by Lions, Perthame and Tadmor in their seminal paper [20]. In the following, we will systematically use the notation $\mathbf{a} = \mathbf{f}'$.

Theorem 5.1 (Kinetic Formulation). *Let $\Omega \subset \mathbb{R}^n$ be an open set, and assume that u is an entropy solution. Let the function χ be defined by*

$$\chi(v, u(x)) := \begin{Bmatrix} 1 & \textit{if } 0 < v \leqslant u(x) \\ 0 & \textit{otherwise} \end{Bmatrix} \quad \forall (v,x) \in \mathbb{R}\times\Omega. \tag{34}$$

Then there exists a nonnegative measure $\mu \in \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R}\times\Omega)$ such that

$$\mathbf{a}(v)\cdot\nabla\chi(v,u(x)) = \tfrac{\partial}{\partial v}\mu \quad \textit{in } \mathscr{D}'(\mathbb{R}\times\Omega). \tag{35}$$

Remark 5.2. A similar construction works also for generalized entropy solutions, for which we only assume that the entropy dissipation is a locally finite measure. Then the measure μ can change sign. We refer to reader to [11].

Proof. Consider the linear map Φ defined by

$$\Phi(\eta,\varphi) := \int_\Omega \mathbf{q}_\eta(u(x)) \cdot \nabla\varphi(x)\, dx \quad \forall(\eta,\varphi) \in \mathscr{D}(\mathbb{R}) \times \mathscr{D}(\Omega), \tag{36}$$

where $\mathbf{q}_\eta$ is related to η by the compatibility condition $\mathbf{q}_\eta'(v) = \eta'(v)\mathbf{a}(v)$ for all $v \in \mathbb{R}$. Notice that the map Φ is indeed well-defined since $\mathbf{q}_\eta$ is unique up to a constant and φ has compact support. We have

$$\eta \text{ linear} \quad \Longrightarrow \quad \Phi(\eta,\varphi) = 0 \quad \forall\varphi \in \mathscr{D}(\Omega),$$

since then $\mathbf{q}_\eta(u) = \alpha\mathbf{f}(u) + \beta$ for some constants α, β, and $\nabla \cdot \mathbf{f}(u) = 0$ in $\mathscr{D}'(\Omega)$. This implies that Φ depends on η only through η''. We also have

$$\eta \text{ convex} \quad \Longrightarrow \quad \Phi(\eta,\varphi) \geqslant 0 \quad \forall\varphi \in \mathscr{D}(\Omega), \varphi \geqslant 0$$

because u is an entropy solution. (It suffices for η to be convex on the unit interval because $u\colon \Omega \longrightarrow (0,1)$.) Recalling that a nonnegative distribution is in fact a measure, we can therefore find $\mu \in \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R} \times \Omega)$ such that

$$\int_\Omega \mathbf{q}_\eta(u(x)) \cdot \nabla\varphi(x)\, dx = \iint_{\mathbb{R}\times\Omega} \eta''(v)\varphi(x)\, d\mu(v,x) \tag{37}$$

for all $(\eta,\varphi) \in \mathscr{D}(\mathbb{R}) \times \mathscr{D}(\Omega)$. Notice that by definition of χ and the compatibility condition for $\mathbf{q}_\eta$ we have that (up to a constant)

$$\mathbf{q}_\eta(u(x)) = \int_\mathbb{R} \chi(v,u(x))\mathbf{q}_\eta'(v)\, dv = \int_\mathbb{R} \chi(v,u(x))\eta'(v)\mathbf{a}(v)\, dv.$$

Hence (37) turns into

$$\int_\Omega \left(\int_\mathbb{R} \chi(v,u(x))\eta'(v)\mathbf{a}(v)\, dv \right) \cdot \nabla\varphi(x)\, dx = \iint_{\mathbb{R}\times\Omega} \eta''(v)\varphi(x)\, d\mu(v,x),$$

which can be rewritten as

$$\iint_{\mathbb{R}\times\Omega} \chi(v,u(x))\nabla\big[\eta'(v)\varphi(x)\big] \cdot \mathbf{a}(v)\, dv\, dx = \iint_{\mathbb{R}\times\Omega} \tfrac{\partial}{\partial v}\big[\eta'(v)\varphi(x)\big]\, d\mu(v,x).$$

Since linear combinations of products $\eta'\varphi$ with $\eta' \in \mathscr{D}(\mathbb{R})$ and $\varphi \in \mathscr{D}(\Omega)$ are dense in $\mathscr{D}(\mathbb{R} \times \Omega)$ (up to a constant), we conclude that indeed

$$\iint_{\mathbb{R}\times\Omega} \chi(v,u(x))\mathbf{a}(v) \cdot \nabla\zeta(v,x)\, dv\, dx = \iint_{\mathbb{R}\times\Omega} \tfrac{\partial}{\partial v}\zeta(v,x)\, d\mu(v,x)$$

for all $\zeta \in \mathscr{D}(\mathbb{R} \times \Omega)$. In general, the measure μ is only locally finite. □

The kinetic formulation can be used to prove the following compactness result for bounded sequences of entropy solutions of (23), see [20].

Theorem 5.3. *Consider a sequence of entropy solutions u_k of* (23) *in some open set $\Omega \subset \mathbb{R}^n$. According to Theorem 5.1 there exists a sequence of non-negative measures $\mu_k \in \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R} \times \Omega)$ such that the pairs (u_k, μ_k) satisfy the kinetic equation* (35). *Assume that the flux $\mathbf{f}\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ is sufficiently smooth and satisfies the nondegeneracy condition* (33). *Assume also that*

$$\text{the measures } \mu_k \text{ are locally uniformly bounded.}$$

Then there exists a subsequence $k_l \longrightarrow \infty$ such that

$$\begin{aligned} u_{k_l} &\longrightarrow u \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n), \\ \mu_{k_l} &\overset{*}{\rightharpoonup} \mu \quad \text{in } \mathrm{M}_{\mathrm{loc}}(\mathbb{R} \times \mathbb{R}^n) \end{aligned}$$

as $l \to \infty$, and the limit (u, μ) satisfies the kinetic equation (35). *In particular, u is a generalized entropy solution of the scalar conservation law* (23).

The first step in proving our Theorem 4.5 is to identify a candidate for the jump set $\mathcal{J} \subset \Omega$. We recall that in the case of an entropy solution u with BV-regularity, the jump set $\mathcal{J}$ is exactly the set where entropy dissipation takes place. By the kinetic formulation, this in turn is related to the support of the kinetic measure μ. Therefore the following definition is natural.

Definition 5.4. *Let u be an entropy solution of* (23) *in $\Omega \subset \mathbb{R}^n$ open, and let μ be the corresponding entropy dissipation measure provided by the kinetic formulation. Let $\nu \in \mathrm{M}^+_{\mathrm{loc}}(\Omega)$ be the x-marginal of μ, defined as*

$$\nu(A) := \mu(\mathbb{R} \times A) \quad \forall \text{Borel sets } A \subset \Omega.$$

We denote by $\mathcal{J}$ the set of points with positive upper $(n-1)$-density of ν:

$$\mathcal{J} := \left\{ y \in \Omega : \limsup_{r \to 0} \frac{\nu(B_r(y))}{r^{n-1}} > 0 \right\}. \tag{38}$$

The main task is then to prove that $\mathcal{J}$ is indeed an $\mathscr{H}^{n-1}$-rectifiable set. For later reference, we first record the following observation.

Lemma 5.5. *Let u be an entropy solution of* (23) *in $\Omega \subset \mathbb{R}^n$ open, and let μ be the entropy dissipation measure provided by the kinetic formulation. Let ν be the x-marginal of μ. Then there exists a constant $C > 0$ such that*

$$\limsup_{r \to 0} \frac{\nu(B_r(y))}{r^{n-1}} \leqslant C \quad \forall y \in \Omega. \tag{39}$$

Proof. Fix a test function $\zeta \in \mathscr{D}(\mathbb{R}^n)$ with $\zeta \geqslant 0$ and $\zeta(x) = 1$ for $x \in B_1(0)$. Choose an entropy $\eta \in \mathscr{D}(\mathbb{R})$ such that $\eta(v) = \frac{1}{2}v^2$ for $v \in [0,1]$, and let $\mathbf{q}$ be the corresponding entropy flux. Using the nonnegativity of μ, we have

$$\begin{aligned}\nu(B_r(y)) &= \iint_{\mathbb{R}\times B_r(y)} \eta''(v)\,d\mu(v,x)\\ &\leqslant \iint_{\mathbb{R}\times\Omega} \eta''(v)\zeta\left(\frac{x-y}{r}\right)d\mu(v,x) \quad \forall B_r(y)\subset\Omega.\end{aligned}$$

Then the kinetic equation (35) yields

$$\begin{aligned}\nu(B_r(y)) &\leqslant \int_\Omega \left(\int_{\mathbb{R}} \eta'(v)\mathbf{a}(v)\,\chi\big(v,u(x)\big)\,dv\right)\cdot\nabla\zeta\left(\frac{x-y}{r}\right)dx\\ &= \int_\Omega \mathbf{q}(u(x))\cdot\nabla\zeta\left(\frac{x-y}{r}\right)dx\\ &\leqslant \|\mathbf{q}(u)\|_{\mathrm{L}^\infty(\Omega)}\|\nabla\zeta\|_{\mathrm{L}^1(\mathbb{R}^n)}r^{n-1} \quad \forall B_r(y)\subset\Omega.\end{aligned}$$

The lemma follows. □

There exist several criteria to check a set for rectifiability. We will use the following well-known result (see Theorem 15.19 in [21]).

Theorem 5.6 (Rectifiability criterion). *Let $\mathcal{J}\subset\mathbb{R}^n$ be a set and assume that there exists a measure $\nu\in\mathrm{M}^+_{\mathrm{loc}}(\mathbb{R}^n)$ with the following properties:*

- *For all $y\in\mathcal{J}$ we have*

$$\liminf_{r\to 0}\frac{\nu(B_r(y))}{r^{n-1}} > 0. \tag{40}$$

- *For all $y\in\mathcal{J}$ there exist an orthonormal coordinate system $x_1,\dots,x_n$ and a cone $\mathrm{C}_y := \big\{x\in\mathbb{R}^n : |x_1|\geqslant c|(x_2,\dots,x_n)|\big\}$ with $c>0$ such that*

$$\lim_{r\to 0}\frac{\nu\big((y+\mathrm{C}_y)\cap B_r(y)\big)}{r^{n-1}} = 0. \tag{41}$$

Then $\mathcal{J}$ is an $\mathscr{H}^{n-1}$-rectifiable set.

We use for ν the x-marginal of the entropy dissipation measure μ. As is suggested by the rectifiability criterion, we study blow-ups: That is, we look at the structure of entropy solutions after "zooming" into a point $y\in\Omega$.

Definition 5.7. *Let u be an entropy solution of (23) in $\Omega\subset\mathbb{R}^n$ open, and let μ be the corresponding entropy dissipation measure provided by the kinetic formulation. Let ν be the x-marginal of μ. For any $y\in\Omega$ and $r>0$ let*

$$\begin{aligned}u^{y,r}(x) &:= u(y+rx),\\ \mu^{y,r}(B\times A) &:= r^{1-n}\mu(B\times(y+rA)),\\ \nu^{y,r}(A) &:= r^{1-n}\nu(y+rA)\end{aligned} \tag{42}$$

for all $x\in\mathbb{R}^n$ and all Borel sets $A\subset\Omega$ and $B\subset\mathbb{R}$. A sequence of rescaled quantities $(u^{y,r},\mu^{y,r},\nu^{y,r})$ for $r\to 0$ will be called a blow-up sequence.

Notice that the measures $\mu^{y,r}$ and $\nu^{y,r}$ are also characterized by

$$\iint_{\mathbb{R}\times\mathbb{R}^n} \zeta(v,x)\,d\mu^{y,r}(v,x) = r^{1-n}\iint_{\mathbb{R}\times\mathbb{R}^n} \zeta\left(v,\frac{x-y}{r}\right)d\mu(v,x) \tag{43}$$

$$\int_{\mathbb{R}^n} \varphi(x)\,d\nu^{y,r}(x) = r^{1-n}\int_{\mathbb{R}^n} \varphi\left(\frac{x-y}{r}\right)d\nu(x) \tag{44}$$

for all $\zeta \in \mathscr{D}(\mathbb{R}\times\mathbb{R}^n)$ and $\varphi \in \mathscr{D}(\mathbb{R}^n)$. For simplicity, we will always assume that μ, ν are extended by zero to $\mathbb{R}^n \setminus \Omega$. The bound (39) translates into

$$\limsup_{r\to 0} \nu^{y,r}(B_1(0)) < \infty \quad \forall y \in \Omega. \tag{45}$$

We know introduce a class of special solutions of (35), for which the entropy dissipation measure μ has a tensor product form.

Definition 5.8. *A split state is a triple (u,h,ν) consisting of*

- *a function $u \in \mathrm{L}^\infty(\mathbb{R}^n)$,*
- *a left-continuous function $h \in \mathrm{BV}(\mathbb{R})$,*
- *a nonnegative measure $\nu \in \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R}^n)$*

such that for every $v \in \mathbb{R}$

$$\mathbf{a}(v)\cdot\nabla\chi(v,u) = h(v)\nu \quad \text{in } \mathscr{D}'(\mathbb{R}^n). \tag{46}$$

The following proposition is a key step in the proof of our Theorem 4.5.

Proposition 5.9 (Blow-ups are Split States). *Let u be an entropy solution of (23) in some open set $\Omega \subset \mathbb{R}^n$. Then there exists a set $E \subset \Omega$ with $\mathscr{H}^{n-1}(E) = 0$ and with the following property: for every point $y \in \Omega\setminus E$ there exists a left-continuous function $h_y \in \mathrm{BV}(\mathbb{R})$ such that*

$$\begin{pmatrix} u^{y,r_k} \longrightarrow u^\infty \text{ in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \\ \nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty \text{ in } \mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n) \end{pmatrix} \quad\Longrightarrow\quad (u^\infty, h_y, \nu^\infty) \text{ is a split state.}$$

For all $y \in \Omega$ there exists at least one $r_k \to 0$ for which (u^{y,r_k}, ν^{y,r_k}) converge.

Remark 5.10. We emphasize that the function h_y only depends on the blow-up point $y \in \Omega \setminus E$, not on the particular blow-up sequence $r_k \to 0$. On the contrary, the limits (u^∞, ν^∞) might depend on the sequence.

Proof. The fact that there always exists a subsequence such that (u^{y,r_k}, ν^{y,r_k}) converge as $r_k \to 0$ follows from Banach-Alaoglu theorem and Theorem 5.3, once we have checked that the rescaled measures are locally uniformly bounded. For $\nu^{y,r}$ this follows from Lemma 5.5, and $\nu^{y,r}$ is the x-marginal of $\mu^{y,r}$. In particular, we have $\nu^{y,r} \overset{*}{\rightharpoonup} 0$ in $\mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n)$ for any blow-up sequence around a point $y \notin \mathscr{J}$. Then also $\mu^{y,r} \overset{*}{\rightharpoonup} 0$ in $\mathrm{M}_{\mathrm{loc}}(\mathbb{R}\times\mathbb{R}^n)$; so we can choose $h_y := 0$ in this case. The argument for $y \in \mathscr{J}$ is slightly more involved.

Step 1. By disintegration of measures (see Theorem 2.28 of [2]) there exists a weakly* ν-measurable map $H\colon \Omega \longrightarrow \mathrm{M}^+(\mathbb{R})$ such that

$$\iint_{\mathbb{R}\times\Omega} \zeta(v,x)\,d\mu(v,x) = \int_\Omega \int_{\mathbb{R}} \zeta(v,x)\,dH_x(v)\,d\nu(x) \quad \forall \zeta \in \mathscr{D}(\mathbb{R}\times\Omega).$$

Select a countable family $\mathcal{S} \subset \mathscr{D}(\mathbb{R})$ which is dense in $\mathscr{D}(\mathbb{R})$ with respect to the uniform topology. For every $\psi \in \mathcal{S}$ we define a map $f_\psi\colon \Omega \longrightarrow \mathbb{R}$ with

$$f_\psi(y) := \int_{\mathbb{R}} \psi(v)\,dH_y(v) \quad \forall y \in \Omega. \tag{47}$$

Then f_ψ is ν-measurable and $f_\psi \in \mathrm{L}^1_{\mathrm{loc}}(\Omega,\nu)$. Let $\mathrm{Leb}(\psi)$ be the set of Lebesgue points of f_ψ. By derivation of measures (see Corollary 2.23 in [2]), we have $\nu(E_\psi) = 0$ for the complement $E_\psi := \Omega \setminus \mathrm{Leb}(\psi)$. We even have $\nu(E_{\mathcal{S}}) = 0$ for $E_{\mathcal{S}} := \bigcup_{\psi\in\mathcal{S}} E_\psi$ since $\mathcal{S}$ is countable. Now fix a point $y \in \Omega \setminus E_{\mathcal{S}}$. For all $r > 0$ we define a linear map $\mathcal{F}_r$ by

$$\mathcal{F}_r(\zeta) := \frac{1}{\nu^{y,r}(B)} \iint_{\mathbb{R}\times B} \zeta(v,x)\left(dH_y(v)\,d\nu^{y,r}(x) - d\mu^{y,r}(v,x)\right)$$

for all $\zeta \in \mathscr{D}(\mathbb{R}\times B)$ with $B := B_1(0)$. Choose $\varphi \in \mathscr{D}(B)$ and $\psi \in \mathcal{S}$. Using definitions (42)–(44), (47) and the decomposition $\mu = H\nu$, we obtain

$$\begin{aligned} |\mathcal{F}_r(\psi\varphi)| &= \left| \frac{1}{\nu(B_r(y))} \iint_{\mathbb{R}\times B_r(y)} \psi(v)\varphi\left(\frac{x-y}{r}\right)\left(dH_y(v) - dH_x(v)\right)d\nu(x) \right| \\ &\leqslant \|\varphi\|_{\mathrm{L}^\infty(\mathbb{R}^n)} \fint_{B_r(y)} |f_\psi(y) - f_\psi(x)|\,d\nu(x) \longrightarrow 0 \quad \text{as } r \to 0, \end{aligned}$$

because y is a Lebesgue point of f_ψ. Since linear combinations of products $\psi\varphi$ with $\psi \in \mathscr{D}(\mathbb{R})$ and $\varphi \in \mathscr{D}(B)$ are dense in $\mathscr{D}(\mathbb{R}\times B)$, and since

$$|\mathcal{F}_r(\zeta) - \mathcal{F}_r(\xi)| \leqslant 2\|\zeta - \xi\|_{\mathrm{L}^\infty(\mathbb{R}\times B)} \quad \forall \zeta,\xi \in \mathscr{D}(\mathbb{R}\times B),$$

we conclude that $\mathcal{F}_r$ vanishes in $\mathscr{D}'(\mathbb{R}\times B)$ as $r \to 0$. With (45) this gives

$$\lim_{r\to 0} \iint_{\mathbb{R}\times B} \zeta(v,x)\left(dH_y(v)\,d\nu^{y,r}(x) - d\mu^{y,r}(v,x)\right) = 0 \tag{48}$$

for all $\zeta \in \mathscr{D}(\mathbb{R}\times B)$. Consider now any subsequence $r_k \to 0$ with

$$\nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty \quad \text{in } \mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n).$$

Extracting another subsequence if necessary we may then assume that also

$$\mu^{y,r_k} \overset{*}{\rightharpoonup} \mu^\infty \quad \text{in } \mathrm{M}_{\mathrm{loc}}(\mathbb{R}\times\mathbb{R}^n).$$

Then (48) implies $\mu^\infty = H_y\nu^\infty$ in $\mathbb{R}\times B$. Obviously, the same argument works for any ball $B_R(0)$ instead of $B_1(0)$. Since the limit is uniquely determined, the sequence μ^{y,r_k} converges whenever ν^{y,r_k} does. If also $u^{y,r_k} \longrightarrow u^\infty$ in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n)$, then the pair $(u^\infty, H_y\nu^\infty)$ satisfies equation (35). Notice that H_y does not depend on the blow-up sequence, but only on $y \in \Omega \setminus E_{\mathcal{S}}$.

Step 2. We use the following implication, which holds for any $\nu \in \mathrm{M}^+_{\mathrm{loc}}(\Omega)$ (see Theorem 2.56 in [2]): For any Borel set $E \subset \Omega$ and any $t \in (0,\infty)$

$$\left(\limsup_{r\to 0} \frac{\nu(B_r(y))}{r^{n-1}} \geqslant t \quad \forall y \in E\right) \quad\Longrightarrow\quad \nu \geqslant t\mathscr{H}^{n-1}\lfloor E. \tag{49}$$

We define $E := \mathcal{J} \cap E_{\mathcal{S}}$ and write $E = \lim_{m\to\infty} E_m$ with increasing sets

$$E_m := \left\{ y \in E : \limsup_{r\to 0} \frac{\nu(B_r(y))}{r^{n-1}} \geqslant 2^{-m} \right\}.$$

Using (49), we obtain $\mathscr{H}^{n-1}(E_m) \leqslant 2^m \nu(E_m) = 0$ for all m because $\nu(E) = 0$, by Step 1. Therefore we also have $\mathscr{H}^{n-1}(E) = 0$. For any $y \in \mathcal{J} \setminus E$ there exists at least one subsequence $r_k \to 0$ such that

$$\begin{aligned} u^{y,r_k} &\longrightarrow u^\infty \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n),\\ \nu^{y,r_k} &\overset{*}{\rightharpoonup} \nu^\infty \quad \text{in } \mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n) \end{aligned}$$

and $\nu^\infty \neq 0$. Let $\varphi \in \mathscr{D}(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \varphi(x)\, d\nu^\infty(x) = 1$ and define

$$h_y(v) := -\mathbf{a}(v) \cdot \int_{\mathbb{R}^n} \nabla\varphi(x)\chi(v, u^\infty(x))\, dx \quad \forall v \in \mathbb{R}. \tag{50}$$

By Step 1, the pair $(u^\infty, H_y\nu^\infty)$ satisfies the kinetic equation (35). Thus

$$\begin{aligned} &\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}} \psi'(v)\, dH_y(v) \right) \varphi(x)\, d\nu^\infty(x) \\ &= \iint_{\mathbb{R}\times\mathbb{R}^n} \psi(v)\mathbf{a}(v) \cdot \nabla\varphi(x)\chi(v, u^\infty(x))\, dx\, dv \\ &= -\int_{\mathbb{R}} \psi(v)\, h_y(v)\, dv \quad \forall \psi \in \mathscr{D}(\mathbb{R}). \end{aligned}$$

By choice of φ, we get $h_y = \frac{\partial}{\partial v} H_y$ in $\mathscr{D}'(\mathbb{R})$. Notice that for any $x \in \mathbb{R}^n$, the map $v \mapsto \chi(v, u^\infty(x))$ is a BV-function, see definition (34). Its total variation is bounded uniformly in x. Moreover, we have by dominated convergence

$$\chi(v - \varepsilon, u^\infty) \longrightarrow \chi(v, u^\infty) \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \text{ as } \varepsilon \to 0+.$$

Therefore the function h_y defined by (50) is left-continuous and in $\mathrm{BV}(\mathbb{R})$, and the kinetic equation holds pointwise in $v \in \mathbb{R}$ as desired. □

6 Classification of Split States

In this section we first give a complete classification of the simplest possible split states: those for which ν is either vanishing or supported on a hyperplane. These results are then used to study general split states.

6.1 Special Split States: no entropy dissipation

If there is no entropy dissipation, then the solution is continuous.

Lemma 6.1. *Let (u,h,ν) be a split state with $h\nu = 0$ in an open set $\Omega \in \mathbb{R}^n$. Then u is continuous and constant along characteristic lines in Ω:*

$$\forall x \in \Omega \quad u(y) = u(x) \begin{cases} \text{for all } y \text{ in the connected component of} \\ \big(x + \mathbb{R}\mathbf{a}(u(x))\big) \cap \Omega \text{ that contains } x. \end{cases} \tag{51}$$

Proof. For any function g, let $\mathrm{Leb}(g)$ denote the set of Lebesgue points of g.

Step 1. We first prove that for every $\varepsilon > 0$ and $u_0 \in [0,1]$ there exists a $\delta > 0$ such that, for any $R > 0$ and $y \in \mathrm{Leb}(u)$ with $B_R(y) \subset \Omega$

$$u(y) \begin{Bmatrix} \geqslant \\ \leqslant \end{Bmatrix} u_0 \quad \Longrightarrow \quad \left(u \begin{Bmatrix} \geqslant \\ \leqslant \end{Bmatrix} u_0 - \varepsilon \quad \text{a.e. in } B_{\delta R}(y) \right). \tag{52}$$

Let $\varepsilon > 0$ and $u_0 \in [0,1]$ be given, and assume without loss of generality that $y = 0$ and $R = 1$. By nondegeneracy of a we can find n real values

$$u_0 > v_1 > v_2 > \ldots > v_n \geqslant u_0 - \varepsilon$$

such that $\mathbb{R}\mathbf{a}(v_1) + \ldots + \mathbb{R}\mathbf{a}(v_n) = \mathbb{R}^n$. Since $0 \in \mathrm{Leb}(u)$ and $u(0) \geqslant u_0 > v_1$, by definition of χ (see (34)) we obtain that $0 \in \mathrm{Leb}(\chi(v_1,u))$ and $\chi(v_1,u(0)) = 1$. The kinetic equation (46) with $v = v_1$ then implies

$$\begin{aligned} &\forall x \in \mathbb{R}\mathbf{a}(v_1) \cap B_1(0) \\ &\qquad x \in \mathrm{Leb}(\chi(v_1,u)) \text{ and } \chi(v_1,u(x)) = 1. \end{aligned}$$

By monotonicity of $v \mapsto \chi(v,u)$, then also

$$\begin{aligned} &\forall x \in \mathbb{R}\mathbf{a}(v_1) \cap B_1(0) \\ &\qquad x \in \mathrm{Leb}(\chi(v_2,u)) \text{ and } \chi(v_2,u(x)) = 1 \end{aligned}$$

since $v_2 \leqslant v_1$. We apply the kinetic equation (46) with $v = v_2$ and find that

$$\begin{aligned} &\forall x \in \big((\mathbb{R}\mathbf{a}(v_1) \cap B_1(0)) + \mathbb{R}\mathbf{a}(v_2)\big) \cap B_1(0) \\ &\qquad x \in \mathrm{Leb}(\chi(v_2,u)) \text{ and } \chi(v_2,u(x)) = 1. \end{aligned}$$

A simple geometric consideration shows that there exists a $\delta_2 > 0$ such that

$$\big((\mathbb{R}\mathbf{a}(v_1) \cap B_1(0)) + \mathbb{R}\mathbf{a}(v_2)\big) \cap B_1(0) \supset \big(\mathbb{R}\mathbf{a}(v_1) + \mathbb{R}\mathbf{a}(v_2)\big) \cap B_{\delta_2}(0).$$

Since by assumption $\mathbb{R}\mathbf{a}(v_1) + \ldots + \mathbb{R}\mathbf{a}(v_n) = \mathbb{R}^n$, by iterating this argument we obtain the existence of a $\delta = \delta_n > 0$ such that

$$\forall x \in B_\delta(0) \quad x \in \mathrm{Leb}(\chi(v_n, u)) \text{ and } \chi(v_n, u(x)) = 1,$$

which implies

$$u \geqslant v_n \geqslant u_0 - \varepsilon \quad \text{a.e. in } B_\delta(0).$$

Notice that the value of δ depends only on a, u_0 and ε. The opposite inequality in (52) can be proved in a similar fashion, so our claim follows.

Step 2. We now prove that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that, for any $R > 0$ and $y \in \mathrm{Leb}(u)$ with $B_R(y) \subset \Omega$

$$|u - u(y)| \leqslant \varepsilon \quad \text{a.e. in } B_{\delta R}(y). \tag{53}$$

This fact follows from Step 1 by a standard compactness argument. Indeed, let $\varepsilon > 0$ be fixed and select finitely many numbers $\{u_k\}_k$ such that

$$[0,1] \subset \bigcup_k [u_k, u_k + \varepsilon/2]. \tag{54}$$

For each k apply Step 1 with u_k and $\varepsilon/2$ instead of u_0 and ε, and let $\delta_k > 0$ be the corresponding radius. Define the minimum $\delta := \min_k \delta_k > 0$. Now fix some point $y \in \mathrm{Leb}(u) \cap \Omega$. By (54), there exists k such that $u(y) \in [u_k, u_k + \varepsilon/2]$, and in particular we have $u(y) \geqslant u_k$. By our choice of δ_k we get

$$u \geqslant u_k - \varepsilon/2 \quad \text{a.e. in } B_{\delta_k R}(y) \supset B_{\delta R}(y).$$

On the other hand, we have $u_k \geqslant u(y) - \varepsilon/2$, so we finally get

$$u \geqslant u(y) - \varepsilon \quad \text{a.e. in } B_{\delta R}(y).$$

The reverse inequality is proved in a similar way. We conclude that there is a locally uniform modulus of continuity in every Lebesgue point of u. Since the Lebesgue points are dense, u admits a continuous representative in Ω.

Step 3. To prove (51), fix $x \in \Omega$ and let y be in the connected component of $\big(x + \mathbb{R}\mathbf{a}(u(x))\big) \cap \Omega$ that contains x. Since Ω is open, there exists a neighborhood $U \ni y$ with $U \subset \Omega$. Consider now the set

$$C^- := \Omega \cap \bigcup_{v < u(x)} x + \mathbb{R}\mathbf{a}(v).$$

For any sequence $y_k \to y$ with $y_k \in C^- \cap U$, let v_k be defined by $y_k \in x + \mathbb{R}\mathbf{a}(v_k)$ for all k. We may assume that the lines connecting x and y_k are all contained in Ω and that $v_k \to u(x)$. From (46) we obtain $\chi(v_k, u(y_k)) = \chi(v_k, u(x)) = 1$, which is equivalent to $u(y_k) > v_k$. The continuity of u in y then yields

$$u(y) = \lim_{k \to \infty} u(y_k) \geqslant \lim_{k \to \infty} v_k = u(x).$$

Similarly, we consider the set

$$C^+ := \Omega \cap \bigcup_{u(x) \leqslant v} x + \mathbb{R}\mathbf{a}(v).$$

For any $y_k \to y$ with $y_k \in C^+ \cap U$, let v_k be defined by $y_k \in x + \mathbb{R}\mathbf{a}(v_k)$ for all k. We may again assume that the segments connecting x and y_k are all contained in Ω and that $v_k \to u(x)$. From (46) we obtain $\chi(v_k, u(y_k)) = \chi(v_k, u(x)) = 0$, which is equivalent to $u(y_k) \leqslant v_k$. The continuity of u in y then yields

$$u(y) = \lim_{k\to\infty} u(y_k) \leqslant \lim_{k\to\infty} v_k = u(x).$$

This proves that indeed u is constant along the characteristic lines in Ω. □

Proposition 6.2 (Liouville Theorem). *Let (u, h, ν) be a split state such that $h\nu = 0$ in all of $\mathbb{R}^n$. Then the function u is constant.*

Proof. From Lemma 6.1 with $\Omega = \mathbb{R}^n$ we already know that u is continuous. Moreover, from Step 2 of the previous proof we find that for any $\varepsilon > 0$ and any point $y \in \mathbb{R}^n$, there exists $\delta > 0$ such that for all $R > 0$

$$|u - u(y)| \leqslant \varepsilon \quad \text{in } B_{\delta R}(y).$$

Sending $R \to \infty$ we obtain the result since ε was arbitrary. □

6.2 Special Split States: ν supported on a hyperplane

These split states are typically obtained from blow-ups at shock points.

Lemma 6.3. *Let (u, h, ν) be a split state with $h \neq 0$ and $\nu = \mathscr{H}^{n-1}\lfloor\mathcal{J}$, where the support $\mathcal{J} \subset \{\eta \cdot x = 0\}$ is relatively open for some unit vector η. Then the conclusion of Lemma 6.1 holds with $\Omega = \mathbb{R}^n \setminus \bar{\mathcal{J}}$. Moreover, there exist strong traces u^+ and u^- that are constant along $\mathcal{J}$, and*

$$h(v) = \mathbf{a}(v) \cdot \eta \left(\chi(v, u^+) - \chi(v, u^-)\right) \quad \forall v \in \mathbb{R} \quad \text{a.e. in } \mathcal{J}. \tag{55}$$

The traces $u^\pm$ and η (up to orientation) are completely determined by h.

Proof. The main step consists in proving the existence of strong traces.

Step 1. Without loss of generality, we may assume that $\eta = (1, 0, \ldots, 0)^T$. We denote by $\hat{}$ the projection onto the last $(n-1)$ components. Because of (46), the function $\chi(v, u)$ is freely transported in the set $\{\eta \cdot x > 0\} = \{x_1 > 0\}$. That is, for any $v \in \mathbb{R}$ with $\mathbf{a}_1(v) \neq 0$ and for a.e. $(x_1, \hat{x}) \in \mathbb{R}_+ \times \mathbb{R}^{n-1}$

$$\chi\big(v, u(x_1, \hat{x})\big) = \chi\left(v, u\left(1, \hat{x} + (1 - x_1)\frac{\hat{\mathbf{a}}(v)}{\mathbf{a}_1(v)}\right)\right). \tag{56}$$

We define

$$\chi^+(v,\hat{x}) := \chi\left(v, u\left(1, \hat{x} + \frac{\hat{\mathbf{a}}(v)}{\mathbf{a}_1(v)}\right)\right) \quad \forall \hat{x} \in \mathbb{R}^{n-1}. \tag{57}$$

By (56) and the Lebesgue lemma, we then obtain that for all $R > 0$

$$\begin{aligned}
&\operatorname*{ess\,lim}_{x_1\to 0+} \int_{\hat{B}_R(0)} \left|\chi\big(v, u(x_1,\hat{x})\big) - \chi^+(v,\hat{x})\right| \,d\hat{x} \\
&\quad = \operatorname*{ess\,lim}_{x_1\to 0+} \int_{\hat{B}_R(0)} \left|\chi\left(v, u\left(1, \hat{x} + (1-x_1)\frac{\hat{\mathbf{a}}(v)}{\mathbf{a}_1(v)}\right)\right)\right. \\
&\qquad\qquad \left. - \chi\left(v, u\left(1, \hat{x} + \frac{\hat{\mathbf{a}}(v)}{\mathbf{a}_1(v)}\right)\right)\right| \,d\hat{x} = 0.
\end{aligned} \tag{58}$$

This shows that (57) is the upper trace of $\chi(v,u)$ in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n-1})$ on the hyperplane $\{x_1 = 0\}$. Since $\chi(v,u)$ takes values in $\{0,1\}$ only, we also have that $\chi^+(v,\cdot) \in \{0,1\}$ a.e. Notice also that the set of $v \in \mathbb{R}$ such that $\mathbf{a}_1(v) = 0$ is a null set since $\mathbf{a}$ is nondegenerate. Thus (58) holds for a.e. $v \in \mathbb{R}$. Let

$$u^+(\hat{x}) := \int_{\mathbb{R}} \chi^+(v,\hat{x}) \,dv \quad \text{for a.e. } \hat{x} \in \mathbb{R}^{n-1}. \tag{59}$$

By (58) and dominated convergence, we then obtain that for all $R > 0$

$$\operatorname*{ess\,lim}_{x_1\to 0+} \int_{\hat{B}_R(0)} \left|u(x_1,\hat{x}) - u^+(\hat{x})\right| \,d\hat{x} = 0. \tag{60}$$

That is, the function (59) is the upper trace of u in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n-1})$ on the hyperplane $\{x_1 = 0\}$. Consider now the set

$$G := \left\{ v \in \mathbb{R} \colon \left|\{\hat{x} \in \mathbb{R}^{n-1} \colon u^+(\hat{x}) = v\}\right| > 0 \right\}$$

and notice that G is at most countable. Then (60) implies that

$$\operatorname*{ess\,lim}_{x_1\to 0+} u(x_1,\hat{x}) = u^+(\hat{x}) \quad \text{for a.e. } \hat{x} \in \mathbb{R}^{n-1},$$

and thus for all $v \in \mathbb{R} \setminus G$

$$\operatorname*{ess\,lim}_{x_1\to 0+} \chi(v, u(x_1,\hat{x})) = \chi(v, u^+(\hat{x})) \quad \text{for a.e. } \hat{x} \in \mathbb{R}^{n-1}.$$

By dominated convergence, we obtain for any $R > 0$ that

$$\operatorname*{ess\,lim}_{x_1\to 0+} \int_{\mathbb{R}} \int_{\hat{B}_R(0)} \left|\chi(v, u(x_1,\hat{x})) - \chi(v, u^+(\hat{x}))\right| \,d\hat{x}\,dv = 0.$$

This shows that the upper trace of $\chi(v,u)$ is in fact given by $\chi(v,u^+)$ a.e. The same reasoning can be applied to prove the existence of a lower trace u^- in $\mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n-1})$ on the hyperplane, with analogous properties.

Step 2. Consider now an even test function $\varphi_1 \in \mathscr{D}(\mathbb{R})$ with $0 \leqslant \varphi_1 \leqslant 1$ and $\varphi_1(0) = 1$, and let $\varepsilon > 0$. Testing the kinetic equation (46) against

$$\varphi_1(x_1/\varepsilon)\, \hat{\varphi}(\hat{x})\psi(v) \quad \text{with } \hat{\varphi} \in \mathscr{D}(\mathbb{R}^{n-1}) \text{ and } \psi \in \mathscr{D}(\mathbb{R}),$$

we obtain in the limit $\varepsilon \to 0$ that

$$\int_{\mathbb{R}} \psi(v)\mathbf{a}(v) \cdot \eta \int_{\mathbb{R}^{n-1}} \big(\chi(v, u^+(\hat{x})) - \chi(v, u^-(\hat{x}))\big)\hat{\varphi}(\hat{x})\, d\hat{x}\, dv \\ = \int_{\mathbb{R}} \psi(v)h(v)\, dv \int_{\hat{\Omega}} \hat{\varphi}(\hat{x})\, d\hat{x},$$

where $\hat{\Omega} := \{\hat{x} \in \mathbb{R}^{n-1} \colon (0, \hat{x}) \in \Omega\}$. Since $\hat{\varphi}$ and ψ were arbitrary,

$$\chi(v, u^+) - \chi(v, u^-) \begin{cases} \text{is constant in } \hat{\Omega} \text{ and} \\ \text{vanishes outside } \hat{\Omega} \end{cases}$$

for a.e. $v \in \mathbb{R}$. This proves (55), which implies $\{u^+, u^-\} = \{\inf \operatorname{spt} h, \sup \operatorname{spt} h\}$. Recall that $h \neq 0$, by assumption. Finally, there exist $v_1, \ldots, v_n \in \operatorname{spt} h$ such that $\mathbb{R}\mathbf{a}(v_1) + \ldots + \mathbb{R}\mathbf{a}(v_n) = \mathbb{R}^n$ because $\mathbf{a}$ is nondegenerate. Therefore η is determined up to orientation by the n conditions

$$h(v_k) = \mathbf{a}(v_k) \cdot \eta\, \big(\chi(v_k, u^+) - \chi(v_k, u^-)\big),$$

each of which determines a hyperplane in $\mathbb{R}^n$. This proves the lemma. □

Proposition 6.4. *Let (u, h, ν) be a split state with $h \neq 0$ and $\nu = \mathscr{H}^{n-1} \lfloor \mathcal{J}$, where the support $\mathcal{J} = \{\eta \cdot x = 0\}$ for some unit vector η. Then the conclusion of Lemma 6.3 holds. Moreover, u is constant on either side of $\mathcal{J}$.*

Proof. From Lemma 6.3 we already know that there exist strong traces u^+ and u^- which are constant along the hyperplane. Now fix some point $y \in \mathcal{J}$ and consider a sequence $y_k \to y$ with $y_k \in \{\eta \cdot x > 0\}$ and $u(y_k) \longrightarrow u^+(y)$. By Lemma 6.1, the function u is continuous and constant along characteristics in $\mathbb{R}^n \setminus \mathcal{J}$. Therefore for any $z \in y + \mathbb{R}\mathbf{a}(u^+(y))$ with $\eta \cdot z > 0$ there exists a sequence $z_k \to z$ such that $z_k \in y_k + \mathbb{R}\mathbf{a}(u(y_k))$ for all k. Then

$$u(z) = \lim_{k\to\infty} u(z_k) = \lim_{k\to\infty} u(y_k) = u^+(y).$$

This shows that $u = u^+$ in the upper halfspace $\{\eta \cdot x > 0\}$ since $y \in \mathcal{J}$ was arbitrary. The same argument applies to the lower halfspace. □

6.3 Special Split States: $\boldsymbol{\nu}$ supported on half a hyperplane

The following result will be used for second blow-ups.

Lemma 6.5. *Let (u,h,ν) be a split state with $h \neq 0$ and $\nu = \mathscr{H}^{n-1}\lfloor\mathcal{J}$,*

$$\mathcal{J} = \{\eta\cdot x = 0,\, \omega\cdot x > 0\}$$

for some pair of orthonormal vectors $\eta \perp \omega$. Then the conclusion of Lemma 6.3 holds. Moreover, the vector ω is fixed in the sense that there exists a cone C, which is not a halfspace and depends only on h, such that $\omega \in C$.

Proof. We explicitly construct the cone C.

Step 1. By Lemma 6.3, the function h has the form

$$h(v) = \mathbf{1}_{(u^-,u^+]}(v)\,\mathbf{a}(v)\cdot\eta \quad \forall v\in\mathbb{R},$$

where the constant traces u^+ and u^- and the unit normal vector η (up to orientation) are completely determined by h. Let $I \subset (u^-,u^+)$ be an interval such that $\mathbf{a}(v)\cdot\eta \neq 0$ for all $v \in I$. Then, for every $\underline{v} < \bar{v}$ in I we have

$$\frac{\mathbf{a}(\underline{v})\cdot\omega}{\mathbf{a}(\underline{v})\cdot\eta} \leqslant \frac{\mathbf{a}(\bar{v})\cdot\omega}{\mathbf{a}(\bar{v})\cdot\eta}. \tag{61}$$

We argue by contradiction. From (46) we deduce that for any $v \in I$

$$\frac{\mathbf{a}(v)}{\mathbf{a}(v)\cdot\eta}\cdot\nabla\chi(v,u) = \mathscr{H}^{n-1}\lfloor\mathcal{J}$$

and therefore

$$\chi(v,u) = \begin{cases} 1 & \text{in } \mathcal{J} + \mathbb{R}_+\dfrac{\mathbf{a}(v)}{\mathbf{a}(v)\cdot\eta}, \\ 0 & \text{in } \mathcal{J} - \mathbb{R}_+\dfrac{\mathbf{a}(v)}{\mathbf{a}(v)\cdot\eta}. \end{cases}$$

Recall that u is continuous and constant along characteristics outside $\mathcal{J}$, by Lemma 6.1. Assume now that we can find $\underline{v} < \bar{v}$ such that (61) does not hold. By the mean value theorem there also exists $v \in (\underline{v},\bar{v})$ such that

$$\frac{\mathbf{a}(\underline{v})\cdot\omega}{\mathbf{a}(\underline{v})\cdot\eta} > \frac{\mathbf{a}(v)\cdot\omega}{\mathbf{a}(v)\cdot\eta} > \frac{\mathbf{a}(\bar{v})\cdot\omega}{\mathbf{a}(\bar{v})\cdot\eta}. \tag{62}$$

Fix any point x such that the line $\mathcal{L} := x + \mathbb{R}\mathbf{a}(v)/(\mathbf{a}(v)\cdot\eta)$ does not intersect $\mathcal{J}$. Then $\chi(v,u)$ is constant along $\mathcal{L}$. Thanks to (62) there exist points

$$\underline{x} \in \mathcal{L}\cap\left(\mathcal{J} + \mathbb{R}_+\frac{\mathbf{a}(\underline{v})}{\mathbf{a}(\underline{v})\cdot\eta}\right) \quad\text{and}\quad \bar{x}\in\mathcal{L}\cap\left(\mathcal{J} - \mathbb{R}_+\frac{\mathbf{a}(\bar{v})}{\mathbf{a}(\bar{v})\cdot\eta}\right).$$

But using the monotonicity of $v \mapsto \chi(v,u)$ we obtain the contradiction

$$1 = \chi(\underline{v},u(\underline{x})) = \chi(v,u(\underline{x})) = \chi(v,u(\bar{x}) = \chi(\bar{v},u(\bar{x})) = 0.$$

We conclude that the map $v \mapsto \big(\mathbf{a}(v)\cdot\omega\big)/(\mathbf{a}(v)\cdot\eta)$ is increasing and

$$0 \leqslant \frac{d}{dv}\left(\frac{\mathbf{a}(v)\cdot\omega}{\mathbf{a}(v)\cdot\eta}\right) = \frac{\Big(\big(\mathbf{a}(v)\cdot\eta\big)\mathbf{a}'(v) - \big(\mathbf{a}'(v)\cdot\eta\big)\mathbf{a}(v)\Big)\cdot\omega}{\big(\mathbf{a}(v)\cdot\eta\big)^2}$$

for all $v \in I$. Notice that the set $\big\{v \in (u^-, u^+) \colon \mathbf{a}(v)\cdot\eta \neq 0\big\}$ is open and dense in the interval (u^-, u^+) since a is continuous and nondegenerate. Therefore

$$\Big(\big(\mathbf{a}(v)\cdot\eta\big)\mathbf{a}'(v) - \big(\mathbf{a}'(v)\cdot\eta\big)\mathbf{a}(v)\Big)\cdot\omega \geqslant 0 \quad \forall v \in (u^-, u^+).$$

Step 2. We now define the cone

$$C^* := \mathbb{R}_+\Big\{\big(\mathbf{a}(v)\cdot\eta\big)\mathbf{a}'(v) - \big(\mathbf{a}'(v)\cdot\eta\big)\mathbf{a}(v) \colon v \in (u^-, u^+)\Big\},$$

and clearly $C^* \subset \{\eta\cdot x = 0\}$. On the other hand, C^* cannot be contained in a proper subspace of $\{\eta\cdot x = 0\}$. That is, there cannot exist a $\xi \perp \eta$ with

$$\Big(\big(\mathbf{a}(v)\cdot\eta\big)\mathbf{a}'(v) - \big(\mathbf{a}'(v)\cdot\eta\big)\mathbf{a}(v)\Big)\cdot\xi = 0 \quad \forall v \in (u^-, u^+), \tag{63}$$

since this would allow us to rewrite (63) in the form

$$\frac{d}{dv}\left(\frac{\mathbf{a}(v)\cdot\xi}{\mathbf{a}(v)\cdot\eta}\right) = 0 \quad \forall v \in I, \tag{64}$$

where $I \subset (u^-, u^+)$ is an open interval with $\mathbf{a}(v)\cdot\eta \neq 0$ for all $v \in I$. From (64) we could find a constant c such that $\mathbf{a}(v)\cdot(\omega - c\eta) = 0$ for all $v \in I$, in contradiction to the nondegeneracy of $\mathbf{a}$. This proves that the cone C^* must be genuinely $(n-1)$-dimensional. Let C be the dual cone to C^* relative to the hyperplane $\{\eta\cdot x = 0\}$. Then C is not a halfspace and $\omega \in C$. □

6.4 Classification of General Split States

Finally, we completely classify general split states.

Proposition 6.6. *Let (u, h, ν) be a split state with $h\nu \neq 0$. Then there exist constants $L, g > 0$ and an orthonormal coordinate system $x_1, \ldots, x_n$ (all depending on h only) with the following property: There exist*

- *a constant $e \in \mathbb{R}$ and*
- *a Lipschitz continuous function $w \colon \mathbb{R}^{n-2} \longrightarrow \mathbb{R}$ with $\mathrm{Lip}(w) \leq L$*

such that $\nu = g\mathscr{H}^{n-1}\lfloor\mathcal{J}$ for some set $\mathcal{J}$ of the form

$$\mathcal{J} = \big\{x_1 = e\big\} \quad \text{or} \quad \mathcal{J} = \big\{x_1 = e, x_n \geqslant w(x_2, \ldots, x_{n-1})\big\}.$$

Proof. Since the proof is rather technical, we only give a sketch of it and refer the reader to [11] for further details. We proceed in four steps.

Step 1. We define the set $\mathcal{J} \subset \mathbb{R}^n$ as above as

$$\mathcal{J} := \left\{ x \in \mathbb{R}^n \colon \limsup_{r \to 0} \frac{\nu(B_r(x))}{r^{n-1}} > 0 \right\}.$$

We can then show that $\mathcal{J}$ is contained in two Lipschitz graphs. In particular, $\mathcal{J}$ is a codimension-one rectifiable set and ν takes the form $\nu = g\mathscr{H}^{n-1}\lfloor\mathcal{J}$, where g is some Borel function that is strictly positive in the set $\mathcal{J}$.

Step 2. We then perform a *second blow-up* around a generic point $x \in \mathcal{J}$. By rectifiability of $\mathcal{J}$, we obtain a special split state of the form described in Lemma 6.3, which implies that both the function value $g(x)$ and the unit normal vector $\eta(x)$ to $\mathcal{J}$ (up to orientation) are independent of x, thus constant along $\mathcal{J}$. They are completely determined by the function h. This allows to conclude that $\mathcal{J}$ is in fact contained in a single Lipschitz graph.

Step 3. Since the unit normal vector η is constant along $\mathcal{J}$, we can show that up to $\mathscr{H}^{n-1}$-negligible sets $\mathcal{J}$ is contained in at most countably many distinct hyperplanes Π_k. Each component $\mathcal{J} \cap \Pi_k$ is given as the intersection of $2n$ Lipschitz supergraphs that are completely determined by h. In particular, each $\mathcal{J} \cap \Pi_k$ is a set of locally finite perimeter.

Step 4. We perform a second blow-up at a generic point $x \in \partial(\mathcal{J} \cap \Pi_k)$. By rectifiability of the boundary, we obtain a split state of the form described in Lemma 6.5, which implies that each $\mathcal{J} \cap \Pi_k$ is contained in a single Lipschitz supergraph with respect to a universal coordinate system $x_1, \ldots, x_n$. This forces $\mathcal{J}$ to be contained in only one hyperplane. The result follows. □

7 Proof of the Main Theorem

In the previous section, we gave a complete classification of split states. In order to conclude the proof of Theorem 4.5, however, we need an extra piece of information that is provided by Proposition 7.1 below. Recall that

$$\mathcal{J} = \left\{ x \in \Omega \colon \limsup_{r \to 0} \frac{\nu\big(B_r(x)\big)}{r^{n-1}} > 0 \right\}, \tag{65}$$

where ν is the x-marginal of the entropy dissipation measure μ. From Proposition 5.9 we know that for a generic point $y \in \mathcal{J}$ there exists $h_y \in \mathrm{BV}(\mathbb{R})$ left-continuous, such that the limits $(u^\infty, \nu^\infty) \in \mathrm{L}^\infty(\mathbb{R}^d) \times \mathrm{M}^+_{\mathrm{loc}}(\mathbb{R}^n)$ of any converging blow-up sequence form a split state: for every $v \in \mathbb{R}$ we have

$$\mathbf{a}(v) \cdot \nabla\chi(v, u^\infty(x)) = h_y(v)\nu^\infty \quad \text{in } \mathscr{D}'(\mathbb{R}^n).$$

According to Proposition 6.6 we then have $\nu^\infty = g_y\mathscr{H}^{n-1}\lfloor\mathcal{J}^\infty$ for a constant $g_y > 0$, and in an appropriate coordinate system $\mathcal{J}^\infty$ has the form

$$\mathcal{J}^\infty = \big\{x_1 = e^\infty\big\} \quad \text{or} \quad \mathcal{J}^\infty = \big\{x_1 = e^\infty, x_n \geqslant w^\infty(x_2, \ldots, x_{n-1})\big\} \tag{66}$$

for a Lipschitz continuous function w^∞ with $\mathrm{Lip}(w^\infty) \leqslant L_y$. All quantities with subscript y only depend on the point $y \in \mathcal{J}$, while the superscript ∞ indicates a dependence on the particular converging blow-up sequence.

Proposition 7.1. *Let $(u^\infty, h_y, \nu^\infty)$ be a split state obtained from a converging blow-up sequence at a point $y \in \mathcal{J}$. In the notation above, we have $0 \in \mathcal{J}^\infty$.*

Proof. The result will follow from the behavior of a discriminating functional along blow-up sequences. Let $x_1, \ldots, x_n$ and $L_y, g_y > 0$ be the coordinate system and constants discussed above. They all depend only on the blow-up point $y \in \mathcal{J}$. Assuming from now on that y is fixed, we do not write the subscript y anymore to simplify notation. We define the wedge

$$W := \big\{x_n \geqslant L|(x_2, \ldots, x_{n-1})|\big\}$$

and the $(n-1)$-dimensional cone $C := W \cap \{x_1 = 0\}$. Since the limit measure ν^∞ of any converging blow-up sequence is supported in a set J^∞ as in (66) with $\mathrm{Lip}(w^\infty) \leqslant L$, we obtain the following implication:

$$x \in \mathcal{J}^\infty \quad \Longrightarrow \quad x + C \subset \mathcal{J}^\infty. \tag{67}$$

Now fix a function $\varphi \in \mathscr{D}([0,\infty))$ with $\varphi(r) > 0$ and $\varphi'(r) < 0$ for $r \in [0,1)$, and $\varphi(r) = 0$ for $r \in [1,\infty)$. Then we define $b := g \int_C \varphi(|x|)\, \mathscr{H}^{n-1}(x)$ and

$$\mathcal{F}(\nu) := \frac{1}{b} \int_W \varphi(|x|)\, d\nu(x) \quad \forall \nu \in \mathrm{M}^+(\mathbb{R}^n).$$

We divide the proof of the proposition into three steps.

Step 1. The functional $\mathcal{F}$ has the following properties:

For any limit $\nu^\infty = g\mathscr{H}^{n-1}\lfloor\mathcal{J}^\infty$ from a converging blow-up sequence

1) $\mathcal{F}(\nu^\infty) \in [0,1]$.
2) $\mathcal{F}(\nu^\infty) = 1$ if and only if $0 \in \mathcal{J}^\infty$.
3) There exists $R > 0$ such that $\mathcal{F}(\nu^\infty) = 0$ implies $\nu^\infty(B_R(0)) = 0$.
4) For the rescaled measure $(\nu^\infty)^{0,s}$ defined as in (42) we have
$$\frac{d}{ds}\bigg|_{s=1} \mathcal{F}\big((\nu^\infty)^{0,s}\big) \geqslant 0,$$
with equality if and only if $\mathcal{F}(\nu^\infty) \in \{0,1\}$.

Indeed, (1) is obvious from the definitions and the shape of ν^∞. If now $0 \in \mathcal{J}^\infty$, then $C \subset \mathcal{J}^\infty$ because of (67), and thus $\mathcal{F}(\nu^\infty) = 1$. On the other hand, if $\mathcal{F}(\nu^\infty) = 1$, then $J^\infty \subset \{x \cdot \eta = 0\}$ since φ is strictly decreasing, and even

$$(W \cap \mathcal{J}^\infty) \cap B_1(0) = C \cap B_1(0),$$

up to $\mathscr{H}^{n-1}$-negligible sets. Since $\mathcal{J}^\infty$ is closed, we obtain that $0 \in \mathcal{J}^\infty$. This proves (2). Assume now that $\mathcal{F}(\nu^\infty) = 0$. Then $(W \cap \mathcal{J}^\infty) \cap B_1(0) = \varnothing$, up to

$\mathscr{H}^{n-1}$-negligible sets, and then $\nu^\infty(B_R(0)) = 0$ for some constant $R > 0$ that does not depend on the particular chosen measure ν^∞. This is statement (3). Finally, by definition (42) of rescaled measures we have

$$\begin{aligned} \mathcal{F}\big((\nu^\infty)^{0,s}\big) &= \frac{1}{b}\int_W \varphi(|x|)\,d\big((\nu^\infty)^{0,s}\big)(x) \\ &= \frac{1}{bs^{n-1}}\int_W \varphi\bigg(\frac{|x|}{s}\bigg)\,d\nu^\infty(x) \\ &= \frac{g}{bs^{n-1}}\int_{W\cap\mathcal{J}^\infty} \varphi\bigg(\frac{|x|}{s}\bigg)\,d\mathscr{H}^{n-1}(x). \end{aligned}$$

We differentiate with respect to s and pass to polar coordinates. This gives

$$\begin{aligned} \frac{d}{ds}\bigg|_{s=1}\mathcal{F}\big((\nu^\infty)^{0,s}\big) &= -\frac{g}{b}\int_{W\cap\mathcal{J}^\infty}\Big((n-1)\varphi(|x|)+\varphi'(|x|)|x|\Big)\,d\mathscr{H}^{n-1}(x) \\ &= -\frac{g}{b}\int_0^1 \frac{d}{dr}\big(r^{n-1}\varphi(r)\big)\omega(r)\,dr, \end{aligned} \tag{68}$$

where

$$\omega(r) = \frac{\mathscr{H}^{n-2}\big((W\cap\mathcal{J}^\infty)\cap\partial B_r(0)\big)}{r^{n-2}}.$$

By some technical argument, for which we refer the reader to [11], one can show that the map $r\mapsto\omega(r)$ is monotone increasing. Integrating by parts in (68), we then obtain the first part of (4). Notice that (68) vanishes if and only if $\omega(r)$ is constant for a.e. $r\in[0,1]$, which means that

$$\begin{aligned} &\text{either} \quad \mathscr{H}^{n-2}\big((W\cap\mathcal{J}^\infty)\cap\partial B_r(0)\big) > 0 \text{ for a.e. } r\in[0,1] \\ &\text{or} \quad \mathscr{H}^{n-2}\big((W\cap\mathcal{J}^\infty)\cap\partial B_r(0)\big) = 0 \text{ for a.e. } r\in[0,1]. \end{aligned}$$

Since $\mathcal{J}^\infty$ is closed, this in turn is equivalent to

$$\text{either} \quad 0\in\mathcal{J}^\infty \quad \text{or} \quad (W\cap\mathcal{J}^\infty)\cap B_1(0) = \varnothing,$$

and then (4) follows by definition of $\mathcal{F}$ and (2).

Step 2. We now consider the behavior of the functional $\mathcal{F}$ under rescaling. We define $f(r) := \mathcal{F}(\nu^{y,r})$ for $r>0$, where the measure $\nu^{y,r}$ is given in (42). If $r_k\to 0$ is a sequence such that $\nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty$ in $\mathrm{M}^+(\mathbb{R}^n)$, then

$$\lim_{k\to\infty} f(r_k) = \mathcal{F}(\nu^\infty), \tag{69}$$

$$\lim_{k\to\infty} r_k f'(r_k) = \frac{d}{ds}\bigg|_{s=1}\mathcal{F}\big((\nu^\infty)^{0,s}\big). \tag{70}$$

Indeed, notice that for the interior $\mathring{W}$ and closure $\bar{W}$ of the wedge W

$$\int_{\mathring{W}} \varphi(|x|)\, d\nu^\infty(x) \leqslant \liminf_{k\to\infty} \int_W \varphi(|x|)\, d\nu^{y,r_k}(x) = \liminf_{k\to\infty} f(r_k),$$
$$\int_{\bar{W}} \varphi(|x|)\, d\nu^\infty(x) \geqslant \limsup_{k\to\infty} \int_W \varphi(|x|)\, d\nu^{y,r_k}(x) = \limsup_{k\to\infty} f(r_k)$$

(see Example 1.63 in [2]). This implies (69) because the limit measure ν^∞ does not concentrate mass on the boundary ∂W of W since

$$\int_{\partial W} \varphi(|x|)\, d\nu^\infty(x) \leqslant g \int_{\partial W \cap \{x_1 = e^\infty\}} \varphi(|x|)\, d\mathscr{H}^{n-1}(x) = 0,$$

see (66). Notice also that $f(sr) = \mathcal{F}((\nu^{y,r})^{0,s})$ for all $r, s > 0$. We compute

$$\begin{aligned} rf'(r) &= \frac{d}{ds}\bigg|_{s=1} \mathcal{F}\big((\nu^{y,r})^{0,s}\big) \\ &= \frac{1}{b}\frac{d}{ds}\bigg|_{s=1} \int_W \varphi(|x|)\, d\big((\nu^{y,r})^{0,s}\big)(x) \\ &= \frac{1}{b}\frac{d}{ds}\bigg|_{s=1} \frac{1}{s^{n-1}} \int_W \varphi\Big(\frac{|x|}{s}\Big)\, d\nu^{y,r}(x) \\ &= -\frac{1}{b}\int_W \Big((n-1)\varphi(|x|) + \varphi'(|x|)|x|\Big)\, d\nu^{y,r}(x). \end{aligned} \tag{71}$$

Repeating the arguments above for (71), we obtain (70).

Step 3. We now prove that for every $\delta > 0$ there exist $\varepsilon > 0$ and $R > 0$ such that for every $r < R$ the following implication holds:

$$f(r) \in [\delta, 1-\delta] \quad \Longrightarrow \quad rf'(r) \geqslant \varepsilon. \tag{72}$$

We argue by contradiction: Assume that there exist $\delta > 0$ and $r_k \to 0$ with $f(r_k) \in [\delta, 1-\delta]$ and $r_k f'(r_k) < 1/k$ for all k. Up to a subsequence, we may suppose that $\nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty$ in $\mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n)$. From Step 2, we obtain

$$\mathcal{F}(\nu^\infty) \in [\delta, 1-\delta] \quad \text{and} \quad \frac{d}{ds}\bigg|_{s=1} \mathcal{F}\big((\nu^\infty)^{0,s}\big) = 0,$$

but this obviously contradicts Statement (4) of Step 1. This proves (72). Now fix any $\delta > 0$ and find $\varepsilon > 0$ and $R > 0$ satisfying (72). Assume that for some $r_0 < R$ we have $f(r_0) \in [\delta, 1-\delta]$. Because of (72), we then have

$$f(r) \leq f(r_0) - \varepsilon \log(r_0/r)$$

for all r with $f([r, r_0]) \subset [\delta, 1-\delta]$. Therefore we can find a number $0 < r_1 < r_0$ such that $f(r_1) < \delta$. Applying (72) again we then conclude that $f(r) < \delta$ for all $r < r_1$. In summary, we have shown the following implication:

$$\liminf_{r\to 0} f(r) < 1-\delta \quad \Longrightarrow \quad \limsup_{r\to 0} f(r) \leqslant \delta.$$

Since $\delta > 0$ was arbitrary, this finally proves that

$$\text{either} \quad \lim_{r\to 0} f(r) = 0 \quad \text{or} \quad \lim_{r\to 0} f(r) = 1.$$

In particular, we have the same limit for all blow-up sequences $r_k \to 0$. Using the Statements (2) and (3) of Step 1, we obtain the following alternative:

$$\begin{array}{rrl} \text{either} & 0 \in \mathcal{J}^\infty & \forall \nu^\infty \\ \text{or} & \nu^\infty(B_R(0)) = 0 & \forall \nu^\infty. \end{array} \tag{73}$$

We want to rule out the second possibility. To achieve this, notice that if ν^∞ is the limit measure of some converging blow-up sequence $r_k \to 0$, then the rescaled measure $(\nu^\infty)^{0,s}$ is the limit of the blow-up sequence $sr_k \to 0$, for any $s > 0$. If now the second possibility in (73) holds, then

$$\big((\nu^\infty)^{0,s}\big)(B_R(0)) = \frac{\nu^\infty(B_{sR}(0))}{s^{n-1}} = 0 \quad \forall s > 0.$$

This implies $\nu^\infty = 0$, in contradiction to our choice of $y \in \mathcal{J}$. □

Proof (of Theorem 4.5). We divide the proof into three steps.

Step 1. We apply Theorem 5.6 to prove that the set $\mathcal{J}$ defined in (65) is codimension-one rectifiable: From Propositions 6.6 and 7.1 we know that for $\mathscr{H}^{n-1}$-a.e. $y \in \mathcal{J}$ there exist constants $L_y, g_y > 0$ and an orthonormal coordinate system $x_1, \ldots, x_n$ such that, for any sequence $r_k \to 0$ with $\nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty$ in $\mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n)$, we have $\nu^\infty = g_y \mathscr{H}^{n-1} \lfloor \mathcal{J}^\infty$ with $0 \in \mathcal{J}^\infty \subset \{x_1 = 0\}$. Now we argue as follows: We define the $(n-1)$-dimensional cone

$$C_y := \big\{x_1 = 0, x_n \geqslant L_y |(x_2, \ldots, x_{n-1})|\big\} \subset \mathcal{J}^\infty.$$

Then we can estimate

$$\begin{aligned} &\liminf_{k\to\infty} \frac{\nu(B_{r_k}(y))}{r_k^{n-1}} \\ &\qquad = \liminf_{k\to\infty} \nu^{y,r_k}(B_1(0)) \geqslant \nu^\infty(B_1(0)) \geqslant g_y \mathscr{H}^{n-1}\big(C_y \cap B_1(0)\big) > 0, \end{aligned}$$

by weak* convergence of $\nu^{y,r_k} \overset{*}{\rightharpoonup} \nu^\infty$ and openness of $B_1(0)$ (see Example 1.63 in [2]). On the other hand, defining $\mathrm{C}_y := \big\{|x_1| \geqslant |(x_2, \ldots, x_n)|\big\}$ we have

$$\begin{aligned} &\limsup_{k\to\infty} \frac{\nu\big((y + \mathrm{C}_y) \cap B_{r_k}(y)\big)}{r_k^{n-1}} \\ &\qquad \leqslant \limsup_{k\to\infty} \nu^{y,r_k}\big(\mathrm{C}_y \cap \bar{B}_1(0)\big) \leqslant \nu^\infty\big(\mathrm{C}_y \cap \bar{B}_1(0)\big) = 0, \end{aligned}$$

by compactness of $\mathrm{C}_y \cap \bar{B}_1(0)$ (see again Example 1.63 in [2]). This gives (41) and (40). Since $r_k \to 0$ and $y \in \mathcal{J}$ were arbitrary, rectifiability follows.

Step 2. Fix a point $y \notin \mathscr{J}$. Then all rescaled measures $\nu^{y,r}$ converge to the zero measure as $r \to 0$, and the only split states $(u^\infty, h_y, \nu^\infty)$ that can be obtained from the blow-up are those satisfying $\nu^\infty = 0$. By Proposition 6.2, the function u^∞ must be constant. We claim that $y \notin \mathscr{J}$ is a point of vanishing mean oscillation. Indeed, let $r_k \to 0$ be a sequence such that

$$u^{y,r_k} \longrightarrow u^\infty \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \tag{74}$$

for some constant u^∞. Substituting $x := y + r_k z$, we can write

$$\int_{B_1(0)} |u^{y,r_k}(z) - u^\infty| \, dz = \omega_n \fint_{B_{r_k}(y)} |u(x) - u^\infty| \, dx, \tag{75}$$

with ω_n the measure of the unit ball in $\mathbb{R}^n$. Then

$$\begin{aligned}
&\fint_{B_{r_k}(y)} \left| u(x) - \fint_{B_{r_k}(y)} u(z)\, dz \right| dx \\
&\qquad \leqslant \fint_{B_{r_k}(y)} |u(x) - u^\infty| \, dx + \left| u^\infty - \fint_{B_{r_k}(y)} u(z)\, dz \right| \\
&\qquad \leqslant 2 \fint_{B_{r_k}(y)} |u(x) - u^\infty| \, dx \longrightarrow 0 \quad \text{as } k \to \infty,
\end{aligned}$$

by (75), (74) and the triangle inequality. This proves the claim.

Step 3. We know from Step 1 that $\mathscr{J}$ is codimension-one rectifiable. Therefore we can decompose $\mathscr{J} = \bigcup_k \mathscr{J}_k$ up to a $\mathscr{H}^{n-1}$-null set, where each $\mathscr{J}_k$ is contained in a Lipschitz graph. This implies in particular that $\mathscr{H}^{n-1} \lfloor \mathscr{J}_k$ is a locally finite measure. From Besicovitch derivation theorem we obtain

$$\nu = g_k \, \mathscr{H}^{n-1} \lfloor J_k + \nu_s, \tag{76}$$

where ν_s and $g_k \mathscr{H}^{n-1} \lfloor \mathscr{J}_k$ are mutually singular nonnegative measures, and the density g_k can be computed for $\mathscr{H}^{n-1}$-a.e. $y \in \mathscr{J}_k$ as the limit

$$g_k(y) = \lim_{r \to 0} \frac{\nu(B_r(y))}{\mathscr{H}^{n-1}(\mathscr{J}_k \cap B_r(y))} \tag{77}$$

(see Theorem 2.22 in [2]). Now notice that by definition of $\mathscr{J} \supset \mathscr{J}_k$

$$\limsup_{r \to 0} \frac{\nu(B_r(y))}{r^{n-1}} > 0 \quad \text{for all } y \in \mathscr{J}_k. \tag{78}$$

On the other hand, the rectifiability of $\mathscr{J}_k$ implies

$$\lim_{r \to 0} \frac{\mathscr{H}^{n-1}(\mathscr{J}_k \cap B_r(y))}{r^{n-1}} = 1 \quad \text{for } \mathscr{H}^{n-1}\text{-a.e. } y \in \mathscr{J}_k$$

(see Theorem 2.83(i) in [2]). Since the limit in (77) exists we conclude that the lim sup in (78) can in fact be replaced by a lim and therefore $g_k(y) > 0$ for $\mathscr{H}^{n-1}$-a.e. $y \in \mathscr{J}_k$. Neglecting the singular part ν_s in (76) we obtain

$$\nu \geqslant g\mathscr{H}^{n-1}\lfloor\mathcal{J} \quad \text{with } g(y) \in (0,\infty) \text{ for } \mathscr{H}^{n-1}\text{-a.e. } y \in \mathcal{J}. \tag{79}$$

Fix a generic point $y \in \mathcal{J}$. According to Propositions 6.6 and 7.1, there exist a unit vector η_y and a constant $g_y > 0$ such that for any limit measure ν^∞ from a converging blow-up sequence we have the representation

$$\nu^\infty = g_y\mathscr{H}^{n-1}\lfloor\mathcal{J}^\infty \quad \text{with } \mathcal{J}^\infty \subset \{\eta_y \cdot x = 0\}. \tag{80}$$

On the other hand, by (79) and rectifiability of $\mathcal{J}$, we also have

$$\nu^\infty \geqslant g(y)\mathscr{H}^{n-1}\lfloor\{\eta(y)\cdot x = 0\}, \tag{81}$$

where $\eta(y)$ is a unit vector normal to $\mathcal{J}$ in y. We conclude that $g(y) = g_y$ and $\eta(y) = \pm\eta_y$. Moreover, we obtain $J^\infty = \{\eta(y)\cdot x = 0\}$, thereby improving (80). Now Proposition 6.4 yields constants $u^+(y)$ and $u^-(y)$ such that

$$u^\infty = \begin{cases} u^+(y) & \text{in } \{\eta(y)\cdot x > 0\} \\ u^-(y) & \text{in } \{\eta(y)\cdot x < 0\} \end{cases}$$

for all limit functions u^∞ from converging blow-up sequences. We claim that $u^+(y)$ and $u^-(y)$ are the strong traces of u in $y \in \mathcal{J}$. Indeed, since

$$u^{y,r} \longrightarrow u^\infty \quad \text{in } \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n) \text{ as } r \to 0$$

by uniqueness of the limit, we can use the substitution $x := y + rz$ to get

$$\begin{aligned} &⨍_{B_r^+(y)} |u(x) - u^+(y)|\,dx + ⨍_{B_r^-(y)} |u(x) - u^-(y)|\,dx \\ &\quad = \frac{1}{\omega_n}\int_{B_1(0)} |u^{y,r}(z) - u^\infty(z)|\,dz \longrightarrow 0 \quad \text{as } r \to 0. \end{aligned}$$

This concludes the proof of the main theorem. □

A Proofs of the Regularity Theorems

In preparation for proving the regularity results stated in Section 4, we first collect some facts about matrices of Vandermonde type: Let $c_1 < \ldots < c_n$ be given numbers and consider for any $(v,\varepsilon) \in \mathbb{R}\times(0,\infty)$ the matrix

$$\begin{aligned} \mathtt{V}(v,\varepsilon) &:= \Big(\mathbf{a}(v + c_1\varepsilon) \ \cdots \ \mathbf{a}(v + c_n\varepsilon)\Big) \\ &= \begin{pmatrix} 1 & \cdots & 1 \\ v + c_1\varepsilon & \cdots & v + c_n\varepsilon \\ \vdots & \ddots & \vdots \\ (v + c_1\varepsilon)^{n-1} & \cdots & (v + c_n\varepsilon)^{n-1} \end{pmatrix}. \end{aligned}$$

Then $\mathtt{V}^{-1}(v,\varepsilon)$ exists and can be computed explicitly. It takes the form

$$\mathtt{V}^{-1}_{kl}(v,\varepsilon) = \varepsilon^{1-n}\frac{p_{kl}(v,\varepsilon)}{\prod_{i\neq k}(c_i-c_k)} \quad \forall k,l\in\{1,\dots,n\}, \tag{82}$$

where the $p_{kl}(v,\varepsilon)$ are suitable polynomials in the arguments $(v+c_i\varepsilon)$. Sharp estimates for the inverse matrix are known: For all $(v,\varepsilon)\in\mathbb{R}\times(0,\infty)$

$$\|\mathtt{V}^{-1}(v,\varepsilon)\| \leqslant \varepsilon^{1-n}\max_k \prod_{i\neq k}\frac{1+|v+c_i\varepsilon|}{|c_i-c_k|}, \tag{83}$$

with $\|\cdot\|$ the maximum absolute row sum norm (see Theorem 3.1 in [15]). Recall, however, that on finite-dimensional vector spaces all norms are equivalent. The invertibility of $\mathtt{V}(v,\varepsilon)$ implies that the vectors $\mathbf{a}(v+c_i\varepsilon)$ span $\mathbb{R}^n$. Another basis is given by the derivatives of the flux $\mathbf{a}$: For any $v\in\mathbb{R}$ let

$$\begin{aligned}\mathtt{W}(v) &:= \Big(\mathbf{a}(v)\ \mathbf{a}'(v)\ \cdots\ \mathbf{a}^{(n-1)}(v)\Big)\\ &= \begin{pmatrix} 1 & & & 0\\ v & 1 & & \\ \vdots & \vdots & \ddots & \\ v^{n-1} & (n-1)v^{n-2} & \cdots & (n-1)! \end{pmatrix}.\end{aligned}$$

Since $\mathbf{a}$ is a polynomial, Taylor expansion gives the identity

$$\mathbf{a}(w) = \sum_{k=0}^{n-1}\tfrac{1}{k!}(w-v)^k\mathbf{a}^{(k)}(v) \quad \forall (w,v)\in\mathbb{R}\times\mathbb{R}. \tag{84}$$

In particular, we can express the matrix $\mathtt{V}(v,\varepsilon)$ in terms of $\mathtt{W}(v)$. We have

$$\mathtt{V}(v,\varepsilon) = \mathtt{W}(v)\begin{pmatrix}1 & & & 0\\ & \varepsilon & & \\ & & \ddots & \\ 0 & & & \frac{1}{(n-1)!}\varepsilon^{n-1}\end{pmatrix}\begin{pmatrix}1 & 1 & \cdots & 1\\ c_1 & c_2 & \cdots & c_n\\ \vdots & \vdots & \ddots & \vdots\\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1}\end{pmatrix}$$

for all $(v,\varepsilon)\in\mathbb{R}\times(0,\infty)$. Since the last factor is just the invertible matrix $\mathtt{V}(0,1)$, for any $k\in\{0,\dots,n-1\}$ we can find numbers β_l^k such that

$$\mathbf{a}^{(k)}(v) = \varepsilon^{-k}\sum_{l=1}^{n}\beta_l^k\mathbf{a}(v+c_l\varepsilon) \quad \forall (v,\varepsilon)\in\mathbb{R}\times(0,\infty). \tag{85}$$

Yet another basis of $\mathbb{R}^n$ can be obtained by rescaling v in $\mathbf{a}(v)$. Let

$$\mathtt{U}(v,\varepsilon) := \Big(\mathbf{a}\big(v(1+c_1\varepsilon)\big) \ \cdots \ \mathbf{a}\big(v(1+c_n\varepsilon)\big)\Big) = \begin{pmatrix} 1 & \cdots & 1 \\ v(1+c_1\varepsilon) & \cdots & v(1+c_n\varepsilon) \\ \vdots & \ddots & \vdots \\ \big(v(1+c_1\varepsilon)\big)^{n-1} & \cdots & \big(v(1+c_n\varepsilon)\big)^{n-1} \end{pmatrix}$$

for $(v,\varepsilon) \in \mathbb{R} \times (0,\infty)$. Then $\mathtt{U}(v,\varepsilon)$ is invertible if $v \neq 0$, and

$$\mathtt{U}(v,\varepsilon) = \begin{pmatrix} 1 & & & 0 \\ & v & & \\ & & \ddots & \\ 0 & & & v^{n-1} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ 1+c_1\varepsilon & \cdots & 1+c_n\varepsilon \\ \vdots & \ddots & \vdots \\ (1+c_1\varepsilon)^{n-1} & \cdots & (1+c_n\varepsilon)^{n-1} \end{pmatrix},$$

where the last factor is just the invertible matrix $\mathtt{V}(1,\varepsilon)$. Let $\mathbf{e}_m$ denote the mth vector of the standard basis of $\mathbb{R}^n$ and recall the representation (82). For any $m \in \{1,\ldots,n\}$ we can then find polynomials $\gamma_l^m(\varepsilon)$ such that

$$\mathbf{e}_m = \varepsilon^{1-n} v^{1-m} \sum_{l=1}^{n} \gamma_l^m(\varepsilon) \mathbf{a}\big(v(1+c_l\varepsilon)\big) \quad \forall (v,\varepsilon) \in \mathbb{R} \times (0,\infty), v \neq 0. \tag{86}$$

Notice that this formula holds for all $v \in \mathbb{R}$ if $m = 1$.

Proof (of Proposition 4.3). Repeating the proof of Theorem 5.1 we obtain

$$\mathbf{a}(v) \cdot \nabla \chi(v, u(x)) = 0 \quad \text{in } \mathscr{D}'(\mathbb{R} \times \Omega) \tag{87}$$

because there is no entropy dissipation in the open set $\Omega \subset \mathbb{R}^n$, by assumption. The function χ is again defined by (34). Then Lemma 6.1 implies that u is continuous in Ω and constant along characteristics. The proof of Proposition 4.3 is just a more quantitative version of the one of Lemma 6.1.

Step 1. We first argue that $u \in \mathrm{C}^{1/(n-1)}_{\mathrm{loc}}(\Omega)$ is equivalent to the following statement: For all compact subsets $K \subset \Omega$ there exists $C_K < \infty$ such that

$$\forall \varepsilon > 0 \quad \forall x \in K \quad \forall y \in \overline{B_{\varepsilon^{n-1}}(x)} \cap K \quad |u(y) - u(x)| \leqslant C_K \varepsilon. \tag{88}$$

The inf over all admissible constants C_K coincides with the $\mathrm{C}^{1/(n-1)}$-norm of u over K. One direction is trivial: Fix some $K \subset \Omega$ compact and choose

$$C_K := \|u\|_{\mathrm{C}^{1/(n-1)}(K)} < \infty.$$

Pick $\varepsilon > 0$ and $x \in K$ arbitrary. Then

$$\forall y \in \overline{B_{\varepsilon^{n-1}}(x)} \cap K \quad |u(x) - u(y)| \leqslant C_K |x-y|^{\frac{1}{n-1}} \leqslant C_K \varepsilon,$$

which is (88). For the converse direction, we argue indirectly. Assume that there exists a compact subset $K \subset \Omega$ such that $\|u\|_{\mathrm{C}^{1/(n-1)}(K)} = \infty$. That is, there exist sequences of numbers $x_k, y_k \in K$ such that

$$|u(x_k) - u(y_k)| > k|x_k - y_k|^{\frac{1}{n-1}} \quad \forall k \in \mathbb{N}.$$

Defining $\varepsilon_k := |x_k - y_k|^{1/(n-1)}$, we obtain that

$$\forall k \in \mathbb{N} \quad \exists \varepsilon_k, x_k \quad \exists y_k \in \overline{B_{\varepsilon_k^{n-1}}(x_k)} \cap K \quad |u(x_k) - u(y_k)| > k\varepsilon_k.$$

Therefore (88) does not hold in this case, which proves the claim.

Step 2. Fix a compact subset $K \subset \Omega$ and let $R := \operatorname{dist}(K, \mathbb{R}^n \setminus \Omega) > 0$. For a given point $x \in K$ and numbers $0 > c_1 > \ldots > c_n > -1$ we define

$$v_k := u(x) + c_k\varepsilon \quad \text{with } k \in \{1, \ldots, n\} \text{ and } \varepsilon > 0. \tag{89}$$

The vectors $\mathbf{a}(v_k)$ form a basis for $\mathbb{R}^n$, therefore the parallelotope

$$\hat{T}(u(x), \varepsilon) := \left\{ \sum_{i=1}^{n} \lambda_i \mathbf{a}(u(x) + c_i\varepsilon) \colon \lambda_i \in (-1, 1) \right\} \tag{90}$$

is genuinely n-dimensional. The rescaled parallelotope

$$T(u(x), \varepsilon) := \hat{T}(u(x), \varepsilon)\ R \left| \sum\nolimits_i \mathbf{a}(u(x) + c_i\varepsilon) \right|^{-1} \tag{91}$$

is contained in $B_R(x)$. We claim that $u(y) \geqslant u(x) - \varepsilon$ for all $y \in x + T(u(x), \varepsilon)$. Indeed, for any such y there exist numbers $\lambda_i \in (-1, 1)$ such that

$$y = x + \alpha \sum_{i=1}^{n} \lambda_i \mathbf{a}(v_i).$$

Since $u(x) > v_1$ we first obtain that $\chi(v_1, u(x)) = 1$, by definition (34). The kinetic equation (87) with $v = v_1$ then implies that

$$\chi(v_1, u(y_1)) = 1 \quad \text{with } y_1 := x + \alpha\lambda_1\mathbf{a}(v_1).$$

By monotonicity of $v \mapsto \chi(v, u)$, then also $\chi(v_2, u(y_1)) = 1$ since $v_2 < v_1$. We apply the kinetic equation (87) with $v = v_2$ and find that

$$\chi(v_2, u(y_2)) = 1 \quad \text{with } y_2 := x + \alpha \sum_{i=1}^{2} \lambda_i \mathbf{a}(v_i).$$

Iterating this argument, we obtain

$$\chi(v_n, u(y)) = 1 \quad \text{with } y = x + \alpha \sum_{i=1}^{n} \lambda_i \mathbf{a}(v_i),$$

which implies $u(y) > v_n > u(x) - \varepsilon$. Similar reasoning gives an upper bound. To finish the proof, we must now estimate the radius of the biggest ball contained in the parallelogram $T(u(x), \varepsilon)$ in terms of ε, uniformly in $u(x)$.

Step 3. Let $K \subset \Omega$ be the compact set of the previous step and let $R > 0$ and c_k be the numbers introduced there. For all $x \in K$ and $\varepsilon > 0$ let

$$\mathtt{V}(u(x), \varepsilon) := \Big(\mathbf{a}(u(x) + c_1\varepsilon) \ \cdots \ \mathbf{a}(u(x) + c_n\varepsilon)\Big).$$

To simplify notation a bit, we occasionally do not write the argument $x \in K$. Let $\operatorname{cof} \mathtt{V}(u, \varepsilon)$ be the cofactor matrix of $\mathtt{V}(u, \varepsilon)$, which is defined by

$$(\operatorname{cof} \mathtt{V})_{ij}(u, \varepsilon) := (-1)^{i+j} \det \hat{\mathtt{V}}_{ij}(u, \varepsilon) \quad \text{with } i, j \in \{1, \ldots, n\}.$$

Here $\hat{\mathtt{V}}_{ij}(u, \varepsilon)$ is obtained from $\mathtt{V}(u, \varepsilon)$ by eliminating the ith row and the jth column. Let $(\operatorname{cof} \mathtt{V})_j(u, \varepsilon)$ be the jth column of $\operatorname{cof} \mathtt{V}(u, \varepsilon)$. By expansion by minors, the scalar product $\mathbf{a}(u + c_i\varepsilon) \cdot (\operatorname{cof} \mathtt{V})_j(u, \varepsilon)$ is just the determinant of the matrix that is obtained from $\mathtt{V}(u, \varepsilon)$ by replacing its jth column by the vector $\mathbf{a}(u + c_i\varepsilon)$. For any j we therefore have

$$\mathbf{a}(u + c_i\varepsilon) \cdot (\operatorname{cof} \mathtt{V})_j(u, \varepsilon) = \begin{cases} \det \mathtt{V}(u, \varepsilon) & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \tag{92}$$

which shows that $(\operatorname{cof} \mathtt{V})_j(u(x), \varepsilon)$ is orthogonal to the hyperplane

$$H_j(x, \varepsilon) := \sum_{i \neq j} \mathbb{R}\mathbf{a}(u(x) + c_i\varepsilon) \quad \forall x \in K.$$

The maximal ball contained between the hyperplanes $\mathbf{a}(u + c_j\varepsilon) \pm H_j(\cdot, \varepsilon)$ therefore has radius given by the scalar product

$$\left|\mathbf{a}(u + c_j\varepsilon) \cdot \frac{(\operatorname{cof} \mathtt{V})_j(u, \varepsilon)}{|(\operatorname{cof} \mathtt{V})_j(u, \varepsilon)|}\right| = \left|\frac{\det \mathtt{V}(u, \varepsilon)}{|(\operatorname{cof} \mathtt{V})_j(u, \varepsilon)|}\right| = \frac{1}{|\mathtt{V}_j^{-1}(u, \varepsilon)|},$$

where $\mathtt{V}_j^{-1}(u, \varepsilon)$ is the jth row of the inverse matrix of $\mathtt{V}(u, \varepsilon)$. We used again identity (92). It follows that the maximal ball contained in the parallelotope defined in (90) has radius given by

$$\min_j \frac{1}{|\mathtt{V}_j^{-1}(u, \varepsilon)|} = \frac{1}{\max_j |\mathtt{V}_j^{-1}(u, \varepsilon)|} \geqslant C \|\mathtt{V}^{-1}(u, \varepsilon)\|^{-1},$$

with $\|\cdot\|$ some matrix norm and C a constant depending only on n. Recall that on a finite-dimensional vector space all norms are equivalent. Consider now the parallelotope $T(u(x), \varepsilon)$ defined in (91). Then the radius $r(x, \varepsilon)$ of the maximal ball contained in $T(u(x), \varepsilon)$ is bounded below by

$$r(x, \varepsilon) \geqslant CR \|\mathtt{V}(u(x), \varepsilon)^{-1}\|^{-1} \left|\sum\nolimits_i \mathbf{a}(u(x) + c_i\varepsilon)\right|^{-1} \geqslant DR\varepsilon^{n-1}$$

for all $x \in K$ and $\varepsilon \in (0, 1)$, where for some constant $\hat{C} > 0$ we defined

$$D := \hat{C}^{-(n-1)}\Big(1 + \|u\|_{\mathrm{L}^\infty(K)}\Big)^{-2(n-1)}.$$

We used (83) and the bound $|\mathbf{a}(v)| \leqslant C(1+|v|)^{n-1}$, which holds for all $v \in \mathbb{R}$, with C some constant. We therefore conclude that

$$\forall \varepsilon \in (0,1) \quad \forall x \in K \quad \forall y \in \overline{B_{DR\varepsilon^{n-1}}(x)} \cap K \quad u(y) - u(x) \geqslant \varepsilon.$$

An upper bound can be proved in the same way, and for simplicity we assume that we obtain the same constants. Let $C_K := (DR)^{1/(n-1)}$ and $\hat{\varepsilon} := C_K \varepsilon$. Recalling the equivalence established in Step 1, we find the estimate

$$\sup_{\substack{(x,y)\in K\times K \\ |x-y| \leqslant C_K^{n-1}}} \frac{|u(x) - u(y)|}{|x-y|^{\frac{1}{n-1}}} \leqslant \hat{C} R^{-\frac{1}{n-1}} \Big(1 + \|u\|_{\mathrm{L}^\infty(K)}\Big)^2. \tag{93}$$

On the other hand, we can use the triangle inequality to get

$$\sup_{\substack{(x,y)\in K\times K \\ |x-y| \geqslant C_K^{n-1}}} \frac{|u(x) - u(y)|}{|x-y|^{\frac{1}{n-1}}} \leqslant 2\hat{C} R^{-\frac{1}{n-1}} \|u\|_{\mathrm{L}^\infty(K)} \Big(1 + \|u\|_{\mathrm{L}^\infty(K)}\Big)^2. \tag{94}$$

Combining (93) and (94) gives the result. The proposition is proved. □

Proof (of Proposition 4.4). We will actually prove a slightly more precise version of the proposition, without the assumption of compact support. By Theorem 5.1 and Remark 5.2 any generalized entropy solution u satisfies

$$\mathbf{a}(v) \cdot \nabla \chi(v, u(x)) = \tfrac{\partial}{\partial v} \mu(v,x) \quad \text{in } \mathscr{D}'(\mathbb{R} \times \Omega),$$

where χ is defined by (34) and μ is a locally finite measure (vanishing outside the support of u). Given some $\varphi \in \mathscr{D}(\Omega)$ we define for all $(v,x) \in \mathbb{R} \times \Omega$

$$\hat{\chi}(v,x) := \varphi(x)\chi(v,u(x)), \tag{95}$$

$$\hat{\mu}(v,x) := \varphi(x)\mu(v,x), \tag{96}$$

$$\hat{r}(v,x) := \Big(\mathbf{a}(v) \cdot \nabla\varphi(x)\Big)\chi(v,u(x)). \tag{97}$$

To simplify notation, we treat measures as if they were functions, and we assume that $\hat{\chi}$, $\hat{\mu}$ and $\hat{r}$ are extended by zero to $\mathbb{R} \times \mathbb{R}^n$. Notice that these terms are all integrable in $\mathbb{R} \times \mathbb{R}^n$. They satisfy the kinetic equation

$$\mathbf{a}(v) \cdot \nabla \hat{\chi}(v,x) = \tfrac{\partial}{\partial v} \hat{\mu}(v,x) + \hat{r}(v,x) \quad \text{in } \mathscr{D}'(\mathbb{R} \times \mathbb{R}^n). \tag{98}$$

For all functions $g \colon \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$ we define the operator

$$\triangle_y g(v,x) := g(v, x+y) - g(v,x) \quad \forall (v,x) \in \mathbb{R} \times \mathbb{R}^n, y \in \mathbb{R}^n.$$

Lemma A.1. *Let* $(\hat{\chi}, \hat{\mu}, \hat{g})$ *be defined by* (95)–(97). *For some* $D > 0$ *let*

$$A := D\Big(\iint_{\mathbb{R}\times\mathbb{R}^n} |\hat{r}|\, dv\, dx \Big)\Big(\iint_{\mathbb{R}\times\mathbb{R}^n} \big|\tfrac{\partial}{\partial v}\hat{\chi}\big|\, dv\, dx \Big)^{-1} \Big(\iint_{\mathbb{R}\times\mathbb{R}^n} |\hat{\mu}|\, dv\, dx \Big)^{-1},$$

$$R_k := D^{k+2}\Big(\iint_{\mathbb{R}\times\mathbb{R}^n} \big|\tfrac{\partial}{\partial v}\hat{\chi}\big|\, dv\, dx \Big)^{-(k+1)} \Big(\iint_{\mathbb{R}\times\mathbb{R}^n} |\hat{\mu}|\, dv\, dx \Big)^{-1}$$

for $k \in \{0, \dots, n-1\}$. *Then there exist constants* $C_k > 0$ *such that*

$$\begin{aligned} &\sup_{|h|\leqslant R_k} |h|^{-\frac{1}{k+2}} \iint_{\mathbb{R}\times\mathbb{R}^n} \big|\triangle_{h\mathbf{a}^{(k)}(v)}\hat{\chi}(v,x)\big|\, dv\, dx \\ &\quad\leqslant C_k(1+A)\Big(\iint_{\mathbb{R}\times\mathbb{R}^n} \big|\tfrac{\partial}{\partial v}\hat{\chi}\big|\, dv\, dx \Big)^{\frac{k+1}{k+2}} \Big(\iint_{\mathbb{R}\times\mathbb{R}^n} |\hat{\mu}|\, dv\, dx \Big)^{\frac{1}{k+2}}. \end{aligned} \tag{99}$$

For simplicity of notation, we do not write the accent ˆ in the following.

Proof (of Lemma A.1). The main difficulty is to prove inequality (99) for $k = 0$. This will be done in the Steps 2 and 3 below. In the first step we show how the case $k \geqslant 1$ can be reduced to the case $k = 0$.

Step 1. Choose numbers $c_1 < \ldots < c_n$ and consider the vectors $\mathbf{a}(v + c_l\varepsilon)$ for $(v, \varepsilon) \in \mathbb{R}\times(0,\infty)$. As noted before, they form a basis of $\mathbb{R}^n$. In particular, for each index $k \in \{1, \dots, n-1\}$ the derivative $\mathbf{a}^{(k)}(v)$ can be expanded in terms of $\mathbf{a}(v + c_l\varepsilon)$, see formula (85). We decompose

$$\begin{aligned} &\triangle_{h\mathbf{a}^{(k)}(v)}\chi(v,x) \\ &\quad= \chi\Big(v, x + h\varepsilon^{-k}\beta_1^k\mathbf{a}(v + c_1\varepsilon)\Big) - \chi(v,x) \\ &\qquad + \dots \\ &\qquad + \chi\Big(v, x + h\varepsilon^{-k}\sum_{l=1}^{n}\beta_l^k\mathbf{a}(v + c_l\varepsilon)\Big) \\ &\qquad\qquad - \chi\Big(v, x + h\varepsilon^{-k}\sum_{l=1}^{n-1}\beta_l^k\mathbf{a}(v + c_l\varepsilon)\Big) \\ &\quad= \sum_{l=1}^{n}\triangle_{h\varepsilon^{-k}\beta_l^k\mathbf{a}(v+c_l\varepsilon)}\chi\Big(v, x + h\varepsilon^{-k}\sum_{j=1}^{l-1}\beta_j^k\mathbf{a}(v + c_j\varepsilon)\Big). \end{aligned}$$

Integrating with respect to x we can use the triangle inequality and the invariance of the $\mathrm{L}^1(\mathbb{R}^n)$-norm under translations to simplify terms. Then

$$\int_{\mathbb{R}^n} |\triangle_{h\mathbf{a}^{(k)}(v)}\chi(v,x)|\, dx \leqslant \sum_{l=1}^{n}\int_{\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k\mathbf{a}(v+c_l\varepsilon)}\chi(v,x)\big|\, dx. \tag{100}$$

In each term of the right-hand side we need to adjust the v-argument in order to be able to use (99) for $k=0$. Using again the invariance of the $\mathrm{L}^1(\mathbb{R}^n)$-norm under translations, we find that for all functions $g\colon \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}$

$$\begin{aligned}
&\int_{\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k \mathbf{a}(v+c_l\varepsilon)} g(v,x)\big|\,dx \\
&\quad\leqslant \int_{\mathbb{R}^n} \Big(\big|g\big(v, x+h\varepsilon^{-k}\beta_l^k\mathbf{a}(v+c_l\varepsilon)\big)\big| + |g(v,x)|\Big)\,dx \\
&\quad= 2\int_{\mathbb{R}^n} |g(v,x)|\,dx. \qquad (101)
\end{aligned}$$

Applying this inequality with $g(v,x) := \chi(v,x) - \chi(v+c_l\varepsilon, x)$ we get

$$\begin{aligned}
&\int_{\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k \mathbf{a}(v+c_l\varepsilon)} \chi(v,x)\big|\,dx \\
&\quad\leqslant 2\int_{\mathbb{R}^n} |\chi(v,x) - \chi(v+c_l\varepsilon,x)|\,dx \\
&\qquad + \int_{\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k \mathbf{a}(v+c_l\varepsilon)} \chi(v+c_l\varepsilon,x)\big|\,dx.
\end{aligned}$$

The map $v \mapsto \chi(v,x)$ has bounded variation uniformly in x. Moreover χ has compact x-support. We integrate with respect to v and obtain

$$\begin{aligned}
&\iint_{\mathbb{R}\times\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k \mathbf{a}(v+c_l\varepsilon)} \chi(v,x)\big|\,dv\,dx \\
&\quad\leqslant 2|c_l|\varepsilon \iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx + \iint_{\mathbb{R}\times\mathbb{R}^n} \big|\triangle_{h\varepsilon^{-k}\beta_l^k \mathbf{a}(v)} \chi(v,x)\big|\,dv\,dx. \qquad (102)
\end{aligned}$$

Assume now that $|h| \leqslant R_k$. We make the ansatz

$$\varepsilon := \alpha_k |h|^{\frac{1}{k+2}} \left(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\right)^{-\frac{1}{k+2}} \left(\iint |\mu|\,dv\,dx\right)^{\frac{1}{k+2}},$$

for some $\alpha_k > 0$ that will be chosen later (see page 48). Here we only assume that α_k is large enough such that $\alpha_k^{-k}\max_l |\beta_l^k| \leqslant 1$. Then

$$\begin{aligned}
|h\varepsilon^{-k}\beta_l^k| &= |h|^{\frac{2}{k+2}} \Big(\alpha_k^{-k}\max_l |\beta_l^k|\Big) \left(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\right)^{\frac{k}{k+2}} \left(\iint |\mu|\,dv\,dx\right)^{-\frac{k}{k+2}} \\
&\leqslant \left(\iint |\chi|\,dv\,dx\right)^2 \left(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\right)^{-1} \left(\iint |\mu|\,dv\,dx\right)^{-1} = R_0.
\end{aligned}$$

Recalling (100) we find a constant $B_k = B_k(c_l, \beta_l^k)$ such that

$$\sup_{|h|\leqslant R_k} |h|^{-\frac{1}{k+2}} \iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h\mathbf{a}^{(k)}(v)}\chi(v,x)|\,dv\,dx$$
$$\leqslant B_k\Bigg\{\alpha_k\Big(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\Big)^{\frac{k+1}{k+2}}\Big(\iint |\mu|\,dv\,dx\Big)^{\frac{1}{k+2}}$$
$$+\,\alpha_k^{-\frac{k}{2}}\Big(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\Big)^{\frac{k}{2(k+2)}}\Big(\iint |\mu|\,dv\,dx\Big)^{-\frac{k}{2(k+2)}}$$
$$\sup_{|\hat h|\leqslant R_0} |\hat h|^{-\frac{1}{2}} \iint_{\mathbb{R}\times\mathbb{R}^n} \big|\triangle_{\hat h\mathbf{a}(v)}\chi(v,x)\big|\,dv\,dx\Bigg\}. \tag{103}$$

The last term can be estimated by (99) with $k=0$. For all $k\geqslant 1$ we get

$$\sup_{|h|\leqslant R_k} |h|^{-\frac{1}{k+2}} \iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h\mathbf{a}^{(k)}(v)}\chi(v,x)|\,dv\,dx$$
$$\leqslant B_k\Big(\alpha_k + C_0(1+A)\alpha_k^{-\frac{k}{2}}\Big)\Big(\iint \big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\Big)^{\frac{k+1}{k+2}}\Big(\iint |\mu|\,dv\,dx\Big)^{\frac{1}{k+2}}.$$

Step 2. Consider now the case $k=0$. Select a test function $\rho\in\mathscr{D}(\mathbb{R})$ with $\int_{\mathbb{R}}\rho(v)\,dv=1$ and $\rho\geqslant 0$. We define the family of mollifiers $\rho_\varepsilon(v):=\varepsilon^{-1}\rho(v/\varepsilon)$ for all $(v,\varepsilon)\in\mathbb{R}\times(0,\infty)$. For any $w\in\mathbb{R}$ we then have the estimate

$$\int_{\mathbb{R}^n} |\triangle_{h\mathbf{a}(w)}\chi(w,x)|\,dx$$
$$=\int_{\mathbb{R}^n}\bigg|\int_{\mathbb{R}}\rho_\varepsilon(v-w)\,\triangle_{h\mathbf{a}(w)}\chi(w,x)\,dv\bigg|\,dx$$
$$\leqslant \iint_{\mathbb{R}\times\mathbb{R}^n}\rho_\varepsilon(v-w)\big|\triangle_{h\mathbf{a}(w)}\big(\chi(w,x)-\chi(v,x)\big)\big|\,dv\,dx$$
$$+\int_{\mathbb{R}^n}\bigg|\int_{\mathbb{R}}\rho_\varepsilon(v-w)\,\triangle_{h\mathbf{a}(w)}\chi(v,x)\,dv\bigg|\,dx. \tag{104}$$

As in (101) we can get rid of the operator $\triangle_{h\mathbf{a}(w)}$ in the first term on the right-hand side by using the triangle inequality and the invariance of the $\mathrm{L}^1(\mathbb{R}^n)$-norm under translations. We integrate (104) with respect to w. Since the function $v\mapsto\chi(v,x)$ has bounded variation uniformly in x we obtain

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\rho_\varepsilon(v-w)\,|\chi(w,x)-\chi(v,x)|\,dv\,dw$$
$$\leqslant \int_{\mathbb{R}}\rho_\varepsilon(z)\Big(\int_{\mathbb{R}}|\chi(w,x)-\chi(w+z,x)|\,dw\Big)\,dz$$
$$\leqslant \Big(\int_{\mathbb{R}}\rho_\varepsilon(z)|z|\,dz\Big)\int_{\mathbb{R}}\big|\tfrac{\partial}{\partial v}\chi(v,x)\big|\,dv \tag{105}$$

for all $x\in\mathbb{R}^n$. There exists a constant $M_1>0$ such that

$$\int_{\mathbb{R}} \rho_\varepsilon(z)|z|\,dz = M_1\varepsilon \quad \forall \varepsilon > 0.$$

We arrive at the following estimate: For all $(h,\varepsilon) \in \mathbb{R} \times (0,\infty)$

$$\begin{aligned} &\iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h\mathbf{a}(w)}\chi(w,x)|\,dw\,dx \\ &\quad \leqslant 2M_1\varepsilon \iint \left|\tfrac{\partial}{\partial v}\chi\right| dv\,dx \\ &\qquad + \iint_{\mathbb{R}\times\mathbb{R}^n} \left| \int_{\mathbb{R}} \rho_\varepsilon(v-w)\,\triangle_{h\mathbf{a}(w)}\chi(v,x)\,dv \right| dw\,dx. \end{aligned} \tag{106}$$

Step 3. To estimate the second term on the right-hand side of (106) we define $R := \frac{\partial}{\partial v}\mu + r$. Without loss of generality we may assume that $h > 0$. Using (84) and (98), we obtain for all $w \in \mathbb{R}$ that in $\mathscr{D}'(\mathbb{R}\times\mathbb{R}^n)$

$$\begin{aligned} &R\left(v, x + \sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\right) \\ &\quad = \mathbf{a}(v) \cdot \nabla\chi\left(v, x + \sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\right) \\ &\quad = \sum_{k=0}^{n-1} \tfrac{1}{k!}(v-w)^k \mathbf{a}^{(k)}(w) \cdot \nabla\chi\left(v, x + \sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\right) \\ &\quad = \sum_{k=0}^{n-1} \tfrac{1}{k!}(v-w)^k \tfrac{\partial}{\partial s_k}\chi\left(v, x + \sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\right). \end{aligned}$$

We average in s over the rectangle $H := [0,h_0] \times \cdots \times [0,h_{n-1}]$ with suitable numbers $h_i > 0$ to be specified later. By Gauss-Green theorem we obtain

$$\begin{aligned} &\sum_{k=0}^{n-1} \tfrac{1}{k!}(v-w)^k\, h_k^{-1} \fint_H \triangle_{h_k\mathbf{a}^{(k)}(w)}\chi\left(v, x + \sum_{l\neq k} s_l \mathbf{a}^{(l)}(w)\right) ds \\ &\quad = \fint_H R\left(v, x + \sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\right) ds. \end{aligned} \tag{107}$$

Our goal is to single out one term in (107) that does not depend on s anymore. To achieve this, we fix $k = 0$ and express χ on the left-hand side as

$$
\begin{aligned}
&\chi\Bigg(v, x+\sum_{l=1}^{n-1} s_l \mathbf{a}^{(l)}(w)\Bigg)\\
&\quad = \chi(v,x)\\
&\qquad + \chi\big(v, x+s_1\mathbf{a}'(w)\big) - \chi(v,x)\\
&\qquad + \dots\\
&\qquad + \chi\Bigg(v, x+\sum_{l=1}^{n-1} s_l \mathbf{a}^{(l)}(w)\Bigg) - \chi\Bigg(v, x+\sum_{l=1}^{n-2} s_l \mathbf{a}^{(l)}(w)\Bigg)\\
&\quad = \chi(v,x) + \sum_{k=1}^{n-1} \triangle_{s_k \mathbf{a}^{(k)}(w)} \chi\Bigg(v, x+\sum_{l=1}^{k-1} s_l \mathbf{a}^{(l)}(w)\Bigg).
\end{aligned}
$$

Recollecting terms, we can now write

$$
\begin{aligned}
&h_0^{-1} \triangle_{h_0 \mathbf{a}(w)} \chi(v,x)\\
&\quad = -\sum_{k=1}^{n-1} \Bigg\{ \tfrac{1}{k!}(v-w)^k\, h_k^{-1} \fint_H \triangle_{h_k \mathbf{a}^{(k)}(w)} \chi\Bigg(v, x+\sum_{l\neq k} s_l \mathbf{a}^{(l)}(w)\Bigg)\, ds\\
&\qquad\qquad + h_0^{-1} \fint_H \triangle_{h_0 \mathbf{a}(w)} \triangle_{s_k \mathbf{a}^{(k)}(w)} \chi\Bigg(v, x+\sum_{l=1}^{k-1} s_l \mathbf{a}^{(l)}(w)\Bigg)\, ds \Bigg\}\\
&\qquad + \fint_H R\Bigg(v, x+\sum_{l=0}^{n-1} s_l \mathbf{a}^{(l)}(w)\Bigg)\, ds.
\end{aligned}
\tag{108}
$$

We first integrate (108) in v against the mollifier $\rho_\varepsilon(v-w)$ and then take the $\mathrm{L}^1(\mathbb{R}^n)$-norm with respect to x. Using the triangle inequality we find

$$
\begin{aligned}
&h_0^{-1} \int_{\mathbb{R}^n} \Bigg| \int_{\mathbb{R}} \rho_\varepsilon(v-w)\, \triangle_{h_0 \mathbf{a}(w)} \chi(v,x)\, dv \Bigg|\, dx\\
&\quad \leqslant \sum_{k=1}^{n-1} h_k^{-1} \int_{\mathbb{R}} \rho_\varepsilon(v-w) \Bigg\{ \tfrac{1}{k!} |v-w|^k \int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(v,x)|\, dx\\
&\qquad\qquad + 2h_0^{-1} \int_0^{h_k} \int_{\mathbb{R}^n} |\triangle_{s_k \mathbf{a}^{(k)}(w)} \chi(v,x)|\, dx\, ds_k \Bigg\}\, dv\\
&\qquad + \int_{\mathbb{R}^n} \Bigg| \int_{\mathbb{R}} \rho_\varepsilon(v-w) R(v,x)\, dv \Bigg|\, dx.
\end{aligned}
\tag{109}
$$

We used the invariance of the $\mathrm{L}^1(\mathbb{R}^n)$-norm under translations to get rid of the operator $\triangle_{h_0 \mathbf{a}(w)}$ on the right-hand side. The same argument gives

$$\begin{aligned}&\int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(v,x)|\, dx \\ &\qquad \leqslant 2 \int_{\mathbb{R}^n} |\chi(v,x) - \chi(w,x)|\, dx + \int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(w,x)|\, dx\end{aligned}$$

and the analogous estimate with h_k replaced by s_k. We use this inequality in (109) and integrate with respect to w in $\mathbb{R}$. Then

$$\begin{aligned}&\int_{\mathbb{R}} \int_{\mathbb{R}} \rho_\varepsilon(v-w)|v-w|^k \bigg(\int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(v,x)|\, dx \bigg)\, dv\, dw \\ &\quad \leqslant 2 \int_{\mathbb{R}^n} \bigg(\iint_{\mathbb{R}\times\mathbb{R}} \rho_\varepsilon(v-w)|v-w|^k\, |\chi(v,x) - \chi(w,x)|\, dv\, dw \bigg)\, dx \\ &\qquad + \iint_{\mathbb{R}\times\mathbb{R}} \rho_\varepsilon(v-w)|v-w|^k \bigg(\int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(w,x)|\, dx \bigg)\, dv\, dw.\end{aligned}$$

For the first term on the right-hand side a similar reasoning as for (105) applies. The second term is a convolution in w, which can be estimated with Young's inequality. Therefore we obtain the following bound

$$\begin{aligned}&\int_{\mathbb{R}} \int_{\mathbb{R}} \rho_\varepsilon(v-w)|v-w|^k \bigg(\int_{\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(v,x)|\, dx \bigg)\, dv\, dw \\ &\quad \leqslant 2 \bigg(\int_{\mathbb{R}} \rho_\varepsilon(z)|z|^{k+1}\, dz \bigg) \iint \big| \tfrac{\partial}{\partial v} \chi \big|\, dv\, dx \\ &\qquad + \bigg(\int_{\mathbb{R}} \rho_\varepsilon(z)|z|^k\, dz \bigg) \iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h_k \mathbf{a}^{(k)}(w)} \chi(w,x)|\, dw\, dx.\end{aligned}$$

Since $\rho \in \mathscr{D}(\mathbb{R})$, there exist constants $M_j > 0$ such that

$$\int_{\mathbb{R}} \rho_\varepsilon(z)|z|^j\, dz = M_j \varepsilon^j \quad \forall \varepsilon > 0$$

for all $j \geqslant 0$. For the corresponding term in (109) with s_k instead of h_k, we can argue in a similar way. Notice that in this case the $|v-w|^k$ do not appear and we obtain different powers in ε. For the last term in (109) we find

$$\begin{aligned}&\int_{\mathbb{R}} \int_{\mathbb{R}^n} \bigg| \int_{\mathbb{R}} \rho_\varepsilon(v-w) R(v,x)\, dv \bigg|\, dx\, dw \\ &\qquad \leqslant C \bigg(\varepsilon^{-1} \iint |\mu|\, dv\, dx + \iint |r|\, dv\, dx \bigg),\end{aligned}$$

with $C > 0$ some constant. Collecting all terms we arrive at

$$
\begin{aligned}
&h_0^{-1}\iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h_0\mathbf{a}(w)}\chi(w,x)|\,dw\,dx\\
&\quad\leqslant \Big(h_0^{-1}\,2M_1\varepsilon+\sum_{k=1}^{n-1}\Big(\tfrac{1}{k!}h_k^{-1}\,2M_{k+1}\varepsilon^{k+1}+2h_0^{-1}\,2M_1\varepsilon\Big)\Big)\iint\big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\\
&\qquad+\sum_{k=1}^{n-1}\Big(\tfrac{1}{k!}h_k^{-1}\,M_k\varepsilon^k+h_0^{-1}\,2\Big)\sup_{s_k\in[0,h_k]}\iint_{\mathbb{R}\times\mathbb{R}^n}|\triangle_{s_k\mathbf{a}^{(k)}(w)}\chi(w,x)|\,dw\,dx\\
&\qquad+C\Big(\varepsilon^{-1}\iint|\mu|\,dv\,dx+\iint|r|\,dv\,dx\Big).
\end{aligned}\tag{110}
$$

For any $k\in\{1,\dots,n-1\}$ we choose h_k in such a way that we obtain the correct homogeneities. With $h_k:=h_0\varepsilon^k$ the inequality (110) simplifies to

$$
\begin{aligned}
&h_0^{-1}\iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h_0\mathbf{a}(w)}\chi(w,x)|\,dw\,dx\\
&\quad\leqslant C\Big\{h_0^{-1}\varepsilon\iint\big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx+\Big(\varepsilon^{-1}\iint|\mu|\,dv\,dx+\iint|r|\,dv\,dx\Big)\\
&\qquad+\sum_{k=1}^{n-1}h_0^{-1}\sup_{s_k\in[0,h_k]}\iint_{\mathbb{R}\times\mathbb{R}^n}|\triangle_{s_k\mathbf{a}^{(k)}(w)}\chi(w,x)|\,dw\,dx\Big\},
\end{aligned}\tag{111}
$$

with $C>0$ some constant. Assume now that $|h_0|\leqslant R_0$. We make the ansatz

$$
\varepsilon:=h_0^{\frac12}\Big(\iint\big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\Big)^{-\frac12}\Big(\iint|\mu|\,dv\,dx\Big)^{\frac12},\tag{112}
$$

which implies the inequalities $|h_k|=|h_0\varepsilon^k|\leqslant R_k$ and

$$
\iint|r|\,dv\,dx\leqslant A\varepsilon^{-1}\iint|\mu|\,dv\,dx,
$$

with A defined above. Multiplying (111) by $h_0^{1/2}$, we get the estimate

$$
\begin{aligned}
&h_0^{-\frac12}\iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h_0\mathbf{a}(w)}\chi(w,x)|\,dw\,dx\\
&\quad\leqslant B_0\Big\{(1+A)\Big(\iint\big|\tfrac{\partial}{\partial v}\chi\big|\,dv\,dx\Big)^{\frac12}\Big(\iint|\mu|\,dv\,dx\Big)^{\frac12}\\
&\qquad+\sum_{k=1}^{n-1}h_0^{-\frac12}h_k^{\frac{1}{k+2}}\sup_{|\hat h|\leqslant R_k}|\hat h|^{-\frac{1}{k+2}}\iint_{\mathbb{R}\times\mathbb{R}^n}\big|\triangle_{\hat h\mathbf{a}^{(k)}(v)}\chi(v,x)\big|\,dv\,dx\Big\},
\end{aligned}\tag{113}
$$

with $B_0=B_0(\rho)$ some constant. Since $h_k=h_0\varepsilon^k$ and by (112), we have

$$h_0^{-\frac{1}{2}} h_k^{\frac{1}{k+2}} = \left(\iint \left|\tfrac{\partial}{\partial v}\chi\right| dv\, dx \right)^{-\frac{k}{2(k+2)}} \left(\iint |\mu|\, dv\, dx \right)^{\frac{k}{2(k+2)}}.$$

The right-hand side of (113) can now be estimated using (103). For each $k \geqslant 1$ we choose $\alpha_k > 0$ large enough such that

$$B_0 B_k \alpha_k^{-k/2} \leqslant \frac{1}{2(n-1)}.$$

Summing up we obtain

$$\begin{aligned} &\sup_{|h_0|\leqslant R_0} |h_0|^{-\frac{1}{2}} \iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{h_0\mathbf{a}(w)}\chi(w,x)|\, dw\, dx \\ &\leqslant B_0 \left((1+A) + \sum_{k=1}^{n-1} B_k \alpha_k \right) \left(\iint \left|\tfrac{\partial}{\partial v}\chi\right| dv\, dx \right)^{\frac{1}{2}} \left(\iint |\mu|\, dv\, dx \right)^{\frac{1}{2}} \\ &\quad + \tfrac{1}{2} \sup_{|\hat{h}|\leqslant R_0} |\hat{h}|^{-\frac{1}{2}} \iint_{\mathbb{R}\times\mathbb{R}^n} |\triangle_{\hat{h}\mathbf{a}(w)}\chi(w,x)|\, dw\, dx. \end{aligned}$$

The last term can then be absorbed into the left-hand side. □

We can now conclude the proof of Proposition 4.4. Recall from (86), that for given numbers $c_1 < \ldots < c_n$ the standard basis vector $\mathbf{e}_1$ can be expanded in terms of the $\mathbf{a}(v(1+c_l\varepsilon))$ for all $(v,\varepsilon) \in \mathbb{R}\times(0,\infty)$. For $\mathbf{e}_m$ with $m \geqslant 2$ the expansion (86) contains negative powers of v. Notice, however, that

$$\mathbf{a}'\big(v(1+c_l\varepsilon)\big) = \begin{pmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & n-1 & 0 \end{pmatrix} \mathbf{a}\big(v(1+c_l\varepsilon)\big)$$

for all $(v,\varepsilon) \in \mathbb{R}\times(0,\infty)$ and $l \in \{1,\ldots,n\}$. By decreasing the dimension by one, we can then use (86) again to find an expansion of $\mathbf{e}_2$ in terms of the derivatives $\mathbf{a}'(v(1+c_l\varepsilon))$ for all $(v,\varepsilon) \in \mathbb{R}\times(0,\infty)$. This argument can be iterated. There exist polynomials $\delta_l^m(\varepsilon)$ such that

$$\mathbf{e}_m = \varepsilon^{m-n} \sum_{l=1}^{n-m+1} \delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}\big(v(1+c_l\varepsilon)\big) \quad \forall (v,\varepsilon) \in \mathbb{R}\times(0,\infty) \tag{114}$$

for all $m \in \{1,\ldots,n\}$. If $m = n$, then $\mathbf{a}^{(n-1)}\big(v(1+c_1\varepsilon)\big) = (n-1)!\mathbf{e}_n$. Now

$$\begin{aligned}
&\triangle_{h\mathbf{e}_m}\chi(v,x)\\
&\quad=\chi\Big(v,x+h\varepsilon^{m-n}\delta_1^m(\varepsilon)\mathbf{a}^{(m-1)}\big(v(1+c_1\varepsilon)\big)\Big)-\chi(v,x)\\
&\qquad+\dots\\
&\qquad+\chi\left(v,x+h\varepsilon^{m-n}\sum_{l=1}^{n-m+1}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}\big(v(1+c_l\varepsilon)\big)\right)\\
&\qquad\qquad-\chi\left(v,x+h\varepsilon^{m-n}\sum_{l=1}^{n-m}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}\big(v(1+c_l\varepsilon)\big)\right)\\
&\quad=\sum_{l=1}^{n-m+1}\triangle_{h\varepsilon^{m-n}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}(v(1+c_l\varepsilon))}\\
&\qquad\qquad\chi\left(v,x+h\varepsilon^{m-n}\sum_{j=1}^{l-1}\delta_j^m(\varepsilon)\mathbf{a}^{(m-1)}\big(v(1+c_j\varepsilon)\big)\right).
\end{aligned}$$

We integrate with respect to x and use the triangle inequality and the invariance of the $\mathrm{L}^1(\mathbb{R}^n)$-norm under translations to simplify terms. Then we integrate with respect to v. We obtain the estimate

$$\begin{aligned}
&\iint_{\mathbb{R}\times\mathbb{R}^n}|\triangle_{h\mathbf{e}_m}\chi(v,x)|\,dv\,dx\\
&\quad\leqslant\sum_{l=1}^{n-m+1}\iint_{\mathbb{R}\times\mathbb{R}^n}\Big|\triangle_{h\varepsilon^{m-n}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}(v(1+c_l\varepsilon))}\chi(v,x)\Big|\,dv\,dx. \qquad (115)
\end{aligned}$$

For each term on the right-hand side we need to adjust the v-argument in order to be able to use (99) with $k=m-1$. Proceeding as before, we get

$$\begin{aligned}
&\iint_{\mathbb{R}\times\mathbb{R}^n}\Big|\triangle_{h\varepsilon^{m-n}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}(v(1+c_l\varepsilon))}\chi(v,x)\Big|\,dv\,dx\\
&\quad\leqslant 2\iint_{\mathbb{R}\times\mathbb{R}^n}\big|\chi(v,x)-\chi\big(v(1+c_l\varepsilon),x\big)\big|\,dv\,dx\\
&\qquad+\iint_{\mathbb{R}\times\mathbb{R}^n}\Big|\triangle_{h\varepsilon^{m-n}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}(v(1+c_l\varepsilon))}\chi\big(v(1+c_l\varepsilon),x\big)\Big|\,dv\,dx. \qquad (116)
\end{aligned}$$

For the first term on the right-hand side recall (95) and (34). Then

$$\begin{aligned}
&\iint_{\mathbb{R}\times\mathbb{R}^n}\big|\chi(v,x)-\chi\big(v(1+c_l\varepsilon),x\big)\big|\,dv\,dx\\
&\quad=\int_{\mathbb{R}^n}|\varphi(x)|\left(\int_{\mathbb{R}}\mathbf{1}_{\left[\frac{u(x)}{1+c_l\varepsilon},u(x)\right]}(v)\,dv\right)dx=\frac{c_l\varepsilon}{1+c_l\varepsilon}\iint|\chi|\,dv\,dx. \qquad (117)
\end{aligned}$$

Without loss of generality let us assume that $|c_l|\leqslant\frac{1}{2}$. We require that $\varepsilon\leqslant 1$, so the right-hand side of (117) is finite. We now make the ansatz

$$\varepsilon := |h|^{\frac{1}{n+1}} \Big(\iint |\chi| \, dv \, dx \Big)^{-\frac{m+1}{n+1}} \Big(\iint \big| \tfrac{\partial}{\partial v} \chi \big| \, dv \, dx \Big)^{\frac{m}{n+1}} \Big(\iint |\mu| \, dv \, dx \Big)^{\frac{1}{n+1}},$$

which implies the bound

$$|h| \leqslant \Big(\iint |\chi| \, dv \, dx \Big)^{m+1} \Big(\iint \big| \tfrac{\partial}{\partial v} \chi \big| \, dv \, dx \Big)^{-m} \Big(\iint |\mu| \, dv \, dx \Big)^{-1} =: L.$$

Then there exists a constant $C > 0$ such that $|h\varepsilon^{m-n}\delta_l^m(\varepsilon)| \leqslant C^{m+1}L$. We want to apply Lemma A.1 to estimate the last term in (116). We have

$$\begin{aligned} &\sup_{|h| \leqslant L} |h|^{-\frac{1}{n+1}} \iint_{\mathbb{R} \times \mathbb{R}^n} \Big| \triangle_{h\varepsilon^{m-n}\delta_l^m(\varepsilon)\mathbf{a}^{(m-1)}(v(1+c_l\varepsilon))} \chi\big(v(1+c_l\varepsilon), x\big) \Big| \, dv \, dx \\ &\quad \leqslant C L^{-\frac{1}{n+1}\frac{m-n}{m+1}} \sup_{|\hat{h}| \leqslant R_{m-1}} |\hat{h}|^{-\frac{1}{m+1}} \iint_{\mathbb{R} \times \mathbb{R}^n} \big| \triangle_{h\mathbf{a}^{(m-1)}(v)} \hat{\chi}(v, x) \big| \, dv \, dx, \end{aligned}$$

where we defined R_{m-1} as in Lemma A.1 with

$$D := C \iint |\chi| \, dv \, dx.$$

Collecting all terms we find a constant $C > 0$ such that

$$\begin{aligned} &\sup_{|h| \leqslant L} |h|^{-\frac{1}{n+1}} \iint_{\mathbb{R} \times \mathbb{R}^n} |\triangle_{h\mathbf{e}_m} \chi(v, x)| \, dv \, dx \\ &\quad \leqslant C(1+A) \Big(\iint |\chi| \, dv \, dx \Big)^{\frac{n-m}{n+1}} \Big(\iint \big| \tfrac{\partial}{\partial v} \chi \big| \, dv \, dx \Big)^{\frac{m}{n+1}} \Big(\iint |\mu| \, dv \, dx \Big)^{\frac{1}{n+1}}, \end{aligned}$$

with A given by Lemma A.1. For large h we use the triangle inequality and the invariance of the $\mathrm{L}^1(\mathbb{R} \times \mathbb{R}^n)$-norm under translations to get

$$\begin{aligned} &\sup_{|h| \geqslant L} |h|^{-\frac{1}{n+1}} \iint_{\mathbb{R} \times \mathbb{R}^n} |\triangle_{h\mathbf{e}_m} \chi(v, x)| \, dv \, dx \\ &\quad \leqslant 2L^{-\frac{1}{n+1}} \Big(\iint |\chi| \, dv \, dx \Big) \\ &\quad = 2\Big(\iint |\chi| \, dv \, dx \Big)^{\frac{n-m}{n+1}} \Big(\iint \big| \tfrac{\partial}{\partial v} \chi \big| \, dv \, dx \Big)^{\frac{m}{n+1}} \Big(\iint |\mu| \, dv \, dx \Big)^{\frac{1}{n+1}}. \end{aligned}$$

We conclude that there exists a universal constant $C > 0$ such that

$$\begin{aligned} &\sup_{|h| \neq 0} |h|^{-\frac{1}{n+1}} \iint_{\mathbb{R} \times \mathbb{R}^n} |\triangle_{h\mathbf{e}_m} \chi(v, x)| \, dv \, dx \\ &\quad \leqslant C(1+A) \Big(\iint |\chi| \, dv \, dx \Big)^{\frac{n-m}{n+1}} \Big(\iint \big| \tfrac{\partial}{\partial v} \chi \big| \, dv \, dx \Big)^{\frac{m}{n+1}} \Big(\iint |\mu| \, dv \, dx \Big)^{\frac{1}{n+1}} \end{aligned}$$

for all $m \in \{1, \dots, n\}$. The proposition now follows easily: If u (and thus μ) has compact support in Ω, then we can choose the cut-off function φ that we used in (95)–(97) equal to one on $\operatorname{spt} u$. Then A vanishes (see the definition in Lemma A.1) and the terms simplify a bit. Since χ has total v-variation equal to two in $\operatorname{spt} u$, we obtain the inequality (29). The proposition is proved. □

References

1. L. Ambrosio, C. De Lellis, and J. Maly. On the chain rule for the divergence of BV like vector fields: Applications, partial results, open problems. In *Perspectives in Nonlinear Partial Differential Equations: in honor of Haim Brezis*. Birkhäuser, 2006.
2. L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.
3. M. Bézard. Régularité L^p précisée des moyennes dans les équations de transport. *Bull. Soc. Math. France*, 122(1):29–76, 1994.
4. F. Bouchut. Hypoelliptic regularity in kinetic equations. *J. Math. Pures Appl. (9)*, 81(11):1135–1159, 2002.
5. F. Bouchut and F. James. Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness. *Comm. Partial Differential Equations*, 24(11-12):2173–2189, 1999.
6. Y. Brenier. Averaged multivalued solutions for scalar conservation laws. *SIAM J. Numer. Anal.*, 21(6):1013–1037, 1984.
7. G.-Q. Chen and H. Frid. Divergence-measure fields and hyperbolic conservation laws. *Arch. Ration. Mech. Anal.*, 147(2):89–118, 1999.
8. G.-Q. Chen and M. Rascle. Initial layers and uniqueness of weak entropy solutions to hyperbolic conservation laws. *Arch. Ration. Mech. Anal.*, 153(3):205–220, 2000.
9. K. S. Cheng. A regularity theorem for a nonconvex scalar conservation law. *J. Differential Equations*, 61(1):79–127, 1986.
10. C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2000.
11. C. De Lellis, F. Otto, and M. Westdickenberg. Structure of entropy solutions for multi-dimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 170(2):137–184, 2003.
12. C. De Lellis and T. Rivière. The rectifiability of entropy measures in one space dimension. *J. Math. Pures Appl. (9)*, 82(10):1343–1367, 2003.
13. R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989.
14. R. J. DiPerna, P.-L. Lions, and Y. Meyer. L^p regularity of velocity averages. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(3-4):271–287, 1991.
15. W. Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numer. Math.*, 4:117–123, 1962.
16. F. Golse, P.-L. Lions, B. Perthame, and R. Sentis. Regularity of the moments of the solution of a transport equation. *J. Funct. Anal.*, 76(1):110–125, 1988.

17. D. Hoff. The sharp form of Oleĭnik's entropy condition in several space variables. *Trans. Amer. Math. Soc.*, 276(2):707–714, 1983.
18. P.-E. Jabin and B. Perthame. Regularity in kinetic formulations via averaging lemmas. *ESAIM Control Optim. Calc. Var.*, 8:761–774 (electronic), 2002.
19. S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.
20. P.-L. Lions, B. Perthame, and E. Tadmor. A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Amer. Math. Soc.*, 7(1):169–191, 1994.
21. P. Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.
22. O. A. Oleĭnik. Discontinuous solutions of non-linear differential equations. *Uspehi Mat. Nauk (N.S.)*, 12(3(75)):3–73, 1957.
23. E. Yu. Panov. Existence of strong traces for generalized solutions of multidimensional scalar conservation laws. *J. Hyperbolic Differ. Equ.*, 2(4):885–908, 2005.
24. E. Tadmor and T. Tao. Velocity averaging, kinetic formulations, and regularizing effects in quasilinear PDEs. *Comm. Pure Appl. Math.*, 2006.
25. A. Vasseur. Strong traces for solutions of multidimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 160(3):181–193, 2001.
26. M. Westdickenberg. Some new velocity averaging results. *SIAM J. Math. Anal.*, 33(5):1007–1032 (electronic), 2002.
27. K. Zumbrun. Decay rates for nonconvex systems of conservation laws. *Comm. Pure Appl. Math.*, 46(3):353–386, 1993.