Lecture Notes of the Unione Matematica Italiana

Luigi Ambrosio
Gianluca Crippa
Camillo De Lellis
Felix Otto
Michael Westdickenberg

5

Transport Equations and Multi-D Hyperbolic Conservation Laws

Springer

UMI

Lecture Notes of
the Unione Matematica Italiana

5

Luigi Ambrosio • Gianluca Crippa
Camillo De Lellis • Felix Otto
Michael Westdickenberg

Transport Equations and Multi-D Hyperbolic Conservation Laws

Editors

Fabio Ancona
Stefano Bianchini
Rinaldo M. Colombo
Camillo De Lellis
Andrea Marson
Annamaria Montanari

Springer

UMI

Authors

Luigi Ambrosio
l.ambrosio@sns.it

Gianluca Crippa
g.crippa@sns.it

Camillo De Lellis
camillo.delellis@math.unizh.ch

Felix Otto
otto@iam.uni-bonn.de

Michael Westdickenberg
mwest@math.gatech.edu

Editors

Fabio Ancona
ancona@ciram.unibo.it

Stefano Bianchini
bianchin@sissa.it

Rinaldo M. Colombo
rinaldo@ing.unibs.it

Camillo De Lellis
camillo.delellis@math.unizh.ch

Andrea Marson
marson@math.unipd.it

Annamaria Montanari
montanar@dm.unibo.it

ISBN 978-3-540-76780-0 e-ISBN 978-3-540-76781-7

DOI 10.1007/978-3-540-76781-7

Lecture Notes of the Unione Matematica Italiana ISSN print edition: 1862-9113
ISSN electronic edition: 1862-9121

Library of Congress Control Number: 2007939405

Mathematics Subject Classification (2000): 35L45, 35L40, 35L65, 34A12, 49Q20, 28A75

Cover design: WMXDesign GmbH

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

This book collects the lecture notes of two courses and one mini-course held in a winter school in Bologna in January 2005. The aim of this school was to popularize techniques of geometric measure theory among researchers and PhD students in hyperbolic differential equations. Though initially developed in the context of the calculus of variations, many of these techniques have proved to be quite powerful for the treatment of some hyperbolic problems. Obviously, this point of view can be reversed: We hope that the topics of these notes will also capture the interest of some members of the elliptic community, willing to explore the links to the hyperbolic world.

The courses were attended by about 70 participants (including post-doctoral and senior scientists) from institutions in Italy, Europe, and North-America. This initiative was part of a series of schools (organized by some of the people involved in the school held in Bologna) that took place in Bressanone (Bolzano) in January 2004, and in SISSA (Trieste) in June 2006. Their scope was to present problems and techniques of some of the most promising and fascinating areas of research related to nonlinear hyperbolic problems that have received new and fundamental contributions in the recent years. In particular, the school held in Bressanone offered two courses that provided an introduction to the theory of control problems for hyperbolic-like PDEs (delivered by Roberto Triggiani), and to the study of transport equations with irregular coefficients (delivered by Francois Bouchut), while the conference hosted in Trieste was organized in two courses (delivered by Laure Saint-Raymond and Cedric Villani) and in a series of invited lectures devoted to the main recent advancements in the study of Boltzmann equation. Some of the material covered by the course of Triggiani can be found in [17, 18, 20], while the main contributions of the conference on Boltzmann will be collected in a forthcoming special issue of the journal DCDS, of title “Boltzmann equations and applications”.

The three contributions of the present volume gravitate all around the theory of BV functions, which play a fundamental role in the subject of hyperbolic conservation laws. However, so far in the hyperbolic community little attention has been paid to some typical problems which constitute an old topic in geometric measure theory: the structure and fine properties of BV functions in more than one space dimension.

The lecture notes of Luigi Ambrosio and Gianluca Crippa stem from the remarkable achievement of the first author, who recently succeeded in extending the so-called DiPerna–Lions theory for transport equations to the BV setting. More precisely, consider the Cauchy problem for a transport equation with variable coefficients

$$\begin{cases} \partial_t u(t,x) + b(t,x)\cdot\nabla u(t,x) \;=\; 0\,, \\ \\ u(0,x) \;=\; u_0(x)\,. \end{cases} \tag{1}$$

When b is Lipschitz, (1) can be explicitly solved via the method of characteristics: a solution u is indeed constant along the trajectories of the ODE

$$\begin{cases} \frac{d\Phi_x}{dt} \;=\; b(t,\Phi_x(t)) \\ \\ \Phi(0,x) \;=\; x\,. \end{cases} \tag{2}$$

Transport equations appear in a wealth of problems in mathematical physics, where usually the coefficient is coupled to the unknowns through some nonlinearities. This already motivates from a purely mathematical point of view the desire to develop a theory for (1) and (2) which allows for coefficients b in suitable function spaces. However, in many cases, the appearance of singularities is a well-established central fact: the development of such a theory is highly motivated from the applications themselves.

In the 1980s, DiPerna and Lions developed a theory for (1) and (2) when $b \in W^{1,p}$ (see [16]). The task of extending this theory to BV coefficients was a long-standing open question, until Luigi Ambrosio solved it in [2] with his Renormalization Theorem. Sobolev functions in $W^{1,p}$ cannot jump along a hypersurface: this type of singularity is instead typical for a BV function. Therefore, not surprisingly, Ambrosio's theorem has found immediate application to some problems in the theory of hyperbolic systems of conservation laws (see [3,5]).

Ambrosio's result, together with some questions recently raised by Alberto Bressan, has opened the way to a series of studies on transport equations and their links with systems of conservation laws (see [4,6–13]). The notes of Ambrosio and Crippa contain an efficient introduction to the DiPerna–Lions theory, a complete proof of Ambrosio's theorem and an overview of the further developments and open problems in the subject.

The first proof of Ambrosio's Renormalization Theorem relies on a deep result of Alberti, perhaps the deepest in the theory of BV functions (see [1]).

Consider a regular open set $\Omega \subset \mathbb{R}^2$ and a map $u : \mathbb{R}^2 \to \mathbb{R}^2$ which is regular in $\mathbb{R}^2 \setminus \partial\Omega$ but jumps along the interface $\partial\Omega$. The distributional derivative of u is then the sum of the classical derivative (which exists in $\mathbb{R}^2 \setminus \partial\Omega$) and a singular matrix-valued radon measure ν, supported on $\partial\Omega$. Let μ be the nonnegative measure on $\mathbb{R}^2$ defined by the property that $\mu(A)$ is the length of $\partial\Omega \cap A$. Moreover, denote by n the exterior unit normal to $\partial\Omega$ and by u^- and u^+, respectively, the interior and exterior traces of u on $\partial\Omega$. As a straightforward application of Gauss' theorem, we then conclude that the measure ν is given by $[(u^+ - u^-) \otimes n]\,\mu$.

Consider now the singular portion of the derivative of *any* BV vector-valued map. By elementary results in measure theory, we can always factorize it into a matrix-valued function M times a nonnegative measure μ. Alberti's Rank-One Theorem states that the values of M are always rank-one matrices. The depth of this theorem can be appreciated if one takes into account how complicated the singular measure μ can be.

Though the most recent proof of Ambrosio's Renormalization Theorem avoids Alberti's result, the Rank-One Theorem is a powerful tool to gain insight in subtle further questions (see for instance [6]). The notes of Camillo De Lellis is a short and self-contained introduction to Alberti's result, where the reader can find a complete proof.

As already mentioned above, the space of BV functions plays a central role in the theory of hyperbolic conservation laws. Consider for instance the Cauchy problem for a scalar conservation law

$$\begin{cases} \partial_t u + \mathrm{div}_x[f(u)] \;=\; 0\,, \\ \\ u(0,\cdot) \;=\; u_0\,. \end{cases} \tag{3}$$

It is a classical result of Kruzhkov that for bounded initial data u_0 there exists a unique entropy solution to (3). Furthermore, if u_0 is a function of bounded variation, this property is retained by the entropy solution.

Scalar conservation laws typically develop discontinuities. In particular jumps along hypersurfaces, the so-called *shock waves*, appear in finite time, even when starting with smooth initial data. These discontinuities travel at a speed which can be computed through the so-called Rankine–Hugoniot condition. Moreover, the admissibility conditions for distributional solutions (often called *entropy conditions*) are in essence devised to rule out certain "non-physical" shocks. When the entropy solution has BV regularity, the structure theory for BV functions allows us to identify a jump set, where all these assertions find a suitable (measure-theoretic) interpretation.

What happens if instead the initial data are merely bounded? Clearly, if f is a linear function, i.e. f'' vanishes, (3) is a transport equation with constant coefficients: extremely irregular initial data are then simply preserved. When we are far from this situation, loosely speaking when the range of f'' is "generic", f is called genuinely nonlinear. In one space dimension an extensively studied case of genuine nonlinearity is that of convex fluxes f. It is then an old result of Oleinik that, under this assumption, entropy solutions are BV functions for any bounded initial data. The assumption of genuine nonlinearity implies a regularization effect for the equation.

In more than one space dimension (or under milder assumptions on f) the BV regularization no longer holds true. However, Lions, Perthame, and Tadmore gave in [19] a kinetic formulation for scalar conservation laws and applied velocity averaging methods to show regularization in fractional Sobolev spaces. The notes of Gianluca Crippa, Felix Otto, and Michael Westdickenberg start with an introduction to entropy solutions, genuine nonlinearity, and kinetic formulations. They then discuss the regularization effects in terms of linear function spaces for a "generalized Burgers" flux, giving optimal results.

From a structural point of view, however, these estimates (even the optimal ones) are always too weak to recover the nice picture available for the BV framework, i.e. a solution which essentially has jump discontinuities behaving like shock waves. Guided by the analogy with the regularity theory developed in [14] for certain variational problems, De Lellis, Otto, and Westdickenberg in [15] showed that this picture is an outcome of an appropriate "regularity theory" for conservation laws. More precisely, the property of being an entropy solution to a scalar conservation law (with a genuinely nonlinear flux f) allows a fairly detailed analysis of the possible singularities. The information gained by this analysis is analogous to the fine properties of a generic BV function, even when the BV estimates fail. The notes of Crippa, Otto, and Westdickenberg give an overview of the ideas and techniques used to prove this result.

Many institutions have contributed funds to support the winter school of Bologna. We had a substantial financial support from the research project GNAMPA (*Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni*) – "Multidimensional problems and control problems for hyperbolic systems"; from CIRAM (*Research Center of Applied Mathematics*) and the *Fund for International Programs* of University of Bologna; and from *Seminario Matematico* and the Department of Mathematics of University of Brescia. We were also funded by the research project INDAM (*Istituto Nazionale di Alta Matematica "F. Severi"*) – "Nonlinear waves and applications to compressible and incompressible fluids". Our deepest thanks to all these institutions which make it possible the realization of this event and as a consequence of the present volume. As a final acknowledgement, we wish to warmly thank *Accademia delle Scienze di Bologna* and the Department of Mathematics of Bologna for their kind hospitality and for all the help and support they have provided throughout the school.

Bologna, Trieste,
Brescia, Zürich,
and Padova,
September 2007

Fabio Ancona
Stefano Bianchini
Rinaldo M. Colombo
Camillo De Lellis
Andrea Marson
Annamaria Montanari

References

1. Alberti, G. *Rank-one properties for derivatives of functions with bounded variations* Proc. Roy. Soc. Edinburgh Sect. A, **123** (1993), 239–274.
2. Ambrosio, L. *Transport equation and Cauchy problem for BV vector fields.* Invent. Math., **158** (2004), 227–260.

3. AMBROSIO, L.; BOUCHUT, F.; DE LELLIS, C. *Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions.* Comm. Partial Differential Equations, **29** (2004), 1635–1651.
4. AMBROSIO, L.; CRIPPA, G.; MANIGLIA, S. *Traces and fine properties of a BD class of vector fields and applications.* Ann. Fac. Sci. Toulouse Math. (6) **14** (2005), no. 4, 527–561.
5. AMBROSIO, L.; DE LELLIS, C. *Existence of solutions for a class of hyperbolic systems of conservation laws in several space dimensions.* Int. Math. Res. Not. **41** (2003), 2205–2220.
6. AMBROSIO L,; DE LELLIS, C.; MALÝ, J. *On the chain rule for the divergence of vector fields: applications, partial results, open problems.* To appear in *Perspectives in Nonlinear Partial Differential Equations: in honor of Haim Brezis* Preprint available at http://cvgmt.sns.it/papers/ambdel05/.
7. AMBROSIO L.; LECUMBERRY, M.; MANIGLIA, S. S. *Lipschitz regularity and approximate differentiability of the DiPerna–Lions flow.* Rend. Sem. Mat. Univ. Padova **114** (2005), 29–50.
8. BRESSAN, A. *An ill posed Cauchy problem for a hyperbolic system in two space dimensions.* Rend. Sem. Mat. Univ. Padova **110** (2003), 103–117.
9. BRESSAN, A. *A lemma and a conjecture on the cost of rearrangements.* Rend. Sem. Mat. Univ. Padova **110** (2003), 97–102.
10. BRESSAN, A. *Some remarks on multidimensional systems of conservation laws.* Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **15** (2004), 225–233.
11. CRIPPA, G.; DE LELLIS, C. *Oscillatory solutions to transport equations.* Indiana Univ. Math. J. **55** (2006), 1–13.
12. CRIPPA, G.; DE LELLIS, C. *Estimates and regularity results for the DiPerna-Lions flow.* To appear in J. Reine Angew. Math. Preprint available at http://cvgmt.sns.it/cgi/get.cgi/papers/cridel06/
13. DE LELLIS, C. *Blow-up of the BV norm in the multidimensional Keyfitz and Kranzer system.* Duke Math. J. **127** (2005), 313–339.
14. DE LELLIS, C.; OTTO, F. *Structure of entropy solutions to the eikonal equation.* J. Eur. Math. Soc. **5** (2003), 107–145.
15. DE LELLIS, C.; OTTO, F.; WESTDICKENBERG, M. *Structure of entropy solutions to scalar conservation laws.* Arch. Ration. Mech. Anal. **170**(2) (2003), 137–184.
16. DIPERNA, R.; LIONS, P. L. *Ordinary differential equations, transport theory and Sobolev spaces.* Invent. Math. **98** (1989), 511–517.
17. LASIECKA, I.; TRIGGIANI, R. *Global exact controllability of semilinear wave equations by a double compactness/uniqueness argument.* Discrete Contin. Dyn. Syst. (2005), suppl., 556–565.
18. LASIECKA, I.; TRIGGIANI, R. *Well-posedness and sharp uniform decay rates at the $L_2(\Omega)$-level of the Schrödinger equation with nonlinear boundary dissipation.* J. Evol. Equ. **6** (2006), no. 3, 485–537.
19. LIONS P.-L.; PERTHAME B.; TADMOR E. *A kinetic formulation of multidimensional scalar conservation laws and related questions.* J. AMS, **7** (1994) 169–191.
20. TRIGGIANI, R. *Global exact controllability on $H^1_{\Gamma_0}(\Omega) \times L_2(\Omega)$ of semilinear wave equations with Neumann $L_2(0,T;L_2(\Gamma_1))$-boundary control.* In: *Control theory of partial differential equations*, 273–336, *Lect. Notes Pure Appl. Math.*, 242, *Chapman & Hall/CRC, Boca Raton, FL*, 2005.

Contents

Part III

Authors

Luigi Ambrosio
Gianluca Crippa
Scuola Normale Superiore
Piazza dei Cavalieri 7
56126 Pisa, Italy
E-mail: l.ambrosio@sns.it
g.crippa@sns.it
URL: http://cvgmt.sns.it/people/ambrosio/

Camillo De Lellis
Institut für Mathematik
Universität Zrich
Winterthurerstrasse 190
CH-8057 Zürich, Switzerland
E-mail: camillo.delellis@math.unizh.ch
URL: http://www.math.unizh.ch/

Felix Otto
Institute for Applied Mathematics
University of Bonn
Wegelerstrae 10
53115 Bonn, Germany
E-mail: otto@iam.uni-bonn.de
URL: http://www-mathphys.iam.uni-bonn.de/∼otto/

Michael Westdickenberg
School of Mathematics
Georgia Institute of Technology
686 Cherry Street
Atlanta, GA 30332-0160, USA
E-mail: mwest@math.gatech.edu
URL: http://www.math.gatech.edu/∼mwest/

Editors

Fabio Ancona
Department of Mathematics
and CIRAM
University of Bologna
Via Saragozza, 8
40123 Bologna, Italy
E-mail: ancona@ciram.unibo.it
URL: http://www.ciram.unibo.it/
ancona/

Stefano Bianchini
SISSA-ISAS,
Via Beirut, 2-4
34014 Trieste, Italy
E-mail: bianchin@sissa.it
URL: http://people.sissa.it/
∼bianchin/

Rinaldo M. Colombo
Department of Mathematics
University of Brescia
Via Branze, 38
25123 Brescia, Italy
E-mail: rinaldo@ing.unibs.it
URL: http://dm.ing.unibs.it/rinaldo/

Camillo De Lellis
Institut für Mathematik
Universität Zürich
Winterthurerstrasse 190
8057 Zürich, Switzerland
E-mail: camillo.delellis@math.unizh.ch
URL: http://www.math.unizh.ch/

Andrea Marson
Department of Pure and Applied
Mathematics
Via Trieste, 63
35131 Padova, Italy
E-mail: marson@math.unipd.it
URL: http://www.math.unipd.it/
∼marson

Annamaria Montanari,
Department of Mathematics
University of Bologna
Piazza di Porta S. Donato, 5
40126 Bologna, Italy
E-mail: montanar@dm.unibo.it
URL: http://www.dm.unibo.it/
∼montanar/

Existence, Uniqueness, Stability and Differentiability Properties of the Flow Associated to Weakly Differentiable Vector Fields

Luigi Ambrosio and Gianluca Crippa

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
e-mail: l.ambrosio@sns.it, g.crippa@sns.it
URL: http://cvgmt.sns.it/people/ambrosio/

1 Introduction

In these notes we discuss some recent progress on the problem of the existence, uniqueness and stability of the flow associated to a weakly differentiable (Sobolev or *BV* regularity with respect to the spatial variables) time-dependent vector field $\boldsymbol{b}(t,x) = \boldsymbol{b}_t(x)$ in $\mathbb{R}^d$. Vector fields with this "low" regularity show up, for instance, in several PDEs describing the motion of fluids, and in the theory of conservation laws.

We are therefore interested to the well posedness of the system of ordinary differential equations

$$\text{(ODE)} \qquad \begin{cases} \dot\gamma(t) = \boldsymbol{b}_t(\gamma(t)) \\ \gamma(0) = x. \end{cases}$$

In some situations one might hope for a "generic" uniqueness of the solutions of (ODE), i.e. for "almost every" initial datum x. But, as a matter of fact, no such uniqueness theorem is presently known in the case when $\boldsymbol{b}(t,\cdot)$ has a Sobolev or BV regularity (this issue is discussed in Sect. 9).

An even weaker requirement is the research of a "canonical selection principle", i.e. a strategy to select for $\mathscr{L}^d$-almost every x a solution $\boldsymbol{X}(\cdot,x)$ in such a way that this selection is stable w.r.t. smooth approximations of $\boldsymbol{b}$.

In other words, we would like to know that, whenever we approximate $\boldsymbol{b}$ by smooth vector fields $\boldsymbol{b}^h$, the classical trajectories $\boldsymbol{X}^h$ associated to $\boldsymbol{b}^h$ satisfy

$$\lim_{h\to\infty} \boldsymbol{X}^h(\cdot,x) = \boldsymbol{X}(\cdot,x) \qquad \text{in } C([0,T];\mathbb{R}^d),\text{ for } \mathscr{L}^d\text{-a.e. } x.$$

The following simple example, borrowed from [8], provides an illustration of the kind of phenomena that can occur.

Example 1.1. Let us consider the autonomous ODE

$$\begin{cases} \dot\gamma(t) = \sqrt{|\gamma(t)|} \\ \gamma(0) = x_0. \end{cases}$$

Then, solutions of the ODE are not unique for $x_0 = -c^2 < 0$. Indeed, they reach the origin in time $2c$, where they can stay for an arbitrary time T, then continuing as $x(t) = \frac{1}{4}(t-T-2c)^2$. Let us consider for instance the Lipschitz approximation (that could easily be made smooth) of $b(\gamma) = \sqrt{|\gamma|}$ given by

$$b_\varepsilon(\gamma) := \begin{cases} \sqrt{|\gamma|} & \text{if } -\infty < \gamma \le -\varepsilon^2; \\ \varepsilon & \text{if } -\varepsilon^2 \le \gamma \le \lambda_\varepsilon - \varepsilon^2 \\ \sqrt{\gamma - \lambda_\varepsilon + 2\varepsilon^2} & \text{if } \lambda_\varepsilon - \varepsilon^2 \le \gamma < +\infty, \end{cases}$$

with $\lambda_\varepsilon - \varepsilon^2 > 0$. Then, solutions of the approximating ODEs starting from $-c^2$ reach the value $-\varepsilon^2$ in time $t_\varepsilon = 2(c-\varepsilon)$ and then they continue with constant speed ε until they reach $\lambda_\varepsilon - \varepsilon^2$, in time $T_\varepsilon = \lambda_\varepsilon/\varepsilon$. Then, they continue as $\lambda_\varepsilon - 2\varepsilon^2 + \frac{1}{4}(t - t_\varepsilon - T_\varepsilon)^2$.

Choosing $\lambda_\varepsilon = \varepsilon T$, with $T > 0$, by this approximation we select the solutions that don't move, when at the origin, exactly for a time T.

Other approximations, as for instance $b_\varepsilon(\gamma) = \sqrt{\varepsilon + |\gamma|}$, select the solutions that move immediately away from the singularity at $\gamma = 0$. Among all possibilities, this family of solutions $x(t,x_0)$ is singled out by the property that $x(t,\cdot)_\#\mathscr{L}^1$ is absolutely continuous with respect to $\mathscr{L}^1$, so no concentration of trajectories occurs at the origin. To see this fact, notice that we can integrate in time the identity

$$0 = x(t,\cdot)_\#\mathscr{L}^1(\{0\}) = \mathscr{L}^1(\{x_0 :\ x(t,x_0) = 0\})$$

and use Fubini's theorem to obtain

$$0 = \int \mathscr{L}^1(\{t:\ x(t,x_0) = 0\})\,dx_0.$$

Hence, for $\mathscr{L}^1$-a.e. x_0, $x(\cdot,x_0)$ does not stay at 0 for a strictly positive set of times.

The theme of existence, uniqueness and stability has been trated in detail in the lectures notes [7], and more recently in [8] (where also the applications to systems of conservation laws [11, 9] and to the semi-geostrophic equation ([53]) are described), so some overlap with the content of these notes is unavoidable. Because of this fact, we decided to put here more emphasis on even more recent results [73, 15, 17, 48], relative to the differentiability properties of $\boldsymbol{X}(t,x)$ with respect to x. This is not a casual choice, as the key idea of the paper [15] was found during the Bologna school.

2 The Continuity Equation

An important tool, in studying existence, uniqueness ad stability of (ODE), is the well-posedness of the Cauchy problem for the homogeneous conservative continuity equation

$$\text{(PDE)} \qquad \frac{d}{dt}\mu_t + D_x\cdot(\boldsymbol{b}\mu_t) = 0 \qquad (t,x)\in I\times\mathbb{R}^d$$

and for the transport equation

$$\frac{d}{dt}w_t + \boldsymbol{b}\cdot\nabla w_t = c_t.$$

We will see that there is a close link between (PDE) and (ODE), first investigated in a nonsmooth setting by DiPerna and Lions in [61].

Let us now make some basic technical remarks on the continuity equation and the transport equation:

Remark 2.1 (Regularity in space of $\boldsymbol{b}_t$ and μ_t). (1) Since the continuity equation (PDE) is in divergence form, it makes sense without *any* regularity requirement on $\boldsymbol{b}_t$ and/or μ_t, provided

$$\int_I\int_A |\boldsymbol{b}_t|\,d|\mu_t|\,dt < +\infty \qquad \forall A\subset\subset\mathbb{R}^d. \tag{1}$$

However, when we consider possibly singular measures μ_t, we must take care of the fact that the product $\boldsymbol{b}_t\mu_t$ is sensitive to modifications of $\boldsymbol{b}_t$ in $\mathscr{L}^d$-negligible sets. In the Sobolev or *BV* case we will consider only measures $\mu_t = w_t\mathscr{L}^d$, so everything is well stated.

(2) On the other hand, due to the fact that the distribution $\boldsymbol{b}_t \cdot \nabla w$ is defined by

$$\langle \boldsymbol{b}_t \cdot \nabla w, \varphi \rangle := -\int_I \int w \boldsymbol{b}_t \cdot \nabla \varphi \, dx dt - \int_I \langle D_x \cdot \boldsymbol{b}_t, w_t \varphi_t \rangle \, dt \qquad \forall \varphi \in C_c^\infty(I \times \mathbb{R}^d)$$

(a definition consistent with the case when w_t is smooth) the transport equation makes sense *only* if we assume that $D_x \cdot \boldsymbol{b}_t = \operatorname{div} \boldsymbol{b}_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$. See also [28, 29] for recent results on the transport equation when $\boldsymbol{b}$ satisfies a one-sided Lipschitz condition.

Next, we consider the problem of the time continuity of $t \mapsto \mu_t$ and $t \mapsto w_t$.

Remark 2.2 (Regularity in time of μ_t). For any test function $\varphi \in C_c^\infty(\mathbb{R}^d)$, condition (1) gives

$$\frac{d}{dt} \int_{\mathbb{R}^d} \varphi \, d\mu_t = \int_{\mathbb{R}^d} \boldsymbol{b}_t \cdot \nabla \varphi \, d\mu_t \in L^1(I)$$

and therefore the map $t \mapsto \langle \mu_t, \varphi \rangle$, for given φ, has a unique uniformly continuous representative in I. By a simple density argument we can find a unique representative $\tilde{\mu}_t$ independent of φ, such that $t \mapsto \langle \tilde{\mu}_t, \varphi \rangle$ is uniformly continuous in I for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. We will always work with this representative, so that μ_t will be well defined *for all t* and even at the endpoints of I. An analogous remark applies for solutions of the transport equation.

There are some other important links between the two equations:

(1) The transport equation reduces to the continuity equation in the case when $c_t = -w_t \operatorname{div} \boldsymbol{b}_t$.

(2) Formally, one can establish a duality between the two equations via the (formal) identity

$$\begin{aligned} \frac{d}{dt} \int w_t \, d\mu_t &= \int \frac{d}{dt} w_t \, d\mu_t + \int \frac{d}{dt} \mu_t w_t \\ &= \int (-\boldsymbol{b}_t \cdot \nabla w_t + c_t) \, d\mu_t + \int \boldsymbol{b}_t \cdot \nabla w_t \, d\mu_t = \int c_t \, d\mu_t . \end{aligned}$$

This duality method is a classical tool to prove uniqueness in a sufficiently smooth setting (but see also [28, 29]).

(3) Finally, if we denote by $\boldsymbol{Y}(t,s,x)$ the solution of the ODE at time t, starting from x at the initial time s, i.e.

$$\frac{d}{dt} \boldsymbol{Y}(t,s,x) = \boldsymbol{b}_t(\boldsymbol{Y}(t,s,x)), \qquad \boldsymbol{Y}(s,s,x) = x,$$

then $\boldsymbol{Y}(t,\cdot,\cdot)$ are themselves solutions of the transport equation: to see this, it suffices to differentiate the semigroup identity

$$\boldsymbol{Y}(t,s,\boldsymbol{Y}(s,l,x)) = \boldsymbol{Y}(t,l,x)$$

w.r.t. s to obtain, after the change of variables $y = \boldsymbol{Y}(s,l,x)$, the equation

$$\frac{d}{ds}\boldsymbol{Y}(t,s,y)+\boldsymbol{b}_s(y)\cdot\nabla\boldsymbol{Y}(t,s,y)=0.$$

This property is used in a essential way in [61] to characterize the flow $\boldsymbol{Y}$ and to prove its stability properties. The approach developed here, based on [6], is based on a careful analysis of the measures transported by the flow, and ultimately on the homogeneous continuity equation only.

3 The Continuity Equation Within the Cauchy–Lipschitz Framework

In this section we recall the classical representation formulas for solutions of the continuity or transport equation in the case when

$$\boldsymbol{b}\in L^1\left([0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)\right).$$

Under this assumption it is well known that solutions $\boldsymbol{X}(t,\cdot)$ of the ODE are unique and stable. A quantitative information can be obtained by differentiation:

$$\begin{aligned}\frac{d}{dt}|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|^2 &= 2\langle\boldsymbol{b}_t(\boldsymbol{X}(t,x))-\boldsymbol{b}_t(\boldsymbol{X}(t,y)),\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)\rangle\\ &\le 2\mathrm{Lip}\,(\boldsymbol{b}_t)|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|^2\end{aligned}$$

(here $\mathrm{Lip}\,(f)$ denotes the least Lipschitz constant of f), so that Gronwall lemma immediately gives

$$\mathrm{Lip}\,(\boldsymbol{X}(t,\cdot))\le\exp\left(\int_0^t\mathrm{Lip}\,(\boldsymbol{b}_s)\,ds\right). \tag{2}$$

Turning to the continuity equation, uniqueness of measure-valued solutions can be proved by the duality method. Or, following the techniques developed in these lectures, it can be proved in a more general setting for positive measure-valued solutions (via the superposition principle) and for signed solutions $\mu_t=w_t\mathscr{L}^d$ (via the theory of renormalized solutions). So in this section we focus only on the existence and the representation issues.

The representation formula is indeed very simple:

Proposition 3.1. *For any initial datum $\bar\mu$ the solution of the continuity equation is given by*

$$\mu_t:=\boldsymbol{X}(t,\cdot)_\#\bar\mu,\quad \textit{i.e.}\quad \int_{\mathbb{R}^d}\varphi\,d\mu_t=\int_{\mathbb{R}^d}\varphi(\boldsymbol{X}(t,x))\,d\bar\mu(x)\quad\forall\varphi\in C_b(\mathbb{R}^d). \tag{3}$$

Proof. Notice first that we need only to check the distributional identity $\frac{d}{dt}\mu_t+D_x\cdot(\boldsymbol{b}_t\mu_t)=0$ on test functions of the form $\psi(t)\varphi(x)$, so that

$$\int_{\mathbb{R}} \psi'(t)\langle \mu_t, \varphi\rangle\, dt + \int_{\mathbb{R}} \psi(t) \int_{\mathbb{R}^d} \langle \boldsymbol{b}_t, \nabla\varphi\rangle\, d\mu_t\, dt = 0.$$

This means that we have to check that $t \mapsto \langle \mu_t, \varphi\rangle$ belongs to $W^{1,1}(0,T)$ for any $\varphi \in C_c^\infty(\mathbb{R}^d)$ and that its distributional derivative is $\int_{\mathbb{R}^d} \langle \boldsymbol{b}_t, \nabla\varphi\rangle\, d\mu_t$.

We show first that this map is absolutely continuous, and in particular $W^{1,1}(0,T)$; then one needs only to compute the pointwise derivative. For every choice of finitely many, say n, pairwise disjoint intervals $(a_i, b_i) \subset [0,T]$ we have

$$\begin{aligned}
\sum_{i=1}^n |\varphi(\boldsymbol{X}(b_i,x)) - \varphi(\boldsymbol{X}(a_i,x))| &\le \|\nabla\varphi\|_\infty \int_{\cup_i (a_i,b_i)} |\dot{\boldsymbol{X}}(t,x)|\, dt \\
&\le \|\nabla\varphi\|_\infty \int_{\cup_i (a_i,b_i)} \sup |\boldsymbol{b}_t|\, dt
\end{aligned}$$

and therefore an integration with respect to $\bar\mu$ gives

$$\sum_{i=1}^n |\langle \mu_{b_i} - \mu_{a_i}, \varphi\rangle| \le \bar\mu(\mathbb{R}^d)\|\nabla\varphi\|_\infty \int_{\cup_i (a_i,b_i)} \sup |\boldsymbol{b}_t|\, dt.$$

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i (b_i - a_i)$ is small. This proves the absolute continuity. For any x the identity $\dot{\boldsymbol{X}}(t,x) = \boldsymbol{b}_t(\boldsymbol{X}(t,x))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0,T]$ the previous identity holds for $\bar\mu$-a.e. x, and therefore

$$\begin{aligned}
\frac{d}{dt}\langle \mu_t, \varphi\rangle &= \frac{d}{dt} \int_{\mathbb{R}^d} \varphi(\boldsymbol{X}(t,x))\, d\bar\mu(x) \\
&= \int_{\mathbb{R}^d} \langle \nabla\varphi(\boldsymbol{X}(t,x)), \boldsymbol{b}_t(\boldsymbol{X}(t,x))\rangle\, d\bar\mu(x) \\
&= \langle \boldsymbol{b}_t \mu_t, \nabla\varphi\rangle
\end{aligned}$$

for $\mathscr{L}^1$-a.e. $t \in [0,T]$. □

In the case when $\bar\mu = \rho\mathscr{L}^d$ we can say something more, proving that the measures $\mu_t = \boldsymbol{X}(t,\cdot)_\#\bar\mu$ are absolutely continuous w.r.t. $\mathscr{L}^d$ and computing *explicitly* their density. Let us start by recalling the classical *area formula*: if $f : \mathbb{R}^d \to \mathbb{R}^d$ is a (locally) Lipschitz map, then

$$\int_A g|Jf|\, dx = \int_{\mathbb{R}^d} \sum_{x \in A \cap f^{-1}(y)} g(x)\, dy$$

for any Borel set $A \subset \mathbb{R}^d$ and any integrable function $g : A \to \mathbb{R}$, where $Jf = \det \nabla f$ (recall that, by Rademacher theorem, Lipschitz functions are differentiable $\mathscr{L}^d$-a.e.). Assuming in addition that f is one to one and onto and that $|Jf| > 0$ $\mathscr{L}^d$-a.e. on A we can set $A = f^{-1}(B)$ and $g = \rho/|Jf|$ to obtain

$$\int_{f^{-1}(B)} \rho\, dx = \int_B \frac{\rho}{|Jf|} \circ f^{-1}\, dy.$$

In other words, we have got a formula for the push-forward:

$$f_\#(\rho \mathscr{L}^d) = \frac{\rho}{|Jf|} \circ f^{-1} \mathscr{L}^d. \tag{4}$$

In our case $f(x) = \boldsymbol{X}(t,x)$ is surely one to one, onto and Lipschitz. It remains to show that $|J\boldsymbol{X}(t,\cdot)|$ does not vanish: in fact, one can show that $J\boldsymbol{X} > 0$ and

$$\exp\left[-\int_0^t \|[\operatorname{div}\boldsymbol{b}_s]^-\|_\infty\, ds\right] \le J\boldsymbol{X}(t,x) \le \exp\left[\int_0^t \|[\operatorname{div}\boldsymbol{b}_s]^+\|_\infty\, ds\right] \tag{5}$$

for $\mathscr{L}^d$-a.e. x, thanks to the following fact, whose proof is left as an exercise.

Exercise 3.2. If $\boldsymbol{b}$ is smooth, we have

$$\frac{d}{dt} J\boldsymbol{X}(t,x) = \operatorname{div}\boldsymbol{b}_t(\boldsymbol{X}(t,x)) J\boldsymbol{X}(t,x).$$

Hint: use the ODE $\frac{d}{dt}\nabla\boldsymbol{X} = \nabla\boldsymbol{b}_t(\boldsymbol{X})\nabla\boldsymbol{X}$.

The previous exercise gives that, in the smooth case, $J\boldsymbol{X}(\cdot,x)$ solves a linear ODE with the initial condition $J\boldsymbol{X}(0,x) = 1$, whence the estimates on $J\boldsymbol{X}$ follow. In the general case the upper estimate on $J\boldsymbol{X}$ still holds by a smoothing argument, thanks to the lower semicontinuity of

$$\Phi(v) := \begin{cases} \|Jv\|_\infty & \text{if } Jv \ge 0 \ \mathscr{L}^d\text{-a.e.} \\ \\ +\infty & \text{otherwise} \end{cases}$$

with respect to the w^*-topology of $W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)$. This is indeed the supremum of the family of $\Phi_p^{1/p}$, where Φ_p are the *polyconvex* (and therefore lower semicontinuous) functionals

$$\Phi_p(v) := \int_{B_p} |\chi(Jv)|^p\, dx.$$

Here $\chi(t)$, equal to ∞ on $(-\infty,0)$ and equal to t on $[0,+\infty)$, is l.s.c. and convex. The lower estimate can be obtained by applying the upper one in a time reversed situation.

Now we turn to the representation of solutions of the transport equation:

Proposition 3.3. *If* $w \in L^1_{\mathrm{loc}}\left([0,T]\times\mathbb{R}^d\right)$ *solves*

$$\frac{d}{dt} w_t + \boldsymbol{b}\cdot\nabla w_t = c \in L^1_{\mathrm{loc}}\left([0,T]\times\mathbb{R}^d\right)$$

then, for $\mathscr{L}^d$-a.e. x, we have

$$w_t(\boldsymbol{X}(t,x)) = w_0(x) + \int_0^t c_s(\boldsymbol{X}(s,x))\,ds \qquad \forall t \in [0,T].$$

The (formal) proof is based on the simple observation that

$$\begin{aligned}\frac{d}{dt} w_t \circ \boldsymbol{X}(t,x) &= \frac{d}{dt} w_t(\boldsymbol{X}(t,x)) + \frac{d}{dt}\boldsymbol{X}(t,x)\cdot \nabla w_t(\boldsymbol{X}(t,x)) \\ &= \frac{d}{dt} w_t(\boldsymbol{X}(t,x)) + \boldsymbol{b}_t(\boldsymbol{X}(t,x))\cdot \nabla w_t(\boldsymbol{X}(t,x)) \\ &= c_t(\boldsymbol{X}(t,x)).\end{aligned}$$

In particular, as $\boldsymbol{X}(t,x) = \boldsymbol{Y}(t,0,x) = [\boldsymbol{Y}(0,t,\cdot)]^{-1}(x)$, we get

$$w_t(y) = w_0(\boldsymbol{Y}(0,t,y)) + \int_0^t c_s(\boldsymbol{Y}(s,t,y))\,ds.$$

We conclude this presentation of the classical theory pointing out two simple local variants of the assumption $\boldsymbol{b} \in L^1\left([0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)\right)$ made throughout this section.

Remark 3.4 (First local variant). The theory outlined above still works under the assumptions

$$\boldsymbol{b} \in L^1\left([0,T];W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right), \qquad \frac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right).$$

Indeed, due to the growth condition on $\boldsymbol{b}$, we still have pointwise uniqueness of the ODE and a uniform local control on the growth of $|\boldsymbol{X}(t,x)|$, therefore we need only to consider a *local* Lipschitz condition w.r.t. x, integrable w.r.t. t.

The next variant will be used in the proof of the superposition principle.

Remark 3.5 (Second local variant). Still keeping the $L^1(W^{1,\infty}_{\mathrm{loc}})$ assumption, and assuming $\mu_t \geq 0$, the second growth condition on $|\boldsymbol{b}|$ can be replaced by a global, but more intrinsic, condition:

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_t|}{1+|x|}\,d\mu_t\,dt < +\infty. \tag{6}$$

Under this assumption one can show that for $\bar{\mu}$-a.e. x the *maximal* solution $\boldsymbol{X}(\cdot,x)$ of the ODE starting from x is defined up to $t = T$ and still the representation $\mu_t = \boldsymbol{X}(t,\cdot)_\#\bar{\mu}$ holds for $t \in [0,T]$.

4 (ODE) Uniqueness Vs. (PDE) Uniqueness

In this section we illustrate some quite general principles, whose application may depend on specific assumptions on $\boldsymbol{b}$, relating the uniqueness for the ODE to the uniqueness for the PDE. The viewpoint adopted in this section is very close in spirit to Young's theory [95] of generalized surfaces and controls (a theory with remarkable applications also to nonlinear PDEs [60, 88] and to Calculus of Variations [19]) and has also some connection with Brenier's weak solutions of incompressible Euler equations [30], with Kantorovich's viewpoint in the theory of optimal transportation [63, 85] and with Mather's theory [80, 81, 20]: in order to study existence, uniqueness and stability with respect to perturbations of the data of solutions to the ODE, we consider suitable measures in the space of continuous maps, allowing for superposition of trajectories. Then, in some special situations we are able to show that this superposition actually does not occur, but still this "probabilistic" interpretation is very useful to understand the underlying techniques and to give an intrinsic characterization of the flow.

The first very general criterion is the following.

Theorem 4.1. *Let $A \subset \mathbb{R}^d$ be a Borel set. The following two properties are equivalent:*

(a) Solutions of the ODE are unique for any $x \in A$.
(b) Nonnegative measure-valued solutions of the PDE are unique for any $\bar{\mu}$ concentrated in A, i.e. such that $\bar{\mu}(\mathbb{R}^d \setminus A) = 0$.

Proof. It is clear that (b) implies (a), just choosing $\bar{\mu} = \delta_x$ and noticing that two different solutions $\boldsymbol{X}(t), \tilde{\boldsymbol{X}}(t)$ of the ODE induce two different solutions of the PDE, namely $\delta_{\boldsymbol{X}(t)}$ and $\delta_{\tilde{\boldsymbol{X}}(t)}$.

The converse implication is less obvious and requires the superposition principle that we are going to describe below, and that provides the representation

$$\int_{\mathbb{R}^d} \varphi \, d\mu_t = \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(\gamma(t)) \, d\eta_x(\gamma) \right) d\mu_0(x) \qquad \forall \varphi \in C_b(\mathbb{R}^d),$$

with η_x probability measures concentrated on the absolutely continuous integral solutions of the ODE starting from x. Therefore, when these are unique, the measures η_x are unique (and are Dirac masses), so that the solutions of the PDE are unique. □

We will use the shorter notation Γ_T for the space $C\left([0,T];\mathbb{R}^d\right)$ and denote by $e_t : \Gamma_T \to \mathbb{R}^d$ the evaluation maps $\gamma \mapsto \gamma(t)$, $t \in [0,T]$.

Definition 4.2 (Superposition solutions). Let $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ be a measure concentrated on the set of pairs (x, γ) such that γ is an absolutely continuous integral solution of the ODE with $\gamma(0) = x$. We define

$$\langle \mu_t^\eta, \varphi \rangle := \int_{\mathbb{R}^d \times \Gamma_T} \varphi(e_t(\gamma)) \, d\eta(x, \gamma) \qquad \forall \varphi \in C_b(\mathbb{R}^d).$$

By a standard approximation argument the identity defining μ_t^η holds for any Borel function φ such that $\gamma \mapsto \varphi(e_t(\gamma))$ is η-integrable (or equivalently for any μ_t^η-integrable function φ).

Under the (local) integrability condition

$$\int_0^T \int_{\mathbb{R}^d \times \Gamma_T} \chi_{B_R}(e_t)|\boldsymbol{b}_t(e_t)|\, d\eta\, dt < +\infty \qquad \forall R > 0 \tag{7}$$

it is not hard to see that μ_t^η solves the PDE with the initial condition $\bar{\mu} := (\pi_{\mathbb{R}^d})_\# \eta$: indeed, let us check first that $t \mapsto \langle \mu_t^\eta, \varphi \rangle$ is absolutely continuous for any $\varphi \in C_c^\infty(\mathbb{R}^d)$. For every choice of finitely many pairwise disjoint intervals $(a_i, b_i) \subset [0,T]$ we have

$$\sum_{i=1}^n |\varphi(\gamma(b_i)) - \varphi(\gamma(a_i))| \le \mathrm{Lip}\,(\varphi) \int_{\cup_i (a_i,b_i)} \chi_{B_R}(|e_t(\gamma)|)\boldsymbol{b}_t(e_t(\gamma))|\, dt$$

for η-a.e. (x,γ), with R such that $\operatorname{supp}\varphi \subset \overline{B}_R$. Therefore an integration with respect to η gives

$$\sum_{i=1}^n |\langle \mu_{b_i}^\eta, \varphi \rangle - \langle \mu_{a_i}^\eta, \varphi \rangle| \le \mathrm{Lip}\,(\varphi) \int_{\cup_i (a_i,b_i)} \int_{\mathbb{R}^d \times \Gamma_T} \chi_{B_R}(e_t)|\boldsymbol{b}_t(e_t)|\, d\eta\, dt.$$

The absolute continuity of the integral shows that the right hand side can be made small when $\sum_i (b_i - a_i)$ is small. This proves the absolute continuity.

It remains to evaluate the time derivative of $t \mapsto \langle \mu_t^\eta, \varphi \rangle$: we know that for η-a.e. (x,γ) the identity $\dot\gamma(t) = \boldsymbol{b}_t(\gamma(t))$ is fulfilled for $\mathscr{L}^1$-a.e. $t \in [0,T]$. Then, by Fubini's theorem, we know also that for $\mathscr{L}^1$-a.e. $t \in [0,T]$ the previous identity holds for η-a.e. (x,γ), and therefore

$$\begin{aligned}
\frac{d}{dt}\langle \mu_t^\eta, \varphi \rangle &= \frac{d}{dt} \int_{\mathbb{R}^d \times \Gamma_T} \varphi(e_t(\gamma))\, d\eta \\
&= \int_{\mathbb{R}^d \times \Gamma_T} \langle \nabla \varphi(e_t(\gamma)), \boldsymbol{b}_t(e_t(\gamma)) \rangle\, d\eta = \langle \boldsymbol{b}_t \mu_t, \nabla \varphi \rangle \quad \mathscr{L}^1\text{-a.e. in } [0,T].
\end{aligned}$$

Remark 4.3. Actually the formula defining μ_t^η does not contain x, and so it involves only the projection of η on Γ_T. Therefore one could also consider measures σ in Γ_T, concentrated on the set of solutions of the ODE (for an arbitrary initial point x). These two viewpoints are basically equivalent: given η one can build σ just by projection on Γ_T, and given σ one can consider the conditional probability measures η_x concentrated on the solutions of the ODE starting from x induced by the random variable $\gamma \mapsto \gamma(0)$ in Γ_T, the law $\bar{\mu}$ (i.e. the push forward) of the same random variable and recover η as follows:

$$\int_{\mathbb{R}^d \times \Gamma_T} \varphi(x,\gamma)\, d\eta(x,\gamma) := \int_{\mathbb{R}^d} \left(\int_{\Gamma_T} \varphi(x,\gamma)\, d\eta_x(\gamma) \right) d\bar{\mu}(x). \tag{8}$$

Our viewpoint has been chosen just for technical convenience, to avoid the use, wherever this is possible, of the conditional probability theorem.

By restricting η to suitable subsets of $\mathbb{R}^d \times \Gamma_T$, several manipulations with superposition solutions of the continuity equation are possible and useful, and these are not immediate to see just at the level of general solutions of the continuity equation. This is why the following result is interesting.

Theorem 4.4 (Superposition principle). *Let $\mu_t \in \mathscr{M}_+(\mathbb{R}^d)$ be a solution of (PDE) and assume that*

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}|_t(x)}{1+|x|}\,d\mu_t\,dt < +\infty.$$

Then μ_t is a superposition solution, i.e. there exists $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ such that $\mu_t = \mu_t^\eta$ for any $t \in [0,T]$.

In the proof we use the *narrow* convergence of positive measures, i.e. the convergence with respect to the duality with continuous and bounded functions, and the *easy* implication in Prokhorov compactness theorem: any tight and bounded family $\mathscr{F}$ in $\mathscr{M}_+(X)$ is (sequentially) relatively compact w.r.t. the narrow convergence. Remember that tightness means:

$$\text{for any } \varepsilon > 0 \text{ there exists } K \subset X \text{ compact s.t. } \mu(X \setminus K) < \varepsilon \ \forall \mu \in \mathscr{F}.$$

A necessary and sufficient condition for tightness is the existence of a *coercive* functional $\Psi : X \to [0,\infty]$ such that $\int \Psi\,d\mu \le 1$ for any $\mu \in \mathscr{F}$.

Proof. **Step 1.** (Smoothing) The smoothing argument which follows has been inspired by [65]. We mollify μ_t w.r.t. the space variable with a kernel ρ having finite first moment M and support equal to the whole of $\mathbb{R}^d$ (a Gaussian, for instance), obtaining smooth and strictly positive functions μ_t^ε. We also choose a function $\psi : \mathbb{R}^d \to [0,+\infty)$ such that $\psi(x) \to +\infty$ as $|x| \to +\infty$ and

$$\int_{\mathbb{R}^d} \psi(x)\mu_0 * \rho_\varepsilon(x)\,dx \le 1 \qquad \forall \varepsilon \in (0,1)$$

and a convex nondecreasing function $\Theta : \mathbb{R}^+ \to \mathbb{R}$ having a more than linear growth at infinity such that

$$\int_0^T \int_{\mathbb{R}^d} \frac{\Theta(|\boldsymbol{b}_t|(x))}{1+|x|}\,d\mu_t dt < +\infty$$

(the existence of Θ is ensured by the Dunford–Pettis Theorem). Defining

$$\mu_t^\varepsilon := \mu_t * \rho_\varepsilon, \qquad \boldsymbol{b}_t^\varepsilon := \frac{(\boldsymbol{b}_t \mu_t) * \rho_\varepsilon}{\mu_t^\varepsilon},$$

it is immediate that

$$\frac{d}{dt}\mu_t^\varepsilon + D_x \cdot (\boldsymbol{b}_t^\varepsilon \mu_t^\varepsilon) = \frac{d}{dt}\mu_t * \rho_\varepsilon + D_x \cdot (\boldsymbol{b}_t \mu_t) * \rho_\varepsilon = 0$$

and that $\boldsymbol{b}^\varepsilon \in L^1\left([0,T]; W^{1,\infty}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$. Therefore Remark 3.5 can be applied and the representation $\mu_t^\varepsilon = \boldsymbol{X}^\varepsilon(t,\cdot)_\#\mu_0^\varepsilon$ still holds. Then, we define

$$\eta^\varepsilon := (x, \boldsymbol{X}^\varepsilon(\cdot,x))_\#\,\mu_0^\varepsilon,$$

so that

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi\,d\mu_t^{\eta_\varepsilon} &= \int_{\mathbb{R}^d\times\Gamma_T}\varphi(\gamma(t))\,d\eta^\varepsilon \\ &= \int_{\mathbb{R}^d}\varphi(\boldsymbol{X}^\varepsilon(t,x))\,d\mu_0^\varepsilon(x) = \int_{\mathbb{R}^d}\varphi\,d\mu_t^\varepsilon.\end{aligned} \tag{9}$$

Step 2. (Tightness) We will be using the inequality

$$((1+|x|)c)*\rho_\varepsilon \le (1+|x|)c*\rho_\varepsilon + \varepsilon c*\tilde\rho_\varepsilon \tag{10}$$

for c nonnegative measure and $\tilde\rho(y) = |y|\rho(y)$, and

$$\Theta(|\boldsymbol{b}_t^\varepsilon(x)|)\mu_t^\varepsilon(x) \le (\Theta(|\boldsymbol{b}_t|)\mu_t)*\rho_\varepsilon(x). \tag{11}$$

The proof of the first one is elementary, while the proof of the second one follows by applying Jensen's inequality with the convex l.s.c. function $(z,t)\mapsto\Theta(|z|/t)t$ (set equal to $+\infty$ if $t<0$, or $t=0$ and $z\neq 0$, and equal to 0 if $z=t=0$) and with the measure $\rho_\varepsilon(x-\cdot)\mathscr{L}^d$.

Let us introduce the functional

$$\Psi(x,\gamma) := \psi(x) + \int_0^T \frac{\Theta(|\dot\gamma|)}{1+|\gamma|}\,dt,$$

set equal to $+\infty$ on $\Gamma_T\setminus AC([0,T];\mathbb{R}^d)$.

Using Ascoli–Arzelá theorem, it is not hard to show that Ψ is coercive (it suffices to show that $\max|\gamma|$ is bounded on the sublevels $\{\Psi\le t\}$). Since

$$\begin{aligned}\int_{\mathbb{R}^d\times\Gamma_T}\int_0^T\frac{\Theta(|\dot\gamma|)}{1+|\gamma|}\,dt\,d\eta^\varepsilon(x,\gamma) &= \int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t^\varepsilon|)}{1+|x|}\,d\mu_t^\varepsilon\,dt \\ \overset{(10),(11)}{\le} (1+\varepsilon M)\int_0^T\int_{\mathbb{R}^d}\frac{\Theta(|\boldsymbol{b}_t|(x))}{1+|x|}\,d\mu_t dt\end{aligned}$$

and

$$\int_{\mathbb{R}^d\times\Gamma_T}\psi(x)\,d\eta^\varepsilon(x,\gamma) = \int_{\mathbb{R}^d}\psi(x)\,d\mu_0^\varepsilon \le 1$$

we obtain that $\int\Psi\,d\eta^\varepsilon$ is uniformly bounded for $\varepsilon\in(0,1)$, and therefore Prokhorov compactness theorem tells us that the family η^ε is narrowly sequentially relatively compact as $\varepsilon\downarrow 0$. If η is any limit point we can pass to the limit in (9) to obtain that $\mu_t = \mu_t^\eta$.

Step 3. (η is concentrated on solutions of the ODE) It suffices to show that

$$\int_{\mathbb{R}^d\times\Gamma_T} \frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{b}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\eta=0 \tag{12}$$

for any $t\in[0,T]$. The technical difficulty is that this test function, due to the lack of regularity of $\boldsymbol{b}$, is not continuous. To this aim, we prove first that

$$\int_{\mathbb{R}^d\times\Gamma_T} \frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{c}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\eta \le \int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s ds \tag{13}$$

for any continuous function $\boldsymbol{c}$ with compact support. Then, choosing a sequence $(\boldsymbol{c}^n)$ converging to $\boldsymbol{b}$ in $L^1(\nu;\mathbb{R}^d)$, with

$$\int \varphi(s,x)\,d\nu(s,x) := \int_0^T\int_{\mathbb{R}^d}\frac{\varphi(s,x)}{1+|x|}\,d\mu_s(x)\,ds$$

and noticing that

$$\int_{\mathbb{R}^d\times\Gamma_T}\int_0^T \frac{|\boldsymbol{b}_s(\gamma(s))-\boldsymbol{c}^n_s(\gamma(s))|}{1+|\gamma(s)|}\,ds d\eta = \int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}^n_s|}{1+|x|}\,d\mu_s ds\to 0,$$

we can pass to the limit in (13) with $\boldsymbol{c}=\boldsymbol{c}^n$ to obtain (12).

It remains to show (13). This is a limiting argument based on the fact that (12) holds for $\boldsymbol{b}^\varepsilon$, η^ε:

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\Gamma_T} \frac{\left|\gamma(t)-x-\int_0^t \boldsymbol{c}_s(\gamma(s))\,ds\right|}{1+\max\limits_{[0,T]}|\gamma|}\,d\eta^\varepsilon\\
&=\int_{\mathbb{R}^d}\frac{\left|\boldsymbol{X}^\varepsilon(t,x)-x-\int_0^t \boldsymbol{c}_s(\boldsymbol{X}^\varepsilon(s,x))\,ds\right|}{1+\max\limits_{[0,T]}|\boldsymbol{X}^\varepsilon(\cdot,x)|}\,d\mu_0^\varepsilon(x)\\
&=\int_{\mathbb{R}^d}\frac{\left|\int_0^t \boldsymbol{b}^\varepsilon_s(\boldsymbol{X}^\varepsilon(s,x))-\boldsymbol{c}_s(\boldsymbol{X}^\varepsilon(s,x))\,ds\right|}{1+\max\limits_{[0,T]}|\boldsymbol{X}^\varepsilon(\cdot,x)|}\,d\mu_0^\varepsilon(x)\le\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}^\varepsilon_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu^\varepsilon_s ds\\
&\le\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}^\varepsilon_s-\boldsymbol{c}^\varepsilon_s|}{1+|x|}\,d\mu^\varepsilon_s ds+\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{c}^\varepsilon_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu^\varepsilon_s ds\\
&\le\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu_s ds+\int_0^t\int_{\mathbb{R}^d}\frac{|\boldsymbol{c}^\varepsilon_s-\boldsymbol{c}_s|}{1+|x|}\,d\mu^\varepsilon_s ds.
\end{aligned}$$

In the last inequalities we added and subtracted $\boldsymbol{c}^\varepsilon_t := (\boldsymbol{c}_t\mu_t)*\rho_\varepsilon/\mu^\varepsilon_t$. Since $\boldsymbol{c}^\varepsilon_t\to\boldsymbol{c}_t$ uniformly as $\varepsilon\downarrow 0$ thanks to the uniform continuity of $\boldsymbol{c}$, passing to the limit in the chain of inequalities above we obtain (13). □

The applicability of Theorem 4.1 is strongly limited by the fact that, on one hand, *pointwise* uniqueness properties for the ODE are known only in very special situations, for instance when there is a Lipschitz or a one-sided Lipschitz (or log-Lipschitz, Osgood...) condition on $\boldsymbol{b}$. On the other hand, also uniqueness for general measure-valued solutions is known only in special situations. It turns out that in many cases uniqueness of the PDE can only be proved in smaller classes $\mathscr{L}$ of solutions, and it is natural to think that this should reflect into a weaker uniqueness condition at the level of the ODE.

We will see indeed that there is uniqueness in the "selection sense". In order to illustrate this concept, in the following we consider a convex class $\mathscr{L}_{\boldsymbol{b}}$ of measure-valued solutions $\mu_t \in \mathscr{M}_+(\mathbb{R}^d)$ of the continuity equation relative to $\boldsymbol{b}$, satisfying the following monotonicity property:

$$0 \le \mu'_t \le \mu_t \in \mathscr{L}_{\boldsymbol{b}} \qquad \Longrightarrow \qquad \mu'_t \in \mathscr{L}_{\boldsymbol{b}} \tag{14}$$

whenever μ'_t still solves the continuity equation relative to $\boldsymbol{b}$, and satisfies the integrability condition

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_t(x)|}{1+|x|}\, d\mu'_t(x)dt < +\infty.$$

The typical application will be with absolutely continuous measures $\mu_t = w_t \mathscr{L}^d$, whose densities satisfy some quantitative and possibly time-depending bound $\big($e.g. $L^\infty(L^1) \cap L^\infty(L^\infty)\big)$.

Definition 4.5 ($\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows). Given the class $\mathscr{L}_{\boldsymbol{b}}$, we say that $\boldsymbol{X}(t,x)$ is a $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu} \in \mathscr{M}_+(\mathbb{R}^d)$ (at time 0) if the following two properties hold:

(a) $\boldsymbol{X}(\cdot,x)$ is absolutely continuous in $[0,T]$ and satisfies

$$\boldsymbol{X}(t,x) = x + \int_0^t \boldsymbol{b}_s(\boldsymbol{X}(s,x))\,ds \qquad \forall t \in [0,T]$$

for $\bar{\mu}$-a.e. x;

(b) $\mu_t := \boldsymbol{X}(t,\cdot)_\# \bar{\mu} \in \mathscr{L}_{\boldsymbol{b}}$.

Heuristically $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows can be thought as suitable selections of the (possibly non unique) solutions of the ODE, made in such a way to produce a density in $\mathscr{L}_{\boldsymbol{b}}$. See Example 1.1 for an illustration of this concept.

We will show that the $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu}$ is unique, modulo $\bar{\mu}$-negligible sets, whenever uniqueness[1] for the PDE holds in the class $\mathscr{L}_{\boldsymbol{b}}$, i.e.

$$\mu_0 = \mu'_0 \qquad \Longrightarrow \qquad \mu_t = \mu'_t \;\; \forall t \in [0,T]$$

whenever μ_t and μ'_t belong to $\mathscr{L}_{\boldsymbol{b}}$.

Before stating and proving the uniqueness theorem for $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows, we state two elementary but useful results. The first one is a simple exercise:

[1] We thank A. Figalli and F. Flandoli for pointing out that the argument works with minor variants when only uniqueness is imposed at the level of the PDE, and not necessarily the comparison principle.

Exercise 4.6. Let $\sigma \in \mathscr{M}_+(\Gamma_T)$ and let $D \subset [0,T]$ be a dense set. Show that σ is a Dirac mass in Γ_T iff its projections $(e(t))_\#\sigma$, $t \in D$, are Dirac masses in $\mathbb{R}^d$.

The second one is concerned with a family of measures η_x:

Lemma 4.7. *Let η_x be a measurable family of positive finite measures in Γ_T with the following property: for any $t \in [0,T]$ and any pair of disjoint Borel sets $E, E' \subset \mathbb{R}^d$ we have*

$$\eta_x(\{\gamma:\ \gamma(t) \in E\})\,\eta_x\big(\{\gamma:\ \gamma(t) \in E'\}\big) = 0 \quad \bar{\mu}\text{-a.e. in } \mathbb{R}^d. \tag{15}$$

Then η_x is a Dirac mass for $\bar{\mu}$-a.e. x.

Proof. Taking into account Exercise 4.6, for a fixed $t \in (0,T]$ it suffices to check that the measures $\lambda_x := \gamma(t)_\#\eta_x$ are Dirac masses for $\bar{\mu}$-a.e. x. Then (15) gives $\lambda_x(E)\lambda_x(E') = 0$ $\bar{\mu}$-a.e. for any pair of disjoint Borel sets $E, E' \subset \mathbb{R}^d$. Let $\delta > 0$ and let us consider a partition of $\mathbb{R}^d$ in countably many Borel sets R_i having a diameter less then δ. Then, as $\lambda_x(R_i)\lambda_x(R_j) = 0$ $\bar{\mu}$-a.e. whenever $i \neq j$, we have a corresponding decomposition of $\bar{\mu}$-almost all of $\mathbb{R}^d$ in Borel sets A_i such that $\operatorname{supp}\lambda_x \subset \overline{R}_i$ for any $x \in A_i$ (just take $\{\lambda_x(R_i) > 0\}$ and subtract from it all other sets $\{\lambda_x(R_j) > 0\}$, $j \neq i$). Since δ is arbitrary the statement is proved. □

Theorem 4.8 (Uniqueness of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows). *Assume that the PDE has the uniqueness property in $\mathscr{L}_{\boldsymbol{b}}$. Then the $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu}$ is unique, i.e. two different selections $\boldsymbol{X}_1(t,x)$ and $\boldsymbol{X}_2(t,x)$ of solutions of the ODE inducing solutions of the continuity equation in $\mathscr{L}_{\boldsymbol{b}}$ satisfy*

$$\boldsymbol{X}_1(\cdot,x) = \boldsymbol{X}_2(\cdot,x) \qquad \text{in } \Gamma_T, \text{ for } \bar{\mu}\text{-a.e. } x.$$

More generally, if $\eta^1, \eta^2 \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ are concentrated on the pairs (x,γ) with γ absolutely continuous solution of the ODE, and if $\mu_t^{\eta^1} = \mu_t^{\eta^2} \in \mathscr{L}_{\boldsymbol{b}}$, then $\eta^1 = \eta^2$ and η^i are concentrated on the graph of a map $x \mapsto \boldsymbol{X}(\cdot,x)$ which is the unique $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow.

Proof. If the first statement were false we could produce a measure η *not* concentrated on a graph inducing a solution $\mu_t^\eta \in \mathscr{L}_{\boldsymbol{b}}$ of the PDE. This is not possible, thanks to the next result (Theorem 4.10). The measure η can be built as follows:

$$\eta := \frac{1}{2}(\eta^1 + \eta^2) = \frac{1}{2}\left[(x, \boldsymbol{X}_1(\cdot,x))_\#\bar{\mu} + (x, \boldsymbol{X}_2(\cdot,x))_\#\bar{\mu}\right].$$

Since $\mathscr{L}_{\boldsymbol{b}}$ is convex we still have $\mu_t^\eta = \frac{1}{2}(\mu_t^{\eta^1} + \mu_t^{\eta^2}) \in \mathscr{L}_{\boldsymbol{b}}$. In a similar way, one can use Theorem 4.10 first to show that η^i are concentrated on graphs, and then the previous combination argument to show that $\eta^1 = \eta^2$. □

Remark 4.9. In the same vein, one can also show that

$$\boldsymbol{X}_1(\cdot,x) = \boldsymbol{X}_2(\cdot,x) \qquad \text{in } \Gamma_T \text{ for } \bar{\mu}_1 \wedge \bar{\mu}_2\text{-a.e. } x$$

whenever $\boldsymbol{X}_1, \boldsymbol{X}_2$ are $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows starting respectively from $\bar{\mu}_1$ and $\bar{\mu}_2$.

We used the following basic result, having some analogy with Kantorovich's and Mather's theories.

Theorem 4.10. *Assume that the PDE has the uniqueness property in $\mathscr{L}_{\boldsymbol{b}}$. Let $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ be concentrated on the pairs (x,γ) with γ absolutely continuous solution of the ODE starting from x, and assume that $\mu_t^\eta \in \mathscr{L}_{\boldsymbol{b}}$. Then η is concentrated on a graph, i.e. there exists a function $x \mapsto X(\cdot,x) \in \Gamma_T$ such that*

$$\eta = \big(x, X(\cdot,x)\big)_\# \bar\mu, \quad \textit{with} \quad \bar\mu := (\pi_{\mathbb{R}^d})_\# \eta = \mu_0^\eta.$$

Proof. We use the representation (8) of η, given by the disintegration theorem, the criterion stated in Lemma 4.7 and argue by contradiction. If the thesis is false then η_x is not a Dirac mass in a set of positive $\bar\mu$ measure and we can find $t \in (0,T]$, disjoint Borel sets $E, E' \subset \mathbb{R}^d$ and a Borel set C with $\bar\mu(C) > 0$ such that

$$\eta_x\big(\{\gamma:\ \gamma(t) \in E\}\big)\eta_x\big(\{\gamma:\ \gamma(t) \in E'\}\big) > 0 \qquad \forall x \in C.$$

Possibly passing to a smaller set having still strictly positive $\bar\mu$ measure we can assume that

$$0 < \eta_x(\{\gamma:\ \gamma(t) \in E\}) \le M\eta_x(\{\gamma:\ \gamma(t) \in E'\}) \qquad \forall x \in C \tag{16}$$

for some constant M. We define measures η^1, η^2 whose disintegrations η_x^1, η_x^2 are given by

$$\eta_x^1 := \chi_C(x)\eta_x \llcorner \{\gamma:\ \gamma(t) \in E\}, \qquad \eta_x^2 := M\chi_C(x)\eta_x \llcorner \{\gamma:\ \gamma(t) \in E'\}$$

and denote by μ_s^i, $s \in [0,t]$, the (superposition) solutions of the continuity equation induced by η^i. Then

$$\mu_0^1 = \eta_x(\{\gamma:\ \gamma(t) \in E\})\bar\mu \llcorner C, \qquad \mu_0^2 = M\eta_x(\{\gamma:\ \gamma(t) \in E'\})\bar\mu \llcorner C,$$

so that (16) yields $\mu_0^1 \le \mu_0^2$. On the other hand, μ_t^1 is orthogonal to μ_t^2: precisely, denoting by η_{tx} the image of η_x under the map $\gamma \mapsto \gamma(t)$, we have

$$\mu_t^1 = \int_C \eta_{tx} \llcorner E\, d\mu(x) \perp M \int_C \eta_{tx} \llcorner E'\, d\mu(x) = \mu_t^2.$$

In order to conclude, let $f : \mathbb{R}^d \to [0,1]$ be the density of μ_0^1 with respect to μ_0^2 and set

$$\tilde\eta_x^2 := Mf(x)\chi_C(x)\eta_x \llcorner \{\gamma:\ \gamma(t) \in E'\}.$$

We define the measure $\tilde\eta^2$ whose disintegration is given by $\tilde\eta_x^2$ and denote by $\tilde\mu_s^2$, $s \in [0,t]$, the (superposition) solution of the continuity equation induced by $\tilde\eta^2$.

Notice also that $\mu_s^i \le \mu_s$ and so the monotonicity assumption (14) on $\mathscr{L}_{\boldsymbol{b}}$ gives $\mu_s^i \in \mathscr{L}_{\boldsymbol{b}}$, and since $\tilde\eta^2 \le \eta^2$ we obtain that $\tilde\mu_s^2 \in \mathscr{L}_{\boldsymbol{b}}$ as well. By construction $\mu_0^1 = \tilde\mu_0^2$, while μ_t^1 is orthogonal to μ_t^2, a measure larger than $\tilde\mu_t^2$. We have thus built two different solutions of the PDE with the same initial condition. $\square$

Now we come to the *existence* of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows.

Theorem 4.11 (Existence of $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flows). *Assume that the PDE has the uniqueness property in $\mathscr{L}_{\boldsymbol{b}}$ and that for some $\bar{\mu} \in \mathscr{M}_+(\mathbb{R}^d)$ there exists a solution $\mu_t \in \mathscr{L}_{\boldsymbol{b}}$ with $\mu_0 = \bar{\mu}$. Then there exists a (unique) $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow starting from $\bar{\mu}$.*

Proof. By the superposition principle we can represent μ_t as $(e_t)_\#\eta$ for some $\eta \in \mathscr{M}_+(\mathbb{R}^d \times \Gamma_T)$ concentrated on pairs (x,γ) solutions of the ODE. Then, Theorem 4.10 tells us that η is concentrated on a graph, i.e. there exists a function $x \mapsto \boldsymbol{X}(\cdot,x) \in \Gamma_T$ such that

$$\big(x, \boldsymbol{X}(\cdot,x)\big)_\#\bar{\mu} = \eta.$$

Pushing both sides via e_t we obtain

$$\boldsymbol{X}(t,\cdot)_\#\bar{\mu} = (e_t)_\#\eta = \mu_t \in \mathscr{L}_{\boldsymbol{b}},$$

and therefore $\boldsymbol{X}$ is a $\mathscr{L}_{\boldsymbol{b}}$-Lagrangian flow. □

5 The Flow Associated to Sobolev or *BV* Vector Fields

Here we discuss the well-posedness of the continuity or transport equations assuming that $\boldsymbol{b}_t(\cdot)$ has a Sobolev regularity, following [61]. Then, the general theory previously developed provides existence and uniqueness of the $\mathscr{L}$-Lagrangian flow, with $\mathscr{L} := L^\infty(L^1) \cap L^\infty(L^\infty)$. We denote by $I \subset \mathbb{R}$ an open interval.

Definition 5.1 (Renormalized solutions). Let $\boldsymbol{b} \in L^1_{\mathrm{loc}}\big(I; L^1_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\big)$ be such that $D \cdot \boldsymbol{b}_t = \operatorname{div} \boldsymbol{b}_t \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in I$, with

$$\operatorname{div} \boldsymbol{b}_t \in L^1_{\mathrm{loc}}\Big(I; L^1_{\mathrm{loc}}(\mathbb{R}^d)\Big).$$

Let $w \in L^\infty_{\mathrm{loc}}\big(I; L^\infty_{\mathrm{loc}}(\mathbb{R}^d)\big)$ and assume that

$$c := \frac{d}{dt} w + \boldsymbol{b} \cdot \nabla w \in L^1_{\mathrm{loc}}(I \times \mathbb{R}^d). \tag{17}$$

Then, we say that w is a renormalized solution of (17) if

$$\frac{d}{dt}\beta(w) + \boldsymbol{b} \cdot \nabla\beta(w) = c\beta'(w) \qquad \forall \beta \in C^1(\mathbb{R}).$$

Equivalently, recalling the definition of the distribution $\boldsymbol{b} \cdot \nabla w$, the definition could be given in a conservative form, writing

$$\frac{d}{dt}\beta(w) + D_x \cdot (\boldsymbol{b}\beta(w)) = c\beta'(w) + \operatorname{div} \boldsymbol{b}_t \beta(w).$$

Notice also that the concept makes sense, choosing properly the class of "test" functions β, also for w not satisfying (17), or not even locally integrable. This is particularly relevant in connection with DiPerna–Lions's existence theorem for the Boltzmann equation [62], or with the case when w is the characteristic of an unbounded vector field $\boldsymbol{b}$.

This concept is also reminiscent of Kruzkhov's concept of *entropy* solution for a scalar conservation law

$$\frac{d}{dt}u + D_x \cdot (\boldsymbol{f}(u)) = 0 \qquad u : (0,+\infty) \times \mathbb{R}^d \to \mathbb{R}.$$

In this case only a distributional one-sided inequality is required:

$$\frac{d}{dt}\eta(u) + D_x \cdot (\boldsymbol{q}(u)) \leq 0$$

for any convex entropy–entropy flux pair $(\eta, \boldsymbol{q})$ (i.e. η is convex and $\eta' \boldsymbol{f}' = \boldsymbol{q}'$).

Remark 5.2 (Time continuity). Using the fact that both $t \mapsto w_t$ and $t \mapsto \beta(w_t)$ have uniformly continuous representatives (w.r.t. the $w^* - L^\infty_{\mathrm{loc}}$ topology) $\bar{w}_t$, σ_t, we obtain that $t \mapsto \bar{w}_t$ is continuous with respect to the strong L^1_{loc} topology for $\mathscr{L}^1$-a.e. t, and precisely for any t such that $\sigma_t = \beta(\bar{w}_t)$. The proof follows by a classical weak-strong convergence argument:

$$f_n \rightharpoonup f, \quad \beta(f_n) \rightharpoonup \beta(f) \qquad \Longrightarrow \qquad f_n \to f$$

provided β is *strictly* convex. In the case of scalar conservation laws there are analogous and more precise results [92, 82]. We remark the fact that, in general, a renormalized solution does *not* need to have a representative which is strongly continuous *for every* t. This can be seen using a variation of an example given by Depauw [59] (see Remark 2.7 of [26]). Depauw's example provides a divergence free vector field $\boldsymbol{a} \in L^\infty([0,1] \times \mathbb{R}^2; \mathbb{R}^2)$, with $\boldsymbol{a}(t,\cdot) \in BV_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{R}^2)$ for $\mathscr{L}^1$-a.e. $t \in [0,1]$ (but $\boldsymbol{a} \notin L^1([0,1]; BV_{\mathrm{loc}})$) such that the Cauchy problem

$$\begin{cases} \partial_t u + \boldsymbol{a} \cdot \nabla u = 0 \\ u(0,\cdot) = 0 \end{cases}$$

has a nontrivial solution, with $|\bar{u}| = 1$ $\mathscr{L}^3$-a.e. in $[0,1] \times \mathbb{R}^2$ and with the property that $\bar{u}(t,\cdot) \rightharpoonup 0$ as $t \downarrow 0$, but this convergence is *not* strong. Now consider a vector field $\boldsymbol{b}$ on $[-1,1] \times \mathbb{R}^2$ defined as Depauw's vector field for $t > 0$, and set $\boldsymbol{b}(t,x) = -\boldsymbol{a}(-t,x)$ for $t < 0$. It is simple to check (as only affine functions $\tilde{\beta}(t) = a + bt$ need to be checked, because for any β there exists an affine $\tilde{\beta}$ such that $\tilde{\beta}(\pm 1) = \beta(\pm 1)$) that the function

$$\bar{w}(t,x) = \begin{cases} \bar{u}(t,x) & \text{if } t > 0 \\ \bar{u}(-t,x) & \text{if } t < 0 \end{cases}$$

is a renormalized solution of $\partial_t w + \boldsymbol{b} \cdot \nabla w = 0$, but this solution is not strongly continuous at $t = 0$.

Remark 5.3. A new insight in the theory of renormalized solutions has been obtained in the recent paper [26]. In particular, it is proved that for a vector field $\boldsymbol{b} \in L^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$ with zero divergence (and without any regularity assumption) the following two conditions are equivalent (the L^2 framework has been considered just for simplicity):

(i) $\boldsymbol{b}$ has the uniqueness property for weak solutions in $C([0,T];w-L^2(\mathbb{R}^d))$ for both the forward and the backward Cauchy problems starting respectively from 0 and T, i.e. the only solutions in $C([0,T];w-L^2(\mathbb{R}^d))$ to the problems

$$\begin{cases}\partial_t u_F + \boldsymbol{b}\cdot\nabla u_F = 0\\ u_F(0,\cdot)=0\end{cases}\quad\text{and}\quad\begin{cases}\partial_t u_B + \boldsymbol{b}\cdot\nabla u_B = 0\\ u_B(T,\cdot)=0\end{cases}$$

are $u_F\equiv 0$ and $u_B\equiv 0$;

(ii) Every weak solution in $C([0,T];w-L^2(\mathbb{R}^d))$ of $\partial_t u + \boldsymbol{b}\cdot\nabla u = 0$ is strongly continuous (i.e. lies in $C([0,T];s-L^2(\mathbb{R}^d))$) and is a renormalized solution.

The proof of this equivalence is obtained through the study of the approximation properties of the solution of the transport equation, with respect to the norm of the graph of the transport operator (see Theorem 2.1 of [26] for the details).

Using the concept of renormalized solution we can prove a comparison principle in the following natural class $\mathscr{L}$:

$$\begin{aligned}\mathscr{L} := \Big\{ & w\in L^\infty\Big([0,T];L^1(\mathbb{R}^d)\Big)\cap L^\infty\Big([0,T];L^\infty(\mathbb{R}^d)\Big):\\ & w\in C\Big([0,T];w^*-L^\infty(\mathbb{R}^d)\Big)\Big\}.\end{aligned}\tag{18}$$

Theorem 5.4 (Comparison principle). *Assume that*

$$\frac{|\boldsymbol{b}|}{1+|x|}\in L^1\Big([0,T];L^\infty(\mathbb{R}^d)\Big)+L^1\Big([0,T];L^1(\mathbb{R}^d)\Big),\tag{19}$$

that $D\cdot\boldsymbol{b}_t = \operatorname{div}\boldsymbol{b}_t\mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t\in[0,T]$, and that

$$[\operatorname{div}\boldsymbol{b}_t]^-\in L^1_{\mathrm{loc}}\Big([0,T)\times\mathbb{R}^d\Big).\tag{20}$$

Setting $\boldsymbol{b}_t\equiv 0$ for $t<0$, assume in addition that any *solution of* (17) *in $(-\infty,T)\times\mathbb{R}^d$ is renormalized. Then the comparison principle for the continuity equation holds in the class $\mathscr{L}$.*

Proof. By the linearity of the equation, it suffices to show that $w\in\mathscr{L}$ and $w_0\le 0$ implies $w_t\le 0$ for any $t\in[0,T]$. We extend first the PDE to negative times, setting $w_t = w_0$. Then, fix a cut-off function $\varphi\in C_c^\infty(\mathbb{R}^d)$ with $\operatorname{supp}\varphi\subset\overline{B}_2(0)$ and $\varphi\equiv 1$ on $B_1(0)$, and the renormalization functions

$$\beta_\varepsilon(t) := \sqrt{\varepsilon^2+(t^+)^2}-\varepsilon\in C^1(\mathbb{R}).$$

Notice that

$$\beta_\varepsilon(t) \uparrow t^+ \quad \text{as } \varepsilon \downarrow 0, \qquad t\beta'_\varepsilon(t) - \beta_\varepsilon(t) \in [0, \varepsilon]. \tag{21}$$

We know that

$$\frac{d}{dt}\beta_\varepsilon(w_t) + D_x \cdot (\boldsymbol{b}\beta_\varepsilon(w_t)) = \operatorname{div}\boldsymbol{b}_t(\beta_\varepsilon(w_t) - w_t\beta'_\varepsilon(w_t))$$

in the sense of distributions in $(-\infty, T) \times \mathbb{R}^d$. Plugging $\varphi_R(\cdot) := \varphi(\cdot/R)$, with $R \geq 1$, into the PDE we obtain

$$\frac{d}{dt}\int_{\mathbb{R}^d} \varphi_R \beta_\varepsilon(w_t)\, dx = \int_{\mathbb{R}^d} \beta_\varepsilon(w_t)\langle \boldsymbol{b}_t, \nabla\varphi_R\rangle\, dx + \int_{\mathbb{R}^d} \varphi_R \operatorname{div}\boldsymbol{b}_t(\beta_\varepsilon(w_t) - w_t\beta'_\varepsilon(w_t))\, dx.$$

Splitting $\boldsymbol{b}$ as $\boldsymbol{b}_1 + \boldsymbol{b}_2$, with

$$\frac{|\boldsymbol{b}_1|}{1+|x|} \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right) \quad \text{and} \quad \frac{|\boldsymbol{b}_2|}{1+|x|} \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right)$$

and using the inequality

$$\frac{1}{R}\chi_{\{R \leq |x| \leq 2R\}} \leq \frac{3}{1+|x|}\chi_{\{R \leq |x|\}}$$

we can estimate the first integral in the right hand side with

$$3\|\nabla\varphi\|_\infty \left\| \frac{\boldsymbol{b}_{1t}}{1+|x|} \right\|_\infty \int_{\{|x| \geq R\}} |w_t|\, dx + 3\|\nabla\varphi\|_\infty \|w_t\|_\infty \int_{\{|x| \geq R\}} \frac{|\boldsymbol{b}_{1t}|}{1+|x|}\, dx.$$

The second integral can be estimated with

$$\varepsilon \int_{\mathbb{R}^d} \varphi_R [\operatorname{div}\boldsymbol{b}_t]^-\, dx.$$

Passing to the limit first as $\varepsilon \downarrow 0$ and then as $R \to +\infty$ and using the integrability assumptions on $\boldsymbol{b}$ and w we get

$$\frac{d}{dt}\int_{\mathbb{R}^d} w_t^+\, dx \leq 0$$

in the distribution sense in $\mathbb{R}$. Since the function vanishes for negative times, this suffices to conclude using Gronwall lemma. □

DiPerna and Lions proved that *all* distributional solutions are renormalized when there is a Sobolev regularity with respect to the spatial variables.

Theorem 5.5. *Let* $\boldsymbol{b} \in L^1_{\mathrm{loc}}\left(I; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R}^d)\right)$ *and let* $w \in L^\infty_{\mathrm{loc}}(I \times \mathbb{R}^d)$ *be a distributional solution of* (17)*. Then* w *is a renormalized solution.*

Proof. We mollify with respect to the spatial variables and we set

$$r^\varepsilon := (\boldsymbol{b}\cdot\nabla w)*\rho_\varepsilon - \boldsymbol{b}\cdot(\nabla(w*\rho_\varepsilon)), \qquad w^\varepsilon := w*\rho_\varepsilon$$

to obtain

$$\frac{d}{dt}w^\varepsilon + \boldsymbol{b}\cdot\nabla w^\varepsilon = c*\rho_\varepsilon - r^\varepsilon.$$

By the smoothness of w^ε w.r.t. x, the PDE above tells us that $\frac{d}{dt}w^\varepsilon_t \in L^1_{\mathrm{loc}}$, therefore $w^\varepsilon \in W^{1,1}_{\mathrm{loc}}(I\times\mathbb{R}^d)$ and we can apply the standard chain rule in Sobolev spaces, getting

$$\frac{d}{dt}\beta(w^\varepsilon) + \boldsymbol{b}\cdot\nabla\beta(w^\varepsilon) = \beta'(w^\varepsilon)c*\rho_\varepsilon - \beta'(w^\varepsilon)r^\varepsilon.$$

When we let $\varepsilon\downarrow 0$, the convergence in the distribution sense of all terms in the identity above is trivial, with the exception of the last one. To ensure its convergence to zero, it seems necessary to show that $r^\varepsilon \to 0$ strongly in L^1_{loc}, or at least that r^ε are equi-integrable (as this would imply their weak L^1 convergence to 0; remember also that $\beta'(w^\varepsilon)$ is locally equibounded w.r.t. ε). This is indeed the case, and it is exactly here that the Sobolev regularity plays a role. □

Proposition 5.6 (Strong convergence of commutators). *If $w \in L^\infty_{\mathrm{loc}}(I\times\mathbb{R}^d)$ and $\boldsymbol{b}\in L^1_{\mathrm{loc}}\left(I;W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$ we have*

$$L^1_{\mathrm{loc}}\text{-}\lim_{\varepsilon\downarrow 0}\,(\boldsymbol{b}\cdot\nabla w)*\rho_\varepsilon - \boldsymbol{b}\cdot(\nabla(w*\rho_\varepsilon)) = 0.$$

Proof. Playing with the definitions of $\boldsymbol{b}\cdot\nabla w$ and of the convolution product of a distribution and a smooth function, one proves first the identity

$$r^\varepsilon(t,x) = \int_{\mathbb{R}^d} w(t,x-\varepsilon y)\frac{(\boldsymbol{b}_t(x-\varepsilon y)-\boldsymbol{b}_t(x))\cdot\nabla\rho(y)}{\varepsilon}\,dy - (w\,\mathrm{div}\,\boldsymbol{b}_t)*\rho_\varepsilon(x). \quad (22)$$

Introducing the commutators in the (easier) conservative form

$$R^\varepsilon := (D_x\cdot(\boldsymbol{b}w))*\rho_\varepsilon - D_x\cdot(\boldsymbol{b}w^\varepsilon)$$

(here we set again $w^\varepsilon := w*\rho_\varepsilon$) it suffices to show that $R^\varepsilon = L^\varepsilon - w^\varepsilon\mathrm{div}\boldsymbol{b}_t$, where

$$L^\varepsilon(t,x) := \int_{\mathbb{R}^d} w(t,z)(\boldsymbol{b}_t(x)-\boldsymbol{b}_t(z))\cdot\nabla\rho_\varepsilon(z-x)\,dz.$$

Indeed, for any test function φ, we have that $\langle R^\varepsilon,\varphi\rangle$ is given by

$$\begin{aligned}
&-\int_I\int w\boldsymbol{b}\cdot\nabla\rho_\varepsilon*\varphi\,dydt - \int_I\int \varphi\boldsymbol{b}\cdot\nabla\rho_\varepsilon*w\,dxdt - \int_I\int w^\varepsilon\varphi\,\mathrm{div}\boldsymbol{b}_t\,dxdt\\
&= -\int_I\int\int w_t(y)\boldsymbol{b}_t(y)\cdot\nabla\rho_\varepsilon(y-x)\varphi(x)\,dxdydt
\end{aligned}$$

$$
\begin{aligned}
&-\int_I\int\int \boldsymbol{b}_t(x)\nabla\rho_\varepsilon(x-y)w_t(y)\varphi(x)dydxdt - \int_I\int w^\varepsilon \varphi \mathrm{div}\,\boldsymbol{b}_t\,dxdt\\
&= \int_I\int L^\varepsilon \varphi\,dxdt - \int_I\int w^\varepsilon \varphi \mathrm{div}\,\boldsymbol{b}_t\,dxdt
\end{aligned}
$$

(in the last equality we used the fact that $\nabla\rho$ is odd).

Then, one uses the strong convergence of translations in L^p and the strong convergence of the difference quotients (a property that *characterizes* functions in Sobolev spaces)

$$
\frac{u(x+\varepsilon z)-u(x)}{\varepsilon} \to \nabla u(x)z \qquad \text{strongly in } L^1_{\mathrm{loc}}, \text{ for } u \in W^{1,1}_{\mathrm{loc}}
$$

to obtain that r^ε strongly converge in $L^1_{\mathrm{loc}}(I\times\mathbb{R}^d)$ to

$$
-w(t,x)\int_{\mathbb{R}^d}\langle\nabla\boldsymbol{b}_t(x)y,\nabla\rho(y)\rangle\,dy - w(t,x)\mathrm{div}\,\boldsymbol{b}_t(x).
$$

The elementary identity

$$
\int_{\mathbb{R}^d} y_i\frac{\partial\rho}{\partial y_j}\,dy = -\delta_{ij}
$$

then shows that the limit is 0 (this can also be derived by the fact that, in any case, the limit of r^ε in the distribution sense should be 0). □

In this context, given $\bar\mu = \rho\mathscr{L}^d$ with $\rho\in L^1\cap L^\infty$, the $\mathscr{L}$-Lagrangian flow starting from $\bar\mu$ (at time 0) is defined by the following two properties:

(a) $\boldsymbol{X}(\cdot,x)$ is absolutely continuous in $[0,T]$ and satisfies

$$
\boldsymbol{X}(t,x) = x + \int_0^t \boldsymbol{b}_s(\boldsymbol{X}(s,x))\,ds \qquad \forall t\in[0,T]
$$

for $\bar\mu$-a.e. x;

(b) $\boldsymbol{X}(t,\cdot)_\#\bar\mu \le C\mathscr{L}^d$ for all $t\in[0,T]$, with C independent of t.

Summing up what we obtained so far, the general theory provides us with the following existence and uniqueness result.

Theorem 5.7 (Existence and uniqueness of $\mathscr{L}$-Lagrangian flows). *Let $\boldsymbol{b}\in L^1\left([0,T];W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be satisfying*

(i) $\dfrac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T];L^1(\mathbb{R}^d)\right) + L^1\left([0,T];L^\infty(\mathbb{R}^d)\right)$*;*

(ii) $[\mathrm{div}\boldsymbol{b}_t]^- \in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right)$.

Then the $\mathscr{L}$-Lagrangian flow relative to $\boldsymbol{b}$ exists and is unique.

Proof. By the previous results, the comparison principle holds for the continuity equation relative to $\boldsymbol{b}$. Therefore the general theory previously developed applies, and Theorem 4.8 provides *uniqueness* of the $\mathscr{L}$-Lagrangian flow.

As for the *existence*, still the general theory (Theorem 4.11) tells us that it can be achieved provided we are able to solve, within $\mathscr{L}$, the continuity equation

$$\frac{d}{dt}w + D_x\cdot(\boldsymbol{b}w) = 0 \tag{23}$$

for any nonnegative initial datum $w_0 \in L^1 \cap L^\infty$. The existence of these solutions can be immediately achieved by a smoothing argument: we approximate $\boldsymbol{b}$ in L^1_{loc} by smooth $\boldsymbol{b}^h$ with a uniform bound in $L^1(L^\infty)$ for $[\mathrm{div}\boldsymbol{b}^h_t]^-$. This bound, in turn, provides a uniform lower bound on $J\boldsymbol{X}^h$ and finally a uniform upper bound on $w^h_t = (w_0/J\boldsymbol{X}^h_t)\circ(\boldsymbol{X}^h_t)^{-1}$, solving

$$\frac{d}{dt}w^h + D_x\cdot(\boldsymbol{b}^h w^h) = 0.$$

Therefore, any weak limit of w^h solves (23). □

Notice also that, choosing for instance a Gaussian, we obtain that the $\mathscr{L}$-Lagrangian flow is well defined up to $\mathscr{L}^d$-negligible sets (and independent of $\bar{\mu} \ll \mathscr{L}^d$, thanks to Remark 4.9).

It is interesting to compare our characterization of Lagrangian flows with the one given in [61]. Heuristically, while the DiPerna–Lions characterization is based on the semigroup of transformations $x \mapsto \boldsymbol{X}(t,x)$, our characterization is based on the properties of the map $x \mapsto \boldsymbol{X}(\cdot,x)$.

Remark 5.8. The definition of the flow in [61] is based on the following three properties:

(a) $\dfrac{\partial \boldsymbol{Y}}{\partial t}(t,s,x) = b\left(t,\boldsymbol{Y}(t,s,x)\right)$ and $\boldsymbol{Y}(s,s,x) = x$ in the distribution sense in $(0,T)\times\mathbb{R}^d$;

(b) The image λ_t of $\mathscr{L}^d$ under $\boldsymbol{Y}(t,s,\cdot)$ satisfies

$$\frac{1}{C}\mathscr{L}^d \le \lambda_t \le C\mathscr{L}^d \qquad \text{for some constant } C>0;$$

(c) For all $s, s', t \in [0,T]$ we have

$$\boldsymbol{Y}\left(t,s,\boldsymbol{Y}(s,s',x)\right) = \boldsymbol{Y}(t,s',x) \qquad \text{for } \mathscr{L}^d\text{-a.e. } x.$$

Then, $\boldsymbol{Y}(t,s,x)$ corresponds, in our notation, to the flow $\boldsymbol{X}^s(t,x)$ starting at time s (well defined even for $t<s$ if one has two-sided L^∞ bounds on the divergence).

In our setting condition (c) can be recovered as a consequence of the following argument: assume to fix the ideas that $s' \le s \le T$ and define

$$\tilde{\boldsymbol{X}}(t,x) := \begin{cases} \boldsymbol{X}^{s'}(t,x) & \text{if } t\in[s',s]; \\ \\ \boldsymbol{X}^{s}\left(t,\boldsymbol{X}^{s'}(s,x)\right) & \text{if } t\in[s,T]. \end{cases}$$

It is immediate to check that $\tilde{\boldsymbol{X}}(\cdot,x)$ is an integral solution of the ODE in $[s',T]$ for $\mathscr{L}^d$-a.e. x and that $\tilde{\boldsymbol{X}}(t,\cdot)_\#\bar{\mu}$ is bounded by $C^2\mathscr{L}^d$. Then, Theorem 5.7 (with s' as initial time) gives $\tilde{\boldsymbol{X}}(\cdot,x) = \boldsymbol{X}(\cdot,s',x)$ in $[s',T]$ for $\mathscr{L}^d$-a.e. x, whence (c) follows.

Let us now discuss the stability properties of $\mathscr{L}$-Lagrangian flows, in the special case when $\mathscr{L}$ is defined as in (18). We need the following lemma.

Lemma 5.9. *Assume that* $\boldsymbol{b}_h : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ *satisfy*

$$\frac{\boldsymbol{b}_h}{1+|x|} \in L^1\left([0,T];L^1(\mathbb{R}^d;\mathbb{R}^d)\right) + L^1\left([0,T];L^\infty(\mathbb{R}^d;\mathbb{R}^d)\right) \tag{24}$$

and that we can write $\boldsymbol{b}_h = \boldsymbol{b}_h^1 + \boldsymbol{b}_h^2$, *with*

$$\frac{|\boldsymbol{b}_h^1|}{1+|x|} \quad \text{bounded and equi-integrable in } L^1\left([0,T]\times\mathbb{R}^d\right), \tag{25}$$

$$\sup_h \left\| \frac{|\boldsymbol{b}_h^2(t,\cdot)|}{1+|x|} \right\|_\infty \in L^1(0,T), \tag{26}$$

$$\boldsymbol{b}_h^1 \to \boldsymbol{b}^1, \quad \boldsymbol{b}_h^2 \to \boldsymbol{b}^2 \qquad \mathscr{L}^{d+1}\text{-a.e. in } (0,T)\times\mathbb{R}^d. \tag{27}$$

Then, setting $\boldsymbol{b} = \boldsymbol{b}^1 + \boldsymbol{b}^2$, *we have*

$$\lim_{h\to\infty} \int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h - \boldsymbol{b}|}{1+|x|^{d+2}}\,dxdt = 0. \tag{28}$$

Proof. Without loss of generality we can assume that $\boldsymbol{b}^1 = \boldsymbol{b}^2 = 0$, hence $\boldsymbol{b} = 0$. The convergence to 0 of

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h^2|}{1+|x|^{d+2}}\,dxdt$$

follows by the standard dominated convergence theorem. The convergence to 0 of

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\boldsymbol{b}_h^1|}{1+|x|}\,dxdt$$

(and a fortiori of the integrals with the factor $1+|x|^{d+2}$) follows by the Vitali dominated convergence theorem. □

Theorem 5.10. *Let* $\boldsymbol{X}_h$ *be* $\mathscr{L}$*-Lagrangian flows relative to vector fields* $\boldsymbol{b}_h$, *starting from* $\bar{\mu} = \bar{\rho}\mathscr{L}^d$ *and satisfying:*

(a) $\boldsymbol{X}_h(t,\cdot)_\#\bar{\mu} \le C\mathscr{L}^d$, *with C independent of h and* $t\in[0,T]$*;*
(b) $\boldsymbol{b}_h = \boldsymbol{b}_h^1 + \boldsymbol{b}_h^2$ *with* $\boldsymbol{b}_h^i$ *satisfying* (25), (26) *and* (27).

Assume that the continuity equation with $\boldsymbol{b} = \boldsymbol{b}^1 + \boldsymbol{b}^2$ *as vector field satisfies the uniqueness property in* $\mathscr{L}$. *Then there exists a unique* $\mathscr{L}$*-Lagrangian flow* $\boldsymbol{X}$ *relative to* $\boldsymbol{b}$ *and* $x \mapsto \boldsymbol{X}^h(\cdot,x)$ *converge to* $x \mapsto \boldsymbol{X}(\cdot,x)$ *in* $\bar{\mu}$*-measure, i.e.*

$$\lim_{h\to\infty}\int_{\mathbb{R}^d} 1\wedge \max_{[0,T]}|\boldsymbol{X}^h(t,x)-\boldsymbol{X}(t,x)|\,d\bar{\mu}(x)=0.$$

Remark 5.11 (Stability with weak convergence in time). We remark the fact that the hypothesis of strong convergence in both time and space of the vector fields in the stability theorem is not natural in view of the applications to the theory of fluid mechanics (see Theorem II.7 in [61] and [75], in particular Theorem 2.5), and it is in contrast with the general philosophy that time regularity is less important than spatial regolarity (only summability with respect to time is necessary for the renormalization property to hold). For instance, this form of the stability theorem does not include the case of weakly converging vector fields which depend *only* on the time variable, while it is clear that in this case the convergence of the flows holds. However, as a consequence of the quantitative estimates presented in Sect. 8, it is possible to show that, under uniform bounds in $L^\infty([0,T];W^{1,p}_{\rm loc})$ for some $p>1$ (and for simplicity under uniform bounds in L^∞), the following form of weak convergence with respect to the time is sufficient to get the thesis:

$$\int_0^T \boldsymbol{b}_h(t,x)\eta(t)\,dt \longrightarrow \int_0^T \boldsymbol{b}(t,x)\eta(t)\,dt \qquad \text{in } L^1_{\rm loc}(\mathbb{R}^d) \text{ for every } \eta\in C^\infty_c(0,T).$$

Indeed, fix a parameter $\varepsilon>0$ and regularize with respect to the spatial variables only using a standard convolution kernel ρ_ε. We can rewrite the difference $\boldsymbol{X}_h(t,x)-\boldsymbol{X}(t,x)$ as

$$\begin{aligned}\boldsymbol{X}_h(t,x)-\boldsymbol{X}(t,x) &= \Big(\boldsymbol{X}_h(t,x)-\boldsymbol{X}^\varepsilon_h(t,x)\Big)+\Big(\boldsymbol{X}^\varepsilon_h(t,x)-\boldsymbol{X}^\varepsilon(t,x)\Big)\\ &\quad+\Big(\boldsymbol{X}^\varepsilon(t,x)-\boldsymbol{X}(t,x)\Big),\end{aligned}$$

where $\boldsymbol{X}^\varepsilon$ and $\boldsymbol{X}^\varepsilon_h$ are the flows relative to the spatial regularizations $\boldsymbol{b}^\varepsilon$ and $\boldsymbol{b}^\varepsilon_h$ respectively. Now, it is simple to check that

- The last term goes to zero with ε, by the stability theorem stated above;
- The first term goes to zero with ε, uniformly with respect to h: this is due to the fact that the difference $\boldsymbol{b}^\varepsilon_h-\boldsymbol{b}_h$ goes to zero in $L^1_{\rm loc}([0,T]\times\mathbb{R}^d)$ uniformly with respect to h (thanks to the uniform control in $W^{1,p}$ of the vector fields $\boldsymbol{b}_h$), hence we can apply the quantitative version of the stability theorem (see Theorem 8.12), and we get the desired convergence;
- The second term goes to zero for $h\to\infty$ when ε is kept fixed, because we are dealing with flows relative to vector fields which are smooth with respect to the space variable, uniformly in time, and weak convergence with respect to the time is enough to get the stability.

Then, in order to conclude, it is enough to let first $h\to\infty$, to eliminate the second term, and then $\varepsilon\to 0$.

It is not clear to us whether this result extends to the $W^{1,1}$ or the BV case: the quantitative estimates are available only for $W^{1,p}$ vector fields, where p is strictly greater than 1, hence the strategy described before does not extend to that case.

Proof of Theorem 5.10. We define η_h as the push forward of $\bar{\mu}$ under the map $x \mapsto (x, \boldsymbol{X}^h(\cdot, x))$ and argue as in the proof of Theorem 4.4.

Step 1. (Tightness of η_h) We claim that the family η_h is tight: indeed, by the remarks made after the statement of Theorem 4.4, it suffices to find a coercive functional $\Psi : \mathbb{R}^d \times \Gamma_T \to [0,\infty)$ whose integrals w.r.t. all measures η_h are uniformly bounded. Since $\bar{\mu}$ has finite mass we can find a function $\varphi : \mathbb{R}^d \to [0,\infty)$ such that $\varphi \in L^1(\bar{\mu})$ and $\varphi(x) \to \infty$ as $|x| \to \infty$. Analogously, we can find a function $\psi : \mathbb{R}^d \to [0,\infty)$ such that

$$\int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b})}{1+|x|}\, dxdt < \infty$$

and $\psi(x) \to \infty$ as $|x| \to \infty$. Then, we define

$$\Psi(x,\gamma) := \varphi(x) + \varphi(\gamma(0)) + \int_0^T \frac{\psi(\dot{\gamma})}{1+|\gamma|}\, dt$$

and notice that the coercivity of Ψ follows immediately from Lemma 5.12 below. Then, using assumption (a), we obtain:

$$\begin{aligned}
\int_{\mathbb{R}^d \times \Gamma_T} \Psi(x,\gamma)\, d\eta_h &= \int_{\mathbb{R}^d} \left(2\varphi(x) + \int_0^T \frac{\psi(\dot{\boldsymbol{X}}^h(t,x))}{1+|\boldsymbol{X}^h(t,x)|}\, dt \right) d\bar{\mu}(x) \\
&= 2\int_{\mathbb{R}^d} \varphi\, d\bar{\mu} + \int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b}^h(t,\boldsymbol{X}^h(t,x)))}{1+|\boldsymbol{X}^h(t,x)|}\, d\bar{\mu}(x) dt \\
&\leq 2\int_{\mathbb{R}^d} \varphi\, d\bar{\mu} + C\int_0^T \int_{\mathbb{R}^d} \frac{\psi(\boldsymbol{b}^h)}{1+|y|} dy\, dt.
\end{aligned}$$

Therefore the integrals of Ψ are uniformly bounded.

Step 2. (The limit flow belongs to $\mathscr{L}$) Let now η be a narrow limit point of η_h along some subsequence that, for notational simplicity, will not be relabelled. Let us show first that $\mu_t^\eta = (e_t)_\# \eta$ is representable as $w_t \mathscr{L}^d$ with w_t belonging to $\mathscr{L}$. Indeed, assumption (a) gives

$$\mu_t^{\eta_h} = \boldsymbol{X}^h(t,\cdot)_\# \bar{\mu} = w_t^h \mathscr{L}^d \quad \text{with} \quad \|w_t^h\|_\infty \leq C. \tag{29}$$

Therefore, as μ_t^η is the narrow limit of $\mu_t^{\eta_h}$, the same property is preserved in the limit for some $w \in L^\infty$. Moreover, the narrow continuity of $t \mapsto \mu_t^\eta$ immediately yields the w^*-continuity of $t \mapsto w_t$ and this proves that $w_t \in \mathscr{L}$.

Step 3. (η is concentrated on solutions of the ODE) Next we show that η is concentrated on the class of solutions of the ODE. Let $\bar{t} \in [0,T]$, $\chi \in C_c^\infty(\mathbb{R}^d)$ with $0 \leq \chi \leq 1$, $\boldsymbol{c} \in L^1\left([0,\bar{t}]; L^\infty(\mathbb{R}^d)\right)$, with $\boldsymbol{c}(t,\cdot)$ continuous in $\mathbb{R}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,\bar{t}]$, and define

$$\Phi_{\boldsymbol{c}}^{\bar{t}}(x,\gamma) := \chi(x) \frac{\left| \gamma(\bar{t}) - x - \int_0^{\bar{t}} \boldsymbol{c}(s,\gamma(s))\, ds \right|}{1 + \sup\limits_{[0,\bar{t}]} |\gamma|^{d+2}}.$$

It is immediate to check that $\Phi_{\boldsymbol{c}}^{\bar{t}} \in C_b(\mathbb{R}^d \times \Gamma_T)$, so that

$$
\begin{aligned}
\int_{\mathbb{R}^d\times\Gamma_T} \Phi_{\boldsymbol{c}}^{\bar t}\, d\eta &= \lim_{h\to\infty}\int_{\mathbb{R}^d\times\Gamma_T} \Phi_{\boldsymbol{c}}^{\bar t}\, d\eta_h \\
&= \lim_{h\to\infty}\int_{\mathbb{R}^d} \chi(x)\frac{\left|\int_0^{\bar t}\boldsymbol{b}_h(s,\boldsymbol{X}^h(s,x))-\boldsymbol{c}(s,\boldsymbol{X}^h(s,x))\,ds\right|}{1+\sup\limits_{[0,\bar t]}|\boldsymbol{X}^h(\cdot,x)|^{d+2}}\, d\bar\mu(x) \\
&\le \limsup_{h\to\infty}\int_{\mathbb{R}^d}\int_0^{\bar t}\frac{\left|\boldsymbol{b}_h(s,\boldsymbol{X}^h(s,x))-\boldsymbol{c}(s,\boldsymbol{X}^h(s,x))\right|}{1+|\boldsymbol{X}^h(s,x)|^{d+2}}\, ds d\bar\mu(x) \\
&\le C\limsup_{h\to\infty}\int_0^{\bar t}\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}_h(s,y)-\boldsymbol{c}(s,y)|}{1+|y|^{d+2}}\, dsdy \\
&= C\int_0^{\bar t}\int_{\mathbb{R}^d}\frac{|\boldsymbol{b}(s,y)-\boldsymbol{c}(s,y)|}{1+|y|^{d+2}}\, dsdy.
\end{aligned}
$$

Now, as (by lower semicontinuity)

$$
\frac{|\boldsymbol{b}|}{1+|x|}\in L^1\left([0,T];L^1(\mathbb{R}^d)\right)+L^\infty\left([0,T];L^\infty(\mathbb{R}^d)\right),
$$

we can find a sequence of vector fields $\boldsymbol{c}_h(t,x)$ continuous with respect to x and satisfying the assumptions of Lemma 5.9. Indeed, writing $\boldsymbol{b}=\boldsymbol{b}_1+\boldsymbol{b}_2=(\boldsymbol{f}_1+\boldsymbol{f}_2)(1+|x|)$, with

$$
|\boldsymbol{f}_1|\in L^1\left([0,T];L^1(\mathbb{R}^d)\right),\qquad |\boldsymbol{f}_2|\in L^1\left([0,T];L^\infty(\mathbb{R}^d)\right),
$$

we define $\boldsymbol{c}=(\boldsymbol{c}_h)_1+(\boldsymbol{c}_h)_2$, with

$$
(\boldsymbol{c}_h)_i := ((1+|x|)\boldsymbol{f}_i)*\rho_{\varepsilon_h}\qquad i=1,2
$$

with $\varepsilon_h\downarrow 0$. It is not hard to show that

$$
|(\boldsymbol{c}_h)_i|\le 2(1+|x|)(|\boldsymbol{f}_i|*\rho_{\varepsilon_h})\qquad i=1,2
$$

if the support of the convolution kernel ρ is contained in the unit ball. Therefore

$$
\frac{|(\boldsymbol{c}_h)_1|}{1+|x|}\text{ is bounded and equi-integrable in } L^1\left([0,T]\times\mathbb{R}^d\right), \tag{30}
$$

$$
\sup_h\left\|\frac{|(\boldsymbol{c}_h)_2(t,\cdot)|}{1+|x|}\right\|_\infty\in L^1(0,T). \tag{31}
$$

In the previous estimate we can now choose $\boldsymbol{c}=\boldsymbol{c}_h$ and use Lemma 5.9 again to obtain

$$
\lim_{h\to\infty}\int_{\mathbb{R}^d\times\Gamma_T}\chi(x)\frac{\left|\gamma(\bar t)-x-\int_0^{\bar t}\boldsymbol{c}_h(s,\gamma(s))\,ds\right|}{1+\sup\limits_{[0,\bar t]}|\gamma|^{d+2}}\, d\eta=0. \tag{32}
$$

Now, using the upper bound (29) and Lemma 5.9 once more we get

$$\limsup_{h\to\infty}\int_{\mathbb{R}^d\times\Gamma_T}\frac{\int_0^{\bar t}|\boldsymbol{c}_h(s,\gamma(s))-\boldsymbol{b}(s,\gamma(s))|\,ds}{1+\sup_{[0,\bar t]}|\gamma|^{d+2}}\,d\eta \tag{33}$$

$$\leq \limsup_{h\to\infty}\int_0^T\int_{\mathbb{R}^d\times\Gamma_T}\frac{|\boldsymbol{c}_h(s,\gamma(s))-\boldsymbol{c}(s,\gamma(s))|}{1+|\gamma(s)|^{d+2}}\,d\eta\,ds$$

$$\leq C\limsup_{h\to\infty}\int_0^T\int_{\mathbb{R}^d}\frac{|\boldsymbol{c}_h-\boldsymbol{b}|}{1+|x|^{d+2}}\,dxds=0.$$

Hence, from (32) and (33) and Fatou's lemma we infer that for $\chi\eta$-a.e. (x,γ) there is a subsequence $\varepsilon_{i(l)}$ such that

$$\lim_{l\to\infty}\left|\gamma(\bar t)-x-\int_0^{\bar t}\boldsymbol{b}^{\varepsilon_{i(l)}}(s,\gamma(s))\,ds\right|+\int_0^{\bar t}|\boldsymbol{b}^{\varepsilon_{i(l)}}(s,\gamma(s))-\boldsymbol{b}(s,\gamma(s)|\,ds=0,$$

so that

$$\gamma(\bar t)=x+\int_0^{\bar t}\boldsymbol{b}(s,\gamma(s))\,ds.$$

Choosing a sequence of cut-off functions χ_R and letting t vary in $\mathbb{Q}\cap[0,T]$ we obtain that (x,γ) solve the ODE in $[0,T]$ for η-a.e. (x,γ).
Step 4. (Conclusion) As we are assuming that the uniqueness property holds in $\mathscr{L}$ for the continuity equation relative to $\boldsymbol{b}$, we are now in the position of applying Theorem 4.10, which says that under these conditions necessarily

$$\eta=(x,\boldsymbol{X}(\cdot,x))_\#\bar\mu$$

for a suitable map $x\mapsto\boldsymbol{X}(\cdot,x)$. Clearly, by the concentration property of η, $\boldsymbol{X}(\cdot,x)$ has to be a solution of the ODE for $\bar\mu$-a.e. x. This proves that $\boldsymbol{X}$ is the $\mathscr{L}$-Lagrangian flow relative to $\boldsymbol{b}$. The convergence in measure of $\boldsymbol{X}_h$ to $\boldsymbol{X}$ follows by a general principle, stated in Lemma 5.13 below. □

Lemma 5.12 (A coercive functional in Γ_T). *Let φ, $\psi:\mathbb{R}^d\to\mathbb{R}$ and let*

$$\Phi(\gamma):=\varphi(\gamma(0))+\int_0^T\frac{\psi(\dot\gamma)}{1+|\gamma|}\,dt$$

be defined on the subspace of Γ_T made by absolutely continuous maps, and set equal to $+\infty$ outside. If $\varphi(x)$, $\psi(x)\to+\infty$ as $|x|\to\infty$ then all sublevel sets $\{\Phi\leq c\}$, $c\in\mathbb{R}$, are compact in Γ_T.

Proof. Let γ_n be such that $\Phi(\gamma_n)$ is bounded and notice that necessarily $|\gamma_n(0)|$ is bounded, by the assumption on φ. By integration of the ODE

$$\frac{d}{dt}\ln(1+|\gamma(t)|)=\frac{\gamma(t)}{|\gamma(t)|}\cdot\frac{\dot\gamma(t)}{1+|\gamma(t)|}$$

one obtains that also $\sup_{[0,T]}|\gamma_n|$ is uniformly bounded. As a consequence the factor $1/(1+|\gamma_n|)$ inside the integral part of Φ can be uniformly estimated from below, and therefore (due to the more than linear growth at infinity of ψ) the sequence $|\dot\gamma_n|$ is equi-integrable in $L^1\left((0,T);\mathbb{R}^d\right)$. As a consequence the sequence (γ_n) is relatively compact in Γ_T. □

Lemma 5.13 (Narrow convergence and convergence in measure). *Let $v_h,\, v : X \to Y$ be Borel maps and let $\bar\mu \in \mathscr{M}_+(X)$. Then $v_h \to v$ in $\bar\mu$-measure iff*

$$(x, v_h(x))_\#\bar\mu \text{ converges to } (x, v(x))_\#\bar\mu \text{ narrowly in } \mathscr{M}_+(X \times Y).$$

Proof. If $v_h \to v$ in $\bar\mu$-measure then $\varphi(x, v_h(x))$ converges in $L^1(\bar\mu)$ to $\varphi(x, v(x))$, and we immediately obtain the convergence of the push-forward measures. Conversely, let $\delta > 0$ and, for any $\varepsilon > 0$, let $w \in C_b(X;Y)$ be such that $\bar\mu(\{v \neq w\}) \le \varepsilon$. We define

$$\varphi(x,y) := 1 \wedge \frac{d_Y(y, w(x))}{\delta} \in C_b(X \times Y)$$

and notice that

$$\bar\mu\left(\{v \neq w\}\right) + \int_{X\times Y} \varphi\, d(x, v_h(x))_\#\bar\mu \ge \bar\mu(\{d_Y(v, v_h) > \delta\}),$$

$$\int_{X\times Y} \varphi\, d(x, v(x))_\#\bar\mu \le \bar\mu(\{w \neq v\}).$$

Taking into account the narrow convergence of the push-forward we obtain that

$$\limsup_{h\to\infty} \bar\mu(\{d_Y(v, v_h) > \delta\}) \le 2\bar\mu(\{w \neq v\}) \le 2\varepsilon\,;$$

since ε is arbitrary the proof is achieved. □

The renormalization Theorem 5.5 has been extended in [6] to the case of a *BV* dependence w.r.t. the spatial variables, but still assuming that

$$D \cdot \boldsymbol{b}_t \ll \mathscr{L}^d \qquad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T). \tag{34}$$

Theorem 5.14. *Let $\boldsymbol{b} \in L^1_{\rm loc}\left((0,T); BV_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be satisfying* (34)*. Then any distributional solution $w \in L^\infty_{\rm loc}\left((0,T)\times\mathbb{R}^d\right)$ of*

$$\frac{d}{dt}w + D_x \cdot (\boldsymbol{b}w) = c \in L^1_{\rm loc}\left((0,T)\times\mathbb{R}^d\right)$$

is a renormalized solution.

A self contained proof of this result, slightly simpler than the one given in the original paper [6], is given in [7] and [8]. The original argument in [6] was indeed based on deep result of G. Alberti [2], saying that for a $BV_{\rm loc}$ function $u : \mathbb{R}^d \to \mathbb{R}^m$ the matrix $M(x)$ in the polar decomposition $Du = M|Du|$ has rank 1 for $|D^s u|$-a.e. x, i.e. there exist unit vectors $\xi(x) \in \mathbb{R}^d$ and $\eta(x) \in \mathbb{R}^m$ such that

$M(x)z = \eta(x)\langle z, \xi(x)\rangle$. However, we observe that in the application of this result to the Keyfitz–Kranzer system [11, 9] the vector field $\boldsymbol{b}$ is of the form $\boldsymbol{f}(\rho)$ with ρ scalar and $\boldsymbol{f} \in C^1$ vectorial, so the rank-one structure of the distributional derivative (as a whole) is easy to check. Analogously, in the case of the semi-geostrophic system considered in [53], the vector field is a monotone map, and for this class of BV functions a much simpler proof of the rank-one property is available [3].

As in the Sobolev case we can now obtain from the general theory given in Sect. 3 *existence and uniqueness* of $\mathscr{L}$-Lagrangian flows, with $\mathscr{L} = L^\infty(L^1) \cap L^\infty(L^\infty)$: we just replace in the statement of Theorem 5.7 the assumption $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$ with $\boldsymbol{b} \in L^1\left([0,T]; BV_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)\right)$, assuming as usual that $D \cdot \boldsymbol{b}_t \ll \mathscr{L}^d$ for $\mathscr{L}^1$-a.e. $t \in [0,T]$.

Analogously, by applying Theorem 5.10 we obtain *stability* of $\mathscr{L}$-Lagrangian flows when $\boldsymbol{b}$ is as in Theorem 5.14.

6 Measure-Theoretic Differentials

In this section we introduce some weak differentiability notions, based on measure-theoretic limits of difference quotients, and we compare them (all results of this section are taken from [17]). An important remark is that none of these concepts gives additional informations on the derivative in the sense of distributions. Conversely, whenever the derivative in the sense of distributions is a measure with locally finite variation, the map is $\mathscr{L}^d$-a.e. approximately differentiable.

We recall that a sequence of measurable maps (f_k) defined on an open set $\Omega \subset \mathbb{R}^d$ is said to converge locally in measure to a measurable map f if

$$\lim_{k\to\infty} \mathscr{L}^d\left(\{x \in K : |f_k(x) - f(x)| > \varepsilon\}\right) = 0$$

for any compact set $K \subset A$ and any $\varepsilon > 0$.

The following simple lemma will be used in many occasions.

Lemma 6.1. *Let $f_k, f : \Omega \times \mathbb{R}^d \to \mathbb{R}$ be measurable functions. Suppose that, for each $y \in \mathbb{R}^d$, $f_k(\cdot, y) \to f(\cdot, y)$ locally in measure in Ω as $k \to \infty$. Then $f_k \to f$ locally in measure in $\Omega \times \mathbb{R}^d$ as $k \to \infty$.*

Proof. Let $K \subset \Omega$ be a compact set. We fix $\gamma > 0$ and set

$$g_k(y) = \mathscr{L}^d\left(\{x \in K : |f_k(x,y) - f(x,y)| \geq \gamma\}\right).$$

Then from the dominated convergence theorem we infer that

$$g_k \to 0 \quad \text{in } L^1_{\mathrm{loc}}(\mathbb{R}^d).$$

In particular, Fubini's theorem gives that $\mathscr{L}^{2d}\left((K \times B) \cap \{|f_k - f| > \gamma\}\right) \to 0$ for any ball $B \subset \mathbb{R}^d$. $\square$

As a consequence of the Lusin theorem, it is easy to check that for any measurable function $f : \Omega \to \mathbb{R}$ the functions $f(x+h)$ converge to f locally in measure as $h \to 0$. Therefore it is natural to study the behaviour, still with respect to local convergence in measure, of the difference quotients. This leads to the following definition.

Definition 6.2 (Fréchet and Gâteaux derivative in measure). Let $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ be a measurable function. We say that $g : \Omega \to \mathbb{R}^d$ is the (Fréchet) derivative in measure of f if

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - g(x)\cdot h}{|h|} = 0 \quad \text{locally in measure in } \Omega.$$

Analogously, g is called Gâteaux derivative in measure if the difference quotients above tend to 0 locally in measure in Ω along all lines passing through the origin.

Another differentiability condition, that we call directional differentiability in measure, involves an averaging procedure also on the direction. It appeared first in [73] in connection with the differentiability properties of the flow associated to Sobolev vector fields (see the next section) and it can be stated as follows.

Definition 6.3 (Directional differentiability in measure). We say that $f : \Omega \to \mathbb{R}$ is directionally differentiable in measure if there exists $W : \Omega \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\frac{f(x+ry) - f(x) - rW(x,y)}{r} \to 0 \qquad \text{locally in measure in } \Omega \times \mathbb{R}^d \text{ as } h \to 0.$$

The following result surprisingly shows that these three concepts are equivalent.

Theorem 6.4. *Let $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then the following assertions are equivalent:*

(i) f is (Fréchet) differentiable in measure in Ω;
(ii) f is Gâteaux differentiable in measure in Ω;
(iii) f is directionally differentiable in measure in Ω.

Moreover, the derivative in measure g is linked to the directional derivative in measure W by

$$g(x)\cdot y = W(x,y) \qquad \text{for } \mathscr{L}^{2d}\text{-a.e. } (x,y) \in \Omega \times \mathbb{R}^d.$$

Proof. We start with some preliminary remarks. It is obvious that if g is the derivative in measure of f then g is also the Gâteaux derivative in measure of f, and that these derivatives in measure are unique, up to $\mathscr{L}^d$-negligible sets.

Moreover, Gâteaux differentiability in measure implies directional differentiability in measure, with $W(x,y) = g(x)y$: this fact is an immediate consequence of Lemma 6.1.

The harder implication is the one from directional differentiability in measure to Fréchet differentiability in measure. We need the following lemma (see [71, 56]): *let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function such that*

$$f(y+z) = f(y) + f(z) \text{ for } \mathscr{L}^{2d}\text{-a.e. } (y,z) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Then there exists $a \in \mathbb{R}^d$ *such that* $f(y) = a \cdot y$ *for* $\mathscr{L}^d$*-a.e.* $y \in \mathbb{R}^d$.

Step 1. Using the above mentioned lemma, we show the "a.e. linearity" of $W(x,\cdot)$. We define

$$\begin{aligned}
F_\delta(x,y,z) &= \frac{f(x+\delta y) - f(x) - \delta W(x,y)}{\delta}, \\
G_\delta(x,y,z) &= \frac{f(x+\delta y+\delta z) - f(x+\delta y) - \delta W(x+\delta y,z)}{\delta}, \\
H_\delta(x,y,z) &= \frac{-f(x+\delta y+\delta z) + f(x) + \delta W(x,y+z)}{\delta}.
\end{aligned}$$

Then

$$F_\delta \to 0 \quad \text{as } \delta \to 0, \tag{35}$$

$$G_\delta \to 0 \quad \text{as } \delta \to 0, \tag{36}$$

$$H_\delta \to 0 \quad \text{as } \delta \to 0 \tag{37}$$

locally in measure with respect to (x,y,z). Indeed, (35) and (37) are easy applications of Lemma 6.1. To verify (36) we need also the following shift argument. Let $\Omega'' \subset\subset \Omega' \subset\subset \Omega$. For small δ we obtain

$$\begin{aligned}
&\mathscr{L}^{3d}\Big(\big\{(x,y,z) \in \Omega'' :\ G_\delta(x,y,z)) \geq \gamma\big\}\Big) \\
&\qquad = \mathscr{L}^{3d}\Big(\big\{(x,y,z) \in \Omega'' :\ |f(x+\delta y+\delta z) - f(x+\delta y) - \delta W(x+\delta y,z)| \geq \gamma\delta\big\}\Big) \\
&\qquad \leq \mathscr{L}^{3d}\Big(\big\{(x',y,z) \in \Omega' : |f(x'+\delta z) - f(x') - \delta W(x',z)| \geq \gamma\delta\big\}\Big) \to 0
\end{aligned}$$

where we used the change of variables $x' = x + \delta y$. This proves (36). From Lemma 6.1 we infer that

$$W(x+\delta y,z) - W(x,z) \to 0 \quad \text{as } \delta \to 0.$$

Adding the three identities we obtain that

$$W(x,y+z) - W(x,y) - W(x,z) = 0 \quad \text{for } \mathscr{L}^{3d}\text{-a.e. } (x,y,z) \in \Omega \times \mathbb{R}^d \times \mathbb{R}^d.$$

Therefore, for $\mathscr{L}^d$-a.e. x, $W(x,\cdot)$ is representable as a linear function. This concludes Step 1.

Step 2. Let α_d be the measure of the unit ball in $\mathbb{R}^d$. We fix a set $\Omega' \subset\subset \Omega$ and $\varepsilon, \gamma > 0$. We find $\delta_1 > 0$ such that

$$B(x, 2\delta_1) \subset \Omega, \qquad x \in \Omega'.$$

Consider the sets

$$A_\delta(x) = \{y \in B(0,2) \colon |f(x+\delta y) - f(x) - \delta g(x)\cdot y| \geq \gamma\delta\},$$
$$0 < \delta < \delta_1, \quad x \in \Omega'.$$

Since $(x,y) \mapsto g(x)\cdot y$ is a directional derivative of f in measure, by the Fubini theorem,

$$\int_{\Omega'} \mathscr{L}^d\left(A_\delta(x)\right) dx \to 0 \quad \text{as } \delta \to 0.$$

Therefore there exists $\delta_2 > 0$ such that for $0 < \delta < \delta_2$ we have

$$\mathscr{L}^d\left(\left\{x \in \Omega' \colon \mathscr{L}^d(A_\delta(x)) \geq \frac{\alpha_d}{4}\right\}\right) \leq \varepsilon.$$

By Lusin theorem, there exists a compact set $K \subset \Omega$ such that $\mathscr{L}^d(\Omega \setminus K) < \varepsilon$ and $g|_K$ is continuous. By the uniform continuity, there exists $\delta_3 > 0$ such that for $0 < \delta < \delta_3$ we have

$$x, x' \in K,\ |x - x'| < \delta \quad \Longrightarrow \quad |g(x) - g(x')| < \gamma.$$

Let $h \in \mathbb{R}^d$, $|h| = r < \frac{1}{2}\min\{\delta_1, \delta_2, \delta_3\}$. Consider the sets

$$\begin{aligned} E &:= \{x \in \Omega' \colon\ \mathscr{L}^d(A_r(x)) \geq \tfrac{1}{4}\alpha_d\},\\ E' &:= \{x \in \Omega' \colon\ \mathscr{L}^d(A_r(x+h)) \geq \tfrac{1}{4}\alpha_d\},\\ F &:= \Omega' \setminus K,\\ F' &:= \{x \in \Omega' \colon\ x+h \notin K\}. \end{aligned}$$

Then

$$|E \cup E' \cup F \cup F'| \leq 4\varepsilon.$$

Let $x \in \Omega \setminus (E \cup E' \cup F \cup F')$. Then both x and $x+h$ belong to K and

$$\begin{aligned} \mathscr{L}^d\left(B(0,1) \cap A_r(x)\right) &\leq \tfrac{1}{4}\alpha_d,\\ \mathscr{L}^d\left(\{y \in B(0,1) \colon\ y - \tfrac{h}{r} \in A_r(x+h)\}\right) &\leq \tfrac{1}{4}\alpha_d. \end{aligned}$$

Thus there exists $y \in B(0,1) \setminus A_r(x)$ such that $y' := y - \frac{h}{r} \notin A_r(x+h)$. Since $|y'| < 2$, the fact that $y' \notin A_r(x+h)$ means that $x + h + ry' \in \Omega$ and

$$|f(x+h+ry') - f(x+h) - rg(x+h)\cdot y'| \leq \gamma r.$$

Since $h + ry' = ry$, we have

$$\begin{aligned} &|f(x+h) - f(x) - g(x)\cdot h|\\ &\leq |f(x+h) - f(x+h+ry') + rg(x+h)\cdot y'|\\ &\quad + |f(x+ry) - f(x) - rg(x)\cdot y| + |r(g(x) - g(x+h))\cdot y'|\\ &\leq \gamma r + \gamma r + \gamma r|y'| \leq 4|h|\gamma. \end{aligned}$$

This shows that
$$\frac{f(x+h)-f(x)-g(x)\cdot h}{|h|} \to 0$$
in measure w.r.t. x as $h \to 0$. □

Now we introduce another more classical weak differentiability property (extensively studied, for instance, in [68]). It has still a measure-theoretic character, but unlike differentiability in measure it has a pointwise meaning.

Definition 6.5 (Approximate differentiability). Let $f : \Omega \to \mathbb{R}$ be a measurable function and let $x \in \Omega$. We say that $a \in \mathbb{R}^d$ is an approximate derivative of f at x if
$$\left\{h\colon \frac{|f(x+h)-f(x)-a\cdot h|}{|h|} > \varepsilon\right\} \quad \text{has zero Lebesgue density at 0 for any } \varepsilon > 0.$$

As we are concerned here with convergence in measure, it is worth mentioning that approximate differentiability at x is equivalent to the convergence
$$\lim_{\delta\to 0} \frac{f(x+\delta y)-f(x)-\delta a\cdot y}{\delta} = 0 \quad \text{locally in measure w.r.t. } y \in \mathbb{R}^d.$$

The following proposition shows that functions which are approximately differentiable on a measurable set A essentially coincide with functions that can be approximated, in the Lusin sense, by Lipschitz maps.

Theorem 6.6 (Lusin theorem for approximately differentiable maps). *Let $f : \Omega \to \mathbb{R}$. Assume that there exists a sequence of measurable sets $A_n \subset \Omega$ such that $\mathscr{L}^d(\Omega \setminus \cup_n A_n) = 0$ and $f|_{A_n}$ is Lipschitz for any n. Then f is approximately differentiable at $\mathscr{L}^d$-a.e. $x \in \Omega$.*
Conversely, if f is approximately differentiable at all points of $\Omega' \subset \Omega$, we can write Ω' as a countable union of sets A_n such that $f|_{A_n}$ is Lipschitz for any n (up to a redefinition in a $\mathscr{L}^d$-negligible set).

Proof. With no loss of generality we can assume that the sets A_n are pairwise disjoint. By Mc Shane Lipschitz extension theorem we can find Lipschitz functions $g_n : \mathbb{R}^d \to \mathbb{R}$ extending $f|_{A_n}$. By Lebesgue differentiation theorem and Rademacher theorem, $\mathscr{L}^d$-a.e. $x \in A_n$ is both a point of density 1 of A_n and a differentiability point of g_n. We claim that at any of these points x the function f is approximately differentiable, with approximate differentiable equal to $\nabla g_n(x)$. Indeed, it suffices to notice that the difference quotients of f and of g_n may differ only on $\Omega \setminus A_n$, that has zero density at x.

Clearly, as $\mathscr{L}^d$-a.e. $x \in \Omega$ has this property for some n, so this proves the $\mathscr{L}^d$-a.e. approximate differentiability of f.

The converse statement is proved in Theorem 3.1.16 of [68]. □

The following theorem shows that, among all the differentiability properties considered in this section, $\mathscr{L}^d$-a.e. approximate differentiability is the stronger one.

Even on the real line, an example built in [17] shows that differentiability in measure does not imply $\mathscr{L}^1$-a.e. approximate differentiability (in fact, the function built in [17] is nowhere approximately differentiable).

Theorem 6.7. *Suppose that $f : \Omega \to \mathbb{R}$ is approximately differentiable $\mathscr{L}^d$-a.e. in Ω. Then f is differentiable in measure.*

Proof. Let g be the approximate derivative of f. By the previous result there exist Lipschitz functions f_j and pairwise disjoint measurable sets $A_j \subset \Omega$ such that

$$f_j = f \text{ on } A_j,$$
$$\nabla f_j = g \text{ on } A_j,$$
$$\mathscr{L}^d\left(\Omega \setminus \bigcup_j A_j\right) = 0.$$

We fix $\Omega' \subset\subset \Omega$ and $\gamma > 0$ and find $\delta > 0$ such that $B(x,\delta) \subset \Omega$ for each $x \in \Omega'$. We set

$$E(h) := \{x \in \Omega' : \, |f(x+h) - f(x) - g(x)h| \geq \gamma |h|\}, \quad h \in B(0,\delta).$$

Then

$$E(h) \subset N \cup \bigcup_j (E_j(h) \cup F_j(h)),$$

where $\mathscr{L}^d(N) = 0$ and

$$E_j(h) = \{x \in \Omega' \cap A_j : \; x + h \notin A_j\},$$
$$F_j(h) = \{x \in \Omega' \cap A_j : \; |f_j(x+h) - f_j(x) - \nabla f_j(x)h| \geq \gamma |h|\}.$$

Then

$$\mathscr{L}^d(E_j(h)) \to 0 \quad \text{as } h \to 0$$

by the L^1_{loc} continuity of translations. Using the differentiability of f_j on A_j, we obtain that also

$$\mathscr{L}^d(F_j(h)) \to 0 \quad \text{as } h \to 0.$$

Choose $\varepsilon > 0$. Since $E_j(h) \cup F_j(h) \subset A_j$ and $\sum_j \mathscr{L}^d(A_j) \leq \mathscr{L}^d(\Omega) < \infty$, we can find an index m independent of h such that

$$\sum_{j=m+1}^{\infty} \mathscr{L}^d\left(E_j(h) \cup F_j(h)\right) \leq \varepsilon.$$

Then

$$\limsup_{h \to 0} \mathscr{L}^d(E(h)) \leq \sum_{j=1}^{m} \limsup_{h \to 0} \mathscr{L}^d\left(E_j(h) \cup F_j(h)\right) + \varepsilon = \varepsilon.$$

It follows that $\mathscr{L}^d(E(h)) \to 0$ as required. □

7 Differentiability of the Flow in the $W^{1,1}$ Case

In this section we discuss the differentiability properties of the $\mathscr{L}$-Lagrangian flow associated to $W^{1,1}$ vector fields, briefly describing the results obtained in [73]. Notice that no differentiability property is presently known in the BV case; on the other hand, in the $W^{1,p}$ case, with $p > 1$, much stronger results are available [48], and we will present them in the next section.

The main theorem in [73] is the following:

Theorem 7.1. *Let $\boldsymbol{b} \in L^1\left([0,T]; W^{1,1}_{\rm loc}(\mathbb{R}^d;\mathbb{R}^d)\right)$ be satisfying*

(i) $\dfrac{|\boldsymbol{b}|}{1+|x|} \in L^1\left([0,T]; L^1(\mathbb{R}^d)\right) + L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right)$*;*

(ii) $[\mathrm{div}\,\boldsymbol{b}_t] \in L^1\left([0,T]; L^\infty(\mathbb{R}^d)\right)$*;*

and let $\boldsymbol{X}(t,x)$ be the corresponding $\mathscr{L}$-Lagrangian flow, given by Theorem 5.7. Then for all $t \in [0,T]$ there exists a measurable function $\boldsymbol{W}_t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\frac{\boldsymbol{X}(t,x+\varepsilon y) - \boldsymbol{X}(t,x) - \varepsilon \boldsymbol{W}_t(x,y)}{\varepsilon} \to 0 \quad \textit{locally in measure in } \mathbb{R}^d_x \times \mathbb{R}^d_y$$

as $\varepsilon \downarrow 0$.

The result actually stated in [73] is slightly stronger, as the convergence above is also shown to be uniform with respect to time. Having in mind the terminology and the results of the previous section, the result can also be rephrased as follows: for all $t \in [0,T]$ there exists a matrix-valued measurable function $\boldsymbol{G}_t : \mathbb{R}^d \to M^{d\times d}$ (i.e. the derivative in measure) such that

$$\frac{\boldsymbol{X}(t,x+h) - \boldsymbol{X}(t,x) - \boldsymbol{G}_t(x)h}{|h|} \to 0 \qquad \text{locally in measure in } \mathbb{R}^d_x \text{ as } h \to 0.$$

The link between $\boldsymbol{G}_t$ and $\boldsymbol{W}_t$ is given by $\boldsymbol{G}_t(x)(y) = \boldsymbol{W}_t(x,y)$ for $\mathscr{L}^{2d}$-a.e. (x,y). So, $\boldsymbol{G}_t$ can be interpreted as the derivative of the flow map $\boldsymbol{X}(t,\cdot)$.

The strategy of the proof in [73] is to look at the behaviour of the $2d$-dimensional flows $\boldsymbol{Y}^\varepsilon$ arising from the difference quotients of $\boldsymbol{X}$:

$$\boldsymbol{Y}^\varepsilon(t,x,y) := \frac{\boldsymbol{X}(t,x+\varepsilon y) - \boldsymbol{X}(t,x)}{\varepsilon}.$$

It is immediate to check (see also [15]) that $\boldsymbol{Y}^\varepsilon$ are $\mathscr{L}$-Lagrangian flows relative to the vector fields $\boldsymbol{B}^\varepsilon$ defined by

$$\boldsymbol{B}^\varepsilon(t,x,y) = \left(\boldsymbol{b}(t,x), \frac{\boldsymbol{b}(t,x+\varepsilon y) - \boldsymbol{b}(t,x)}{\varepsilon}\right).$$

Therefore, it is natural to expect that their limit $\boldsymbol{Y}$ (if any) should be a flow relative to the limit vector field

$$\boldsymbol{B}(t,x,y) = (\boldsymbol{b}(t,x), \nabla_x \boldsymbol{b}(t,x)y)\,.$$

Notice that $\operatorname{div}\boldsymbol{B}_t(x,y) = 2\operatorname{div}\boldsymbol{b}_t(x)$, and therefore

$$\operatorname{div}\boldsymbol{B}_t \in L^1\left([0,T];L^\infty(\mathbb{R}^{2d})\right).$$

On the other hand, the last d components of $\boldsymbol{B}$ have no regularity with respect to the x-variable. However, in this special case, an anisotropic smoothing argument (by a regularization in the x variable much faster than the one in the y variable), based on the special structure of this vector field (see also [25, 74]), still guarantees that bounded solutions to the continuity equation with velocity field $\boldsymbol{B}$ are renormalizable. Moreover, using the renormalization property, a variant of the argument used in Theorem 5.4 shows that the continuity equation with $\boldsymbol{B}$ as vector field satisfies the comparison principle in the class

$$\mathscr{L}^* := \mathscr{L} \cap L^\infty\left([0,T];L^\infty_{\mathrm{loc}}(\mathbb{R}^d_x;L^1(\mathbb{R}^d_y))\right),$$

where $\mathscr{L}$ is defined as in (18) (in $2d$ space dimensions). The reason for this restriction to the smaller space $\mathscr{L}^*$ is the fact that $|\boldsymbol{B}|/(1+|x|+|y|)$ in general does not belong to

$$L^1\left([0,T];L^1(\mathbb{R}^d_x\times\mathbb{R}^d_y)\right) + L^1\left([0,T];L^\infty(\mathbb{R}^d_x\times\mathbb{R}^d_y)\right),$$

because the last d components do not tend to 0 as $|y|\to\infty$ while x is kept fixed (and their limit is possibly unbounded as a function of x). If $\boldsymbol{b}$ satisfies condition (ii) above, the last d components $\boldsymbol{B}^2$ of $\boldsymbol{B}$ satisfy instead

$$\frac{|\boldsymbol{B}^2|}{1+|y|} \in L^1\left([0,T];L^1_{\mathrm{loc}}(\mathbb{R}^d_x, L^1(\mathbb{R}^d_y)+L^\infty(\mathbb{R}^d_y))\right).$$

For this reason, a weaker growth condition on $\boldsymbol{B}$ turns into a stronger growth condition on w.

Then, the renormalization property ensures the well posedness of the continuity equation and therefore (much like as in Theorem 4.8 and Theorem 5.4) that the $\mathscr{L}^*$-Lagrangian flow relative to $\boldsymbol{B}$ is unique. A smoothing argument (see [73] for details) proves its existence even within the smaller class $\mathscr{L}^*$. So, denoting by $\boldsymbol{Y}$ the $\mathscr{L}^*$-Lagrangian flow relative to $\boldsymbol{B}$, we can represent it as

$$\boldsymbol{Y}(t,x,y) = (\boldsymbol{X}(t,x), \boldsymbol{W}(t,x,y))$$

for some map $\boldsymbol{W}$. Finally, the same argument used for the existence of $\boldsymbol{Y}$ shows that $\boldsymbol{Y}$ is the limit, with respect to local convergence in measure (uniform w.r.t. time) of $\boldsymbol{Y}^\varepsilon$: this is due to the fact that $\boldsymbol{B}^\varepsilon$ have properties analogous to the ones of $\boldsymbol{B}$ (in particular they have uniformly bounded divergences) and converge to $\boldsymbol{B}$.

This leads to the proof of Theorem 7.1.

8 Differentiability and Compactness of the Flow in the $W^{1,p}$ Case

In this section we present some recent results, obtained in [48], relative to the approximate differentiability and to the Lipschitz properties of regular Lagrangian flows associated to $W^{1,p}$ vector fields, with $p > 1$. The first results relative to the approximative differentiability of the flow have been obtained in [15]. An important fact that we want to remark from the beginning is that the approach of [15] and [48] is completely different from the one of [73] described in the previous section: these two papers are not based on the theory of renormalized solutions, but new kind of estimates are introduced. The general idea, which we are going to explain with all the details in the following, is trying to find estimates for the spatial gradient of the flow in terms of bounds on the derivative of the vector field.

We start by recalling some basic facts about the theory of *maximal functions*. These tools will be used throughout all this section. We begin with the definition of the local maximal function.

Definition 8.1 (Maximal function). Let μ be a (vector-valued) measure with locally finite total variation. For every $\lambda > 0$, we define the *(λ-local) maximal function* of μ as

$$M_\lambda \mu(x) = \sup_{0<r<\lambda} \frac{|\mu|(B_r(x))}{\mathscr{L}^d(B_r(x))} \qquad x \in \mathbb{R}^d.$$

When $\mu = f\mathscr{L}^d$, where f is a function in $L^1_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^m)$, we will use the notation $M_\lambda f$ for $M_\lambda \mu$.

In the following we will use a lot of times the following two lemmas about maximal functions. Their proof is classical and can be found for example in [87]. The first lemma shows that it is possible to control the L^p norm of a maximal function with the L^p norm of the function itself, in the case $p > 1$; however, this estimate is *false* in the case $p = 1$. The second lemma will be used to estimate the difference quotients of the vector field by means of the maximal function of the spatial derivative of the vector field.

Lemma 8.2. *The local maximal function of μ is finite for $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ and, for every $R > 0$, we have*

$$\int_{B_R(0)} M_\lambda f(y)\,dy \le c_{d,R} + c_d \int_{B_{R+\lambda}(0)} |f(y)| \log(2+|f(y)|)\,dy \quad \forall \lambda > 0.$$

For $p > 1$ we have

$$\int_{B_R(0)} (M_\lambda f(y))^p\,dy \le c_{d,p} \int_{B_{R+\lambda}(0)} |f(y)|^p\,dy \qquad \forall \lambda > 0.$$

Lemma 8.3. *If $u \in BV(\mathbb{R}^d;\mathbb{R}^m)$ then there exists an $\mathscr{L}^d$-negligible set $N \subset \mathbb{R}^d$ such that*

$$|u(x)-u(y)| \leq c_d|x-y|\,(M_\lambda Du(x)+M_\lambda Du(y))$$

for $x,y \in \mathbb{R}^d \setminus N$ *with* $|x-y| \leq \lambda$.

We also recall Chebyshev inequality for a measurable function $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$:

$$\mathscr{L}^d\big(\{|f|>t\}\big) \leq \frac{1}{t}\int_{\{|f|>t\}} |f(x)|\,dx \leq \frac{\mathscr{L}^d\big(\{|f|>t\}\big)^{1/q}}{t}\|f\|_{L^p(\Omega)},$$

which immediately implies

$$\mathscr{L}^d\big(\{|f|>t\}\big)^{1/p} \leq \frac{\|f\|_{L^p(\Omega)}}{t}. \tag{38}$$

In all this section we will make the following assumptions on the vector field $\boldsymbol{b} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$:

(A) $\boldsymbol{b} \in L^1([0,T];W^{1,p}(\mathbb{R}^d;\mathbb{R}^d))$ for some $p>1$;
(B) $\boldsymbol{b} \in L^\infty([0,T]\times\mathbb{R}^d;\mathbb{R}^d)$;
(C) $[\mathrm{div}_x\boldsymbol{b}]^- \in L^1([0,T];L^\infty(\mathbb{R}^d))$.

We remark that condition (B) could be relaxed to

$$\frac{|\boldsymbol{b}(t,x)|}{1+|x|} \in L^1([0,T];L^1(\mathbb{R}^d)) + L^1([0,T];L^\infty(\mathbb{R}^d)),$$

getting some slightly weaker results (due to the possible unboundedness of the velocity field), but in this introductory presentation we prefer to avoid these technicalities and focus on the case of a uniformly bounded vector field. We also notice that, under this assumption, the global $W^{1,p}$ hypothesis is not essential: thanks to the finite speed of propagation, we could truncate our vector field out of a ball, then getting the same results with just a $W^{1,p}_{\mathrm{loc}}$ hypothesis. However, we prefer to assume this global condition, mainly in order to simplify typographically some estimates. We refer to [48] for these more general hypotheses, the main modification being an estimate of the superlevels of the flow.

We recall that in this context the results of Sect. 5 read as follow. For every vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C) there exists a unique regular Lagrangian flow $\boldsymbol{X}$, that is a measurable map $\boldsymbol{X} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ which satisfies the following two properties:

(i) For $\mathscr{L}^d$-a.e. $x \in \mathbb{R}^d$ the map $t \mapsto \boldsymbol{X}(t,x)$ is an absolutely continuous solution of $\dot\gamma(t) = \boldsymbol{b}(t,\gamma(t))$ with $\gamma(0)=x$ (the ODE is fulfilled in the integral sense or, equivalently, in the a.e. sense);
(ii) There exists a constant C independent of t such that $\boldsymbol{X}(t,\cdot)_\#\mathscr{L}^d \leq C\mathscr{L}^d$ for every $t \in [0,T]$.

For every vector field satisfying assumption (C), the quantity

$$C = \exp\left(\int_0^T \|[\mathrm{div}_x\boldsymbol{b}(t,\cdot)]^-\|_{L^\infty(\mathbb{R}^d)}\,dt\right)$$

satisfies the inequality in the second property of the regular Lagrangian flow. However, we remark the fact that all our estimates will depend on the *compressibility constant* L of $\boldsymbol{X}$, i.e. the best constant C for which (ii) holds, rather than on the $L^1(L^\infty)$ norm of $[\mathrm{div}_x \boldsymbol{b}]^-$. This is an important remark in the context of the compactness theorem (see Theorem 8.7): some compactness results under bounds on the divergence were already available in [61], while one of the merits of this new approach is this weaker requirement.

The starting point of the estimates given in [48] is already present in [15] (and, at least in a formal way, in [61]): in a smooth context, we can control the time derivative $\frac{d}{dt}\log(|\nabla \boldsymbol{X}|)$ with $|\nabla \boldsymbol{b}|(\boldsymbol{X})$. The strategy of [15] allows to make this remark rigorous: it is possible to consider some integral quantities which contain a discretization of the space gradient of the flow and to prove some estimates along the flow. Then, the application of Egorov theorem allows the passage from integral estimates to pointwise estimates on big sets, and from this it is possible to recover Lipschitz regularity on big sets. However, the application of Egorov theorem implies a loss of quantitative informations: this strategy does not allow a control of the Lipschitz constant in terms of the size of the "neglected" set.

Starting from this results, the main point of [48] is a modification of the estimates in such a way that quantitative informations are not lost. We define (for $p > 1$ and $R > 0$) the following integral quantity:

$$A_p(R,\boldsymbol{X}) := \left[\int_{B_R(0)} \left(\sup_{0\le t\le T} \sup_{0<r<2R} \right.\right.$$
$$\left.\left. \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r} + 1 \right) dy \right)^p dx \right]^{1/p}. \tag{39}$$

Heuristically, one can view $|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|/r$ as a "discrete gradient" of the flow $\boldsymbol{X}$. Then the quantity $A_p(R,\boldsymbol{X})$ is constructed averaging this discrete gradient over balls of radius r, asking some uniformity with respect to t and r and finally integrating on a bounded set with respect to the second variable. We are now going to give some quantitative estimates of the quantity $A_p(R,\boldsymbol{X})$ in the case when the map $\boldsymbol{X}$ is the regular Lagrangian flow associated to a vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C). The estimate will depend only on the $L^1(L^p)$ norm of the derivative of $\boldsymbol{b}$ and on the compressibility constant L. In all the following computations we will denote by $c_{q_1,\dots,q_n}$ universal constants which depend only on the parameters $q_1,\dots,q_n$ and which can change from line to line. To simplify the notation we will also denote by L^p and $L^1(L^p)$ the global spaces $L^p(\mathbb{R}^d)$ and $L^1([0,T];L^p(\mathbb{R}^d))$, respectively. Out of the smooth setting, all the computations are easily justified by condition (i) in the definition of regular Lagrangian flow; however, the reader could check the estimates in the smooth case, and then obtain the general case simply by an approximation procedure (based on Theorem 5.10).

Proposition 8.4. *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C). Denote by $\boldsymbol{X}$ its regular Lagrangian flow and let L be the compressibility constant of the flow. Then we have*

$$A_p(R,\boldsymbol{X}) \le C\left(R, L, \|D_x \boldsymbol{b}\|_{L^1(L^p)}\right).$$

Proof. For $0 \le t \le T$, $0 < r < 2R$ and $x \in B_R(0)$ define

$$Q(t,x,r) := \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right) dy.$$

With some easy computations we get

$$\begin{aligned}\frac{dQ}{dt}(t,x,r) \le & \fint_{B_r(x)} \left|\frac{d\boldsymbol{X}}{dt}(t,x)-\frac{d\boldsymbol{X}}{dt}(t,y)\right| (|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|+r)^{-1}\, dy \\ = & \fint_{B_r(x)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x))-\boldsymbol{b}(t,\boldsymbol{X}(t,y))|}{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|+r}\, dy. \end{aligned} \tag{40}$$

We now set $\tilde{R} = 4R + 2T\|\boldsymbol{b}\|_\infty$. Since we clearly have $|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)| \le \tilde{R}$, applying Lemma 8.3 we can estimate

$$\begin{aligned}&\frac{dQ}{dt}(t,x,r)\\ &\le c_d \fint_{B_r(x)} (M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,x))+M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,y)))\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|+r}\, dy \\ &\le c_d M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) + c_d \fint_{B_r(x)} M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dy.\end{aligned}$$

Integrating with respect to the time, we estimate

$$\sup_{0\le t\le T} Q(t,x,r) \le c + c_d \int_0^T M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,x))\, dt + c_d \int_0^T \fint_{B_r(x)} M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dydt. \tag{41}$$

Passing to the supremum for $0 < r < 2R$ we obtain

$$\begin{aligned}&\sup_{0\le t\le T}\sup_{0<r<2R} Q(t,x,r)\\ &\le c + c_d \int_0^T M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,x))\, dt + c_d \int_0^T \sup_{0<r<2R} \fint_{B_r(x)} M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dydt.\end{aligned}$$

Taking the L^p norm over $B_R(0)$ we get

$$\begin{aligned}A_p(R,\boldsymbol{X}) = & \left\|\sup_{0\le t\le T}\sup_{0<r<2R} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right) dy\right\|_{L^p_x(B_R(0))} \\ \le & \; c_{p,R} + c_d \left\|\int_0^T M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,x))\, dt\right\|_{L^p_x(B_R(0))} \\ & + c_d \left\|\int_0^T \sup_{0<r<2R} \fint_{B_r(x)} M_{\tilde{R}}D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dydt\right\|_{L^p_x(B_R(0))}.\end{aligned}$$

Recalling Lemma 8.2 and the definition of the compressibility constant L, the first integral can be estimated with

$$
\begin{aligned}
c_d \int_0^T \|M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x))\|_{L^p_x(B_R(0))}\,dt &\le c_d L^{1/p} \int_0^T \|M_{\tilde{R}} D\boldsymbol{b}(t,x)\|_{L^p_x(B_{R+T\|\boldsymbol{b}\|_\infty}(0))}\,dt \\
&\le c_{d,p} L^{1/p} \int_0^T \|D\boldsymbol{b}(t,x)\|_{L^p_x(B_{R+\tilde{R}+T\|\boldsymbol{b}\|_\infty}(0))}\,dt.
\end{aligned}
$$

The second integral can be estimated in a similar way with

$$
\begin{aligned}
& c_d \int_0^T \left\| \sup_{0<r<2R} \fint_{B_r(x)} \left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t,\boldsymbol{X}(t,\cdot))\right](y)\,dy \right\|_{L^p_x(B_R(0))} dt \\
&= c_d \int_0^T \|M_{2R}\left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t,\boldsymbol{X}(t,\cdot))\right](x)\|_{L^p_x(B_R(0))}\,dt \\
&\le c_{d,p} \int_0^T \|\left[(M_{\tilde{R}} D\boldsymbol{b}) \circ (t,\boldsymbol{X}(t,\cdot))\right](x)\|_{L^p_x(B_{3R}(0))}\,dt \\
&= c_{d,p} \int_0^T \|(M_{\tilde{R}} D\boldsymbol{b}) \circ (t,\boldsymbol{X}(t,x))\|_{L^p_x(B_{3R}(0))}\,dt \\
&\le c_{d,p} L^{1/p} \int_0^T \|M_{\tilde{R}} D\boldsymbol{b}(t,x)\|_{L^p_x(B_{3R+T\|\boldsymbol{b}\|_\infty}(0))}\,dt \\
&\le c_{d,p} L^{1/p} \int_0^T \|D\boldsymbol{b}(t,x)\|_{L^p_x(B_{3R+T\|\boldsymbol{b}\|_\infty+\tilde{R}}(0))}\,dt.
\end{aligned}
$$

Then we obtain the desired estimate for $A_p(R,\boldsymbol{X})$. □

This estimate implies in an easy way a Lusin-type approximation of the flow with Lipschitz maps. As we observed before the proof of the proposition, the main point is that the estimate is now quantitative, and this will imply a precise control of the Lipschitz constant, in terms of the measure of the set we are "neglecting" (this explicit control was one of the open problems stated in [15]). This will be the key point in the proof of the compactness theorem.

Theorem 8.5 (Lipschitz estimates). *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C) and denote by $\boldsymbol{X}$ its regular Lagrangian flow. Then, for every ε, $R > 0$, we can find a set $K \subset B_R(0)$ such that*

- $\mathscr{L}^d(B_R(0) \setminus K) \le \varepsilon$;
- *For any $0 \le t \le T$ we have*

$$
\mathrm{Lip}\,(\boldsymbol{X}(t,\cdot)|_K) \le \exp \frac{c_d A_p(R,\boldsymbol{X})}{\varepsilon^{1/p}},
$$

with $A_p(R,\boldsymbol{X})$ satisfying the estimate of Proposition 8.4.

Proof. Fix $\varepsilon > 0$ and $R > 0$. We apply Proposition 8.4 and equation (38) to obtain a constant

$$M = M(\varepsilon, p, A_p(R, \boldsymbol{X})) = \frac{A_p(R, \boldsymbol{X})}{\varepsilon^{1/p}}$$

and a set $K \subset B_R(0)$ with $\mathscr{L}^d(B_R(0) \setminus K) \leq \varepsilon$ and

$$\sup_{0 \leq t \leq T} \sup_{0 < r < 2R} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dy \leq M \qquad \forall x \in K.$$

This clearly means that

$$\fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dy \leq M$$
$$\text{for every } x \in K,\, t \in [0,T] \text{ and } r \in]0, 2R[.$$

Now fix $x, y \in K$. Clearly $|x - y| < 2R$. Set $r = |x - y|$ and compute

$$\begin{aligned}
\log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) &= \fint_{B_r(x) \cap B_r(y)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dz \\
\leq \fint_{B_r(x) \cap B_r(y)} &\log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,z)|}{r} + 1\right) + \log\left(\frac{|\boldsymbol{X}(t,y) - \boldsymbol{X}(t,z)|}{r} + 1\right) dz \\
\leq c_d \fint_{B_r(x)} &\log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,z)|}{r} + 1\right) dz \\
+ c_d \fint_{B_r(y)} &\log\left(\frac{|\boldsymbol{X}(t,y) - \boldsymbol{X}(t,z)|}{r} + 1\right) dz \\
\leq c_d M = \frac{c_d A_p(R, \boldsymbol{X})}{\varepsilon^{1/p}}&.
\end{aligned}$$

This implies that

$$|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)| \leq \exp\left(\frac{c_d A_p(R, \boldsymbol{X})}{\varepsilon^{1/p}}\right) |x - y| \qquad \text{for every } x, y \in K.$$

Therefore

$$\mathrm{Lip}(\boldsymbol{X}(t,\cdot)|_K) \leq \exp \frac{c_d A_p(R, \boldsymbol{X})}{\varepsilon^{1/p}}.$$

□

Recalling Theorem 6.6, the following corollary is an immediate consequence of the Lipschitz estimates.

Corollary 8.6 (Approximate differentiability of the flow). *Let $\boldsymbol{b}$ be a vector field satisfying assumptions (A), (B) and (C) and denote by $\boldsymbol{X}$ its regular Lagrangian flow. Then $\boldsymbol{X}(t,\cdot)$ is approximately differentiable $\mathscr{L}^d$-a.e. in $\mathbb{R}^d$, for every $t \in [0,T]$.*

Now we are going to present some new compactness results. The strategy of the proof is quite elementary: thanks to the Lipschitz estimates, the flows will be equi-continuous on large sets and this, together with the equi-boundedness in L^∞, will allow the use of Ascoli–Arzelà theorem "on big sets". From this, the compactness in L^1_{loc} easily follows.

The following theorem gives a partial answer to a conjecture raised by Alberto Bressan in [34]: in fact, Bressan's conjecture is the L^1 version of the L^p result that we are going to present. Presently this conjecture is still unsolved.

Theorem 8.7 (Compactness of the flow). *Let $\boldsymbol{b}_n$ be a sequence of smooth vector fields. Denote by $\boldsymbol{X}_n$ their regular Lagrangian flows and let L_n be the compressibility constant of the flow $\boldsymbol{X}_n$. Suppose that*

- *$|\boldsymbol{b}_n|$ are equi-bounded in $L^\infty([0,T]\times\mathbb{R}^d)$,*
- *$|D_x\boldsymbol{b}_n|$ are equi-bounded in $L^1(L^p)$ for some $p>1$,*
- *$\{L_n\}_n$ is a bounded sequence (in $\mathbb{R}$).*

Then the sequence $\{\boldsymbol{X}_n\}_n$ is relatively compact in $L^1_{\mathrm{loc}}([0,T]\times\mathbb{R}^n)$.

Proof. Fix $R>0$ and notice that, since the vector fields $\{\boldsymbol{b}_n\}_n$ are uniformly bounded in L^∞, the flows $\{\boldsymbol{X}_n\}_n$ are uniformly bounded in $L^\infty([0,T]\times B_R(0))$; denote by C_R a common bound in L^∞ for the flows. Applying Proposition 8.4 we obtain that the quantities $A_p(R,\boldsymbol{X}_n)$ are uniformly bounded with respect to n. Now, fix $j\in\mathbb{N}$ and apply Theorem 8.5 with $\varepsilon=1/j$ to obtain, for every $n\in\mathbb{N}$, a set $K_n\subset B_R(0)$ with $\mathscr{L}^d(B_R(0)\setminus K_n)\le 1/j$ such that

$$\mathrm{Lip}(\boldsymbol{X}_n(t,\cdot)|_{K_n}) \qquad \text{is uniformly bounded w.r.t. } n \text{ uniformly in } t.$$

Since $\frac{d}{dt}\boldsymbol{X}_n(t,x)=\boldsymbol{b}_n(t,\boldsymbol{X}_n(t,x))$, thanks to the equi-boundedness in L^∞ of the vector fields we also get that

$$\mathrm{Lip}(\boldsymbol{X}_n|_{[0,T]\times K_n}) \qquad \text{is uniformly bounded w.r.t. } n.$$

Now, by applying some classical results on the extension of Lipschitz maps, we can extend every $\boldsymbol{X}_n|_{[0,T]\times K_n}$ to a map $\tilde{\boldsymbol{X}}_{n,j}$ defined on $[0,T]\times B_R(0)$ in such a way that

$$\mathrm{Lip}(\tilde{\boldsymbol{X}}_{n,j}|_{[0,T]\times B_R(0)}) \le c_d \mathrm{Lip}(\boldsymbol{X}_n|_{[0,T]\times K_n}) \tag{42}$$

and

$$\|\tilde{\boldsymbol{X}}_{n,j}\|_{L^\infty([0,T]\times B_R(0))} \le \|\boldsymbol{X}_n\|_{L^\infty([0,T]\times B_R(0))}. \tag{43}$$

Applying Ascoli–Arzelà theorem and a standard diagonal argument we can find a subsequence (not relabeled) such that for each j the sequence $\{\tilde{\boldsymbol{X}}_{n,j}\}_n$ converges uniformly (and hence strongly in $L^1([0,T]\times B_R(0))$) to a map $\tilde{\boldsymbol{X}}_{\infty,j}$.

Notice that

$$\|\tilde{\boldsymbol{X}}_{n,j}-\boldsymbol{X}_n\|_{L^1([0,T]\times B_R(0))} \le \frac{2}{j}T\mathscr{L}^d(B_\rho(0))C_R.$$

Next, for any given $\varepsilon>0$ select j such that

$$\frac{2}{j}T\mathscr{L}^d(B_R(0))C_R \le \varepsilon/3,$$

and then $N > 0$ such that

$$\|\tilde{\boldsymbol{X}}_{i,j} - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \leq \varepsilon/3 \qquad \text{for all } i,k > N.$$

Hence for every $i,k > N$ we get

$$\begin{aligned}\|\boldsymbol{X}_i - \boldsymbol{X}_k\|_{L^1([0,T]\times B_R(0))} \leq \|\boldsymbol{X}_i - \tilde{\boldsymbol{X}}_{i,j}\|_{L^1([0,T]\times B_R(0))} + \|\tilde{\boldsymbol{X}}_{i,j} - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \\ + \|\boldsymbol{X}_k - \tilde{\boldsymbol{X}}_{k,j}\|_{L^1([0,T]\times B_R(0))} \leq \varepsilon.\end{aligned}$$

Hence $\{\boldsymbol{X}_n\}_n$ is a Cauchy sequence in $L^1([0,T] \times B_R(0))$. A second diagonal argument yields a subsequence, not relabeled, which is a Cauchy sequence in $L^1([0,T] \times B_l(0))$ for every $l \in \mathbb{N}$. □

Remark 8.8 (Existence of the regular Lagrangian flow). As a corollary we obtain a new proof of the existence of a regular Lagrangian flow associated to a bounded vector field $\boldsymbol{b}$ satisfying assumptions (A), (B) and (C). Indeed, approximating $\boldsymbol{b}$ by convolution with a positive convolution kernel, we get a sequence which satisfies the assumptions of the previous theorem, hence in the limit we get a flow associated to $\boldsymbol{b}$, thanks to the compactness in the strong topology. An analogous remark applies to Theorem 8.11.

With similar techniques it is also possible to show directly compactness in the $W^{1,1}$ case, under the assumption that the maximal functions of the derivatives of the vector fields $\boldsymbol{b}_n$ are uniformly bounded in $L^1(L^1)$. We start defining an integral quantity similar to the previous one, but without the supremum with respect to the radius r. For $R > 0$ and $0 < r < R/2$ fixed we set

$$a(r,R,\boldsymbol{X}) = \int_{B_R(0)} \sup_{0\leq t\leq T} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x) - \boldsymbol{X}(t,y)|}{r} + 1\right) dydx.$$

We first give a quantitative estimate for the quantity $a(r,R,\boldsymbol{X})$, similar to the previous one for $A_p(R,\boldsymbol{X})$.

Proposition 8.9. *Let $\boldsymbol{b}$ be a vector field satisfying the assumptions (A) and (C) and such that*

$$M_\lambda D\boldsymbol{b} \in L^1(L^1) \qquad \text{for every } \lambda > 0.$$

Denote by $\boldsymbol{X}$ its regular Lagrangian flow and let L be the compressibility constant of the flow. Then we have

$$a(r,R,\boldsymbol{X}) \leq C\left(R, L, \|M_{\tilde{R}} D_x \boldsymbol{b}\|_{L^1(L^1)}\right),$$

where $\tilde{R} = 3R/2 + 2T\|\boldsymbol{b}\|_\infty$.

Proof. We start as in the proof of Proposition 8.4, obtaining inequality (41) (but this time it is sufficient to set $\tilde{R} = 3R/2 + 2T\|\boldsymbol{b}\|_\infty$). Integrating with respect to x over $B_R(0)$, we obtain

$$\begin{aligned} a(r,R,\boldsymbol{X}) &= \int_{B_R(0)} \sup_{0\le t\le T} \fint_{B_r(x)} \log\left(\frac{|\boldsymbol{X}(t,x)-\boldsymbol{X}(t,y)|}{r}+1\right) dydx \\ &\le c_R + c_d \int_{B_R(0)} \int_0^T M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x))\, dtdx \\ &\quad + c_d \int_{B_R(0)} \int_0^T \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dydtdx. \end{aligned}$$

As in the previous computations, the first integral can be estimated with

$$c_d L \|M_{\tilde{R}} D\boldsymbol{b}\|_{L^1([0,T];L^1(B_{R+T\|\boldsymbol{b}\|_\infty}(0)))},$$

but this time we cannot bound the norm of the maximal function with the norm of the derivative. To estimate the last integral we compute

$$\begin{aligned} & c_d \int_{B_R(0)} \int_0^T \fint_{B_r(x)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,y))\, dydtdx \\ &= c_d \int_{B_R(0)} \int_0^T \fint_{B_r(0)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x+z))\, dzdtdx \\ &\le c_d \fint_{B_r(0)} \int_0^T \int_{B_R(0)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x+z))\, dxdtdz \\ &\le c_d \fint_{B_r(0)} \int_0^T L \int_{B_{3R/2+T\|\boldsymbol{b}\|_\infty}(0)} M_{\tilde{R}} D\boldsymbol{b}(t,w)\, dwdtdx \\ &= c_d L \|M_{\tilde{R}} D\boldsymbol{b}\|_{L^1([0,T];L^1(B_{3R/2+T\|\boldsymbol{b}\|_\infty}(0)))}. \end{aligned}$$

This concludes the proof of the estimate for $a(r,R,\boldsymbol{X})$. □

We recall the following classical criterion for strong compactness in L^p, since it will be used in the proof of the compactness theorem. For the proof of the lemma we refer for example to [35].

Lemma 8.10 (Riesz–Fréchet–Kolmogorov compactness criterion). *Let $\mathscr{F}$ be a bounded subset of $L^p(\mathbb{R}^N)$ for some $1 \le p < \infty$. Suppose that*

$$\lim_{|h|\to 0} \|f(\cdot - h) - f\|_{L^p} = 0 \qquad \textit{uniformly in } f \in \mathscr{F}.$$

Then $\mathscr{F}$ is relatively compact in $L^p_{\mathrm{loc}}(\mathbb{R}^N)$.

Theorem 8.11. *Let $\boldsymbol{b}_n$ be a sequence of smooth vector fields. Denote by $\boldsymbol{X}_n$ their regular Lagrangian flows and let L_n be the compressibility constant of the flow $\boldsymbol{X}_n$. Suppose that*

- $|\boldsymbol{b}_n|$ *are equi-bounded in* $L^\infty([0,T]\times\mathbb{R}^d)$,
- $M_\lambda D_x\boldsymbol{b}_n$ *are equi-bounded in* $L^1(L^1)$ *for every* $\lambda>0$,
- $\{L_n\}_n$ *is a bounded sequence (in* $\mathbb{R}$*).*

Then the sequence $\{\boldsymbol{X}_n\}_n$ *is relatively compact in* $L^1_{\mathrm{loc}}([0,T]\times\mathbb{R}^d)$.

Proof. We apply Proposition 8.9 to obtain that, under the assumptions of the Theorem, the quantities $a(r,R,\boldsymbol{X}_n)$ are uniformly bounded with respect to n. Now observe that, for $0\le z\le\bar R$ (with $\bar R=3R/2+2T\|\boldsymbol{b}\|_\infty$ as in Proposition 8.9), thanks to the concavity of the logarithm we have

$$\log\left(\frac{z}{r}+1\right)\ge\frac{\log\left(\frac{\bar R}{r}+1\right)}{\bar R}z.$$

Since $|\boldsymbol{X}_n(t,x)-\boldsymbol{X}_n(t,y)|\le\bar R$ this implies that

$$\begin{aligned}&\int_{B_R(0)}\sup_{0\le t\le T}\ \fint_{B_r(x)}|\boldsymbol{X}_n(t,x)-\boldsymbol{X}_n(t,y)|\,dydx\\ &\le\frac{\bar R}{\log\left(\frac{\bar R}{r}+1\right)}C\left(R,L_n,\|M_{\bar R}D_x\boldsymbol{b}_n\|_{L^1(L^1)}\right)\le g(r),\end{aligned}$$

where the function $g(r)$ does not depend on n and satisfies $g(r)\downarrow 0$ for $r\downarrow 0$. Changing the integration order this implies

$$\fint_{B_r(0)}\int_{B_R(0)}|\boldsymbol{X}_n(t,x)-\boldsymbol{X}_n(t,x+z)|\,dxdz\le g(r),$$

uniformly with respect to t and n.

Now notice the following elementary fact. There exists a dimensional constant $\alpha_d>0$ with the following property: if $A\subset B_1(0)$ is a measurable set with $\mathscr{L}^d(B_1(0)\setminus A)\le\alpha_d$, then $A+A\supset B_{1/2}(0)$. Then fix α_d as above and apply Chebyshev inequality for every n to obtain, for every $0<r<R/2$, a measurable set $K_{r,n}\subset B_r(0)$ with $\mathscr{L}^d(B_r(0)\setminus K_{r,n})\le\alpha_d\mathscr{L}^d(B_r(0))$ and

$$\int_{B_R(0)}|\boldsymbol{X}_n(t,x+z)-\boldsymbol{X}_n(t,x)|\,dx\le\frac{g(r)}{\alpha_d}\qquad\text{for every } z\in K_{r,n}.$$

For such a set $K_{r,n}$, thanks to the previous remark, we have that $K_{r,n}+K_{r,n}\supset B_{r/2}(0)$. Now let $v\in B_{r/2}(0)$ be arbitrary. For every n we can write $v=z_{1,n}+z_{2,n}$ with $z_{1,n},z_{2,n}\in K_{r,n}$. We can estimate the increment in the spatial directions as follows:

$$\begin{aligned}&\int_{B_{R/2}(0)}|\boldsymbol{X}_n(t,x+v)-\boldsymbol{X}_n(t,x)|\,dx\\ &=\int_{B_{R/2}(0)}|\boldsymbol{X}_n(t,x+z_{1,n}+z_{2,n})-\boldsymbol{X}_n(t,x)|\,dx\end{aligned}$$

$$
\begin{aligned}
&\le \int_{B_{R/2}(0)} |\boldsymbol{X}_n(t,x+z_{1,n}+z_{2,n}) - \boldsymbol{X}_n(t,x+z_{1,n})| + |\boldsymbol{X}_n(t,x+z_{1,n}) - \boldsymbol{X}_n(t,x)|\,dx \\
&\le \int_{B_R(0)} |\boldsymbol{X}_n(t,y+z_{2,n}) - \boldsymbol{X}_n(t,y)|\,dy + \int_{B_R(0)} |\boldsymbol{X}_n(t,x+z_{1,n}) - \boldsymbol{X}_n(t,x)|\,dx \\
&\le \frac{2g(r)}{\alpha_d}.
\end{aligned}
$$

Now notice that, by definition of regular Lagrangian flow, we have

$$
\frac{d\boldsymbol{X}_n}{dt}(t,x) = \boldsymbol{b}_n(t,\boldsymbol{X}_n(t,x)).
$$

Then we can estimate the increment in the time direction in the following way

$$
\begin{aligned}
|\boldsymbol{X}_n(t+h,x) - \boldsymbol{X}_n(t,x)| &\le \int_0^h \left| \frac{d\boldsymbol{X}_n}{dt}(t+s,x) \right| ds \\
&= \int_0^h |\boldsymbol{b}_n(t+s,\boldsymbol{X}_n(t+s,x))|\,ds \le h\|\boldsymbol{b}_n\|_\infty.
\end{aligned}
$$

Combining these two informations, for $(t_0,t_1) \subset\subset [0,T]$, $R > 0$, $v \in B_{r/2}(0)$ and $h > 0$ sufficiently small we can estimate

$$
\begin{aligned}
&\int_{t_0}^{t_1} \int_{B_{R/2}(0)} |\boldsymbol{X}_n(t+h,x+v) - \boldsymbol{X}_n(t,x)|\,dxdt \\
&\le \int_{t_0}^{t_1} \int_{B_{R/2}(0)} |\boldsymbol{X}_n(t+h,x+v) - \boldsymbol{X}_n(t+h,x)| + |\boldsymbol{X}_n(t+h,x) - \boldsymbol{X}_n(t,x)|\,dxdt \\
&\le T\frac{2g(r)}{\alpha_d} + \int_{t_0}^{t_1} \int_{B_{R/2}(0)} h\|\boldsymbol{b}_n\|_\infty\,dxdt \le T\frac{2g(r)}{\alpha_d} + c_d T R^d h \|\boldsymbol{b}_n\|.
\end{aligned}
$$

The thesis follows applying the Riesz–Fréchet–Kolmogorov compactness criterion, recalling that $\boldsymbol{b}_n$ are uniformly bounded in L^∞. □

We conclude this section showing another result obtained in [48] with techniques which are very similar to the ones described so far. It is a result of quantitative stability for regular Lagrangian flows: it improves the stability result given in Theorem 5.10 in the sense that it gives a rate of convergence of the flows in terms of convergence of the vector fields.

Theorem 8.12 (Quantitative stability). *Let $\boldsymbol{b}$ and $\tilde{\boldsymbol{b}}$ be vector fields satisfying assumptions (A), (B) and (C). Denote by $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ the respective regular Lagrangian flows and let L and $\tilde{L}$ be the compressibility constants of the two flows. Then, for every time $\tau \in [0,T]$, we have*

$$
\|\boldsymbol{X}(\tau,\cdot) - \tilde{\boldsymbol{X}}(\tau,\cdot)\|_{L^1(B_r(0))} \le C \left| \log\left(\|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_{L^1([0,\tau]\times B_R(0))} \right) \right|^{-1},
$$

where $R = r + T\|\boldsymbol{b}\|_\infty$ and the constant C only depends on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$.

Proof. Set $\delta := \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|_{L^1([0,T]\times B_R(0))}$ and consider the function

$$g(t) := \int_{B_r(0)} \log\left(\frac{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)|}{\delta} + 1\right) dx.$$

Clearly $g(0) = 0$ and after some standard computations we get

$$\begin{aligned}
g'(t) &\le \int_{B_r(0)} \left|\frac{d\boldsymbol{X}(t,x)}{dt} - \frac{d\tilde{\boldsymbol{X}}(t,x)}{dt}\right| \big(|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta\big)^{-1} dx \\
&= \int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \tilde{\boldsymbol{b}}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta} dx \\
&\le \frac{1}{\delta}\int_{B_r(0)} |\boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x)) - \tilde{\boldsymbol{b}}(t,\tilde{\boldsymbol{X}}(t,x))|\, dx \\
&\quad + \int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta} dx. \qquad (44)
\end{aligned}$$

We set $\tilde{R} = 2r + T(\|\boldsymbol{b}\|_\infty + \|\tilde{\boldsymbol{b}}\|_\infty)$ and we apply Lemma 8.3 to estimate the last integral as follows:

$$\begin{aligned}
&\int_{B_r(0)} \frac{|\boldsymbol{b}(t,\boldsymbol{X}(t,x)) - \boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))|}{|\boldsymbol{X}(t,x) - \tilde{\boldsymbol{X}}(t,x)| + \delta} dx \\
&\quad \le c_d \int_{B_r(0)} M_{\tilde{R}} D\boldsymbol{b}(t,\boldsymbol{X}(t,x)) + M_{\tilde{R}} D\boldsymbol{b}(t,\tilde{\boldsymbol{X}}(t,x))\, dx.
\end{aligned}$$

Inserting this estimate in (44), setting $\tilde{r} = r + T\max\{\|\boldsymbol{b}\|_\infty, \|\tilde{\boldsymbol{b}}\|_\infty\}$, changing variables in the integrals and using Lemma 8.2 we get

$$\begin{aligned}
g'(t) &\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + \big(\tilde{L} + L\big)\int_{B_{\tilde{r}}(0)} M_{\tilde{R}} D\boldsymbol{b}(t,y)\, dy \\
&\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + c_d \tilde{r}^{n-n/p}\big(\tilde{L} + L\big)\|M_{\tilde{R}} D\boldsymbol{b}(t,\cdot)\|_{L^p_x} \\
&\le \frac{\tilde{L}}{\delta}\int_{B_{r+T\|\tilde{\boldsymbol{b}}\|_\infty}(0)} |\boldsymbol{b}(t,y) - \tilde{\boldsymbol{b}}(t,y)|\, dy + c_{d,p} \tilde{r}^{n-n/p}\big(\tilde{L} + L\big)\|D\boldsymbol{b}(t,\cdot)\|_{L^p_x}.
\end{aligned}$$

For any $\tau \in [0,T]$, integrating the last inequality between 0 and τ we get

$$g(\tau) = \int_{B_r(0)} \log\left(\frac{|\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|}{\delta} + 1\right) dx \le C_1, \qquad (45)$$

where the constant C_1 depends on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$.

Next we fix a second parameter $\eta > 0$ to be chosen later. Using Chebyshev inequality we find a measurable set $K \subset B_r(0)$ such that $\mathscr{L}^d(B_r(0) \setminus K) \leq \eta$ and

$$\log\left(\frac{|\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|}{\delta} + 1\right) \leq \frac{C_1}{\eta} \qquad \text{for } x \in K.$$

Therefore we can estimate

$$\begin{aligned}
&\int_{B_r(0)} |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\,dx \\
&\leq \eta\left(\|\boldsymbol{X}(\tau,\cdot)\|_{L^\infty(B_r(0))} + \|\tilde{\boldsymbol{X}}(\tau,\cdot)\|_{L^\infty(B_r(0))}\right) + \int_K |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\,dx \\
&\leq \eta C_2 + c_d r^n \delta\left(\exp(C_1/\eta)\right) \leq C_3\left(\eta + \delta \exp(C_1/\eta)\right), \qquad (46)
\end{aligned}$$

with C_1, C_2 and C_3 which depend only on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$. Without loss of generality we can assume $\delta < 1$. Setting $\eta = 2C_1|\log\delta|^{-1} = 2C_1(-\log\delta)^{-1}$, we have $\exp(C_1/\eta) = \delta^{-1/2}$. Thus we conclude

$$\int_{B_r(0)} |\boldsymbol{X}(\tau,x) - \tilde{\boldsymbol{X}}(\tau,x)|\,dx \leq C_3\left(2C_1|\log\delta|^{-1} + \delta^{1/2}\right) \leq C|\log\delta|^{-1}, \qquad (47)$$

where C depends only on T, r, $\|\boldsymbol{b}\|_\infty$, $\|\tilde{\boldsymbol{b}}\|_\infty$, L, $\tilde{L}$, and $\|D_x\boldsymbol{b}\|_{L^1(L^p)}$. This completes the proof. □

Remark 8.13 (Uniqueness of the regular Lagrangian flow). We observe that the previous theorem also gives a new proof of the uniqueness of the regular Lagrangian flow associated to a vector field $\boldsymbol{b}$ which satisfies assumptions (A), (B) and (C).

9 Bibliographical Notes and Open Problems

Section 3. The material contained in this section is classical. Good references are [64], Chap. 8 of [14], [27] and [61]. For the proof of the area formula, see for instance [13, 67, 68]. The proof of the second local variant, under the stronger assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|\,d\mu_t dt < \infty$, is given in Proposition 8.1.8 of [14]. The same proof works under the weaker assumption (6).

Section 4. Many ideas of this section, and in particular the idea of looking at measures in the space of continuous maps to characterize the flow and prove its stability, are borrowed from [6], dealing with *BV* vector fields. Later on, the arguments have been put in a more general form, independent of the specific class of vector fields under consideration, in [7]. Here we present the same version of [8].

The idea of a probabilistic representation is of course classical, and appears in many contexts (particularly for equations of diffusion type); to our knowledge the first reference in the context of conservation laws and fluid mechanics is [30], where a similar approach is proposed for building generalized geodesics in the space $\mathbf{G}$ of measure-preserving diffeomorphisms; this is related to Arnold's interpretation of the incompressible Euler equation as a geodesic in $\mathbf{G}$ (see also [31, 32, 33]): in this case the compact (but neither metrizable, nor separable) space $X^{[0,T]}$, with $X \subset \mathbb{R}^d$ compact, has been considered.

This approach is by now a familiar one also in optimal transport theory, where transport maps and transference plans can be thought in a natural way as measures in the space of minimizing geodesics [85], and in the so called irrigation problems, a nice variant of the optimal transport problem [23]. See also [20] for a similar approach within Mather's theory. The lecture notes [94] (see also the Appendix of [78]) contain, among several other things, a comprehensive treatment of the topic of measures in the space of action-minimizing curves, including at the same time the optimal transport and the dynamical systems case (this unified treatment was inspired by [22]). Another related reference is [58].

The superposition principle is proved, under the weaker assumption $\int_0^T \int_{\mathbb{R}^d} |\boldsymbol{b}_t|^p \, d\mu_t dt < +\infty$ for some $p > 1$, in Theorem 8.2.1 of [14], see also [79] for the extension to the case $p = 1$ and to the nonhomogeneous continuity equation. Very closely related results, relative to the representation of a vector field as the superposition of "elementary" vector fields associated to curves, appear in [86, 20].

Section 5. The definition of renormalized solution and the strong convergence of commutators are entirely borrowed from [61]. See also [62] for the relevance of this concept in connection with the existence theory for Boltzmann equation. The proof of the comparison principle assuming only an $L^1(L^1_{\mathrm{loc}})$ bound (instead of an $L^1(L^\infty)$ one, as in [61, 6]) on the divergence was suggested to us by G. Savaré. See also [40] for a proof, using radial convolution kernels, of the renormalization property for vector fields satisfying $D_i\boldsymbol{b}^j + D_j\boldsymbol{b}^i \in L^1_{\mathrm{loc}}$.

No general existence result for Sobolev (or even BV) vector fields seems to be known in the infinite-dimensional case: the only reference we are aware of is [24], dealing with vector-fields having an exponentially integrable derivative, extending previous results in [49, 50, 51]. Also the investigation of nonEuclidean geometries, e.g. Carnot groups and horizontal vector fields, could provide interesting results.

Finally, notice that the theory has a natural invariance, namely if $\boldsymbol{X}$ is a flow relative to $\boldsymbol{b}$, then $\boldsymbol{X}$ is a flow relative to $\tilde{\boldsymbol{b}}$ whenever $\{\tilde{\boldsymbol{b}} \neq \boldsymbol{b}\}$ is $\mathscr{L}^{1+d}$-negligible in $(0,T) \times \mathbb{R}^d$. So a natural question is whether the uniqueness "in the selection sense" might be enforced by choosing a canonical representative $\tilde{\boldsymbol{b}}$ in the equivalence class of $\boldsymbol{b}$: in other words we may think that, for a suitable choice of $\tilde{\boldsymbol{b}}$, the ODE $\dot{\gamma}(t) = \tilde{\boldsymbol{b}}_t(\gamma(t))$ has a unique absolutely continuous solution starting from x for $\mathscr{L}^d$-a.e. x.

Concerning the case of BV vector fields, the main idea of the proof of Theorem 5.14, i.e. the adaptation of the convolution kernel to the local behaviour of the vector field, has been used at various level of generality in [25, 77, 45] (see also [41, 42] for related results independent of this technique), until the general result obtained in [6].

The optimal regularity condition on $\boldsymbol{b}$ ensuring the renormalization property, and therefore the validity of the comparison principle in $\mathscr{L}_{\boldsymbol{b}}$, is still not known. New results, both in the Sobolev and in the BV framework, are presented in [10, 73, 74].

In [12] we investigate in particular the possibility to prove the renormalization property for nearly incompressible $BV_{\rm loc} \cap L^\infty$ fields $\boldsymbol{b}$: they are defined by the property that there exists a positive function ρ, with $\ln \rho \in L^\infty$, such that the space-time field $(\rho, \rho\boldsymbol{b})$ is divergence free. As in the case of the Keyfitz-Kranzer system, the existence a function ρ with this property seems to be a natural replacement of the condition $D_x \cdot \boldsymbol{b} \in L^\infty$ (and is actually implied by it); as explained in [9], a proof of the renormalization property in this context would lead to a proof of a conjecture, due to Bressan, on the compactness of flows associated to a sequence of vector fields bounded in $BV_{t,x}$.

Section 6. The material of this section is entirely taken from [17]. See Chap. 3 of [68] for a deep study of approximate differentiability and calculus with the class of approximately differentiable maps.

Section 7. Here we have presented the main result in [73].

Section 8. The first progress in the direction of proving approximate differentiability of the flow $\boldsymbol{X}(t,x)$ with respect to x has been achieved in [15]. Later on, these results have been substantially improved in [47], and for this reason we have chosen to present only the latter results and proofs in this section.

References

1. M.Aizenman: *On vector fields as generators of flows: a counterexample to Nelson's conjecture.* Ann. Math., **107** (1978), 287–296.
2. G.Alberti: *Rank-one properties for derivatives of functions with bounded variation.* Proc. Roy. Soc. Edinburgh Sect. A, **123** (1993), 239–274.
3. G.Alberti & L.Ambrosio: *A geometric approach to monotone functions in $\mathbb{R}^n$.* Math. Z., **230** (1999), 259–316.
4. G.Alberti & S.Müller: *A new approach to variational problems with multiple scales.* Comm. Pure Appl. Math., **54** (2001), 761–825.
5. F.J.Almgren: *The theory of varifolds – A variational calculus in the large,* Princeton University Press, 1972.
6. L.Ambrosio: *Transport equation and Cauchy problem for BV vector fields.* Inventiones Mathematicae, **158** (2004), 227–260.
7. L.Ambrosio: *Lecture notes on transport equation and Cauchy problem for BV vector fields and applications.* Preprint, 2004 (available at http://cvgmt.sns.it).
8. L.Ambrosio: *Lecture notes on transport equation and Cauchy problem for non-smooth vector fields and applications.* Preprint, 2005 (available at http://cvgmt.sns.it).
9. L.Ambrosio, F.Bouchut & C.De Lellis: *Well-posedness for a class of hyperbolic systems of conservation laws in several space dimensions.* Comm. PDE, **29** (2004), 1635–1651.
10. L.Ambrosio, G.Crippa & S.Maniglia: *Traces and fine properties of a BD class of vector fields and applications.* Ann. Sci. Toulouse, **XIV** (4) (2005), 527–561.
11. L.Ambrosio & C.De Lellis: *Existence of solutions for a class of hyperbolic systems of conservation laws in several space dimensions.* International Mathematical Research Notices, **41** (2003), 2205–2220.
12. L.Ambrosio, C.De Lellis & J.Malý: *On the chain rule for the divergence of BV like vector fields: applications, partial results, open problems.* To appear in the forthcoming book by the AMS series in contemporary mathematics "Perspectives in Nonlinear Partial Differential Equations: in honor of Haim Brezis" (available at http://cvgmt.sns.it).

13. L.AMBROSIO, N.FUSCO & D.PALLARA: *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.
14. L.AMBROSIO, N.GIGLI & G.SAVARÉ: *Gradient flows in metric spaces and in the Wasserstein space of probability measures.* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.
15. L.AMBROSIO, M.LECUMBERRY & S.MANIGLIA: *Lipschitz regularity and approximate differentiability of the DiPerna–Lions flow.* Rendiconti del Seminario Fisico Matematico di Padova, **114** (2005), 29–50.
16. L.AMBROSIO, S.LISINI & G.SAVARÉ: *Stability of flows associated to gradient vector fields and convergence of iterated transport maps.* Manuscripta Math., **121** (2006), 1–50.
17. L.AMBROSIO & J.MALÝ: *Very weak notions of differentiability.* Proceedings of the Royal Society of Edinburgh, **137**A (2007), 447–455.
18. E.J.BALDER: *New fundamentals of Young measure convergence.* CRC Res. Notes in Math. **411**, 2001.
19. J.BALL & R.JAMES: *Fine phase mixtures as minimizers of energy.* Arch. Rat. Mech. Anal., **100** (1987), 13–52.
20. V.BANGERT: *Minimal measures and minimizing closed normal one-currents.* Geom. funct. anal., **9** (1999), 413–427.
21. J.-D.BENAMOU & Y.BRENIER: *Weak solutions for the semigeostrophic equation formulated as a couples Monge-Ampere transport problem.* SIAM J. Appl. Math., **58** (1998), 1450–1461.
22. P.BERNARD & B.BUFFONI: *Optimal mass transportation and Mather theory.* J. Eur. Math. Soc. (JEMS), **9** (2007), 85–121.
23. M.BERNOT, V.CASELLES & J.M.MOREL: *Traffic plans.* Publ. Mat., **49** (2005), 417–451.
24. V.BOGACHEV & E.M.WOLF: *Absolutely continuous flows generated by Sobolev class vector fields in finite and infinite dimensions.* J. Funct. Anal., **167** (1999), 1–68.
25. F.BOUCHUT: *Renormalized solutions to the Vlasov equation with coefficients of bounded variation.* Arch. Rational Mech. Anal., **157** (2001), 75–90.
26. F.BOUCHUT & G.CRIPPA: *Uniqueness, Renormalization, and Smooth Approximations for Linear Transport Equations.* SIAM J. Math. Anal., **38** (2006), 1316–1328.
27. F.BOUCHUT, F.GOLSE & M.PULVIRENTI: *Kinetic equations and asymptotic theory.* Series in Appl. Math., Gauthiers-Villars, 2000.
28. F.BOUCHUT & F.JAMES: *One dimensional transport equation with discontinuous coefficients.* Nonlinear Analysis, **32** (1998), 891–933.
29. F.BOUCHUT, F.JAMES & S.MANCINI: *Uniqueness and weak stability for multi-dimensional transport equations with one-sided Lipschitz coefficients.* Annali Scuola Normale Superiore, Ser. 5, **4** (2005), 1–25.
30. Y.BRENIER: *The least action principle and the related concept of generalized flows for incompressible perfect fluids.* J. Amer. Mat. Soc., **2** (1989), 225–255.
31. Y.BRENIER: *The dual least action problem for an ideal, incompressible fluid.* Arch. Rational Mech. Anal., **122** (1993), 323–351.
32. Y.BRENIER: *A homogenized model for vortex sheets.* Arch. Rational Mech. Anal., **138** (1997), 319–353.
33. Y.BRENIER: *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations.* Comm. Pure Appl. Math., **52** (1999), 411–452.
34. A.BRESSAN: *An ill posed Cauchy problem for a hyperbolic system in two space dimensions.* Rend. Sem. Mat. Univ. Padova, **110** (2003), 103–117.
35. H.BREZIS: *Analyse fonctionnelle. Théorie et applications.* Masson, Paris, 1983.
36. L.A.CAFFARELLI: *Some regularity properties of solutions of Monge Ampère equation*, Comm. Pure Appl. Math., **44** (1991), 965–969.
37. L.A.CAFFARELLI: *Boundary regularity of maps with convex potentials*, Comm. Pure Appl. Math., **45** (1992), 1141–1151.
38. L.A.CAFFARELLI: *The regularity of mappings with a convex potential.* J. Amer. Math. Soc., **5** (1992), 99–104.
39. L.A.CAFFARELLI: *Boundary regularity of maps with convex potentials.* Ann. of Math., **144** (1996), 453–496.

40. I.CAPUZZO DOLCETTA & B.PERTHAME: *On some analogy between different approaches to first order PDE's with nonsmooth coefficients.* Adv. Math. Sci Appl., **6** (1996), 689–703.
41. A.CELLINA: *On uniqueness almost everywhere for monotonic differential inclusions.* Nonlinear Analysis, TMA, **25** (1995), 899–903.
42. A.CELLINA & M.VORNICESCU: *On gradient flows.* Journal of Differential Equations, **145** (1998), 489–501.
43. F.COLOMBINI, G.CRIPPA & J.RAUCH: *A note on two dimensional transport with bounded divergence.* Comm. Partial Differential Equations, **31** (2006), 1109–1115.
44. F.COLOMBINI & N.LERNER: *Uniqueness of continuous solutions for BV vector fields.* Duke Math. J., **111** (2002), 357–384.
45. F.COLOMBINI & N.LERNER: *Uniqueness of L^∞ solutions for a class of conormal BV vector fields.* Contemp. Math. **368** (2005), 133–156.
46. F.COLOMBINI, T. LUO & J.RAUCH: *Nearly Lipschitzean diverge free transport propagates neither continuity nor BV regularity.* Commun. Math. Sci., **2** (2004), 207–212.
47. G.CRIPPA & C.DE LELLIS: *Oscillatory solutions to transport equations.* Indiana Univ. Math. J., **55** (2006), 1–13.
48. G.CRIPPA & C.DE LELLIS: *Estimates and regularity results for the DiPerna–Lions flow.* Preprint, 2006 (available at http://cvgmt.sns.it). Accepted by J. Reine Angew. Math.
49. A.B.CRUZEIRO: *Équations différentielles ordinaires: non explosion et mesures quasi-invariantes.* J. Funct. Anal., **54** (1983), 193–205.
50. A.B.CRUZEIRO: *Équations différentielles sur l'espace de Wiener et formules de Cameron-Martin non linéaires.* J. Funct. Anal., **54** (1983), 206–227.
51. A.B.CRUZEIRO: *Unicité de solutions d'équations différentielles sur l'espace de Wiener.* J. Funct. Anal., **58** (1984), 335–347.
52. M.CULLEN: *On the accuracy of the semi-geostrophic approximation.* Quart. J. Roy. Metereol. Soc., **126** (2000), 1099–1115.
53. M.CULLEN & M.FELDMAN: *Lagrangian solutions of semigeostrophic equations in physical space.* SIAM J. Math. Anal., **37** (2006), 1371–1395.
54. M.CULLEN & W.GANGBO: *A variational approach for the 2-dimensional semi-geostrophic shallow water equations.* Arch. Rational Mech. Anal., **156** (2001), 241–273.
55. C.DAFERMOS: *Hyperbolic conservation laws in continuum physics.* Springer Verlag, 2000.
56. N.G.DE BRUIJN: *On almost additive functions.* Colloq. Math. **15** (1966), 59–63.
57. C.DE LELLIS: *Blow-up of the BV norm in the multidimensional Keyfitz and Kranzer system.* Duke Math. J., **127** (2004), 313–339.
58. L.DE PASCALE, M.S.GELLI & L.GRANIERI: *Minimal measures, one-dimensional currents and the Monge-Kantorovich problem.* Calc. Var. Partial Differential Equations, **27** (2006), 1–23.
59. N.DEPAUW: *Non unicité des solutions bornées pour un champ de vecteurs BV en dehors d'un hyperplan.* C.R. Math. Sci. Acad. Paris, **337** (2003), 249–252.
60. R.J.DIPERNA: *Measure-valued solutions to conservation laws*. Arch. Rational Mech. Anal., **88** (1985), 223–270.
61. R.J.DIPERNA & P.L.LIONS: *Ordinary differential equations, transport theory and Sobolev spaces.* Invent. Math., **98** (1989), 511–547.
62. R.J.DIPERNA & P.L.LIONS: *On the Cauchy problem for the Boltzmann equation: global existence and weak stability.* Ann. of Math., **130** (1989), 312–366.
63. L.C.EVANS: *Partial Differential Equations and Monge–Kantorovich Mass Transfer.* Current Developments in Mathematics (1997), 65–126.
64. L.C.EVANS: *Partial Differential Equations.* Graduate studies in Mathematics, **19** (1998), American Mathematical Society.
65. L.C.EVANS & W.GANGBO: *Differential equations methods for the Monge-Kantorovich mass transfer problem.* Memoirs AMS, **653**, 1999.
66. L.C.EVANS, W.GANGBO & O.SAVIN: *Diffeomorphisms and nonlinear heat flows.* SIAM J. Math. Anal., **37** (2005), 737–751.
67. L.C.EVANS & R.F.GARIEPY: *Lecture notes on measure theory and fine properties of functions,* CRC Press, 1992.

68. H.FEDERER: *Geometric measure theory*, Springer, 1969.
69. M.HAURAY: *On Liouville transport equation with potential in $BV_{\rm loc}$*. Comm. Partial Differential Equations, **29** (2004), 207–217.
70. M.HAURAY: *On two-dimensional Hamiltonian transport equations with $L^p_{\rm loc}$ coefficients.* Ann. IHP Nonlinear Anal. Non Linéaire, **20** (2003), 625–644.
71. W.B.JURKAT: *On Cauchy's functional equation.* Proc. Amer. Math. Soc., **16** (1965), 683–686.
72. B.L.KEYFITZ & H.C.KRANZER: *A system of nonstrictly hyperbolic conservation laws arising in elasticity theory.* Arch. Rational Mech. Anal. **1980**, *72*, 219–241.
73. C.LE BRIS & P.L.LIONS: *Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications.* Annali di Matematica, **183** (2003), 97–130.
74. N.LERNER: *Transport equations with partially BV velocities.* Ann. Sc. Norm. Super. Pisa Cl. Sci., **3** (2004), 681–703.
75. P.L.LIONS: *Mathematical topics in fluid mechanics, Vol. I: incompressible models.* Oxford Lecture Series in Mathematics and its applications, **3** (1996), Oxford University Press.
76. P.L.LIONS: *Mathematical topics in fluid mechanics, Vol. II: compressible models.* Oxford Lecture Series in Mathematics and its applications, **10** (1998), Oxford University Press.
77. P.L.LIONS: *Sur les équations différentielles ordinaires et les équations de transport.* C. R. Acad. Sci. Paris Sér. I, **326** (1998), 833–838.
78. J.LOTT & C.VILLANI: *Weak curvature conditions and functional inequalities.* J. Funct. Anal., in press.
79. S.MANIGLIA: *Probabilistic representation and uniqueness results for measure-valued solutions of transport equations.* J. Math. Pures Appl., **87** (2007), 601–626.
80. J.N.MATHER: *Minimal measures.* Comment. Math. Helv., **64** (1989), 375–394.
81. J.N.MATHER: *Action minimizing invariant measures for positive definite Lagrangian systems.* Math. Z., **207** (1991), 169–207.
82. E.Y.PANOV: *On strong precompactness of bounded sets of measure-valued solutions of a first order quasilinear equation.* Math. Sb., **186** (1995), 729–740.
83. G.PETROVA & B.POPOV: *Linear transport equation with discontinuous coefficients.* Comm. PDE, **24** (1999), 1849–1873.
84. F.POUPAUD & M.RASCLE: *Measure solutions to the liner multidimensional transport equation with non-smooth coefficients.* Comm. PDE, **22** (1997), 337–358.
85. A.PRATELLI: *Equivalence between some definitions for the optimal transport problem and for the transport density on manifolds.* Ann. Mat. Pura Appl., **184** (2005), 215–238.
86. S.K.SMIRNOV: *Decomposition of solenoidal vector charges into elementary solenoids and the structure of normal one-dimensional currents.* St. Petersburg Math. J., **5** (1994), 841–867.
87. E.M.STEIN: *Singular integrals and differentiability properties of functions.* Princeton University Press, 1970.
88. L. TARTAR: *Compensated compactness and applications to partial differential equations.* Research Notes in Mathematics, Nonlinear Analysis and Mechanics, ed. R. J. Knops, vol. **4**, Pitman Press, New York, 1979, 136–211.
89. R.TEMAM: *Problémes mathématiques en plasticité.* Gauthier-Villars, Paris, 1983.
90. J.I.E.URBAS: *Global Hölder estimates for equations of Monge-Ampère type*, Invent. Math., **91** (1988), 1–29.
91. J.I.E.URBAS: *Regularity of generalized solutions of Monge-Ampère equations*, Math. Z., **197** (1988), 365–393.
92. A.VASSEUR: *Strong traces for solutions of multidimensional scalar conservation laws.* Arch. Ration. Mech. Anal., **160** (2001), 181–193.
93. C.VILLANI: *Topics in mass transportation.* Graduate Studies in Mathematics, **58** (2004), American Mathematical Society.
94. C.VILLANI: *Optimal transport: old and new.* Lecture Notes of the 2005 Saint-Flour Summer school.
95. L.C.YOUNG: *Lectures on the calculus of variations and optimal control theory,* Saunders, 1969.

A Note on Alberti's Rank-One Theorem

Camillo De Lellis

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057
Zürich, Switzerland
e-mail: delellis@math.unizh.ch
URL: http://www.math.unizh.ch/

1 Introduction

The aim of these notes is to illustrate a proof of the following remarkable Theorem of Alberti (first proved in [1]). Here, when μ is a Radon measure on $\Omega \subset \mathbb{R}^n$, we denote by μ^a its absolutely continuous part (with respect to the Lebesgue measure $\mathscr{L}^n$), by $\mu^s := \mu - \mu^a$ its singular part, and by $|\mu|$ its total variation measure. Clearly, $|\mu|^a = |\mu^a|$ and $|\mu|^s = |\mu^s|$. When $\mu = Du$ for some $u \in BV(\Omega, \mathbb{R}^k)$, we will write $D^s u$ and $D^a u$. If ν is a nonnegative measure, μ/ν will denote the Radon–Nykodim derivative of μ with respect to ν. Finally we recall the polar decomposition of Radon measures, namely the identity $\mu = \frac{\mu}{|\mu|}|\mu|$, which implies that the vetor Borel map $\mu/|\mu|$ has modulus 1 μ-a.e.

Theorem 1.1. *Let* $u \in BV(\Omega, \mathbb{R}^k)$ *for some open set* $\Omega \subset \mathbb{R}^n$. *Then* $\mathrm{rank}(Du/|Du|(x)) = 1$ *for* $|D^s u|$*-a.e.* $x \in \Omega$.

We start by discussing what can be inferred from the "standard theory" of BV functions without much effort. A first conclusion can be drawn from the BV Structure Theorem (see Sect. 3.6, Theorem 3.77, and Proposition 3.92 of [3]) for which we first need some terminology. Given an L^1 function u we say that u is approximately continuous at x if there exists $\tilde{u}(x) \in \mathbb{R}^k$ such that $\lim_r r^{-n} \int_{B_r(x)} |u(y) - \tilde{u}(x)|\, dy = 0$. We denote by S_u the set of points where u is not approximately continuous and we say that $x \in S_u$ is an approximate jump point if there exists $\nu(x) \in \mathbf{S}^{n-1}$

and $u^{\pm}(x) \in \mathbb{R}^k$ such that

$$\lim_{r\downarrow 0} \frac{1}{r^n}\left(\int_{B_r^+(x)} |u(y)-u^+(x)|\,dy + \int_{B_r^-(x)} |u(y)-u^-(x)|\,dy\right) = 0,$$

where $B_r^{\pm}(x) = \{y \in B_r(x) : \pm(y-x)\cdot\nu(x) > 0\}$. The triple $(\nu(x), u^+(x), u^-(x))$ is unique up to a change of sign of $\nu(x)$ and a permutation of $u^+(x)$ and $u^-(x)$. The set of approximate jump points is denoted by J_u.

Finally, we recall that an $(n-1)$-dimensional rectifiable set $R \subset \mathbb{R}^n$ is a Borel set which can be covered $\mathscr{H}^{n-1}$-almost all by a countable family of C^1 $(n-1)$-dimensional surfaces. Here, $\mathscr{H}^k$ denotes the k-dimensional Hausdorff measure.

Theorem 1.2 (Structure Theorem for *BV* functions). *If $\Omega \subset \mathbb{R}^n$ is open and $u \in BV(\Omega,\mathbb{R}^k)$, then J_u is a rectifiable $(n-1)$-dimensional set, $\mathscr{H}^{n-1}(S_u \setminus J_u) = |Du|(S_u \setminus J_u) = 0$ and $D^s u$ can be decomposed as $D^c u + D^j u$, where*

- *$|D^c u|(E) = 0$ for every Borel set E with $\mathscr{H}^{n-1}(E) < \infty$;*
- *$D^j u = (u^+ - u^-) \otimes \nu \mathscr{H}^{n-1} \llcorner J_u$.*

Here and in what follows, given a measure μ and a Borel set E we denote by $\mu \llcorner E$ the measure given by $\mu \llcorner E(A) = \mu(E \cap A)$. Following [5], we call $D^c u$ and $D^j u$ respectively Cantor part and Jump part of the measure Du. Thus, Theorem 1.2 implies the statement of Theorem 1.1 when we replace $|D^s u|$ with $|D^j u|$.

A second fact that can be inferred from the "standard theory" of *BV* functions is the following dimensional reduction:

Proposition 1.3. *Theorem 1.1 holds if and only if it holds for $\Omega = B_1(0) \subset \mathbb{R}^2$ and $\mathbb{R}^k = \mathbb{R}^2$.*

This proposition will be proved in Sect. 2. Thus, the key point of Theorem 1.1 is to show that M has rank one $|D^c u|$-a.e. when u is a *BV* planar map. A first heuristic idea of why this property indeed holds is given in Sect. 3. The key remark of that section is the following lemma, which has a quite simple proof.

Lemma 1.4. *Let $\Omega \subset \mathbb{R}^2$ be connected and $u \in BV(\Omega,\mathbb{R}^2)$ be such that $Du/|Du|$ is a constant matrix M of rank 2. Then, $Du = cM\mathscr{L}^2 \llcorner \Omega$ for some $c > 0$.*

Building on this lemma and on a "blow-up" argument, we prove in Sect. 3 a particular case of Theorem 1.1. However, this simple strategy cannot prove Theorem 1.1 in its full generality (see Sect. 3, in particular Proposition 3.3). Alberti's strategy relies on replacing Lemma 1.4 with Lemma 1.5 below. From now on a set $C \subset \mathbb{R}^2$ will be called a closed convex cone if there exist $e \in \mathbf{S}^1$ and $0 < a < 1$ such that $C = C(e,a) := \{x : x \cdot e \geq a|x|\}$.

Lemma 1.5. *Let C_1 and C_2 be two closed convex cones such that $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$. Let $\Omega \subset \mathbb{R}^2$ be open and $v_1, v_2 \in BV(\Omega)$ be two scalar functions such that $Dv_i/|Dv_i|(x) \in C_i$ for $|Dv_i|$-a.e. x. If $\mu \geq 0$ is a measure such that $\mu << |Dv_i|$ for $i = 1,2$, then $\mu << \mathscr{L}^2 \llcorner \Omega$.*

This lemma will be proved in Sect. 4. We want to stress here the analogies with Lemma 1.4. Set $v = (v_1, v_2)$. By the polar factorization, the main assumption of Lemma 1.5 could be restated as *$Dv/|Dv|$ belongs ($|Dv|$-almost everywhere) to a suitably small neighborhood of a constant matrix M of rank 2*. Moreover the last sentence is equivalent to $|Dv| << \mathscr{L}^2 \llcorner \Omega$. Thus, we can consider Lemma 1.4 as a rigidity result and Lemma 1.5 as its quantitative counterpart.

Now consider $u \in BV(\Omega, \mathbb{R}^2)$ and the Borel set $E := \{x : \mathrm{rank}\,(Du/|Du|(x)) = 2\}$. Standard arguments show that E can be decomposed in countably many Borel pieces E_i where $Du/|Du|$ is very close to a single constant matrix M_i. Thus the relaxed assumption of Lemma 1.5 suggests that we could use a "decomposition" approach, in contrast with the "blow-up" argument which builds on the rigidity Lemma 1.4. More precisely, we will show in Sect. 5 that the decomposition in Borel pieces E_is can be chosen so that

- If we fix any i and set $\mu := |Du| \llcorner E_i$, then there are two BV scalar functions v_1 and v_2 such that v_1, v_2 and μ satisfy the hypotheses of Lemma 1.5.

Clearly, the decomposition stated above and Lemma 1.5 show that μ is absolutely continuous, i.e. they prove Theorem 1.1. The construction of the v_is is the second key idea of Alberti's proof. The argument combines a simple geometric consideration on the level sets of the u_is together with a clever use of the coarea formula for BV scalar functions.

Recently, Alberti, Csorniey and Preiss, (see [2]) have proposed a different proof of the Rank-One Theorem. This new proof uses as well the coarea formula, but it avoids Lemma 1.5, and relies instead on a general covering result for Lebesgue-null sets of the plane. Let us mention, in passing, that this last result has many other deep implications in real analysis and geometric measure theory; see [2].

2 Dimensional Reduction

Proof of Proposition 1.3. Assume that Theorem 1.1 holds for maps $u \in BV(B_1(0), \mathbb{R}^2)$ with $B_1(0) \subset \mathbb{R}^2$. Clearly, by translating and rescaling, we immediately conclude the theorem when $u \in BV(B, \mathbb{R}^2)$ for any two-dimensional ball B. The statement of Theorem 1.1 is trivially true if $\Omega \subset \mathbb{R}$ or if $k = 1$ Moreover, any open set $\Omega \subset \mathbb{R}^n$ can be written as countable union of balls. Hence it suffices to prove the theorem when Ω is a ball of $\mathbb{R}^n$, $n \geq 2$, and $k \geq 2$.

From $n = 2$ to n generic. Here we prove Theorem 1.1 for maps $u \in BV(B, \mathbb{R}^2)$ whenever B is an n-dimensional ball. We argue by contradiction and let $u \in BV(B, \mathbb{R}^2)$ be such that $\mathrm{rank}\,(Du/|Du|(x)) = 2$ on some set E with $|D^s u|(E) > 0$. Set $M = Du/|Du|$ and choose coordinates $x_1, \ldots, x_n$ on B and u_1, u_2 on $\mathbb{R}^2$. Clearly, M has $n(n-1)/2$ different minors, corresponding to the choice of coordinates x_i, x_j with $1 \leq i < j \leq n$: We denote them by M^{ij}. If we set $E_{ij} := \{x : \mathrm{rank}\,(M^{ij}(x)) = 2\}$, then $E = \bigcup_{ij} E_{ij}$, and hence $|D^s u|(E_{ij}) > 0$ for some i and j. Without loss of generality we assume $i = 1$ and $j = 2$. Consider the matrix valued measure $(\mu)_{l\alpha} = (\partial_{x_l} u_\alpha)_{l\alpha}$ with $l, \alpha = 1,2$. Then, $\operatorname{rank}(\mu/|\mu|(x)) = 2$ for $|\mu|$-a.e. $x \in E_{12}$ and $|\mu^s|(E_{12}) > 0$.

For any $y \in \mathbb{R}^{n-2}$ we define $B_y = \{(x_1,x_2) \in \mathbb{R}^2 : (x_1,x_2,y) \in B\}$. Clearly, B_y is either empty or it is an open two-dimensional ball. Moreover, we define

$$v_y : B_y \to \mathbb{R}^2 \qquad \text{by} \qquad v_y(x_1,x_2) = u(x_1,x_2,y).$$

By the slicing theory of BV functions (see Theorem 3.103, Theorem 3.107, and Theorem 3.108 of [3]) we have:

(a) $v_y \in BV(B_y, \mathbb{R}^2)$ for $\mathscr{L}^{n-2}$-a.e. $y \in \mathbb{R}^{n-2}$;
(b) $\mu = Dv_y \otimes \mathscr{L}^{n-2}$ and $|\mu| = |Dv_y| \otimes \mathscr{L}^{n-2}$.

(Here, when α is a measure on Y and $y \mapsto \beta_y$ a weakly measurable map from Y into the space $\mathscr{M}(X)$ of Radon measures on X, the symbol $\beta_y \otimes \alpha$ denotes the measure γ on $X \times Y$ which satisfies

$$\int_{X\times Y} \varphi(x,y)\, d\gamma(x,y) = \int_Y \int_X \varphi(x,y)\, d\beta_y(x)\, d\alpha(y)$$

for every $\varphi \in C_c(X \times Y)$.)

(b) implies two things. First of all,

$$\frac{Dv_y}{|Dv_y|}(x_1,x_2) = \frac{\mu}{|\mu|}(x_1,x_2,y) \quad \text{for } \mathscr{L}^{n-2}\text{-a.e. } y \text{ and } |Dv_y|\text{-a.e. } (x_1,x_2). \tag{1}$$

Second, if for every y we set $E_y := \{(x_1,x_2) : (x_1,x_2,y) \in E\}$, then

$$\int_{\mathbb{R}^{n-2}} |Dv_y^s|(E_y)\, d\mathscr{L}^{n-2}(y) = |\mu^s|(E) > 0. \tag{2}$$

Thus, from (a), (1) and (2), we conclude that there exists a y such that $v_y \in BV(B_y, \mathbb{R}^2)$, $|Dv_y^s|(E_y) > 0$, and $\operatorname{rank}(Dv_y/|Dv_y|(x)) = 2$ for $|Dv_y|$-a.e. $x \in E_y$. Such v_y contradicts our assumption that Theorem 1.1 holds for maps $u \in BV(B_y, \mathbb{R}^2)$.

From $k = 2$ to k generic. Fix any $u \in BV(B, \mathbb{R}^k)$, with $k \geq 2$ and B n-dimensional ball, and choose coordinates $u_1, \ldots, u_k$ on $\mathbb{R}^k$. For any pair of integers $1 \leq i < j \leq k$, consider the map $u_{ij} := (u_i, u_j) \in BV(B, \mathbb{R}^2)$. If $M = Du/|Du|$ and M_{ij} is the corresponding $2 \times n$ minor, then $Du_{ij} = M_{ij}|Du|$. Thus, by the previous step, $\operatorname{rank}(M_{ij}(x)) \leq 1$ for $|D^s u_{ij}|$-a.e. x, and hence for $|D^s u|$-a.e. x. Set

$$E_{ij} := \big\{x : \operatorname{rank}(M_{ij}(x)) \leq 1\big\} \qquad \text{and} \qquad E := \bigcap_{1 \leq i < j \leq k} E_{ij}.$$

Then, $|D^s u|(\mathbb{R}^n \setminus E) = 0$ and $\operatorname{rank}(M(x)) \leq 1$ for every $x \in E$. This concludes the proof. $\square$

3 A Blow-Up Argument Leading to a Partial Result

We start this section by proving Lemma 1.4.

Proof of Lemma 1.4. We let M be the constant matrix $Du/|Du|$ and $\mu = |Du|$. By standard arguments, it suffices to prove the lemma when Ω is the unit ball $B_1(0)$. Denote by u_1 and u_2 the two components of u. Then $Du_i = v_i\mu$, where $v_1, v_2 \in \mathbb{R}^2$ are two linearly independent vectors. Let $\{\varphi_\varepsilon\}_{\varepsilon>0}$ be a standard family of mollifiers supported in $B_\varepsilon(0)$ and consider the mollifications $u_i * \varphi_\varepsilon$ in $B_{1-\varepsilon}(0)$. Notice that $D(u_i * \varphi_\varepsilon) = v_i\mu * \varphi_\varepsilon$, and hence $u_i * \varphi_\varepsilon$ is constant on the direction orthogonal to v_i. Therefore the density of the absolutely continuous measure $\mu * \varphi_\varepsilon$ is a function f_ε which is constant along two linearly independent directions. Thus, f_ε is constant. Letting $\varepsilon \downarrow 0$ we complete the proof. □

This simple remark leads to a partial answer to Theorem 1.1, given in Proposition 3.2.

Definition 3.1. Let μ be a measure on $\Omega \subset \mathbb{R}^2$ and for any x in the support of μ and any $r \in]0, \mathrm{dist}\,(x, \partial\Omega)[$ consider the measures $\mu_{x,r}$ on $B_1(0)$ given by

$$\mu_{x,r}(A) = \mu(x + rA)/|\mu|(B_r(x)) \qquad \text{for any Borel set } A \subset B_1(0).$$

We say that a measure μ_0 is tangent to μ at x if for some sequence $r_n \downarrow 0$ we have $\mu_{x,r_n} \rightharpoonup^* \mu_0$.

A nonnegative measure μ on $\Omega \subset \mathbb{R}^2$ is said to have only trivial blow-ups at x, if every tangent measure to μ at x is of the form $c\mathscr{L}^2 \llcorner B_1(0)$. For $u \in BV(\Omega, \mathbb{R}^2)$ we denote by T the set of points where $|D^s u|$ has only trivial blow-ups.

This definition of tangent measure is very similar to that introduced by Preiss in the fundamental paper [6]. We are now ready to state our

Proposition 3.2. *Let* $u \in BV(\Omega, \mathbb{R}^2)$*. Then* $\mathrm{rank}\,(Du/|Du|(x)) = 1$ *for* $|D^s u|$*-a.e.* $x \notin T$.

Proof. We argue by contradiction and assume that the proposition is false for some u. Denote by μ the measure $|D^s u|$. Then, by standard measure-theoretic arguments, it is possible to find a point $x \notin T$ and a sequence $r_n \downarrow 0$ such that the following properties hold:

(i) $\mu_{x,r_n} \rightharpoonup^* \mu_0$, and $\mu_0 \neq c\mathscr{L}^n \llcorner B_1(0)$;
(ii) $||Du| - \mu|(B_r(x)) = o(\mu(B_r(x)))$;
(iii) $M = Du/|Du|(x)$ is a matrix of rank > 1 and

$$\lim_{r\downarrow 0} \frac{1}{|Du|(B_r(x))} \int_{B_r(x)} \big|Du/|Du|(y) - M\big|\, d|Du|(y) = 0.$$

Let $\overline{u}_r$ be the average of u on $B_r(x)$ and define the function $u_r \in BV(B_1(0), \mathbb{R}^2)$ as

$$u_r(y) = \frac{r^{n-1}(u(x+ry) - \overline{u}_r)}{|Du|(B_r(x))}.$$

It follows that $Du_r = [Du]_{x,r}$, and hence $|Du_r|(B_1(0)) = 1$. Moreover, since the average of u_r is 0, the Poincaré inequality gives $\|u_r\|_{L^1} \leq C$. Thus, we can assume that a subsequence, not relabeled, of $\{u_{r_n}\}$ converges to some $u_0 \in BV(B_1(0), \mathbb{R}^2)$ strongly in L^1. Now, from (ii) we get $|Du|_{x,r} - \mu_{x,r} \rightharpoonup^* 0$ and from (iii) we conclude $[Du]_{x,r} - M|Du|_{x,r} \rightharpoonup^* 0$. Therefore, by (i), $Du_r = [Du]_{x,r} \rightharpoonup^* M\mu_0$. This implies $Du_0 = M\mu_0$, because u_{r_n} converges to u_0. Applying Lemma 1.4 we conclude $\mu_0 = c\mathscr{L}^2 \llcorner B_1(0)$, which contradicts (i).

Unfortunately, we cannot hope to prove Theorem 1.1 by showing that singular parts of BV functions have necessarily nontrivial blow-ups. More precisely we have

Proposition 3.3. *There exist BV maps u such that* $|D^s u|(T) > 0$.

Proof. The example 5.8(1) of [6] gives a nonnegative measure μ on a bounded interval I which is singular and such that $\mu_{x,r} \rightharpoonup^* \frac{1}{2}\mathscr{L}^1 \llcorner [-1,1]$ for μ-a.e. x. Clearly, any primitive of μ is a bounded BV function which satisfies the requirements of the proposition. □

4 The Fundamental Lemma

Before coming to the proof of the lemma, let us explain its basic ingredients. Assume for the moment that the v_is of the lemma are regular, and that $\mu = f\mathscr{L}^2 \leq C|\nabla v_i|$. Consider the map $v = (v_1, v_2)$. Since the gradients ∇v_i belong everywhere to the cones C_i and $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$, a simple algebraic consideration shows that $\det \nabla v$ controls, up to some constant depending on the C_is, the product $|\nabla v_1||\nabla v_2|$, and hence f^2. Thus we can bound the L^2 norm of f by the integral of $\det \nabla v$. A second key remark is that the geometric constraints on the C_is imply that v is almost injective (more precisely, v would be injective if $\nabla v_i \in C_i \setminus \{0\}$). Thus, $\int \det \nabla v$ can be computed using the area formula. This means that $\int f^2$ can be bound in terms, for instance, of the L^∞ norm of v, but independently of ∇v. In the proof below we will extend such a priori estimate to the general case, using truncations and a suitable regularization procedure.

Remark 4.1. In the rest of these notes we will use extensively the following elementary fact. Let $C = C(e,a) = \{x : x \cdot e \geq a|x|\}$ be a closed convex cone, $\Omega \subset \mathbb{R}^2$ an open set and $v \in BV(\Omega, \mathbb{R})$. Then, it follows easily from the polar decomposition of measures that $Dv/|Dv|(x) \in C$ for $|Dv|$-a.e. x if and only if $\partial_e v \geq a|Dv|$.

Proof of Lemma 1.5. We can assume without loss of generality that $v_1, v_2 \in L^\infty$. Indeed for every $k \in \mathbb{N}$ set $v_i^k = \min(\max(v_i, -k), k)$ and $E_k = \{|v_1| < k\} \cap \{|v_2| < k\}$. Then, by the locality of $|Dv|$ (see Remark 3.93 of [3]):

- v_1^k, v_2^k are bounded BV functions which satisfy the assumptions of the lemma;
- $\mu(\Omega \setminus \bigcup_k E_k) = 0$ and $\mu \llcorner E_k << |Dv_i| \llcorner E_k = |Dv_i^k| \llcorner E_k \leq |Dv_i^k|$.

Therefore, if the lemma holds for bounded BV functions, then we conclude that $\mu \llcorner E_k << \mathscr{L}^2 \llcorner \Omega$, and hence that $\mu << \mathscr{L}^2 \llcorner \Omega$. In addition, since every open set Ω can be covered by a countable family of convex subsets, we will assume that Ω is convex. Finally, we can assume, without loss of generality, that $\mu \leq N|Dv_i|$ for some constant N. Indeed, for any $N > 0$ let E_N be the set of points x where the Radon–Nykodim derivatives $\mu/|Dv_i|(x) \leq N$. Then $\mu(\mathbb{R}^2 \setminus \bigcup_N E_N) = 0$ and $\mu \llcorner E_N \leq N|Dv_i|$.

Let any such v_is and Ω satisfy all these assumptions, and let C_1 and C_2 be the cones of the lemma. Recall that $C_i = C(e_i, a_i)$ for some $1 > a_i > 0$ and $e_i \in \mathbf{S}^1$. Given two vectors $z_1, z_2 \in \mathbb{R}^2$ we measure the angle $\theta(z_1, z_2)$ between z_1 and z_2 in counterclockwise direction. By possibly exchanging the indices we can assume $\theta(e_1, e_2) < \pi$. Then, the assumptions $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$ translate into the existence of a constant $\delta_0 > 0$ such that $\delta_0 \leq \theta(f_1, f_2) \leq \pi - \delta_0$ for every pair $(f_1, f_2) \in C_1 \times C_2$. Therefore, for $\delta = \sin\delta_0 > 0$,

$$\det(f_1, f_2) = |f_1||f_2| \sin\theta(f_1, f_2) \geq \delta |f_1||f_2| \qquad \forall (f_1, f_2) \in C_1 \times C_2. \tag{3}$$

By Remark 4.1, $\partial_{e_i} v_i \geq a_i |Dv_i|$. Set $w_i(x) = v_i(x) + \arctan(x \cdot e_i)$ and $w = (w_1, w_2)$ and note that

(a) $\partial_{e_i} w_i \geq a_i |Dw_i|$;
(b) $[\partial_{e_i} w_i](B_r(x)) > 0$ for every ball $B_r(x) \subset \Omega$;
(c) $\mu \leq N|Dv_i| \leq N a_i^{-1} \partial_{e_i} v_i \leq N a_i^{-1} \partial_{e_i} w_i$.

Let $\{\varphi_\varepsilon\}$ be a standard family of nonnegative mollifiers supported in $B_\varepsilon(0)$ and consider the mollifications $w * \varphi_\varepsilon$ in the open sets $\Omega_\varepsilon := \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) > \varepsilon\}$. We claim that

(a') $\nabla(w_i * \varphi_\varepsilon)(x) \in C_i$ for any i and any $x \in \Omega_\varepsilon$;
(b') $w * \varphi_\varepsilon : \Omega_\varepsilon \to \mathbb{R}^2$ is injective;
(c') $\mu * \varphi_\varepsilon \leq N a_i^{-1} \partial_{e_i}(w_i * \varphi_\varepsilon)$.

From (a) we get $\partial_{e_i}(w_i * \varphi_\varepsilon) \geq a_i |Dw_i| * \varphi_\varepsilon \geq a_i |D(w_i * \varphi_\varepsilon)|$, which, by Remark 4.1 and the smoothness of $w_i * \varphi_\varepsilon$, implies (a'). (c') follows from $\mu \leq N a_i^{-1} \partial_{e_i} w_i$. We now come to (b'). Note that, by (b), $\partial_{e_i}(w_i * \varphi_\varepsilon) > 0$. So $\nabla w_i * \varphi_\varepsilon(x) \neq 0$ for every $x \in \Omega_\varepsilon$, and hence belongs to $C_i \setminus \{0\}$. Let $x \neq y \in \Omega_\varepsilon$, and set $f := (x-y)/|x-y|$. We claim that, for some i,

$$|f \cdot z| > 0 \text{ for all } z \in C_i \setminus \{0\}. \tag{4}$$

Otherwise, there are $z_1 \in C_1$ and $z_2 \in C_2$ with $|z_i| = 1$ and $z_i \perp f$. Therefore, either $z_1 = z_2$ or $z_1 = -z_2$, contradicting $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$. Next, write

$$w_i * \varphi_\varepsilon(y) - w_i * \varphi_\varepsilon(x) \; = \; \int_0^{|y-x|} \nabla w_i * \varphi_\varepsilon (x + \sigma f) \cdot f \, d\sigma. \tag{5}$$

Recall that $\nabla w_i * \varphi_\varepsilon\,(x+\sigma f) \in C_i \setminus \{0\}$. Moreover, since $C_i \setminus \{0\}$ is connected, (4) implies that the integrand in (5) is either strictly positive, or strictly negative. In any case, $w_i * \varphi_\varepsilon(y) \neq w_i * \varphi_\varepsilon(x)$, which gives (b').

We are now ready for the final step. (a'), (b'), (c') and the area formula give

$$\begin{aligned}\|w_1\|_\infty \|w_2\|_\infty &\geq \|w_1 * \varphi_\varepsilon\|_\infty \|w_2 * \varphi_\varepsilon\|_\infty \;\geq\; \mathscr{L}^2(w * \varphi_\varepsilon(\Omega_\varepsilon))\\ &\overset{(b')}{=} \int_{\Omega_\varepsilon} \det(\nabla(w * \varphi_\varepsilon)(x))\,dx\\ &\overset{(a')+(3)}{\geq} \delta \int_{\Omega_\varepsilon} |\nabla(w_1 * \varphi_\varepsilon)(x)||\nabla(w_2 * \varphi_\varepsilon)(x)|\,dx\\ &\geq \delta \int_{\Omega_\varepsilon} [\partial_{e_1}(w_1 * \varphi_\varepsilon)](x)\,[\partial_{e_2}(w_2 * \varphi_\varepsilon)](x)\,dx\\ &\overset{(c')}{\geq} \delta N^{-2} a_1 a_2 \int_{\Omega_\varepsilon} (\mu * \varphi_\varepsilon(x))^2\,dx.\end{aligned}$$

Hence, $\|\mu * \varphi_\varepsilon\|^2_{L^2(\Omega_\varepsilon)} \leq N^2(a_1 a_2 \delta)^{-1}\|w_1\|_\infty \|w\|_\infty$, which, letting $\varepsilon \downarrow 0$, gives $\mu = f\mathscr{L}^2$ for some $f \in L^2(\Omega)$. □

5 Proof of Theorem 1.1 in the Planar Case

We will argue by contradiction, and hence in a different way with respect to what said in Sect. 1. However, this is only to make the presentation more transparent: The ideas presented in this section can be easily adapted to prove the general decomposition property claimed at the end of the introduction.

So, let $u = (u_1, u_2) \in BV(B, \mathbb{R}^2)$ where B is a two-dimensional disk. Define

$$E \;:=\; \big\{x : \operatorname{rank}(Du/|Du|(x)) = 2\big\}, \tag{6}$$

and assume that $|D^s u|(E) > 0$. Without loss of generality, we can assume $u \in L^\infty$. Indeed, for every k truncate u_1 and u_2 by setting $u_i^k = \min\{\max\{u_i, -k\}, k\}$, and define

$$u^k := (u_1^k, u_2^k) \qquad \text{and} \qquad E_k := \big\{x : \operatorname{rank}(Du^k/|Du^k|(x)) = 2\big\}.$$

Then, $|D^s u^k|(E_k) \to |D^s u|(E)$ as $k \uparrow \infty$.

Hence, from now on we assume that $u \in BV \cap L^\infty$. For each point $x \in E$, we set $w_i(x) := Du_i/|Du|(x)$, which must be nonzero vectors. Thus, we can define $e_i(x) := w_i(x)/|w_i(x)|$, which is paralell to $Du_i/|Du_i|(x)$ and pointing in the same direction. Next, let

- $\mathscr{F}_k$ be the set of pairs $(f_1, f_2) \in \mathbf{S}^1 \times \mathbf{S}^1$ which form an angle $\geq 1/k$ and $\leq \pi - 1/k$;
- $F_k := \{x \in E : (e_1(x), e_2(x)) \in \mathscr{F}_k\}$.

Since $E = \bigcup_k F_k$, obviously $|D^s u|(F_k) > 0$ for some k. Fix any such k and for any $(f_1, f_2) \in \mathscr{F}_k$ and any $\varepsilon > 0$ define

$$F(f_1, f_2, \varepsilon) := \Big\{x \in F_k : e_1(x) \in C(f_1, 1-\varepsilon), e_2(x) \in C(f_2, 1-\varepsilon)\Big\}.$$

We claim that there exist $(f_1, f_2) \in \mathscr{F}_k$ such that $|D^s u|(F(f_1, f_2, \varepsilon)) > 0$ for every $\varepsilon > 0$. Otherwise, by compactness of $\mathscr{F}_k$, we can find N pairs (f_1^j, f_2^j) and N positive numbers $\varepsilon_j > 0$ such that

$$\mathscr{F}_k \subset \bigcup_{j=1}^N C(f_1^j, 1-\varepsilon_j) \times C(f_2^j, 1-\varepsilon_j)$$

and $|D^s u|(F(f_1^j, f_2^j, \varepsilon_j)) = 0$. This would give $|D^s u|(F_k) \leq \sum_j |D^s u|(F(f_1^j, f_2^j, \varepsilon_j)) = 0$.

Therefore, fix $(f_1, f_2) \in \mathscr{F}_k$ such that $|D^s u|(F(f_1, f_2, \varepsilon)) > 0$ for every positive ε. Note that, since f_1 and f_2 are linearly independent, for ε sufficiently small the closed convex cones $C_i = C(f_i, 1-\varepsilon)$ satisfy $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$. We choose such an ε and we define

$$F' := \left\{x : \frac{Du_i}{|Du_i|}(x) \in C_i \qquad \text{for both } i\text{s}\right\}. \tag{7}$$

Theorem 1.1 is then implied by the following

Proposition 5.1. *Let $C = C(e, a)$ be a closed convex cone, $v \in BV \cap L^\infty(B, \mathbb{R})$ and*

$$G := \left\{x \; : \; \frac{Dv}{|Dv|}(x) \in C\right\}. \tag{8}$$

For any convex cone $C' = C(e, a')$ with $a' < a$ there exists $w \in BV \cap L^\infty(B, \mathbb{R})$ such that $|Dv| \llcorner G << |Dw|$, and

$$\frac{Dw}{|Dw|}(x) \in C' \qquad \textit{for } |Dw|\textit{-a.e. } x. \tag{9}$$

Proof of Theorem 1.1. We recall that we argue by contradiction. The discussion above gives a bounded BV map $u : B \to \mathbb{R}^2$ and two closed convex cones C_1 and C_2 such that

- $C_1 \cap C_2 = (-C_1) \cap C_2 = \{0\}$;
- If E and F' are defined as in (6) and (7), then $|D^s u|(E \cap F') > 0$.

Now, by definition of E, $|D^s u| \llcorner E << |Du_i|$ for both $i = 1, 2$. Thus, if we set $\mu := |D^s u| \llcorner (E \cap F')$, then μ is a singular measure such that $\mu << |Du_i| \llcorner F'$ for both $i = 1, 2$.

Next choose two larger closed convex cones C_1' and C_2' so that $C_1' \cap C_2' = (-C_1') \cap C_2' = \{0\}$. Apply Proposition 5.1 to find v_1 and v_2 such that $Dv_i/|Dv_i|(x) \in C_i'$ for

$|Dv_i|$-a.e. x, and $|Du_i|\llcorner F' << |Dv_i|$. Thus, we have $\mu << |Dv_i|$ for both $i = 1,2$. Applying Lemma 1.5 we conclude that μ is absolutely continuous, which is the desired contradiction. □

Therefore, we are left with the task of proving Proposition 5.1. A special case of this proposition is when v is the indicator function of a set (which therefore is a Caccioppoli set). This case turns out to be an elementary geometric remark, but it is the key to prove the proposition in its full generality, via the coarea formula.

Proof of Proposition 5.1 when v is the indicator function of a set A. Since v is a BV function, A is a Caccioppoli set. We denote by $\partial^* A$ its reduced boundary (see Sect. 3.5 of [3] for the definition) and by η the approximate exterior unit normal to $\partial^* A$. Since $Dv = \eta \mathscr{H}^1 \llcorner \partial^* A$, the set G is given by $\{x \in \partial^* A : \eta(x) \in C\}$. Since $\partial^* A$ is rectifiable (cp. with Theorem 3.59 of [3]), G can be decomposed as $G_0 \cup \bigcup_{i=1}^\infty G_i$, where:

- $\mathscr{H}^1(G_0) = 0$ and for $i \geq 1$ each G_i is the subset of a C^1 curve γ_i;
- $\eta|_{G_i}$ coincides with the normal to the curve γ_i.

Step 1. For each i we claim that there are Lipschitz open sets $\{S_{i,j}\}_{j\in\mathbb{N}}$ such that: the exterior normal to $\partial S_{i,j}$ belongs $\mathscr{H}^1$-a.e. to C' and $\{\partial S_{i,j}\}_j$ is a covering of G_i.

Recall that $C' = C(e,a')$, and choose coordinates x_1, x_2 in $\mathbb{R}^2$ in such a way that $e = (0,1)$. For any $x \in G_i$, the normal $\nu_i(x)$ belongs to $C(e,a)$, and thus it is transversal to $(1,0)$. Since γ_i is C^1, this implies that we can choose an open ball B_x centered at x such that $\gamma_i \cap B_x$ is the graph $\{(x_1, f(x_1))\}$ of a C^1 function $f : I \to \mathbb{R}$, where I is some bounded open interval of $\mathbb{R}$. Moreover, by continuity of the normal ν_i, we can choose B_x so that $\nu_i(y) \in C'$ for every $y \in \gamma_i \cap B_x$.

Fix any such y. Note that the angle θ between e and $\nu_i(y)$ is equal to the angle between $(0,1)$ and the tangent to γ_i at y. Since $\nu_i(y) \in C(e,a')$, we conclude that

$$\theta \;=\; \arccos(\nu_i(y)\cdot e) \leq \arccos(1/a').$$

Thus $|f'| \leq \tan(\arccos(1/a')) \leq \sqrt{a'^2 - 1}$, and hence f is a Lipschitz function with constant less than $\sqrt{a'^2-1}$. It is an elementary well-known fact that f can be extended to a function $\tilde{f} : \mathbb{R} \to \mathbb{R}$ with the same Lipschitz constant. If we define $S_x := \{(x_1,x_2) : x_2 < \tilde{f}(x_1)\}$, then S_x is a Lipschitz open set, the normal to S_x belongs everywhere to the cone C', and ∂S_x covers $B_x \cap \gamma_i$ (Fig. 1). Since we can cover γ_i with a countable family of these balls B_x, the corresponding S_x form the desired countable covering $\{S_{i,j}\}_j$

Step 2. Consider the sets $\{S_{i,j}\}_{i,j}$. Their boundaries have all finite lengths, which we denote by $\ell_{i,j}$ and they cover $\mathscr{H}^1$-a.e. G. Let $\lambda_{i,j}$ be a collection of positive numbers such that $\sum_{i,j} \lambda_{i,j} \leq 1$ and $\sum_{i,j} \lambda_{i,j}\ell_{i,j} \leq 1$. Let w be the function

$$\sum_{i,j} \lambda_{i,j} \mathbf{1}_{S_{i,j}}.$$

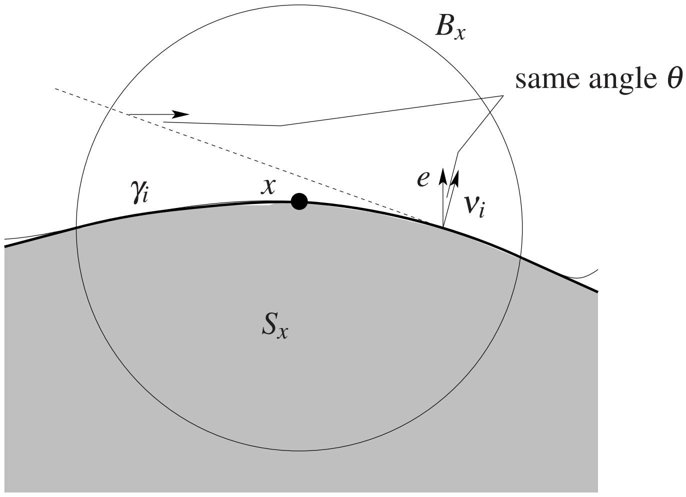

Fig. 1 The set S_x and the ball B_x

First of all, $\|w\|_\infty \leq \sum_{i,j}\lambda_{i,j} \leq 1$. Second, $w \in BV$ and $|Dw|$ is the nonnegative measure $\sum_{i,j}\lambda_{i,j}\mathscr{H}^1 \llcorner \partial S_{i,j}$. Thus, $|Dv| \llcorner G = \mathscr{H}^1 \llcorner G << |Dw|$. Finally,

$$\|w\|_{BV} = \|w\|_{L^1} + |Dw|(B) \leq 2\pi + \sum_{i,j}\lambda_{i,j}\ell_{i,j} \leq 2\pi + 1.$$

$\square$

Proof of Proposition 5.1. Fix v, C, G and C' as in the statement, set $c := \|v\|_\infty$ and for every $t \in [-c,c]$ consider the function $v_t := \mathbf{1}_{\{v>t\}}$. Then, it follows from the coarea formula (see Theorem 3.40 of [3]) that:

(i) v_t is a BV function for $\mathscr{L}^1$-a.e. t, i.e. $\{v > t\}$ is a Caccioppoli set, and we denote by ν_t its exterior unit normal;
(ii) $\nu_t(x) = Dv/|Dv|(x)$ for $\mathscr{L}^1$-a.e. t and $\mathscr{H}^1$-a.e. $x \in \partial^*\{v > t\}$;
(iii) $|Dv| = \int_{-c}^{c} |Dv_t|\, d\mathscr{L}^1(t)$.

(Here, when α is a measure on Y and $y \mapsto \beta_y$ a weakly measurable map from Y into the space $\mathscr{M}(X)$ of Radon measures on X, the symbol $\int \beta_y\, d\alpha(y)$ denotes the measure γ on X which satisfies

$$\int \varphi(x)\, d\gamma(x) = \int_Y \int_X \varphi(x)\, d\beta_y(x)\, d\alpha(y)$$

for every $\varphi \in C_c(X)$.)

Therefore, for $\mathscr{L}^1$-a.e. t, v_t, C, G, and C' satisfy the hypotheses of the proposition. We denote by w_t the corresponding BV function given by the special case of this proposition, proved above. We will show below that w_t can be selected in such a way that the map $t \mapsto w_t \in L^\infty$ is weakly* measurable, i.e. that $t \mapsto \int \varphi w_t$ is measurable for every $\varphi \in L^1(B)$. Having a map with this property, we choose $\lambda \in L^1([-c,c])$ such that $\lambda > 0$ and

$$\int_{-c}^{c} \lambda(t)\big(\|w_t\|_\infty + \|w_t\|_{BV}\big)\, dt < \infty.$$

Assuming this fact, we set $w(x) := \int_{-c}^{c} \lambda(t) w_t(x)\, d\mathscr{L}^1(t)$. Then w is bounded, $|Dw|$ is a measure and $|Dw| \le \int_{-c}^{c} \lambda(t)|Dw_t|d\mathscr{L}^1(t)$, which is a finite measure. Therefore $w \in BV \cap L^\infty$. Next, recall that $C' = C(e,a')$ for some real a' and some $e \in \mathbf{S}^1$. By Remark 4.1, $\partial_e w_t \ge a'|Dw_t|$. Thus $\partial_e w_t$ is a nonnegative measure for $\mathscr{L}^1$-a.e. t. From this we conclude

$$\partial_e w = \int_{-c}^{c} \lambda(t)\partial_e w_t d\mathscr{L}^1(t) \ge a' \int_{-c}^{c} \lambda(t)|Dw_t|d\mathscr{L}^1(t) \ge a'|Dw|,$$

which, by Remark 4.1, gives $Dw/|Dw|(x) \in C'$ for $|Dw|$-a.e. x. Finally, $|Dv_t|\llcorner G << |Dw_t|$, from which we get

$$\begin{aligned} |Dv|\llcorner G = \int_{-c}^{c} |Dv_t|\llcorner G d\mathscr{L}^1(t) &<< \int_{-c}^{c} \lambda(t)|Dw_t|d\mathscr{L}^1(t) \\ &\le a' \int_{-c}^{c} \lambda(t)\partial_e w_t d\mathscr{L}^1(t) = a'\,\partial_e w \le a'\,|Dw|. \end{aligned}$$

Thus w satisfies the requirements of the proposition.

Proof of the existence of a measurable selection $t \mapsto w_t$. In order to show the existence of such a selection, we will use a general Measurable Selection Theorem due to Aumann (see Theorem III.2 in [4]). More precisely, consider the set S of functions z such that $z \in BV \cap L^\infty$ and $Dz/|Dz|(x) \in C'$ for $|Dz|$-a.e. x. We endow S with the L^∞ weak* topology.

Next, set $F_t := \{z \in S : |Dv_t|\llcorner G << |Dz|\}$ if $v_t \in BV$, and $F_t = \emptyset$ otherwise. In order to apply Aumann's Theorem we need that:

- S and $[-c,c]$ are both locally compact and separable;
- The set $F := \{(t,u) : u \in F_t\} \subset [-c,c] \times S$ is a Borel set;
- $F_t \neq \emptyset$ for $\mathscr{L}^1$-a.e. t.

This last condition has been already shown. Moreover, $[-c,c]$ is compact and separable. Thus it remains to show that S is locally compact and separable and that F is a Borel set.

S is locally compact and separable. For every $N \in \mathbb{N}$ consider the set $S_N := S \cap \{\|z\|_\infty \le N\}$. Since on bounded sets the L^∞ weak* topology is metrizable, clearly S_N is separable. Therefore, S is separable. We next show that S_N is compact, which implies that S is locally compact. Indeed consider any sequence $\{z_n\} \subset S_N$. By weak* compactness we can assume that $z_n \rightharpoonup^* z$ for some $z \in L^\infty$: Our task is to show that $z \in S$. Recall that $a'\partial_e z_n \ge |Dz_n|$. Thus $\{\partial_e z_n\}$ is a sequence of nonnegative measures which converge distributionally to $\partial_e z$. Therefore, these measures

are uniformly bounded, i.e. $\|z_n\|_{BV}$ is uniformly bounded. Thus $Dz_n \rightharpoonup^* Dz$. Up to extraction of a subsequence we can assume that $|Dz_n|$ converges in the sense of measures to some ν. Then,

$$|Dz| \leq \nu \;=\; w^* \lim_n |Dz_n| \;\leq\; a' w^* \lim_n \partial_e z_n \;=\; a' \partial_e z.$$

This implies that $z \in S$.

F is a Borel set. Denote by $\mathscr{M}^2$ the set of $\mathbb{R}^2$-valued Radon measures on B and by $\mathscr{M}^+$ the set of nonnegative Radon measures. Define $T : \mathscr{M}^+ \times \mathscr{M}^2 \to \mathbb{R}$ by

$$T(\nu,\mu) \;:=\; \int \frac{\nu}{|\mu|}(x)\, d|\mu|(x).$$

Note that $\nu << |\mu|$ if and only if $T(\nu,\mu) = \nu(B)$. Thus,

$$F \;=\; \big\{(t,z) \in [-c,c] \times S : T(|Dv_t| \llcorner G, Dz) = |Dv_t|(B \cap G)\big\}.$$

Since the map $t \mapsto |Dv_t|$ can be chosen Borel-measurable, in order to prove that F is a Borel set it suffices to show that T is a Borel function.

First of all, note that

$$T(\nu,\mu) \;=\; \sup_{n\in\mathbb{N}} \int \min\left\{n, \frac{\nu}{|\mu|}(x)\right\} d|\mu|(x) \;=\; \sup_{n\in\mathbb{N}} n \int \min\left\{1, \frac{\nu/n}{|\mu|}(x)\right\} d|\mu|(x).$$

Therefore, it suffices to show that the map $\tilde{T} : \mathscr{M}^+ \times \mathscr{M}^2 \to \mathbb{R}$ given by

$$\tilde{T}(\alpha,\mu) \;=\; \int \min\left\{1, \frac{\alpha}{|\mu|}(x)\right\} d|\mu|(x)$$

is Borel measurable. Note that $\tilde{T}(\alpha,\mu) = \inf\{\alpha(A) + |\mu|(B \setminus A) : A \subset B$ is measurable$\}$. Therefore,

$$\begin{aligned}
\tilde{T}(\alpha,\mu) &= \inf_{f\in C_c(B),\, 0<f<1} \left[\int (1-f)\, d\alpha + \int f\, d|\mu|\right] \\
&= \inf_{f\in C_c(B),\, 0<f<1} \left[\int (1-f)\, d\alpha + \sup_{g\in C_c(B,\mathbb{R}^2),\, 0\leq |g|<f} \int g \cdot d\mu\right].
\end{aligned}$$

Let $\mathscr{F}_1$ be a countable dense subset of $\{f \in C_c(B) : 0 < f < 1\}$ and $\mathscr{F}_2$ a countable dense subset of $C_c(B,\mathbb{R}^2)$. Then

$$\tilde{T}(\alpha,\mu) \;=\; \inf_{f\in\mathscr{F}_1} \sup_{g\in\mathscr{F}_2,\, 0\leq|g|<f} \left[\int (1-f)\, d\alpha + \int g \cdot d\mu\right]. \tag{10}$$

Since for each $(f,g) \in \mathscr{F}_1 \times \mathscr{F}_2$ the map

$$(\alpha,\mu) \mapsto \int (1-f)\, d\alpha + \int g \cdot d\mu$$

is weakly* continuous, (10) implies that $\tilde{T}$ is a Borel function. □

Acknowledgements The author has been supported by a grant of the Swiss National Foundation. Moreover, he wishes to thank Giovanni Alberti and Matteo Focardi for reading the manuscript and suggesting many improvements.

References

1. ALBERTI, G. *Rank-one properties for derivatives of functions with bounded variations.* Proc. Roy. Soc. Edinburgh Sect. A, **123** (1993), 239–274.
2. ALBERTI, G.; CSORNIEY, M.; PREISS, D. *Structure of null-sets in the plane and applications* European congress of mathematics, Stockholm, June 27–July 2, 2004. A. Laptev (ed.). European Mathematical Society, Zürich 2005.
3. AMBROSIO, G.; FUSCO, N.; PALLARA, D. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, 2000.
4. CASTAING, C.; VALADIER, M. *Convex analysis and measurable multifunctions.* Lecture Notes in Mathematics 580. Springer, Berlin, 1977.
5. DE GIORGI, E.; AMBROSIO, L. *Un nuovo funzionale del calcolo delle variazioni.* Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (8) Mat. Appl. **82** (1988), 199–210.
6. PREISS, D. *Geometry of measures in $\mathbb{R}^n$: distribution, rectifiability, and densities.* Ann. of Math. **125** (1987), 537–643.

Regularizing Effect of Nonlinearity in Multidimensional Scalar Conservation Laws

Gianluca Crippa[1], Felix Otto[2], and Michael Westdickenberg[3]

[1]Scuola Normale Superiore, Piazza dei Cavalieri 7, I–56126 Pisa, Italy
e-mail: g.crippa@sns.it
[2]Institut für Angewandte Mathematik, Universität Bonn, Wegelerstraße 10, D–53115 Bonn, Germany
e-mail: otto@iam.uni-bonn.de
[3]School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, Georgia 30332-0160, U.S.A.
e-mail: mwest@math.gatech.edu
URL: http://www.mathphys.iam.unibonn.de/∼otto/
URL: http://www.math.gatech.edu/∼mwest/

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 Sect. 2, there is a lack of existence of smooth solutions, even for one-dimensional 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 possible 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 non-linear 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 Sect. 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 Sect. 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 Sect. 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 one-dimensional 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 one-dimensional 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 $\mathscr{J} \subset \mathbb{R}_+ \times \mathbb{R}^n$ that is timelike in the sense that in every point $(t,x) \in \mathscr{J}$, the normal vector $\mu(t,x) \in \mathbb{R}^{n+1}$ is not parallel to the timeaxis. This implies that the normal vector can be written (up to some factor) 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 $\mathscr{J}$ 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 $(t,x) \in \mathscr{J}$: 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 $\mathscr{J}$), 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 $\mathscr{J}$.

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 only keeps those weak solutions 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 that 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 $\varepsilon \to 0$, since $\eta''(u_\varepsilon) \geq 0$ 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 Sect. 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 $\mathscr{J} \subset \mathbb{R}_+ \times \mathbb{R}^n$ that 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 $\mathscr{J}$. This last statement follows from a generalization of the *chain rule* for functions of bounded variation (see Sect. 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 one-dimensional 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 Sect. 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 the 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 $\mathbf{f}'(v)$. We can therefore not expect any regularizing effect: At any positive time, the solution is not more regular the initial data. A 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

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

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

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

which we will study in more detail later on. If (15) holds, then any sequence of entropy solutions contains a subsequence that 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 $\mathscr{J} \subset \Omega$ is called $\mathscr{H}^n$-rectifiable if $\mathscr{J} = E \cup \bigcup_{k\in\mathbb{N}} E_k$, 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 $\mathscr{J} \subset \Omega$ be an $\mathscr{H}^n$-rectifiable set. An orientation of $\mathscr{J}$ is a Borel vector field $\mu\colon \mathscr{J} \longrightarrow \mathbb{S}^n$ with the property that for $\mathscr{H}^n$-a.e. $\mathbf{x} \in \mathscr{J}$, the vector $\mu(\mathbf{x})$ is normal to $\mathscr{J}$.

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

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

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 Sect. 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

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

Here $\mathscr{J} \subset \Omega$ is an $\mathscr{H}^n$-rectifiable set oriented by a unit normal vector field $\mu\colon \mathscr{J} \longrightarrow \mathbb{S}^n$, and the functions $u^\pm\colon \mathscr{J} \longrightarrow \mathbb{R}$ are the traces of u on $\mathscr{J}$. 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

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

where the rectifiable set $\mathscr{J}$ and the functions μ and $u^\pm$ are the same as in (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

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

  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

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

These considerations show that BV-regularity is a very desirable property for entropy solutions of scalar conservation laws. As we will discuss in Sect. 4, 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

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

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 $\mathscr{J}$. We refer the reader to Sect. 3.9 in [2].

4 Structure of Entropy Solutions

As explained in Sect. 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 $\eta : \mathbb{R} \longrightarrow \mathbb{R}$ let the corresponding entropy flux $\mathbf{q}_\eta : \mathbb{R} \longrightarrow \mathbb{R}^n$ 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 nonclassical (entropy violating) shocks.

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

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

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) *we assume that*

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

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

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

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 first part of Proposition 4.3 to the Appendix, see page 114. 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 $m \in \{1, \ldots, n\}$ we then have*

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

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 117. 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

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

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 solution of (23), and prove first that the Rankine–Hugoniot condition is satisfied along each discontinuity. Indeed, consider

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

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

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

using that $\mathbf{f}(0) = 0$. To check assumption (24), consider for any convex entropy $\eta : \mathbb{R} \longrightarrow \mathbb{R}$ 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

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

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 $\mathscr{J} \subset \Omega$ such that:

- *For all $y \in \mathscr{J}$ the function u has strong traces on $\mathscr{J}$.*
- *For all $y \notin \mathscr{J}$ 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 $\mathscr{J}$ (see however the result by De Lellis and Rivière [12]). Another open problem is to prove that there is no entropy dissipation outside of $\mathscr{J}$. 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 $\mathscr{J}$ has the correct structure: If $\nu\colon \mathscr{J} \longrightarrow \mathbb{S}^{n-1}$ denotes a normal vector field along $\mathscr{J}$, then we have for all convex entropy/entropy flux pairs $(\eta, \mathbf{q}_\eta)$ that

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

To prove that $\nabla \cdot \mathbf{q}_\eta(u) = 0$ outside $\mathscr{J}$ 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. $\square$

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 nonnegative 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 the measures μ_k 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 $\mathscr{J} \subset \Omega$. We recall that in the case of an entropy solution u with BV-regularity, the jump set $\mathscr{J}$ 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 $\mathscr{J}$ the set of points with positive upper $(n-1)$-density of ν:

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

The main task is then to prove that $\mathscr{J}$ 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 $\mathscr{J} \subset \mathbb{R}^n$ 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 $y \in \mathscr{J}$ we have*
$$\liminf_{r\to 0} \frac{\nu(B_r(y))}{r^{n-1}} > 0. \tag{40}$$
- *For all $y \in \mathscr{J}$ 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 $\mathscr{J}$ 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^∞) may 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 $\nu^{y,r} \xrightarrow{*} 0$ in $\mathrm{M}_{\mathrm{loc}}(\mathbb{R}^n)$ for any blow-up sequence around a point $y \notin \mathscr{J}$. Then also $\mu^{y,r} \xrightarrow{*} 0$ 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 $\mathscr{S} \subset \mathscr{D}(\mathbb{R})$ which is dense in $\mathscr{D}(\mathbb{R})$ with respect to the uniform topology. For every $\psi \in \mathscr{S}$ we define a map $f_\psi \colon \Omega \longrightarrow \mathbb{R}$ by setting

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

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 $\nu(E_\mathscr{S}) = 0$ for $E_\mathscr{S} := \bigcup_{\psi\in\mathscr{S}} E_\psi$ since $\mathscr{S}$ is countable. Now fix a point $y \in \Omega \setminus E_\mathscr{S}$. For all $r > 0$ we define a linear map $\mathscr{F}_r$ by

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

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

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

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

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

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

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

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 $y\in\Omega\setminus E_{\mathscr{S}}$.

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 $E:=\mathscr{J}\cap E_{\mathscr{S}}$ 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 $y\in\mathscr{J}\setminus E$ there exists at least one subsequence $r_k\to 0$ such that

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

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. $\square$

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

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

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. $(x_1, \hat{x}) \in \mathbb{R}_+ \times \mathbb{R}^{n-1}$

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

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 $R > 0$

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

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 $R > 0$

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

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

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

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

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

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

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

where $\hat{\Omega} := \{\hat{x} \in \mathbb{R}^{n-1} : (0,\hat{x}) \in \Omega\}$. 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

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

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 $\nu = \mathscr{H}^{n-1} \lfloor \mathscr{J}$, where the support $\mathscr{J} = \{\eta \cdot x = 0\}$ for some unit vector η. Then the conclusion of Lemma 6.3 holds. Moreover, u is constant on either side of $\mathscr{J}$.*

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 $y \in \mathscr{J}$ 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 $\mathbb{R}^n \setminus \mathscr{J}$. 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 $y \in \mathscr{J}$ was arbitrary. The same argument applies to the lower halfspace. □

6.3 Special Split States: ν 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 $\nu = \mathscr{H}^{n-1} \lfloor \mathscr{J}$,*

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

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

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

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 $v \in I$

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

and therefore

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

Recall that u is continuous and constant along characteristics outside $\mathscr{J}$, 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 $\mathscr{L} := x + \mathbb{R}\mathbf{a}(v)/(\mathbf{a}(v) \cdot \eta)$ does not intersect $\mathscr{J}$. Then $\chi(v,u)$ is constant along $\mathscr{L}$. Thanks to (62) there exist points

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

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 $x_1,\dots,x_n$ (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 $\mathrm{Lip}(w) \le L$*

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

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

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 $\mathscr{J} \subset \mathbb{R}^n$ as above as

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

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

Step 2. We then perform a *second blow-up* around a generic point $x \in \mathscr{J}$. By rectifiability of $\mathscr{J}$, 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 $\mathscr{J}$ (up to orientation) are independent of x, thus constant along $\mathscr{J}$. They are completely determined by the function h. This allows to conclude that $\mathscr{J}$ is in fact contained in a single Lipschitz graph.

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

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

7 Proof of the Main Theorem

In Sect. 6, 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

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

where ν is the x-marginal of the entropy dissipation measure μ. From Proposition 5.9 we know that for a generic point $y \in \mathscr{J}$ 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 $\nu^\infty = g_y \mathscr{H}^{n-1} \llcorner \mathscr{J}^\infty$ for a constant $g_y > 0$, and in an appropriate coordinate system $\mathscr{J}^\infty$ has the form

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

for a Lipschitz continuous function w^∞ with $\mathrm{Lip}(w^\infty) \leqslant L_y$. All quantities with subscript y only depend on the point $y \in \mathscr{J}$, 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 $y \in \mathscr{J}$. In the notation above, we have $0 \in \mathscr{J}^\infty$.*

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 $y \in \mathscr{J}$. Assuming from now on that y is fixed, we do not write the subscript y anymore to simplify notation. We define the wedge

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

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 $\operatorname{Lip}(w^\infty) \leqslant L$, we obtain the following implication:

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

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

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

We divide the proof of the proposition into three steps.

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

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

1) $\mathscr{F}(\nu^\infty) \in [0,1]$.
2) $\mathscr{F}(\nu^\infty) = 1$ if and only if $0 \in \mathscr{J}^\infty$.
3) There exists $R > 0$ such that $\mathscr{F}(\nu^\infty) = 0$ implies $\nu^\infty(B_R(0)) = 0$.
4) For the rescaled measure $(\nu^\infty)^{0,s}$ defined as in (42) we have

$$\left.\frac{d}{ds}\right|_{s=1} \mathscr{F}\left((\nu^\infty)^{0,s}\right) \geqslant 0,$$

with equality if and only if $\mathscr{F}(\nu^\infty) \in \{0,1\}$.

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

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

up to $\mathscr{H}^{n-1}$-negligible sets. Since $\mathscr{J}^\infty$ is closed, we obtain that $0 \in \mathscr{J}^\infty$. This proves (2). Assume now that $\mathscr{F}(\nu^\infty) = 0$. Then $(W \cap \mathscr{J}^\infty) \cap B_1(0) = \varnothing$, 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 measure ν^∞. This is statement (3). Finally, by definition (42) of rescaled measures we have

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

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

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

where

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

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

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

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

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

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

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

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

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

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 $f(sr) = \mathscr{F}((\nu^{y,r})^{0,s})$ for all $r,s > 0$. We compute

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

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

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

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:

$$\begin{aligned} \text{either} \qquad\qquad 0 \in \mathscr{J}^\infty \quad & \forall \nu^\infty \\ \text{or } \nu^\infty(B_R(0)) = 0 \quad & \forall \nu^\infty. \end{aligned} \tag{73}$$

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 $y \in \mathscr{J}$. □

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

Step 1. We apply Theorem 5.6 to prove that the set $\mathscr{J}$ defined in (65) is codimension-one rectifiable: From Propositions 6.6 and 7.1 we know that for $\mathscr{H}^{n-1}$-a.e. $y \in \mathscr{J}$ there exist constants $L_y, g_y > 0$ and an orthonormal coordinate system $x_1, \dots, x_n$ 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 $\nu^\infty = g_y \mathscr{H}^{n-1} \llcorner \mathscr{J}^\infty$ with $0 \in \mathscr{J}^\infty \subset \{x_1 = 0\}$. Now we argue as follows: We define the $(n-1)$-dimensional cone

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

Then we can estimate

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

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 $\mathrm{C}_y := \big\{|x_1| \geqslant |(x_2, \dots, x_n)|\big\}$ we have

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

by compactness of $C_y \cap \bar{B}_1(0)$ (see again Example 1.63 in [2]). This gives (41) and (40). Since $r_k \to 0$ and $y \in \mathscr{J}$ 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

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

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

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

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 $\limsup$ 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

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

Fix a generic point $y \in \mathscr{J}$. 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

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

On the other hand, by (79) and rectifiability of $\mathscr{J}$, 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 $\mathscr{J}$ 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 $y \in \mathscr{J}$. 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

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

This concludes the proof of the main theorem. □

8 Proofs of the Regularity Theorems

In preparation for proving the regularity results stated in Sect. 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

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

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

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

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]). 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

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

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 $k\in\{0,\ldots,n-1\}$ we can find numbers β^k_l 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

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

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 $m \in \{1,\dots,n\}$ 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 $\|u\|_{C^{1/(n-1)}(K)} = \infty$. 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

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

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

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

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

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

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

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

Proof (of Lemma 8.1). 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 $c_1<\dots<c_n$ 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

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

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

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

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 $g\colon \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}$

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

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

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

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

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

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

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

for some $\alpha_k > 0$ that will be chosen later (see page 124). Here we only assume that α_k is large enough such that $\alpha_k^{-k}\max_l |\beta_l^k| \leqslant 1$. Then

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

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

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

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

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

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

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

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

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} \rho_\varepsilon(v-w)\,|\chi(w,x)-\chi(v,x)|\,dv\,dw \\
&\quad \leqslant \int_{\mathbb{R}} \rho_\varepsilon(z)\left(\int_{\mathbb{R}} |\chi(w,x)-\chi(w+z,x)|\,dw\right)dz \\
&\quad \leqslant \left(\int_{\mathbb{R}} \rho_\varepsilon(z)|z|\,dz\right)\int_{\mathbb{R}} \left|\tfrac{\partial}{\partial v}\chi(v,x)\right|dv
\end{aligned} \tag{105}$$

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 $(h,\varepsilon)\in\mathbb{R}\times(0,\infty)$

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

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 $\mathscr{D}'(\mathbb{R}\times\mathbb{R}^n)$

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

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

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

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

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

Recollecting terms, we can now write

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

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

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

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

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

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

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

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

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

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

$$\int_{\mathbb{R}}\int_{\mathbb{R}^n}\left|\int_{\mathbb{R}} \rho_\varepsilon(v-w)R(v,x)\,dv\right|dx\,dw \leqslant C\left(\varepsilon^{-1}\iint|\mu|\,dv\,dx+\iint|r|\,dv\,dx\right),$$

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

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

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

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

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

$$
\varepsilon := h_0^{\frac{1}{2}} \left(\iint \left|\tfrac{\partial}{\partial v}\chi\right| dv\,dx\right)^{-\frac{1}{2}} \left(\iint |\mu|\,dv\,dx\right)^{\frac{1}{2}},
\tag{112}
$$

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

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

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

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

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

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

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

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

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

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

which implies the bound

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

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 8.1 to estimate the last term in (116). We have

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

where we defined R_{m-1} as in Lemma 8.1 with

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

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

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

with A given by Lemma 8.1. 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

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

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

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

for all $m \in \{1,\ldots,n\}$. 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 8.1) 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.

Index

Editor in Chief: Franco Brezzi

Editorial Policy

1. The UMI Lecture Notes aim to report new developments in all areas of mathematics and their applications - quickly, informally and at a high level. Mathematical texts analysing new developments in modelling and numerical simulation are also welcome.

2. Manuscripts should be submitted (preferably in duplicate) to
   Redazione Lecture Notes U.M.I.
   Dipartimento di Matematica
   Piazza Porta S. Donato 5
   I – 40126 Bologna
   and possibly to one of the editors of the Board informing, in this case, the Redazione about the submission. In general, manuscripts will be sent out to external referees for evaluation. If a decision cannot yet be reached on the basis of the first 2 reports, further referees may be contacted. The author will be informed of this. A final decision to publish can be made only on the basis of the complete manuscript, however a refereeing process leading to a preliminary decision can be based on a pre-final or incomplete manuscript. The strict minimum amount of material that will be considered should include a detailed outline describing the planned contents of each chapter, a bibliography and several sample chapters.

3. Manuscripts should in general be submitted in English. Final manuscripts should contain at least 100 pages of mathematical text and should always include
   - a table of contents;
   - an informative introduction, with adequate motivation and perhaps some historical remarks: it should be accessible to a reader not intimately familiar with the topic treated;
   - a subject index: as a rule this is genuinely helpful for the reader.

4. For evaluation purposes, manuscripts may be submitted in print or electronic form (print form is still preferred by most referees), in the latter case preferably as pdf- or zipped ps- files. Authors are asked, if their manuscript is accepted for publication, to use the LaTeX2e style files available from Springer's web-server at
   ftp://ftp.springer.de/pub/tex/latex/svmonot1/ for monographs
   and at
   ftp://ftp.springer.de/pub/tex/latex/svmultt1/ for multi-authored volumes

5. Authors receive a total of 50 free copies of their volume, but no royalties. They are entitled to a discount of 33.3% on the price of Springer books purchased for their personal use, if ordering directly from Springer.

6. Commitment to publish is made by letter of intent rather than by signing a formal contract. Springer-Verlag secures the copyright for each volume. Authors are free to reuse material contained in their LNM volumes in later publications: A brief written (or e-mail) request for formal permission is sufficient.