Velocity Averaging, Kinetic Formulations, and Regularizing Effects in Quasi-Linear PDEs

EITAN TADMOR
University of Maryland, College Park

AND

TERENCE TAO
University of California, Los Angeles

To Peter Lax and Louis Nirenberg on their 80th birthdays
With friendship and admiration

Abstract

We prove in this paper new velocity-averaging results for second-order multidimensional equations of the general form $\mathcal{L}(\nabla_x, v)f(x,v) = g(x,v)$ where $\mathcal{L}(\nabla_x, v) := \mathbf{a}(v)\cdot\nabla_x - \nabla_x^\top \cdot \mathbf{b}(v)\nabla_x$. These results quantify the Sobolev regularity of the averages, $\int_v f(x,v)\phi(v)dv$, in terms of the nondegeneracy of the set $\{v : |\mathcal{L}(i\xi, v)| \le \delta\}$ and the mere integrability of the data, $(f,g) \in (L^p_{x,v}, L^q_{x,v})$. Velocity averaging is then used to study the *regularizing effect* in quasi-linear second-order equations, $\mathcal{L}(\nabla_x, \rho)\rho = S(\rho)$, which use their underlying kinetic formulations, $\mathcal{L}(\nabla_x, v)\chi_\rho = g_S$. In particular, we improve previous regularity statements for nonlinear conservation laws, and we derive completely new regularity results for convection-diffusion and elliptic equations driven by degenerate, nonisotropic diffusion.

1 Introduction

We study the regularity of solutions to multidimensional quasi-linear scalar equations of the form

$$\sum_{j=0}^{d} \frac{\partial}{\partial x_j} A_j(\rho) - \sum_{j,k=1}^{d} \frac{\partial^2}{\partial x_j \partial x_k} B_{jk}(\rho) = S(\rho) \tag{1.1}$$

where $\rho : \mathbb{R}^{1+d} \to \mathbb{R}$ is the unknown field and A_j, B_{jk}, and S are given functions from $\mathbb{R}$ to $\mathbb{R}$.

This class of equations governs time-dependent solutions $\rho(t = x_0, x)$ of nonlinear conservation laws where $A_0(\rho) = \rho$ and $B_{jk} \equiv 0$, time-dependent solutions of degenerate diffusion and convection-diffusion equations where $\{B'_{jk}\} \ge 0$, and spatial solutions $\rho(x)$ of degenerate elliptic equations where $A_j \equiv 0$. The notion of solution should be interpreted here in an appropriate weak sense, since we focus our attention on the degenerate diffusion case, which is too weak to enforce the

Communications on Pure and Applied Mathematics, Vol. LX, 1488–1521 (2007)

smoothness required for a notion of a strong solution. Instead, a common feature of such problems is the (limited) regularity of their solutions, which is dictated by the *nonlinearity* of the governing equations. A prototype example is provided by discontinuous solutions of nonlinear conservation laws. In [36], Lions, Perthame, and Tadmor have shown that entropy solutions of such laws admit a regularizing effect of a fractional order, dictated by the order of nondegeneracy of the equations. In this paper we extend this result for the general class of second-order equations (1.1). In particular, we improve the regularity statement in [36] for nonlinear conservation laws, and we derive completely new regularity results for convection-diffusion and elliptic equations driven by degenerate, nonisotropic diffusion.

The derivation of these regularity results employs a *kinetic formulation* of (1.1). To describe this formulation let us proceed formally, seeking an equation that governs the indicator function associated with ρ, $\chi_{\rho(x)}(v) := \mathrm{sgn}(v)(|\rho| - |v|)_+$; it depends on an auxiliary, so-called velocity variable $v \in \mathbb{R}$, borrowing the terminology from the classical kinetic framework. To this end, we consider the distribution $g = g(x, v)$, defined via its velocity derivative $\partial_v g$ using the formula

$$\partial_v g(x, v) := \big(\mathbf{a}(v) \cdot \nabla_x - \nabla_x^\top \cdot \mathbf{b}(v)\nabla_x + S(v)\partial_v\big)\chi_{\rho(x)}(v), \tag{1.2}$$

where the convection vector $\mathbf{a}$ and diffusion matrix $\mathbf{b}$ are given by $\mathbf{a}_j := A_j'$ and $\mathbf{b}_{jk} := B_{jk}' \geq 0$. Observe that the nonlinear quantities $\Phi(\rho)$ can be expressed as the v-moments of χ_ρ, $\Phi(\rho) \equiv \int_v \Phi'(v)\chi_\rho(v)dv$, $\Phi(0) = 0$. Therefore, by velocity averaging of (1.2), we recover (1.1). Moreover, for a proper notion of weak solution ρ, one augments (1.1) with additional conditions on the behavior of $\Phi(\rho)$ for a large enough family of entropies Φ. These additional entropy conditions imply that g is in fact a positive distribution $g = m \in \mathcal{M}^+$ that measures the entropy dissipation of the nonlinear equation.

We arrive at the kinetic formulation of (1.1)

$$\mathcal{L}(\nabla_x, v)f(x, v) = \partial_v m - S(v)\partial_v \chi_{\rho(x)}(v), \quad f = \chi_\rho,\ m \in \mathcal{M}^+, \tag{1.3}$$

where $\mathcal{L}$ is identified with the linear symbol $\mathcal{L}(i\xi, v) := \mathbf{a}(v) \cdot i\xi - \langle \mathbf{b}(v)\xi, \xi\rangle$. We recall that $\rho(x)$ itself can be recovered by velocity averaging of $f(x, v) = \chi_\rho(v)$ via the identity $\rho(x) = \int f(x, v)dv$. In Section 2 we discuss the regularity gained by such velocity averaging.

There is a relatively short yet intense history of such regularity results, commonly known as "velocity-averaging lemmas." We mention the early works of [20, 24] and their applications, in the context of nonlinear conservation laws, in [29, 35, 36, 37]; a detailed list of references can be found in [39] and is revisited in Section 2 below. Almost all previous averaging results dealing with (1.3) are restricted to the first-order transport equation, $\deg \mathcal{L}(i\xi, \cdot) = 1$. In Section 2.2 we present an extension to general symbols $\mathcal{L}$ that satisfy the so-called *truncation property*; luckily, as shown in Section 2.4 below, all $\mathcal{L}$ with $\deg \mathcal{L}(i\xi, \cdot) \leq 2$ satisfy this property. If $\mathcal{L}$ satisfies the truncation property and it is *nondegenerate* in the sense that there exist $\alpha \in (0,1)$ and $\beta > 0$ such that[1]

$$\sup_{|\xi|\sim J} |\{v : \ |\mathcal{L}(i\xi, v)| \leq \delta\}| \lesssim \left(\frac{\delta}{J^\beta}\right)^\alpha,$$

then, by velocity averaging of the kinetic solution f in (1.3), $\rho(x) = \int_v f(x,v)dv$ is shown to have $W^{s,r}$-regularity of order $s < \beta\alpha/(3\alpha+2)$ with an appropriate $r = r_s > 1$; consult (2.25) below. In the particular case of a truncation property satisfying $\mathcal{L}$ with homogeneous symbol of order k, the nondegeneracy requirement amounts to having an $\alpha \in (0,1)$ such that

$$\sup_{|\xi|=1} |\{v : |\mathcal{L}(i\xi, v)| \leq \delta\}| \lesssim \delta^\alpha,$$

so we may take $\beta = k$, and velocity averaging implies a $W^{s,r}_{loc}$-regularity with exponent $s < k\alpha/(3\alpha+2)$. The main results are summarized in Averaging Lemmas 2.2 and 2.3.

In Section 3 we turn to the first application of these averaging results in the context of nonlinear conservation laws, $\rho_t + \nabla_x \cdot A(\rho) = 0$, subject to L^∞-initial data ρ_0. If the equation is nondegenerate of order α in the sense that for all $\tau^2 + |\xi|^2 = 1$, there holds

$$|\{v : |\tau + \mathbf{a}(v)\cdot\xi| \leq \delta\}| \lesssim \delta^\alpha \quad \text{and} \quad \sup_{\{v : |\tau+\mathbf{a}(v)\cdot\xi|\leq\delta\}} |\mathbf{a}'(v)\cdot\xi| \lesssim \delta^{1-\alpha},$$

then for $t > 0$, the corresponding entropy solution $\rho(t,\cdot) \in W^{s,1}_{\mathrm{loc}}(\mathbb{R}^d_x)$ gains Sobolev regularity of order $s < s_\alpha = \alpha/(2\alpha+1)$. This improves the Sobolev-regularity exponent of order $\alpha/(\alpha+2)$ derived at [36] (while facing the same barrier of $s_1 = \frac{1}{3}$ discussed in [15]).

Section 4 is devoted to convection-diffusion equations. We begin, in Section 4.1, with second-order degenerate diffusion $\rho_t - \sum \partial^2_{x_j x_k} B_{jk}(\rho) = 0$. The emphasis here is on nonisotropic diffusion, beyond the prototype case of the porous medium equation (which corresponds to the case when B_{jk} is a scalar multiple of the identity $B_{jk} = B\delta_{jk}$). The regularizing effect is determined by the smallest nonzero eigenvalue $\lambda(v) = \lambda(\mathbf{b}(v)) \not\equiv 0$ of $\mathbf{b}(v) := B'(v)$ such that

$$|\{v : 0 \leq \lambda(v) \leq \delta\}| \lesssim \delta^\alpha \quad \text{and} \quad \sup_{|\xi|=1} \sup_{\{v:\ \lambda(v)\leq\delta\}} |\langle \mathbf{b}'(v)\xi, \xi\rangle| \lesssim \delta^{1-\alpha}. \tag{1.4}$$

Starting with initial conditions $\rho_0 \in L^\infty$, the corresponding kinetic solution $\rho(t > 0, \cdot)$ gains $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < 2\alpha/(2\alpha+1)$. In Section 4.2 we take into account the additional effect of nonlinear convection. The resulting convection-diffusion equations, coupling degenerate and possibly nonisotropic diffusion with nonconvex convection governing capillarity effects, are found in a variety of applications.

[1] We use $X \lesssim Y$ to denote the estimate $X \leq CY$ where C is a constant that can depend on exponents such as α, β, and p and on symbols such as $\mathcal{L}$, A_j, and B_{jk} but is independent of fields such as ρ, coordinates such as x, t, and v, and scale parameters such as δ. We use $X \sim Y$ to denote the assertion that $X \lesssim Y \lesssim X$.

Consider the prototype one-dimensional case

$$\rho_t + A(\rho)_x - B(\rho)_{xx} = 0, \quad A'(\rho) \sim \rho^\ell,\ B'(\rho) \sim |\rho|^n.$$

The regularizing effect is dictated by the strength of the degenerate diffusion versus the convective degeneracy. If $n \le \ell$ we find $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < 2/(n+2)$, which is the same Sobolev-regularity exponent we find with the "purely diffusive" porous-medium equation, i.e., when $A = 0$. On the other hand, if the diffusion is too weak so that $n \ge 2\ell$, we then conclude with a Sobolev regularity exponent of order $s < 1/(\ell + 2)$, which is dominated by the convective part of the equation. In Section 4.2 we present similar results for multidimensional convection-diffusion equations with increasing degree of degeneracy.

In particular, consider the two-dimensional equation

$$\rho_t + (\partial_{x_1} + \partial_{x_2})A(\rho) - (\partial_{x_1} - \partial_{x_2})^2 B(\rho) = 0.$$

If we set $A = 0$, the equation has a strong, rank-1 parabolic degeneracy with no regularizing effect coming from its purely diffusion part, since $\langle \mathbf{b}(v)\xi, \xi\rangle \equiv 0$ $\forall \xi_1 - \xi_2 = 0$, indicating the persistence of steady oscillations $\rho_0(x + y)$; moreover, if $B = 0$ then the equation has no regularization coming from its purely convection part, since $\mathbf{a}(v) \cdot \xi \equiv 0\ \ \forall \xi_1 + \xi_2 = 0$ indicates the persistence of steady oscillations $\rho_0(x - y)$. Nevertheless, the combined *convection-diffusion* with $A(\rho) \sim \rho^{\ell+1}$ and $B(\rho) \sim |\rho|^n\rho$ does have a $W^{s,1}$-regularizing effect of order $s < 6/(2 + 2n - \ell)$ for $n \ge 2\ell$; consult Corollary 4.5 below.

Finally, in Section 5 we consider degenerate elliptic equations,

$$-\sum \partial^2_{x_j x_k} B_{jk}(\rho) = S(\rho).$$

Assuming that the nondegeneracy condition (1.4) holds, then the kinetic formulation of bounded solutions for such equations, $\rho \in W^{s,1}(D)$, have interior regularity of order $s < \min(\alpha, 2\alpha/(2\alpha + 1))$. We conclude by noting that it is possible to adapt our arguments in more general setups, for example, when a suitable source term $S(\rho)$ is added to the time-dependent problem when dealing with degenerate *temporal* fluxes, $\partial_t A_0(\rho),\ A_0' \ge 0$, or when lower-order convective terms $\nabla_x \cdot A(\rho)$ are added in the elliptic case.

2 The Averaging Lemma

We are concerned with the regularity of averages of solutions for differential equations of the form

$$\mathcal{L}(\nabla_x, v) f = \Lambda_x^\eta \partial_v^N g, \quad \Lambda_x := (-\Delta_x^2)^{1/2}. \tag{2.1}$$

Here, $f = f(x, v) \in W^{\sigma,p}_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$ and $g = g(x, v) \in L^q_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$ are real-valued functions of the spatial variables $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and an additional parameter $v \in \mathbb{R}$, called *velocity* by analogy with the kinetic framework, and $\mathcal{L}(\nabla_x, v)$ is a differential operator on $\mathbb{R}^d_x$ of order $\le k$, whose coefficients are smooth functions of v.

The velocity-averaging lemma asserts that if $\mathcal{L}(\cdot, v)$ is *nondegenerate* in the sense that its null set is sufficiently small (to be made precise below), then the v-moments of $f(x, \cdot)$,

$$\overline{f}(x) := \int_v f(x, v)\phi(v)dv, \quad \phi \in C_0^\infty,$$

are smoother than the usual regularity associated with the data of $f(x, \cdot)$ and $g(x, \cdot)$. That is, by averaging over the so-called microscopic v-variable, there is a *gain* of regularity in the macroscopic x-variables.

There is a relatively short yet intense history of such results, motivated by kinetic models such as Boltzmann, Vlasov, radiative transfer, and similar equations where the v-moments of f represent macroscopic quantities of interest. We refer to the early works of Agoshkov [1] and Golse, Lions, Perthame, and Sentis [24, 25] treating first-order transport operators with f and g integrability of order $p = q > 1$. The work of DiPerna, Lions, and Meyer [20] provided the first treatment of the general case $p \neq q$, followed by Bézard [4], the improved regularity in [34], and an optimal Besov regularity result of DeVore and Petrova [16].

Extensions to more general streaming operators were treated by DiPerna and Lions in [18, 19] with applications to Boltzmann and Vlasov-Maxwell equations, and by Lions, Perthame, and Tadmor in [36] with applications to nonlinear conservation laws and related parabolic equations. Gérard [22] together with Golse [23] provided an L^2-treatment of general differential operators. A different line of extensions consists of velocity-averaging lemmas that take into account different orders of integrability in x and v, leading to sharper velocity regularization results sought in various applications. We refer to the results in [20] for $f \in L^{p_1}(\mathbb{R}_v, L^{p_2}(\mathbb{R}_x^d))$, $g \in L^{q_1}(\mathbb{R}_v, L^{q_2}(\mathbb{R}_x^d))$ and to Jabin and Vega [30] for $f \in W^{N_1,p_1}(\mathbb{R}_v, L^{p_2}(\mathbb{R}_x^d))$, $g \in W^{N_2,q_1}(\mathbb{R}_v, L^{q_2}(\mathbb{R}_x^d))$. Westdickenberg [54] analyzed a general case of the form $f \in B^\sigma_{p_1,p_2}(\mathbb{R}_x^d, L^{r_1}(\mathbb{R}_v))$, $g \in B^\sigma_{q_1,q_2}(\mathbb{R}_x^d, L^{r_2}(\mathbb{R}_v))$. Jabin, Perthame, and Vega [29, 31] used a mixed integrability of

$$f \in L^p(\mathbb{R}_x^d, W^{N_1,p}(\mathbb{R}_v)), \quad g \in L^q(\mathbb{R}_x^d, W^{N_2,q}(\mathbb{R}_v)),$$

to improve the regularizing results for nonlinear conservation laws [36, 37, 35], Ginzburg-Landau, and other nonlinear models. Golse and Saint-Raymond [26] showed that a minimal requirement of *equi-integrability*, say $f \in L^1(\mathbb{R}_x^d, L^\Phi(\mathbb{R}_v))$ and $g \in L^1(\mathbb{R}_x^d \times \mathbb{R}_v)$ measured in Orlicz space L^Φ with superlinear Φ, is sufficient for relative compactness of the averages $\overline{f}$, which otherwise might fail for mere L^1-integrability [24].

The derivation of velocity averaging in the above works was accomplished by various methods. The main approach, which we use below, is based on decomposition in Fourier space, carefully tracking $\widehat{f}(\xi, v)$ in the "elliptic" region where $\{v : \mathcal{L}(i\xi, v) \neq 0\}$ and the complement region, which is made sufficiently small by a nondegeneracy assumption. Other approaches include the use of microlocal defect measures and H-measures [22, 48], wavelet decomposition [16], and "real-space methods"—in time [5, 52] and in space-time using Radon transform [12, 54], X-transform [30, 31], and duality-based dispersion estimates [26]. Almost all of these results are devoted to the phenomena of velocity averaging in the context of transport equations, $k = 1$.

Our study of velocity averaging applies to a large class of $\mathcal{L}$'s satisfying the so-called *truncation property*: in Section 2.4 we show that all $\mathcal{L}$'s of order $k \leq 2$ satisfy this truncation property. In particular, we improve the regularity statement for first-order velocity averaging and extend the various velocity-averaging results of the above works from first-order transport to general second-order transport-diffusion and elliptic equations. The results are summarized in Averaging Lemmas 2.1 and 2.2 for homogeneous symbols and in Averaging Lemma 2.3 for general, truncation-property-satisfying $\mathcal{L}$'s. Our derivation is carried out in Fourier space using Littlewood-Paley decompositions of $f \in W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$ and $g \in L^q_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$. To avoid an overload of indices, we leave for future work possible extensions for more general data with mixed (x, v)–integrability of f and g.

2.1 The Truncation Property

We now come to a fundamental definition.

DEFINITION 2.1 Let $m(\xi)$ be a complex-valued Fourier multiplier. We say that m has the *truncation property* if, for any locally supported bump function ψ on $\mathbb{C}$ and any $1 < p < \infty$, the multiplier with symbol $\psi(m(\xi)/\delta)$ is an L^p-multiplier uniformly in $\delta > 0$, that is, its L^p-multiplier norm depends solely on the support and C^ℓ size of ψ (for some large ℓ that may depend on m) but otherwise is independent of δ.

In Section 2.4 we will describe some examples of multipliers with this truncation property. Equipped with this notion of a truncation property, we turn to discuss the L^p-size of *parametrized multipliers*. Let

$$\mathsf{M}_\psi f(x, v) := \mathcal{F}_x^{-1} \psi\left(\frac{m(\xi, v)}{\delta}\right) \widehat{f}(\xi, v)$$

denote the operator associated with the complex-valued multiplier $m(\,\cdot\,, v)$, truncated at level δ, where $\widehat{f}(\xi, v) = \mathcal{F}_x f(\xi, v) := \int_{\mathbb{R}^d} e^{-2\pi i \xi \cdot x} f(x, v) dx$ is the Fourier transform of $f(\,\cdot\,, v)$. Our derivation of the averaging lemmas below is based on the following straightforward estimate for such parameterized multipliers.

LEMMA 2.2 (Basic L^p-Estimate) *Let I be a finite interval, $I \subset \mathbb{R}_v$, and assume $m(\xi, v)$ satisfies the truncation property uniformly in $v \in I$. Let $1 < p \leq 2$. Let $\overline{\mathsf{M}_\psi}$ denote the velocity-averaged Fourier multiplier*

$$\overline{\mathsf{M}_\psi f}(x) := \int_I \mathsf{M}_\psi f(x, v) dv = \int_I \mathcal{F}_x^{-1} \psi\left(\frac{m(\xi, v)}{\delta}\right) \mathcal{F}_x f(x, v) dv.$$

For each $\xi \in \mathbb{R}^d$ and $\delta > 0$, let $\Omega_m(\xi;\delta) \subset I$ be the velocity set

$$\Omega_m(\xi;\delta) := \left\{ v \in I : \frac{m(\xi, v)}{\delta} \in \operatorname{supp}\psi \right\}.$$

Then we have the L^p-multiplier estimate

$$\|\overline{M_\psi f}(x)\|_{L^p(\mathbb{R}^d_x)} \lesssim \sup_\xi |\Omega_m(\xi;\delta)|^{1/p'} \cdot \|f\|_{L^p(\mathbb{R}^d_x \times \mathbb{R}_v)}. \tag{2.2}$$

PROOF: For $p = 2$, the claim (2.2) follows from Plancherel's theorem and the Cauchy-Schwarz inequality, while for p close to 1_+ the claim follows (ignoring the bounded $|\Omega_m(\xi;\delta)|$ factor) from the assumption that $m(\xi, v)$ satisfies the truncation property uniformly in v (and, in fact, the endpoint $p = 1$, with the usual $\mathcal{H}^1$ replacement of L^1, can be treated by a refined argument along the lines of [20]). The general case of $1 < p < 2$ follows by interpolation. $\square$

Remark 2.3. Clearly, if $m(\xi, \cdot)$ and $c(\xi)$ are L^p-multipliers, then so is their product, and in particular, if m has the truncation property, then (2.2) applies for

$$\psi\left(\frac{m(\xi, v)}{\delta}\right) c(\xi).$$

2.2 Averaging Lemma for Homogeneous Symbols

We will present several versions of the averaging lemma. The later versions will supersede the former, but for pedagogical reasons we will start with the simpler case of homogeneous symbols. In this case it is convenient to use polar coordinates $\xi = |\xi|\xi'$, where $\xi' \in S^{d-1}$ is defined for all nonzero frequencies ξ by $\xi' := \xi/|\xi|$. We begin with the following:

AVERAGING LEMMA 2.1 *Let $1 \le q \le 2$ and let $g \in L^q_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$ if $q > 1$ or let g be a locally bounded measure, $g \in \mathcal{M}(\mathbb{R}^d_x \times \mathbb{R}_v)$, if $q = 1$. Let $\eta, N \ge 0$ and let $f \in W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$, $\sigma \ge 0$, $1 < p \le 2$, solve the equation*

$$\mathcal{L}(\nabla_x, v) f(x, v) = \Lambda^\eta_x \partial^N_v g(x, v) \quad \text{in } \mathcal{D}'(\mathbb{R}^d_x \times \mathbb{R}_v). \tag{2.3}$$

Here $\mathcal{L}(\nabla_x, \cdot)$ is a differential operator that has sufficiently smooth coefficients, $\mathcal{L}(\cdot, v) \in C^{N,\varepsilon>0}$, and let $\mathcal{L}(i\xi, v)$ be the corresponding symbol. We assume that $\mathcal{L}(i\xi, v)$ is homogeneous in ξ of order k, $k > \sigma + \eta$, that the modified symbol $\mathcal{L}(i\xi', v)$ obeys the truncation property uniformly in v, and that it is nondegenerate in the sense that there exists an α, $0 < \alpha < (N+1)q'$, such that

$$\sup_{\substack{\xi \in \mathbb{R}^d \\ |\xi|=1}} |\Omega_{\mathcal{L}}(\xi;\delta)| \lesssim \delta^\alpha, \quad \Omega_{\mathcal{L}}(\xi;\delta) := \{v \in I : |\mathcal{L}(i\xi, v)| \le \delta\}. \tag{2.4}$$

Then there exist $\theta = \theta_\alpha \in (0,1)$ and $s_\alpha = s(\theta_\alpha) > \sigma$ such that for all bump functions $\phi \in C^\infty_0(I)$, the averages $\overline{f}(x) := \int f(x, v)\phi(v)dv$ belong to the Sobolev space $W^{s,r}_{\text{loc}}(\mathbb{R}^d_x)$ for all $s \in (\sigma, s_\alpha)$ and the following estimate holds:

$$\|\overline{f}\|_{W^{s,r}_{\text{loc}}(\mathbb{R}^d_x)} \lesssim \|g\|^\theta_{L^q_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)} \cdot \|f\|^{1-\theta}_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)}, \quad s \in (\sigma, s_\alpha). \tag{2.5}$$

Here $s_\alpha := (1-\theta_\alpha)\sigma + \theta_\alpha(k-\eta)$, *where* $\theta \equiv \theta_\alpha(p,q,N)$ *and* r *are given by*

$$\theta = \frac{\alpha/p'}{\alpha(1/p'-1/q')+N+1}, \quad \frac{1}{r} = \frac{1-\theta}{p} + \frac{\theta}{q}, \quad 0<\theta<1. \tag{2.6}$$

Remark 2.4. It would be more natural to assume that the symbol $\mathcal{L}(i\xi, v)$ itself, rather than the modified symbol $\mathcal{L}(i\xi', v)$, obeyed the truncation property, as it is typically easier to verify the truncation property for the unmodified symbol. Indeed, when we turn to more advanced versions of the averaging lemma (which rely on Littlewood-Paley theory and do not assume homogeneity) we will work with the truncation property for the unmodified symbol. However, we choose to work here with the modified symbol as it simplifies the argument slightly.

PROOF: We start with a smooth partition of unity, $1 \equiv \sum \psi_j(2^{-j}z)$, such that ψ_0 is a bump function supported inside the disc $|z| \le 2$ and the other ψ_j are bump functions supported on the annulus $\frac{1}{2} < |z| < 2$ (we note in passing that the other ψ_j's can be taken to be equal, so the index j merely serves to signal their "action" on the shells, $2^{j-1} \le |z| \le 2^{j+1}$). We set

$$f_j(x,v) := \mathcal{F}_x^{-1}\psi_j\left(\frac{\mathcal{L}(i\xi', v)}{2^j\delta}\right)\widehat{f}(\xi, v), \quad j = 0,1,2,\dots,$$

recalling that $\xi' := \xi/|\xi|$, and we consider the corresponding decomposition $f = f_0 + \sum_{j\ge1} f_j$. We distinguish between two pieces, $f = f^{(0)} + f^{(1)}$ where $f^{(0)} := f_0$ and $f^{(1)} := \sum_{j\ge1} f_j$. Observe that the v-support of $\widehat{f^{(0)}}$ is restricted to the degenerate set $\Omega_{\mathcal{L}}(\xi; 2\delta)$, whereas $\widehat{f^{(1)}} = \sum_{j\ge1}\widehat{f_j}$ offers a decomposition of the nondegenerate complement, $\Omega^c_{\mathcal{L}}(\xi;\delta)$. The free parameter δ is to be chosen later.

We first note that $f^{(0)} = f_0$ is associated with the multiplier $\psi_0(\mathcal{L}(i\xi', v)/\delta)$. Since $\mathcal{L}(i\xi', v)$ satisfies the truncation property, we can use Lemma 2.2 and the nondegeneracy assumption (2.4) to obtain

$$\begin{aligned}\|\overline{f^{(0)}}\|_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x)} &\lesssim \sup_{|\xi|=1}\left|\Omega_{\mathcal{L}}(\xi;2\delta)\right|^{1/p'} \cdot \|f\|_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)} \\ &\lesssim \delta^{\alpha/p'}\|f\|_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)}.\end{aligned} \tag{2.7}$$

We turn to the other averages $\overline{f_j}$, $j \ge 1$, which make $f^{(1)}$. Since $\mathcal{L}(i\xi,\cdot)$ is homogeneous of order k, equation (2.3) states that

$$\widehat{\Lambda_x^{k-\eta} f}(\xi, v) = |\xi|^{k-\eta}\frac{|\xi|^\eta}{\mathcal{L}(i\xi, v)}\,\partial_v^N\widehat{g}(\xi,v) = \frac{1}{\mathcal{L}(i\xi', v)}\,\partial_v^N\widehat{g}(\xi,v),$$

and thus

$$\Lambda_x^{k-\eta}\overline{f_j} = \frac{1}{2^j\delta}\mathcal{F}_x^{-1}\int_v \widetilde{\psi}_j\left(\frac{\mathcal{L}(i\xi', v)}{2^j\delta}\right)\partial_v^N\widehat{g}(\xi,v)\phi(v)dv, \quad j = 1,2,\dots,$$

where $\widetilde{\psi}_j(z) := \psi_j(z)/z$ is a bump function much like ψ_j is. Integration by parts then yields

$$\Lambda_x^{k-\eta}\overline{f_j} = \frac{1}{(2^j\delta)^{N+1}}\mathcal{F}_x^{-1}\int_v \widetilde{\psi}_j^{(N)}\left(\frac{\mathcal{L}(i\xi', v)}{2^j\delta}\right)\mathcal{L}_v^N(i\xi', v)\widehat{g}(\xi, v)\phi(v)dv \\ + \text{lower-order or similar terms.} \tag{2.8}$$

We can safely neglect the lower- or similar-order terms that involve (powers of) the bounded multipliers $\partial_v^\ell \mathcal{L}(i\xi', v)$, $\ell < N$, and we focus on the leading term in (2.8) associated with the multipliers

$$\widetilde{\psi}_j^{(N)}\left(\frac{\mathcal{L}(i\xi', v)}{2^j\delta}\right)\mathcal{L}_v^N(i\xi', v). \tag{2.9}$$

By the Hörmander-Mikhlin or Marcinkiewicz multiplier theorems, $\mathcal{L}_v^N(i\xi', v)$ are bounded multipliers and hence (2.8) are upper-bounded by

$$\|\Lambda_x^{k-\eta}\overline{f_j}\|_{L^q_{\text{loc}}(\mathbb{R}_x^d)} \lesssim \frac{1}{(2^j\delta)^{N+1}}\|\overline{\mathsf{M}_j g}\|_{L^q_{\text{loc}}(\mathbb{R}_x^d)}.$$

Here $\mathsf{M}_j = \mathsf{M}_{\widetilde{\psi}_j^{(N)}}$ are the Fourier multipliers with symbol $\widetilde{\psi}_j^{(N)}(\mathcal{L}(i\xi', v)/2^j\delta)$.

We now fix $q > 1$. By our assumption, $\mathcal{L}(i\xi', v)$ satisfies the truncation property, and Lemma 2.2 implies $\|\overline{\mathsf{M}_j g}\|_{L^q_{\text{loc}}(\mathbb{R}_x^d)} \lesssim (2^{j+1}\delta)^{\alpha/q'}\|g\|_{L^q_{\text{loc}}(\mathbb{R}_x^d)}$. Adding together all the f_j's, we find that $f^{(1)} = \sum_{j\geq 1} f_j$ satisfies

$$\begin{aligned}\|\Lambda_x^{k-\eta}\overline{f^{(1)}}\|_{L^q_{\text{loc}}(\mathbb{R}_x^d)} &\lesssim \sum_{j\geq 1}\frac{1}{(2^j\delta)^{N+1}}(2^{j+1}\delta)^{\alpha/q'}\|g\|_{L^q_{\text{loc}}(\mathbb{R}_x^d\times\mathbb{R}_v)} \\ &\lesssim \delta^{\alpha/q'-(N+1)}\|g\|_{L^q_{\text{loc}}(\mathbb{R}_x^d\times\mathbb{R}_v)}.\end{aligned} \tag{2.10}$$

Thus, if we fix $t > 0$ and choose δ to equilibrate the bounds in (2.7) and (2.10),

$$\delta^{\alpha(1/p'-1/q')+N+1} \sim \frac{t\|g(x, v)\|_{L^q_{\text{loc}}}}{\|f(x, v)\|_{W^{\sigma,p}_{\text{loc}}}},$$

then this tells us that

$$\begin{aligned}\inf_{\overline{f^{(0)}}+\overline{f^{(1)}}=\overline{f}} &\left[\|\overline{f^{(0)}}\|_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}_x^d)} + t\|\overline{f^{(1)}}\|_{\dot{W}^{k-\eta,q}_{\text{loc}}(\mathbb{R}_x^d)}\right] \\ &\lesssim t^\theta\cdot\|g(x, v)\|^\theta_{L^q_{\text{loc}}(\mathbb{R}_x^d\times\mathbb{R}_v)}\cdot\|f(x, v)\|^{1-\theta}_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}_x^d\times\mathbb{R}_v)},\end{aligned}$$

and the desired $W^{s,r}_{\text{loc}}$-bound follows for $s < (1-\theta)\sigma + \theta(k-\eta)$ with θ given in (2.6). The remaining case of $q = 1$ can be converted into the previous situation using Sobolev embedding. In this case, g being a measure, it belongs to $W^{-\epsilon,q_\epsilon}$ for all (ϵ, q_ϵ) such that

$$g \in W^{-\epsilon,q_\epsilon}, \quad \frac{d+2}{q'_\epsilon} < \epsilon < 1 < q_\epsilon < \frac{d+2}{d+1},$$

and hence (2.5) applies for $s < (1-\theta)\sigma + \theta(k-\eta-\epsilon)$ and $\theta = \theta_\alpha(p, q_\epsilon, N)$; we then let ϵ approach 0_+ so that q_ϵ approaches arbitrarily close to 1_+ to recover (2.5) with s_α and $\theta_\alpha(p, 1, N)$. □

Remark 2.1. As an example, consider a (possibly pseudo-) differential operator $\mathcal{L}(\nabla_x, \cdot)$ of order k and let $f(x, v) \in W^{\sigma,2}_{\mathrm{loc}}$ such that $\mathcal{L}(\nabla_x, v) f \in W^{\sigma-k+1,2}_{\mathrm{loc}}$. Assume that $\mathcal{L}_v(i\xi, v) \neq 0$ so that the nondegeneracy condition (2.4) holds with $\alpha = 1$. Application of Averaging Lemma 2.1 with $p = q = 2$, $N = 0$, and $\eta = k - \sigma - 1$ then yields the gain of half a derivative, $\overline{f}(x) \in W^{s,2}_{\mathrm{loc}}$ with $s < \sigma/2 + (\sigma+1)/2 = \sigma + 1/2$, in agreement with [23, theorem 2.1]. The main aspect here is going beyond the L^2-framework, while allowing for general and possibly different orders of integrability, $(f, g) \in (W^{\sigma,p}, L^q)$.

Remark 2.2. The limiting case of the interpolation estimate (2.5), $\theta = 1$, $s = s_\alpha$, corresponds to Besov regularity $\overline{f} \in B^{s,r}_{t=\infty}(\mathbb{R}^d_x)$. This regularity can be worked out using a more precise bookkeeping of the Littlewood-Paley blocks. For the transport case, $k = 1$, it was carried out first in [20, theorem 3] and improved in [4], and a final refinement with a secondary index $t = p$ can be found in [16]. This limiting case is encountered in the particular situation when $k = \eta$, so that the interval (σ, s_α) "survives" at $\theta = 1$. Here one cannot expect a regularizing effect, but there is a persistence of relative compactness of the mapping $g(x, v) \mapsto \overline{f}(x)$ [41].

To gain better insight into the last averaging lemma, we focus our attention on the case where f is a $W^{\sigma,p}_{\mathrm{loc}}$-solution of

$$\begin{gathered} \mathcal{L}(\nabla_x, v) f(x, v) = \partial_v g(x, v), \\ g \in L^q_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v), \;\; f \in W^{\sigma,p}_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v), \end{gathered} \tag{2.11}$$

corresponding to the special case $\eta = 0$ and $N = 1$ in (2.3). This case will suffice to cover all the single-valued applications we have in mind for the discussion in Sections 3, 4, and 5 without the burden of carrying out an excessive amount of indices. Averaging Lemma 2.1 implies that $\overline{f}(x)$ has Sobolev regularity of order $s < (1-\theta)\sigma + \theta k$,

$$\begin{gathered} \|\overline{f}\|_{W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x)} \lesssim \|g\|^{\theta}_{L^q_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)} \cdot \|f\|^{1-\theta}_{L^p_{\mathrm{loc}}(\mathbb{R}^d_x)}, \\ \theta \equiv \theta_\alpha := \frac{\alpha/p'}{\alpha(1/p' - 1/q') + 2}. \end{gathered} \tag{2.12}$$

The last regularity statement can be improved. To this end, we revisit the dyadic multipliers in (2.9), $\widetilde{\psi}_j^{(N)}(\mathcal{L}(i\xi', v)/2^j\delta)\mathcal{L}^N_v(i\xi', v)$. The key observation is that $\mathcal{L}_v(i\xi', v)$ acts only on the *subset* of v's—those that belong to $v \in \Omega_{\mathcal{L}}(\xi'; 2^{j+1}\delta)$. Linking the size of $\mathcal{L}_v(i\xi', v)$ to that of $\mathcal{L}(i\xi', v)$, we arrive at the following improved averaging regularity lemma:

AVERAGING LEMMA 2.2 *Let $f \in W^{\sigma,p}_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$, $\sigma \geq 0$, $1 < p \leq 2$, solve the equation*

$$\mathcal{L}(\nabla_x, v) f(x,v) = \partial_v g(x,v), \quad g \in \begin{cases} L^q(\mathbb{R}^d_x \times \mathbb{R}_v), & 1 < q \leq 2, \\ \mathcal{M}(\mathbb{R}^d_x \times \mathbb{R}_v), & q = 1. \end{cases} \tag{2.13}$$

Let $\mathcal{L}(i\xi, v)$ be the corresponding symbol. We assume that $\mathcal{L}(i\xi, v)$ is homogeneous in $i\xi$ of order k, $k > \sigma$, that the modified symbol $\mathcal{L}(i\xi', v)$ satisfies the truncation property uniformly in v, and that it is nondegenerate in the sense that there exists an α, $0 < \alpha < q'$, such that (2.4) *holds. Moreover, assume that there exists a $\mu \in [0,1]$ such that*

$$\sup_{|\xi|=1} \sup_{v \in \Omega_{\mathcal{L}}(\xi;\delta)} |\mathcal{L}_v(i\xi, v)| \lesssim \delta^{\mu}, \quad \Omega_{\mathcal{L}}(\xi;\delta) := \{v \in I : |\mathcal{L}(i\xi, v)| \leq \delta\}. \tag{2.14}$$

Then there exist $\theta = \theta_\alpha \in (0,1)$ and r given by

$$\theta := \frac{\alpha/p'}{\alpha(1/p' - 1/q') + 2 - \mu}, \quad \frac{1}{r} := \frac{1-\theta}{p} + \frac{\theta}{q}, \tag{2.15}$$

such that, for all bump functions $\phi \in C_0^\infty(I)$, the averages

$$\overline{f}(x) := \int f(x,v)\phi(v)\,dv$$

belong to the Sobolev space $W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x)$ for all $s \in (\sigma, s_\alpha)$, $s_\alpha = (1-\theta_\alpha)\sigma + \theta_\alpha k$, and (2.5) *holds.*

PROOF: For the proof we revisit (2.8) with $\eta = 0$, $N = 1$,

$$\Lambda^k_x \overline{f_j} = \frac{1}{(2^j\delta)^2} \mathcal{F}^{-1}_x \int_v \partial_z \widetilde{\psi}_j\left(\frac{\mathcal{L}(i\xi', v)}{2^j\delta}\right) \mathcal{L}_v(i\xi', v)\widehat{g}(\xi, v)\phi(v)\,dv.$$

Its bound in (2.10) can now be improved by the extra factor of $(2^{j+1}\delta)^{\mu}$, which follows from (2.14), yielding

$$\begin{aligned}\left\|\Lambda^k_x \overline{f^{(1)}}\right\|_{L^q_{\mathrm{loc}}(\mathbb{R}^d_x)} &\lesssim \sum_{j\geq 1} \frac{1}{(2^j\delta)^2} (2^{j+1}\delta)^{\frac{\alpha}{q'}+\mu} \|g\|_{L^q_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)} \\ &\lesssim \delta^{\frac{\alpha}{q'} - (2-\mu)} \|g\|_{L^q_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)},\end{aligned}$$

and we conclude by arguing along the lines of Averaging Lemma 2.1. □

Remark 2.3. In the generic case of a symbol $\mathcal{L}(i\xi, v)$ that is analytic in $i\xi$ uniformly in v, the "degeneracy of order α" in (2.4) implies that

$$\sup\{|\mathcal{L}_v(i\xi, v)| : |\xi| = 1, |\mathcal{L}(i\xi, v)| \leq \delta\} \lesssim \delta^{1-\alpha}.$$

Thus, (2.14) holds with $\mu = 1 - \alpha$, and in this case, the velocity averaging holds with Sobolev-regularity exponent

$$\overline{f}(x) \in W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x), \quad s < s_\alpha := (1-\theta_\alpha)\sigma + \theta_\alpha k, \ \theta_\alpha := \frac{\alpha/p'}{\alpha(1/p' + 1/q) + 1}.$$

Remark 2.4. Averaging Lemma 2.2 can be extended for $f \in L^p(\mathbb{R}^d_x \times \mathbb{R}_v)$ and $g \in L^q_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$ in the full range of $1 < p < \infty$ and $1 \le q \le \infty$. The dual claim to (2.2), based on (L^2, BMO)–interpolation for $2 \le p < \infty$, reads

$$\|\overline{\mathsf{M}_\psi f}(x)\|_{L^p(\mathbb{R}^d_x)} \lesssim \sup_\xi |\Omega_m(\xi;\delta)|^{1/p} \cdot \|f\|_{L^p(\mathbb{R}^d_x \times \mathbb{R}_v)}, \quad 2 \le p < \infty. \tag{2.16}$$

This yields the same $W^{s,r}$-regularity as before, namely, for all $s \in (\sigma, (1-\theta_\alpha)\sigma + \theta_\alpha k)$ we have

$$\|\overline{f}\|_{W^{s,r}_{\text{loc}}(\mathbb{R}^d_x)} \lesssim \|g\|^{\theta}_{L^q_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)} \cdot \|f\|^{1-\theta}_{W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)}, \tag{2.17}$$

with $\theta = \theta_\alpha(\overline{p}, \overline{q})$ given by

$$\theta := \frac{\alpha/\overline{p}}{\alpha(1/\overline{p} - 1/\overline{q}) + 2 - \mu}, \quad \frac{1}{r} := \frac{1-\theta}{\overline{p}} + \frac{\theta}{\overline{q}},$$

$$\overline{p} := \max(p, p'), \quad \overline{q} := \max(q, q').$$

2.3 An Averaging Lemma for General Symbols

We now turn our attention to averages involving general, not necessarily homogeneous, symbols; as such the polar coordinate representation $\xi = |\xi|\xi'$ is no longer useful and will be discarded. We focus on equations of the form

$$\mathcal{L}(\nabla_x, v) f(x,v) = \partial_v g(x,v),$$

corresponding to $(\eta, N) = (0,1)$ in (2.3).

AVERAGING LEMMA 2.3 *Let $f \in W^{\sigma,p}_{\text{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)$, $\sigma \ge 0$, $1 < p \le 2$, solve the equation*

$$\mathcal{L}(\nabla_x, v) f(x,v) = \partial_v g(x,v), \quad g \in \begin{cases} L^q(\mathbb{R}^d_x \times \mathbb{R}_v), & 1 < q \le 2, \\ \mathcal{M}(\mathbb{R}^d_x \times \mathbb{R}_v), & q = 1. \end{cases} \tag{2.18}$$

Let $\mathcal{L}(i\xi, \cdot)$ be the corresponding symbol of degree $\le k$ with sufficiently smooth v-dependent coefficients and assume it obeys the truncation property. Denote

$$\omega_{\mathcal{L}}(J;\delta) := \sup_{\substack{\xi \in \mathbb{R}^d \\ |\xi| \sim J}} |\Omega_{\mathcal{L}}(\xi;\delta)|, \quad \Omega_{\mathcal{L}}(\xi;\delta) := \{v : |\mathcal{L}(i\xi, v)| \le \delta\},$$

and suppose $\mathcal{L}$ is nondegenerate in the sense that there exist $\alpha, \beta > 0$ such that

$$\omega_{\mathcal{L}}(J;\delta) \lesssim \left(\frac{\delta}{J^\beta}\right)^\alpha \quad \forall \delta > 0,\ J \gtrsim 1. \tag{2.19}$$

Moreover, assume that there exist $\lambda \ge 0$ and $\mu \in [0,1]$ such that

$$\sup_{|\xi| \sim J} \sup_{v \in \Omega_{\mathcal{L}}(\xi;\delta)} |\mathcal{L}_v(i\xi, v)| \lesssim J^{\beta\lambda} \delta^\mu. \tag{2.20}$$

Then for all bump functions $\phi \in C^\infty_0(I)$, the average $\overline{f}(x) := \int f(x,v)\phi(v)dv$ belongs to the Sobolev space $W^{s,r}_{\text{loc}}(\mathbb{R}^d_x)$ for $s \in (\sigma, s_{\alpha,\beta})$ and the following estimate holds:

$$\left\| \int f(x,v)\phi(v)dv \right\|_{W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x)} \lesssim \|f(x,v)\|_{W^{\sigma,p}_{\mathrm{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)} + \|g(x,v)\|_{L^q_{\mathrm{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)}. \tag{2.21}$$

Here, $s_{\alpha,\beta} := (1-\theta_\alpha)\sigma + \theta_\alpha\beta(2-\mu-\lambda)$ *where* $\theta \equiv \theta_\alpha$ *and* r *are given by*

$$\theta := \frac{\alpha/p'}{\alpha(1/p'-1/q')+2-\mu}, \quad \frac{1}{r} := \frac{1-\theta}{p} + \frac{\theta}{q}, \quad 0 < \theta < 1. \tag{2.22}$$

Remark 2.5. How does the last averaging lemma compare with the previous ones? We note that since $\deg \mathcal{L}_v(i\xi,\cdot) \le k$, the additional assumption (2.20) always holds with $\mu = 0$ and $\beta\lambda = k$. Hence, Averaging Lemma 2.3 with just the nondegeneracy condition (2.19) yields the regularity $\overline{f}(x) \in W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x)$ of (the reduced) order $s \in (\sigma, s_{\alpha,\beta})$, where

$$s_{\alpha,\beta} := (1-\theta_\alpha)\sigma + \theta_\alpha(2\beta-k), \quad \theta_\alpha := \frac{\alpha/p'}{\alpha(1/p'-1/q')+2}. \tag{2.23}$$

Now, if in particular $\mathcal{L}(i\xi,\cdot)$ is homogeneous of degree k, then

$$\omega_{\mathcal{L}}(\xi;\delta) = \omega_{\mathcal{L}}\left(\xi'; \frac{\delta}{|\xi|^k}\right), \quad \xi' := \frac{\xi}{|\xi|};$$

this shows that if the nondegeneracy condition (2.4) of Averaging Lemma 2.1 holds, $\omega_{\mathcal{L}}(J;\delta) \sim (\delta/J^k)^\alpha$, then it implies (2.19) with $\beta = k$, and we recover the homogeneous Averaging Lemma 2.1 with $(\eta, N) = (0,1)$; namely, the averages $\bar{f}$ gain regularity of order $s < (1-\theta)\sigma + \theta(2\beta-k) = (1-\theta)\sigma + \theta k$. The only difference is that now the truncation property is assumed on the unmodified symbol $\mathcal{L}(i\xi, v)$ rather than the modified one $\mathcal{L}(i\xi', v)$.

As for Averaging Lemma 2.2, we first note that in the generic case of a homogeneous $\mathcal{L}(\cdot, v)$, the additional assumption (2.20) holds with $\lambda = \alpha$ and $\mu = 1-\alpha$,

$$\sup_{|\xi|\sim J}\ \sup_{\{v\in I:|\mathcal{L}(i\xi,v)|\le\delta\}} |\mathcal{L}_v(i\xi,v)| \lesssim J^{\beta\alpha}\delta^{1-\alpha}. \tag{2.24}$$

Indeed, all the homogeneous examples discussed in Sections 3, 4, and 5 below employ Averaging Lemma 2.2 with these parameters that yield $W^{s,r}$-regularity of order $s < (1-\theta_\alpha)\sigma + \theta_\alpha\beta$,

$$\overline{f}(x) \in W^{s,r}_{\mathrm{loc}}(\mathbb{R}^d_x), \quad s < (1-\theta_\alpha)\sigma + \theta_\alpha\beta, \quad \theta_\alpha = \frac{\alpha/p'}{\alpha(1/p'+1/q)+1}. \tag{2.25}$$

In the particular case of $\mathcal{L}$ being homogeneous of order k, then $\beta = k$ and we recover Averaging Lemma 2.2 (except that the truncation hypothesis is now assumed on the unmodified symbol).

PROOF: We begin by noting that we can safely replace the nondegeneracy condition (2.19) with a slightly weaker one; namely, there exist $\alpha, \beta > 0$ and $\varepsilon > 0$

such that

$$\omega_{\mathcal{L}}(J;\delta) \lesssim \left(\frac{\delta^{1+\varepsilon}}{J^{\beta}}\right)^{\alpha}, \quad \forall\delta > 0,\ J \gtrsim 1, \tag{2.26}$$

and still retain the same gain of regularity of order $s < s_{\alpha,\beta}$. This can be achieved by replacing the values of α in (2.19) by $\alpha/(1+\varepsilon)$ and then absorbing ε into a slightly smaller order of regularity, $(1-\varepsilon)s_{\alpha,\beta}$ dictated by $\theta_{\alpha/(1+\varepsilon)}$. The extra ε-power of δ will be needed below to insure simple summability, which probably could be eliminated by a more refined argument involving Besov spaces, along the lines of [20].

Next, we break up f into Littlewood-Paley pieces [44, sec. VI, 4]

$$f = f_0 + \sum_{\text{dyadic } J\text{'s} \gtrsim 1} f_J,$$

so that $\widehat{f_J}(\xi, v)$, the spatial Fourier transform of $f_J(x, v)$, is supported for frequencies $|\xi| \sim J$, and $\hat{f}_0$ has support in $|\xi| \lesssim 1$. Since f_0 is a smooth average of f at unit scales, the contribution of f_0 is easily seen to be acceptable. By giving up an ε in the index $s_{\alpha,\beta}$ we may thus reduce (2.21) to a single value of J. It thus suffices to show that

$$J^s \|\overline{f_J}\|_{L^r_{\text{loc}}(\mathbb{R}^d)} \lesssim J^{\sigma}\|f_J\|_{L^p_{\text{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)} + \|g\|_{L^q_{\text{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)},$$
$$\overline{f_J} := \int_v f_J(x, v)\phi(v)dv,$$

for each $J \gtrsim 1$.

Fix J. Because of the local nature of Littlewood-Paley projections when $J \gtrsim 1$, we may replace the localized L^p-norms with global norms. Actually we may replace L^r by weak L^r since we may pay another ε in the index $s_{\alpha,\beta}$ to improve this. By the duality of weak L^r and $L^{r',1}$, it thus suffices to show that

$$|\langle\overline{f_J}, \chi_E\rangle| \lesssim J^{-s}|E|^{1/r'} \tag{2.27}$$

for all sets $E \subset \mathbb{R}^d$ of finite measure, where we have normalized $\|f_J\|_{L^p(\mathbb{R}^d_x\times\mathbb{R}_v)} \lesssim J^{-\sigma}$ and $\|g\|_{L^q_{\text{loc}}(\mathbb{R}^d_x\times\mathbb{R}_v)} \lesssim 1$.

We now decompose the action in v-space of each of the Littlewood-Paley pieces (rather than a decomposition f itself used in Lemma 2.1),

$$f_J(x, v) = \sum_{\text{dyadic } \delta\text{'s} \lesssim J^k} \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(x, v),$$

where $\psi(\mathcal{L}(\nabla_x, v)/\delta) := \mathcal{F}_x^{-1}\psi(\mathcal{L}(i\xi, v)/\delta)\mathcal{F}_x$. Here, $\psi(z)$ is a bump function on $\mathbb{C}$ supported on the region $|z| \sim 1$. It suffices to estimate

$$\left\langle \int \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(\,\cdot\,, v)\phi(v)dv, \chi_E \right\rangle \tag{2.28}$$

with a summable decay as the dyadic $\delta \to 0$ so that (2.27) holds.

By our assumption, $\mathcal{L}(i\xi, v)$ satisfies the truncation property uniformly in v; hence by Lemma 2.2 we see that for all $1 < p \le 2$,

$$\left\| \int \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(x, v)\phi(v)dv \right\|_{L^p_{\mathrm{loc}}(\mathbb{R}^d_x)} \lesssim \omega_{\mathcal{L}}(J; \delta)^{1/p'} \|f_J\|_{L^p_{\mathrm{loc}}(\mathbb{R}^d_x \times \mathbb{R}_v)}. \tag{2.29}$$

From (2.29) and Hölder we may thus estimate (2.28) by

$$\left|\left\langle \int \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(\,\cdot\,, v)\phi(v)dv,\, \chi_E \right\rangle\right| \lesssim J^{-\sigma}\omega_{\mathcal{L}}(J; \delta)^{1/p'}|E|^{1/p'}. \tag{2.30}$$

On the other hand, thanks to equation (2.18) we can write

$$\psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(x, v) = \widetilde{\psi}\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right)\frac{1}{\delta}\frac{\partial}{\partial v} g_J(x, v)$$

where $\widetilde{\psi}(z) := \psi(z)/z$ and the g_J's are the corresponding Littlewood-Paley dyadic pieces of g. We thus have

$$\int \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(x, v)\phi(v)dv = \frac{1}{\delta}\int \widetilde{\psi}\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right)\frac{\partial}{\partial v} g_J(x, v)\phi(v)dv. \tag{2.31}$$

We now integrate by parts to move the ∂/∂_v derivative somewhere else. We will assume that the derivative hits $\widetilde{\psi}(\mathcal{L}(\nabla_x, v)/\delta)$, since the case when the derivative hits the bump function $\phi(v)$ is much better. We are thus led to estimate

$$\frac{1}{\delta^2}\left|\left\langle \int \widetilde{\psi}_z\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right)\mathcal{L}_v(\nabla_x, v) g_J(x, v)\phi(v)dv,\, \chi_E \right\rangle\right|. \tag{2.32}$$

Since g_J is localized to frequencies $\sim J$, then by (2.20), the multiplier $\mathcal{L}_v(i\xi, v)$ acts like a constant of order $\mathcal{O}(J^{\beta\lambda}\delta^{\mu})$. Also, $\widetilde{\psi}_z$ is a bump function much like ψ. Thus we may modify (2.29) with p replaced by q (assuming that $q > 1$ and using the modified argument for the case $q = 1$ as before), ψ replaced by $\widetilde{\psi}_z$, and f_J replaced by $\mathcal{L}_v(\nabla_x, v)g_J$, to estimate (2.31) by

$$\left|\left\langle \int \psi\left(\frac{\mathcal{L}(\nabla_x, v)}{\delta}\right) f_J(\,\cdot\,, v)\phi(v)dv,\, \chi_E \right\rangle\right| \lesssim \delta^{-(2-\mu)} J^{\beta\lambda}\omega_{\mathcal{L}}(J; \delta)^{1/q'}|E|^{1/q'}. \tag{2.33}$$

Interpolating this bound with (2.30), we may bound (2.28) by

$$\delta^{-\theta(2-\mu)} J^{(1-\theta)(-\sigma)+\theta\beta\lambda}\omega_{\mathcal{L}}(J; \delta)^{1/r'}|E|^{1/r'}.$$

The parametrization in (2.22) dictates $\alpha/r' = \theta(2-\mu)$. Finally, we put the extra ε power into use: by (2.26), $\omega_{\mathcal{L}}(J; \delta)^{1/r'} \lesssim (\delta^{1+\varepsilon}J^{-\beta})^{\theta(2-\mu)}$ and hence the last quantity is bounded by

$$\delta^{\theta(2-\mu)\varepsilon} J^{-s_{\alpha,\beta}} |E|^{1/r'}, \quad s_{\alpha,\beta} = (1-\theta)\sigma - \theta\beta(2-\mu-\lambda).$$

Summing in δ and using the hypothesis that $s < s_{\alpha,\beta}$, we obtain (2.27). □

2.4 Velocity Averaging for First- and Second-Order Symbols

To apply velocity Averaging Lemma 2.3 we need to find out which multipliers $m(\xi)$ have the truncation property. Fortunately, there are large classes of such multipliers.

First of all, it is clear that the multipliers $m(\xi) = \xi \cdot e_1$ and $m(\xi) = |\xi|^2$ have the truncation property, as in these cases the Fourier multipliers are just convolutions with finite measures. Now, observe that if $m(\xi)$ has the truncation property, then so does $m(L(\xi))$ for any invertible linear transformation L on $\mathbb{R}^d$, with a bound that is uniform in L. This is because the L^p-multiplier class is invariant under linear transformations.

Because of this, we see that the multipliers

$$m_1(\xi) = \mathbf{a}(v) \cdot i\xi \quad \text{and} \quad m_2(\xi) = \langle \mathbf{b}(v)\xi, \xi \rangle$$

have the truncation property uniformly in v, where $\mathbf{a}(v)$ are arbitrary real coefficients and $\mathbf{b}(v)$ is an arbitrary elliptic bilinear form with real coefficients.

From the Hörmander-Mikhlin or Marcinkeiwicz multiplier theorems and the linear transformation argument one can also show that $m_1(\xi')$ has the truncation property uniformly in v. These arguments go back to the discussion of [20]. The situation with $m_2(\xi')$ is less clear, but fortunately we will not need to verify that these second-order modified multipliers obey the truncation property since our averaging lemmas also work with a truncation property hypothesis on the unmodified multiplier.

Now we observe that if $m_1(\xi)$ and $m_2(\xi)$ are real multipliers with the truncation property, then the complex multiplier $m_1(\xi) + im_2(\xi)$ also has the truncation property. The basic observation is that one can use Fourier series to write any symbol of the form

$$\psi\left(\frac{m_1(\xi) + im_2(\xi)}{\delta}\right)$$

as

$$\sum_{j,k\in\mathbb{Z}} \widehat{\psi}(j,k)\widetilde{\psi}_j\left(\frac{m_1(\xi)}{\delta}\right)\widetilde{\psi}_k\left(\frac{m_2(\xi)}{\delta}\right)$$

where $\widetilde{\psi}_j(x) := e^{2\pi ijx}\widetilde{\psi}$ and $\widetilde{\psi}$ is some bump function that equals 1 on the one-dimensional projections of the support of ψ. Since the C^ℓ-norm of $\widetilde{\psi}_j$ grows polynomially in j and $\widehat{\psi}(j,k)$ decays rapidly in j and k (if ψ is sufficiently smooth), we are done since the product of two L^p-multipliers is still an L^p-multiplier.

3 Nonlinear Hyperbolic Conservation Laws

Having developed our averaging lemmas, we now present some applications to nonlinear PDEs. We begin with the study of (real-valued) solutions $\rho(t,x) = \rho(t,x_1,\dots,x_d) \in L^\infty((0,\infty)\times\mathbb{R}^d_x)$ of multidimensional scalar conservation laws,

$$\frac{\partial}{\partial t}\rho(t,x) + \sum_{j=1}^{d}\frac{\partial}{\partial x_j}A_j(\rho(t,x)) = 0 \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x). \tag{3.1}$$

We abbreviate (3.1) as $\rho_t + \nabla_x\cdot\mathbf{A}(\rho) = 0$ where $\mathbf{A}$ is the vector of $C^{2,\epsilon}$-spatial fluxes, $\mathbf{A} := (A_1, A_2, \dots, A_d)$. Let $\chi_\gamma(v)$ denote the velocity indicator function

$$\chi_\gamma(v) = \begin{cases} 1 & \text{if } 0 < v \le \gamma \\ -1 & \text{if } \gamma \le v < 0 \\ 0 & \text{otherwise.}\end{cases}$$

We say that $\rho(t,x)$ is a *kinetic solution* of the conservation law (3.1) if the corresponding distribution function $\chi_{\rho(t,x)}(v)$ satisfies the transport equation

$$\partial_t\chi_{\rho(t,x)}(v) + \mathbf{a}(v)\cdot\nabla_x\chi_{\rho(t,x)}(v) = \partial_v m(t,x,v) \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x\times\mathbb{R}_v), \tag{3.2}$$

for some nonnegative measure, $m(t,x,v) \in \mathcal{M}^+((0,\infty)\times\mathbb{R}^d_x\times\mathbb{R}_v)$. Here, $\mathbf{a}(v)$ is the vector of transport velocities, $\mathbf{a}(v) := (a_1(v),\dots,a_d(v))$ where $a_j(\cdot) := A_j'(\cdot)$, $j = 1,\dots,d$. The regularizing effect associated with the proper notion of *nonlinearity* of the conservation law (3.1) was explored in [36] through the averaging properties of an underlying kinetic formulation. For completeness, we include here a brief description that will serve our discussion on nonlinear parabolic and elliptic equations in the following sections, and we refer to [36] for a complete discussion.

The starting point is the *entropy inequalities* associated with (3.1),

$$\partial_t\eta(\rho(t,x)) + \nabla_x\cdot\mathbf{A}^\eta(\rho(t,x)) \le 0 \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x).$$

Here, η is an arbitrary *entropy function* (a convex function from $\mathbb{R}$ to $\mathbb{R}$) and $\mathbf{A}^\eta := (A_1^\eta,\dots,A_d^\eta)$ is the corresponding vector of entropy fluxes, $A_j^\eta(\rho) := \int^\rho \eta'(s)A_j'(s)ds$, $j = 1,\dots,d$. A function $\rho \in L^\infty$ is an *entropy solution* if it satisfies the entropy inequalities for *all* pairs $(\eta,\mathbf{A}^\eta)$ induced by convex entropies η. Entropy solutions are precisely those solutions that are *realizable* as vanishing viscosity limit solutions and are uniquely determined by their $L^\infty\cap L^1$–initial data $\rho_0(x)$, prescribed at $t = 0$, e.g., [33]. A decisive role is played by the one-parameter family of *Kružkov entropy pairs*, $(\eta(\rho;v),\mathbf{A}^\eta(\rho;v))$, parametrized by $v\in\mathbb{R}$,

$$\eta(\rho;v) := |\rho - v|, \quad A_j^\eta(\rho;v) := \operatorname{sgn}(\rho - v)(A_j(\rho) - A_j(v)).$$

Kružkov entropy pairs lead to a complete L^1-theory of existence, uniqueness, and stability of first-order quasi-linear conservation laws [32].

We turn to the kinetic formulation. We define the distribution $m(t,x,v) = m_{\rho(t,x)}(v)$ by the formula

$$(3.3)\quad m(t,x,v) := -\left[\partial_t \frac{\eta(\rho;v)-\eta(0;v)}{2} + \nabla_x\cdot\left(\frac{\mathbf{A}^\eta(\rho;v)-\mathbf{A}^\eta(0;v)}{2}\right)\right].$$

The entropy inequalities tell us that the distribution $m = m_\rho$ is in fact a nonnegative measure, $m(t,x,v)\in\mathcal{M}^+((0,\infty)\times\mathbb{R}^d_x\times\mathbb{R}_v)$. Next, we differentiate (3.3) with respect to v: a straightforward computation yields that $\chi_{\rho(t,x)}(v)$ satisfies the kinetic transport equation (3.2). This reveals the interplay between Kružkov entropy inequalities and the underlying kinetic formulation; for nonlinear conservation laws, kinetic solutions coincide with the entropy solutions [36, 40, 43]. Observe that by velocity averaging we recover the macroscopic quantities associated with the entropy solution ρ,

$$\int_v \chi_\rho(v)\phi(v)dv = \Phi(\rho),$$

where $\Phi(\rho) := \int_{s=0}^{\rho}\phi(s)ds$ is the primitive of ϕ. In particular, one can recover ρ itself by setting $\phi(v) = 1_{[-M,M]}(v)$, $M = \|\rho\|_{L^\infty}$.

We now use Averaging Lemma 2.2 to study the regularity of ρ. To this end we first extend (3.2) over the full $\mathbb{R}_t\times\mathbb{R}^d_x$–space, using a $C_0^\infty(0,\infty)$–cutoff function, $\psi \equiv 1$ for $t\geq\epsilon$, so that $f(t,x,v) := \chi_{\rho(t,x)}(v)\psi(t)$ and $g(t,x,v) := m_v(t,x,v)\psi(t) + \chi_{\rho(t,x)}(v)\partial_t\psi(t)$ satisfy

$$\partial_t f(t,x,v) + \mathbf{a}(v)\cdot\nabla_x f(t,x,v) = \partial_v g(t,x,v) \quad \text{in } \mathcal{D}'(\mathbb{R}_t\times\mathbb{R}^d_x\times\mathbb{R}_v),$$

with $g\in\mathcal{M}(\mathbb{R}_t\times\mathbb{R}^d_x\times\mathbb{R}_v)$. Set $I := [\inf\rho_0, \sup\rho_0]$ and assume that the first-order symbol is nondegenerate so that (2.4), (2.14) hold; namely, there exist $\alpha\in(0,1)$ and $\mu\in[0,1]$ such that

$$(3.4)\qquad \sup_{\tau^2+|\xi|^2=1} |\Omega_{\mathbf{a}}(\xi;\delta)| \lesssim \delta^\alpha, \quad \Omega_{\mathbf{a}}(\xi;\delta) := \{v\in I : |\tau+\mathbf{a}(v)\cdot\xi|\leq\delta\},$$

and

$$(3.5)\qquad \sup_{|\xi|=1}\ \sup_{\Omega_{\mathbf{a}}(\xi;\delta)} |\mathbf{a}'(v)\cdot\xi| \lesssim \delta^\mu.$$

We apply Averaging Lemma 2.2 for first-order symbols, $k=1$, with $q=1$, $p=2$, and $\sigma=0$, to find that $\overline{f}(x)$ and hence $\rho(x)$ belongs to $W^{s,r}_{\text{loc}}(\mathbb{R}_t\times\mathbb{R}^d_x)$,

$$\rho(t,x)\in W^{s,r}_{\text{loc}}((\epsilon,\infty)\times\mathbb{R}^d_x), \quad s<\theta_\alpha := \frac{\alpha}{\alpha+4-2\mu}, \quad r := \frac{\alpha+4-2\mu}{\alpha+2-\mu}.$$

At this stage we invoke the monotonicity property of entropy solutions, which implies that for $s<1$, $\|\rho(t,\cdot)\|_{W^{s,1}_{\text{loc}}(\mathbb{R}^d_x)}$ is nonincreasing, and we deduce that $\rho(t,\cdot)\in W^{s,1}(\mathbb{R}^d_x)$ for $t>\epsilon$. We conclude that the entropy solution operator associated with the nonlinear conservation law (3.1),(3.4), $\rho_0(\cdot) \mapsto \rho(t, \cdot)$, has a *regularizing effect*, mapping $L^\infty(\mathbb{R}_x^d)$ into $W_{\text{loc}}^{s,1}(\mathbb{R}_x^d)$,

$$\forall t \geq \epsilon > 0 \quad \rho(t, \cdot) \in W_{\text{loc}}^{s,1}(\mathbb{R}_x^d), \quad s < s_1, \; s_1 := \theta_\alpha = \frac{\alpha}{\alpha + 4 - 2\mu}.$$

Next, we use the bootstrap argument of [36, sec. 3] to deduce an improved regularizing effect. The $W_{\text{loc}}^{s,1}(\mathbb{R}_t \times \mathbb{R}_x^d)$–regularity of $\rho(t,x)\psi(t)$ implies that $f(t,x,v) = \chi_{\rho(t,x)}(v)\psi(t)$ belongs to $L^1(W^{s,1}(\mathbb{R}_t \times \mathbb{R}_x^d), \mathbb{R}_v)$; moreover, since $\partial_v \chi_\rho(v)$ is a bounded measure, $f \in L^1(W^{s,1}(\mathbb{R}_t \times \mathbb{R}_x^d), \mathbb{R}_v) \cap L^1(\mathbb{R}_t \times \mathbb{R}_x^d, W^{s,1}(\mathbb{R}_v))$ and hence

$$f \in W_{\text{loc}}^{s,1}(\mathbb{R}_t \times \mathbb{R}_x^d \times \mathbb{R}_v) \quad \forall s < s_1.$$

Interpolation with the obvious L^∞-bound of f then yields that $f \in W_{\text{loc}}^{s,2}(\mathbb{R}_t \times \mathbb{R}_x^d \times \mathbb{R}_v)$ for all $s < s_1/2$. Therefore, Averaging Lemma 2.2 applies to $f = \chi_{\rho(t,x)}(v)\psi(t)$ with $q = 1$, $p = 2$, and $\sigma = s_1/2$, implying that $\rho(t,\cdot)$ has improved $W_{\text{loc}}^{s,1}$-regularity of order $s < s_2 = (1-\theta_\alpha)s_1/2 + \theta_\alpha$. Reiterating this argument yields the fixed point $s_k \uparrow s_\infty = 2\theta_\alpha/(1+\theta_\alpha)$ and we conclude with a regularizing effect

$$\forall t \geq \epsilon > 0 \quad \rho_0 \in L^\infty \cap L^1(\mathbb{R}_x^d) \mapsto \rho(t,\cdot) \in W_{\text{loc}}^{s,1}(\mathbb{R}_x^d), \quad s < \frac{\alpha}{\alpha + 2 - \mu}.$$

As indicated earlier in Remark 2.5, in the generic case, $\mu = 1 - \alpha$,

$$\sup_{|\xi|=1} \sup_{\{v \in I : |\tau + \mathbf{a}(v)\cdot\xi| \leq \delta\}} |\mathbf{a}'(v)\cdot\xi| \lesssim \delta^{1-\alpha}, \tag{3.6}$$

which yields a $W^{s,1}$-regularizing effect of order $s < \alpha/(2\alpha+1)$. This improves the previous regularity result [36, theorem 4] of order $s < \alpha/(\alpha+2)$, corresponding to $\mu = 0$.

We can extend the last statement for general L_{loc}^p initial data. Recall that the entropy solution operator associated with (3.1) is L^1-contractive. We now invoke a general *nonlinear* interpolation argument of J.-L. Lions, e.g., [49, "Interpolation Theory," lecture 8]; namely, if a possibly nonlinear T is Lipschitz on X with a Lipschitz constant L_X and maps boundedly $Y_1 \mapsto Y_2$ with a bound B_Y, then one verifies that the corresponding K-functional satisfies $K(Tx, t; X, Y_2) \leq L_X K(x, tB_Y/L_X; X, Y_1)$, and hence T maps $[X, Y_1]_{\theta,q} \mapsto [X, Y_2]_{\theta,q}$. Consequently, the entropy solution operator maps $[L^1, L^\infty]_{\theta,q} \mapsto [L^1, W_{\text{loc}}^{s,1}]_{\theta,q}$, $0 < \theta < 1 < q$, and we conclude the following:

COROLLARY 3.1 *Consider the nonlinear conservation law* (3.1) *subject to* $L^p \cap L^1$*–initial data,* $\rho(0,x) = \rho_0(x)$. *Assume the nondegeneracy condition of order* α*, namely* (3.4)*,* (3.6)*, hold over arbitrary finite intervals* I. *Then* $\rho(t,x)$ *gains a regularity of order* s/p'*,*

$$\forall t \geq \epsilon > 0 \quad \rho_0 \in L^p \cap L^1(\mathbb{R}_x^d) \mapsto \rho(t,\cdot) \in W_{\text{loc}}^{s,1}(\mathbb{R}_x^d), \quad s < \frac{\alpha}{(2\alpha+1)p'}.$$

The study of regularizing effects in one- and two-dimensional *nonlinear* conservation laws has been studied by a variety of different approaches; an incomplete list of references includes [21, 38, 45, 46, 47].

We close this section with three examples. Let $\ell \geq 1$ and consider the one-dimensional conservation law

$$\frac{\partial}{\partial t}\rho(t,x) + \frac{\partial}{\partial x}\left\{\frac{1}{\ell+1}\rho^{\ell+1}(t,x)\right\} = 0, \quad \rho_0 \in [-M, M]. \tag{3.7}$$

It satisfies the nondegeneracy condition (3.4) with $\alpha = \frac{1}{\ell}$; hence $\rho(t,\cdot)_{|t>\epsilon} \in W^{s,1}_{\mathrm{loc}}$ with $s < \alpha/(2\alpha+1) = 1/(\ell+2)$. It is well-known, however, that the entropy solution operator of the inviscid Burgers equation corresponding to $\ell = 1$ maps $L^\infty \mapsto \mathrm{BV}$ [38]. This shows that the regularizing effect of order $\alpha/(2\alpha+1)$ stated in Corollary 3.1 is not sharp; consult [15] following [29]. Accordingly, it was conjectured in [36] that (3.4) yields a regularizing effect of order α.

Next, let $\ell, m \geq 1$ and consider the two-dimensional conservation law

$$\frac{\partial}{\partial t}\rho(t,x) + \frac{\partial}{\partial x_1}\left\{\frac{1}{\ell+1}\rho^{\ell+1}(t,x)\right\} + \frac{\partial}{\partial x_2}\left\{\frac{1}{m+1}\rho^{m+1}(t,x)\right\} = 0,$$

subject to initial condition $\rho_0 \in [-M, M]$. If $\ell \neq m$ then (3.4) is satisfied with $\alpha = \min\{\frac{1}{\ell}, \frac{1}{m}\}$ and we conclude $\rho(t \geq \epsilon, \cdot) \in W^s_{\mathrm{loc}}(L^1)$ with $s < \min\{\frac{1}{\ell+2}, \frac{1}{m+2}\}$. If $\ell = m$, however, then there is *no regularizing effect* since $\tau' + v^\ell \xi_1' + v^m \xi_2' \equiv 0$ for $\tau = 0, \xi_1 + \xi_2 = 0$; indeed, $\rho_0(x-y)$ are steady solutions that allow oscillations to persist along $x - y = \text{const}$. Other cases can be worked out based on their polynomial degeneracy; for example,

$$\frac{\partial}{\partial t}\rho(t,x) + \frac{\partial}{\partial x_1}\sin(\rho(t,x)) + \frac{\partial}{\partial x_2}\left\{\frac{1}{3}\rho^3(t,x)\right\} = 0, \quad \rho_0 \in [-M, M],$$

has a nondegeneracy of order $\alpha = \frac{1}{4}$, yielding $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < \frac{1}{6}$.

4 Nonlinear Degenerate Parabolic Equations

We are concerned with second-order, possibly degenerate parabolic equations in conservative form:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(t,x) &+ \sum_{j=1}^d \frac{\partial}{\partial x_j}A_j(\rho(t,x)) \\ &- \sum_{j,k=1}^d \frac{\partial^2}{\partial x_j \partial x_k}B_{jk}(\rho(t,x)) = 0 \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x).\end{aligned} \tag{4.1}$$

We abbreviate, $\rho_t + \nabla_x \cdot \mathbf{A}(\rho) - \mathrm{tr}(\nabla_x \otimes \nabla_x \mathbf{B}(\rho)) = 0$ where $\mathbf{B}$ is the matrix $\mathbf{B} := \{B_{jk}\}_{j,k=1}^d$. By degenerate parabolicity we mean that the matrix $\mathbf{B}'(\cdot)$ is nonnegative, $\langle \mathbf{B}'(\cdot)\xi, \xi\rangle \geq 0 \ \forall \xi \in \mathbb{R}^d$. Our starting point is the *entropy inequalities* associated with (4.1) such that for *all* convex η's,

$$\begin{aligned} &\partial_t \eta(\rho(t,x)) + \nabla_x \cdot \mathbf{A}^\eta(\rho(t,x)) \\ &\qquad - \operatorname{tr}\left(\nabla_x \otimes \nabla_x \mathbf{B}^\eta(\rho(t,x))\right) \le 0 \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x). \end{aligned} \tag{4.2}$$

Here $\mathbf{A}^\eta$ is the same vector of hyperbolic entropy fluxes we had before, $\mathbf{A}^\eta = (A^\eta_1, \dots, A^\eta_d)$, and $\mathbf{B}^\eta$ is the matrix of parabolic entropy fluxes, $\mathbf{B}^\eta := (B^\eta_{jk})^d_{j,k=1}$ where $B^\eta_{jk}(\rho) := \int^\rho \eta'(s) B'_{jk}(s)ds$. We turn to the kinetic formulation. Utilizing the Kružkov entropies $\eta(\rho; v) := |\rho - v|$, we define the distribution, $m(t,x,v) = m_{\rho(t,x)}(v)$,

$$\begin{aligned} m(t,x,v) := &-\left[\partial_t \frac{\eta(\rho;v)-\eta(0;v)}{2} + \nabla_x \cdot \left(\frac{\mathbf{A}^\eta(\rho;v)-\mathbf{A}^\eta(0;v)}{2}\right)\right] \\ &+ \operatorname{tr}\left(\nabla_x \otimes \nabla_x \frac{\mathbf{B}^\eta(\rho;v)-\mathbf{B}^\eta(0;v)}{2}\right). \end{aligned} \tag{4.3}$$

The entropy inequalities tell us that $m(t,x,v) \in \mathcal{M}^+((0,\infty)\times\mathbb{R}^d_x\times\mathbb{R}_v)$ and differentiation with respect to v yields the kinetic formulation

$$\partial_t \chi_{\rho(t,x)}(v) + \mathbf{a}(v)\cdot\nabla_x \chi_{\rho(t,x)}(v) - \nabla_x^\top \cdot \mathbf{b}(v)\nabla_x \chi_{\rho(t,x)}(v) = \partial_v m(t,x,v) \tag{4.4}$$

for some nonnegative $m \in \mathcal{M}^+$ that measures entropy + dissipation production. Here $\mathbf{a}$ is the same vector of velocities we had before, $\mathbf{a} = \mathbf{A}'$, and $\mathbf{b}$ is the nonnegative diffusion matrix $\mathbf{b} := \mathbf{B}' \ge 0$. The representation $\eta(\rho) - \eta(0) = \int \eta'(s)\chi_\rho(s)ds$ shows that the kinetic formulation (4.4) is in fact the equivalent dual statement of the entropy inequalities (4.2). But neither of these statements settles the question of uniqueness, except for certain special cases, such as the isotropic diffusion, $B_{jk}(\rho) = B(\rho)\delta_{jk} \ge 0$, e.g., [8], or special cases with mild singularities, e.g., a porous-media-type one-point degeneracy [17, 51]. The extension of Kružkov theory to the present context of general parabolic equations with possibly nonisotropic diffusion was completed only recently in [11], after the pioneering work [53]. Observe that the entropy production measure m consists of contributions from the hyperbolic entropy dissipation and the parabolic dissipation of the equation $m = m_{\mathbf{A}} + m_{\mathbf{B}}$. The solutions sought by Chen and Perthame in [11], $\rho \in L^\infty$, require that their corresponding distribution function χ_ρ satisfy (4.4) with a *restricted* form of parabolic defect measure $m_{\mathbf{B}}$: the restriction imposed on $m_{\mathbf{B}}$ reflects a certain renormalization property of the mixed derivatives of ρ (or, more precisely, the primitive of $\sqrt{\mathbf{b}(\rho)}$). Accordingly, we can refer to these Chen-Perthame solutions as *renormalized solutions* with a kinetic formulation (4.4). These renormalized kinetic solutions admit an equivalent interpretation as entropy solutions [11] and as dissipative solutions [42]. A general L^1-theory of existence, uniqueness, and stability can be found in [10].

For a recent overview with a more complete list of references on such convection-diffusion equations in divergence form, we refer the reader to [9]. The regularizing effect of such equations, however, is less understood. In [36, §5] we used the kinetic formulation (4.4) to prove that the solution operator, $\rho_0 \mapsto \rho(t,\cdot)$ is relatively compact under a generic nondegeneracy condition

$$\sup_{|\xi|=1} |\Omega_{\mathcal{L}}(\xi;0)| = 0, \quad \Omega_{\mathcal{L}}(\xi;0) := \{v : \tau + \mathbf{a}(v)\cdot\xi = 0,\ \langle \mathbf{b}(v)\xi,\xi\rangle = 0\}.$$

A general compactness result in this direction can be found in [22]. We turn to *quantify* the regularizing effect associated with such kinetic solutions. We emphasize that our regularity results are based on the "generic" kinetic formulation (4.4), but otherwise they are *independent* of the additional information on the renormalized Chen-Perthame solutions encoded in their entropy production measure m. The extra restrictions of the latter will likely yield even better regularity results than those stated below. We divide our discussion into two stages in order to highlight different aspects of degenerate diffusion in Section 4.1 and the coupling with nonlinear convection in Section 4.2.

4.1 Nonisotropic Degenerate Diffusion

We consider the parabolic equation

$$\frac{\partial}{\partial t}\rho(t,x) - \sum_{j,k=1}^{d} \frac{\partial^2}{\partial x_j \partial x_k} B_{jk}(\rho(t,x)) = 0 \quad \text{in } \mathcal{D}'((0,\infty)\times\mathbb{R}^d_x). \tag{4.5}$$

Here we ignore the hyperbolic part and focus on the effect of nonisotropic diffusion. The corresponding kinetic formulation (4.4) extended to the full $\mathbb{R}_t \times \mathbb{R}^d_x \times \mathbb{R}_v$ reads

$$\partial_t f(t,x,v) - \nabla_x^\top \cdot \mathbf{b}(v)\nabla_x f(t,x,v) = \partial_v m(t,x,v), \quad f := \chi_\rho \psi(t),$$

with $m \in \mathcal{M}^+(\mathbb{R}_t \times \mathbb{R}^d_x \times \mathbb{R}_v)$. Set $I := [\inf\rho_0, \sup\rho_0]$. The corresponding symbol is $\mathcal{L}(i\tau, i\xi, v) = i\tau - \langle \mathbf{b}(v)\xi,\xi\rangle$, and it suffices to make the nondegeneracy assumption (2.4) on the second-order homogeneous part of the symbol $\mathcal{L}(0,i\xi,v) = -\langle \mathbf{b}(v)\xi,\xi\rangle$. We make the assumption of nondegeneracy; namely, there exist $\alpha \in (0,1)$ and $\mu \in [0,1]$ such that

$$\sup_{|\xi|=1} |\Omega_{\mathbf{b}}(\xi;\delta)| \lesssim \delta^\alpha, \quad \Omega_{\mathbf{b}}(\xi;\delta) := \{v \in I : 0 \le \langle \mathbf{b}(v)\xi,\xi\rangle \le \delta\} \tag{4.6}$$

and

$$\sup_{|\xi|=1} \sup_{\Omega_{\mathbf{b}}(\xi;\delta)} |\langle \mathbf{b}'(v)\xi,\xi\rangle| \lesssim \delta^\mu. \tag{4.7}$$

We apply Averaging Lemma 2.3 with $q=1$, $p=2$, $\sigma=0$, and $k=2$ to find that $\overline{f}(t,x)$ and hence $\rho(t,x)$ belong to $W^{s,r}_{\mathrm{loc}}(\mathbb{R}_t \times \mathbb{R}^d_x)$,

$$\rho(t,\cdot) \in W^{s,r}_{\mathrm{loc}}((\epsilon,\infty)\times\mathbb{R}^d_x), \quad s < 2\theta_\alpha,\ \theta_\alpha := \frac{\alpha}{\alpha+4-2\mu},\ r := \frac{\alpha+4-2\mu}{\alpha+2-\mu}.$$

We follow the hyperbolic arguments. The kinetic solution operator associated with (4.1) is L^1-contractive, hence $\|\rho(t,\cdot)\|_{W^{s,1}_{\mathrm{loc}}(\mathbb{R}^d_x)}$ is nonincreasing and we conclude that $\forall t > \varepsilon$, $\rho(t,\cdot)$ has $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < s_1$, $s_1 := 2\theta_\alpha = 2\alpha/(\alpha + 4 -$

2μ). We then bootstrap. Since $\chi_\rho(v)\psi(t) \in W^{s,2}_{\mathrm{loc}}(\mathbb{R}_t \times \mathbb{R}^d_x \times \mathbb{R}_v)$ for all $s < s_1/2$, we can apply Averaging Lemma 2.1 with $\sigma = s_1/2$ leading to $W^{s,1}_{\mathrm{loc}}$-regularity of order $s_2 := (1-\theta_\alpha)s_1/2 + 2\theta_\alpha$ with fixed point $s_k \uparrow s_\infty = 4\theta_\alpha/(1+\theta_\alpha)$,

$$\text{(4.8)}\quad \forall t \geq \epsilon > 0 \quad \rho_0 \in L^\infty \cap L^1(\mathbb{R}^d_x) \mapsto \rho(t,\cdot) \in W^{s,1}_{\mathrm{loc}}(\mathbb{R}^d_x),\ s < \frac{2\alpha}{\alpha+2-\mu}.$$

We distinguish between two different types of degenerate parabolicity, summarized in the following two corollaries:

COROLLARY 4.1 (Degenerate Parabolicity I: The Case of a Full Rank) *Consider the degenerate parabolic equation* (4.1) *subject to $L^p \cap L^1$–initial data, $\rho(0,x) = \rho_0$. Let $\lambda_1(v) \geq \cdots \geq \lambda_d(v) \geq 0$ be the eigenvalues of $\mathbf{b}(v)$ and assume that $\lambda_d(v) \not\equiv 0$ over $I = [\inf \rho_0, \sup \rho_0]$. Then $\langle \mathbf{b}(v)\xi, \xi\rangle \geq \lambda_d(v)|\xi|^2$ and $\rho(t,x)$ has a regularizing of order $s < 2\alpha/(\alpha+2-\mu)p'$. In particular,* (4.8) *holds with α and μ dictated by $\lambda_d(v) \equiv \lambda_d(\mathbf{b}(v))$, $|\Omega_\lambda(\delta)| \lesssim \delta^\alpha$, and*

$$\sup_{|\xi|=1} \sup_{v \in \Omega_\lambda(\delta)} |\langle \mathbf{b}'(v)\xi, \xi\rangle| \lesssim \delta^\mu,$$

where $\Omega_\lambda(\delta) := \{v \in I : 0 \leq \lambda_d(v) \leq \delta\}$.

Corollary 4.1 applies to the special case of *isotropic* diffusion,

$$\text{(4.9)}\qquad \frac{\partial}{\partial t}\rho(t,x) - \Delta B(\rho(t,x)) = 0, \quad B'(v) \geq 0,$$

subject to $L^\infty \cap L^1$–initial data, $\rho(0,\cdot) = \rho_0$. If $b(\cdot) := B'(v)$ is degenerate of order α in the sense that

$$|\Omega_b(\delta)| \lesssim \delta^\alpha \ \text{ and } \sup_{v \in \Omega_b(\delta)} |b'(v)| \lesssim \delta^{1-\alpha}, \quad \Omega_b(\delta) := \{v \in I : 0 \leq b(v) \leq \delta\},$$

then Corollary 4.1 implies, $\forall t \geq \epsilon$, $\rho(t,\cdot) \in W^{s,1}_{\mathrm{loc}}$, $s < 2\alpha/(2\alpha+1)$. For $L^p \cap L^1$–data ρ_0, the corresponding solution $\rho(t,\cdot)$ gains $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < 2\alpha/(2\alpha+1)p'$ and we conjecture, in analogy with the hyperbolic case, that the nondegeneracy (4.6) yields an improved regularizing effect of order $2\alpha/p'$. Existence, uniqueness, and regularizing effects of the isotropic equation (4.9) were studied earlier in [2, 3, 6]. The prototype is provided by the porous media equation

$$\text{(4.10)}\qquad \begin{gathered} \frac{\partial}{\partial t}\rho(t,x) - \Delta\left\{\frac{1}{n+1}|\rho^n(t,x)|\rho(t,x)\right\} = 0, \\ \rho(0,x) = \rho_0(x) \geq 0,\ \rho_0 \in L^\infty. \end{gathered}$$

The velocity averaging yields $W^{s,1}$-regularity of order $2/(n+2)$ and, as in the hyperbolic case, it does not recover the optimal Hölder continuity in this case, e.g., [17, 50]. In fact, the kinetic arguments do not yield continuity. Instead, our main contribution here is to the *nonisotropic* case where we conjecture the same gain of regularity driven by $\lambda_d(\mathbf{b}(v))$ as the isotropic regularity driven by $b(v)$.

We continue with the subtler case where $\mathbf{b}(\cdot)$ does not have a full rank so that

$$\exists \ell \ 1 \le \ell < d : \lambda_1(v) \ge \cdots \ge \lambda_\ell(v) \ge 0, \quad \lambda_{\ell+1}(v) \equiv \cdots \equiv \lambda_d(v) \equiv 0.$$

Despite this stronger degeneracy, there is still some regularity that can be "saved." To demonstrate our point, we consider the two-dimensional case.

COROLLARY 4.2 (Degenerate Parabolicity II: The Case of a Partial Rank) *We consider the two-dimensional degenerate equation*

$$\frac{\partial}{\partial t}\rho(t,x) - \left\{ \frac{\partial^2}{\partial x_1^2} B_{11}(\rho(t,x)) + \frac{\partial^2}{\partial x_1 \partial x_2} B_{12}(\rho(t,x)) + \frac{\partial^2}{\partial x_2^2} B_{22}(\rho(t,x)) \right\} = 0, \tag{4.11}$$

subject to $L^\infty \cap L^1$–initial data $\rho(0,x) = \rho_0$. Assume strong degeneracy, $b_{12}^2(v) \equiv 4b_{11}(v)b_{22}(v)$, so that $\lambda_2(\mathbf{b}(v)) \equiv 0 \ \forall v \in I = [\inf \rho_0, \sup_0]$. In this case, $\langle \mathbf{b}(v)\xi, \xi\rangle = (\sqrt{b_{11}}(v)\xi_1 + \sqrt{b_{22}}(v)\xi_2)^2$ and $\rho(t,\cdot)$ admits a $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < 2\alpha/(\alpha + 2 - \mu)$, which is dictated by the nondegeneracy condition

$$\sup_{|\xi|=1} |\Omega_{\mathbf{b}}(\xi;\delta)| \lesssim \delta^\alpha \tag{4.12}$$

and

$$\sup_{|\xi|=1} \sup_{v \in \Omega_{\mathbf{b}}(\xi;\delta)} \left| \frac{b'_{11}(v)}{\sqrt{b_{11}(v)}}\xi_1 + \frac{b'_{22}(v)}{\sqrt{b_{22}(v)}}\xi_2 \right| \lesssim \delta^{\mu - 1/2} \tag{4.13}$$

where $\Omega_{\mathbf{b}}(\xi;\delta) := \{v \in I : |\sqrt{b_{11}(v)}\xi_1 + \sqrt{b_{22}(v)}\xi_2|^2 \le \delta\}$.

We distinguish between two extreme scenarios.

(1) If $|b_{11}(v)| \gg |b_{22}(v)| \ \ \forall v \in I$, then the regularizing effect (4.8) holds with (α, μ) dictated by $b_{22}(v)$, namely,

$$|\Omega_{b_{22}}(\delta)| \lesssim \delta^\alpha \quad \text{and} \quad \sup_{v \in \Omega_{b_{22}}(\delta)} |b'_{22}(v)| \lesssim \delta^\mu,$$

where $\Omega_{b_{22}}(\delta) := \{v \in I : 0 \le b_{22}(v) \le \delta\}$.

(2) If $b_{11}(v) \equiv b_{22}(v) \ \ \forall v \in I$, then there is *no* regularizing effect since the symbol $(\sqrt{b_{11}}(v)\xi_1 + \sqrt{b_{22}}(v)\xi_2)^2$ vanishes for all $\xi_1 \pm \xi_2 = 0$ (so that (4.12) is fulfilled with $\alpha = 0$). Indeed, equation (4.11), with $b_{11}(v) = b_{22}(v) =: B'(v)$, takes the form

$$\frac{\partial}{\partial t}\rho(t,x) - \left\{ \frac{\partial^2}{\partial x_1^2} \pm 2\frac{\partial^2}{\partial x_1 \partial x_2} + \frac{\partial^2}{\partial x_2^2} \right\} B(\rho(t,x)) = 0,$$

and we observe that $\rho_0(x \mp y)$ are steady solutions that allow for oscillations to persist along $x \mp y = \text{const}$.

4.2 Convection-Diffusion Equations

We begin with the one-dimensional case

$$\frac{\partial}{\partial t}\rho(t,x) + \frac{\partial}{\partial x}A(\rho(t,x)) - \frac{\partial^2}{\partial x^2}B(\rho(t,x)) = 0. \tag{4.14}$$

We look at the prototype example of high-order Burgers-type nonlinearity, $a(v) := v^\ell$, $\ell \geq 1$, combined with porous medium diffusion $b(v) = |v|^n$, $n \geq 1$. The corresponding symbol is given by $\mathcal{L}((i\tau, i\xi), v) = i\tau + v^\ell i\xi + |v|^n\xi^2$. We study the regularity of this convection-diffusion equation using Averaging Lemma 2.3, which employs the size of the set

$$\Omega_{\mathcal{L}}\big(J;\delta\big) := \{v \mid J|\tau + v^\ell\xi| + J^2|v|^n\xi^2 \leq \delta\}, \quad \tau^2+\xi^2 = 1,\ J \gtrsim 1,\ \delta \lesssim 1.$$

Comparing diffusion-versus-nonlinear convection effects, we can distinguish here between three different cases. Clearly $\Omega_{\mathcal{L}}(J;\delta) \subset \Omega_b := \{v : |v|^n \leq \delta/J^2\}$, hence $\omega_{\mathcal{L}}(J;\delta) \lesssim (\delta/J^2)^{1/n}$ and (2.19) holds with $\alpha_b = 1/n$ and $\beta_b = 2$. We shall use this bound whenever $n \leq \ell$, which is the case dominated by the parabolic part of (4.14). Indeed, in this case we have $(\delta/J)^{1/\ell} \gtrsim (\delta/J^2)^{1/n}$, which in turn yields

$$\sup_{v\in\Omega_{\mathcal{L}}(J;\delta)} |\mathcal{L}_v((i\tau, i\xi), v)| \lesssim \sup_{v\in\Omega_b}(J|v|^{\ell-1} + J^2|v|^{n-1}) \lesssim J^{2/n}\delta^{1-1/n}.$$

This shows that (2.20) holds with $(\lambda_b, \mu_b) = (\alpha_b, 1-\alpha_b)$, and velocity averaging implies a $W^{s,1}_{\text{loc}}$-regularizing effect with Sobolev exponent $s < \beta_b\alpha_b/(3\alpha_b+2)$,

$$s < s_n = \frac{2}{2n+3}.$$

We also have $\Omega_{\mathcal{L}}(J;\delta) \subset \Omega_a := \{v :\ |v|^\ell \lesssim \delta/J\}$, so that $\omega_{\mathcal{L}}(J;\delta) \lesssim (\delta/J)^{1/\ell}$; i.e., (2.19) holds with $\alpha_a = 1/\ell$ and $\beta_a = 1$. We shall use this bound whenever $n \geq 2\ell$, which is the case driven by the hyperbolic part (4.14). In this case, $(\delta/J)^{1/\ell} \lesssim (\delta/J^2)^{1/n}$, hence

$$\sup_{v\in\Omega_{\mathcal{L}}(J;\delta)} |\mathcal{L}_v((i\tau, i\xi), v)| \lesssim \sup_{v\in\Omega_a}(J|v|^{\ell-1} + J^2|v|^{n-1}) \lesssim J^{1/\ell}\delta^{1-1/\ell},$$

implying that (2.20) is fulfilled with $(\lambda_a, \mu_a) = (\alpha_a, 1-\alpha_a)$. The corresponding Sobolev exponent, $s < \beta_a\alpha_a/(3\alpha_a+2)$, is then given by

$$s < s_\ell := \frac{1}{2\ell+3}.$$

Finally, for intermediate n's, $\ell < n < 2\ell$, we interpolate the previous two $\omega_{\mathcal{L}}$-bounds (which are valid for all n's),

$$\omega_{\mathcal{L}}(J;\delta) \lesssim (\delta/J)^{(1-\zeta)/\ell}(\delta/J^2)^{\zeta/n}$$

for some $\zeta \in [0,1]$, which we choose as $\zeta := (n/\ell) - 1$, so that (2.19) holds with $\alpha = (1-\zeta)/\ell + \zeta/n$ and $\beta\alpha = (1-\zeta)/\ell + 2\zeta/n$, and (2.20) holds with $(\lambda, \mu) = (\alpha, 1-\alpha)$. This then yields the Sobolev-regularity exponent, $s < \beta\alpha/(3\alpha+2)$,

$$s < \frac{n + (2\ell - n)\zeta}{3n + 3(\ell-n)\zeta + 2n\ell}, \quad \zeta := \frac{n}{\ell} - 1,\ \ell < n < 2\ell.$$

An additional bootstrap argument improves this Sobolev exponent, $s < \beta\alpha/(2\alpha + 1)$, and we summarize the three different cases in the following corollary:

COROLLARY 4.3 *The convection-diffusion equation*

$$\frac{\partial}{\partial t}\rho(t,x)+\frac{\partial}{\partial x}\left\{\frac{1}{\ell+1}\rho^{\ell+1}(t,x)\right\}-\frac{\partial^2}{\partial x^2}\left\{\frac{1}{n+1}|\rho^n(t,x)|\rho(t,x)\right\}=0, \tag{4.15}$$

subject to initial conditions $\rho_0 \in [-M, M]$, *has a regularizing effect,*

$$\rho_0 \in L^\infty(\mathbb{R}_x) \mapsto \rho(t > \varepsilon, \cdot) \in W^{s,1}_{\mathrm{loc}}(\mathbb{R}_x),$$

of order $s < s_{\ell,n}$ *given by*

$$s_{\ell,n} = \frac{n + (2\ell - n)\zeta_{\ell,n}}{2n + 2(\ell - n)\zeta_{\ell,n} + n\ell}, \quad \zeta_{\ell,n} := \begin{cases} 0, & n \le \ell, \\ (n/\ell) - 1, & \ell < n < 2\ell, \\ 1, & n \ge 2\ell. \end{cases}$$

We note that when $n \le \ell$, then $\Omega_b \subset \Omega_a$ and (4.15) is dominated by degenerate diffusion with a regularizing effect of order $s_{\ell,n} = s_n = 2/(n+2)$. Thus, we recover the same order of regularity we met with the "purely diffusive" porous medium equation (4.10). If $n \ge 2\ell$, however, then $\Omega_a \subset \Omega_b$, and it is the hyperbolic part that dominates diffusion, driving the overall regularizing effect of (4.15) with order $s_{\ell,n} = s_\ell = 1/(\ell + 2)$; we recover regularity with the same order that we met with the "purely convective" hyper-Burgers equation (3.7). Finally, in the intermediate "mixed cases," $\ell < n < 2\ell$, we find a regularity of a (nonoptimal) order

$$s_{\ell,n} = \frac{n\ell + (2\ell - n)(n - \ell)}{2n\ell - 2(n - \ell)^2 + n\ell^2}, \quad \ell < n < 2\ell.$$

We turn to the multidimensional case (4.1). The regularizing effect is determined by the size of the set

$$\Omega(J;\delta) = \{v \mid J|\tau + \mathbf{a}(v)\cdot\xi| + J^2\langle \mathbf{b}(v)\xi, \xi\rangle \le \delta\}, \quad \tau^2 + |\xi|^2 = 1,\ J \gtrsim 1,\ \delta \lesssim 1.$$

Assume that the degenerate parabolic part of the equation has a full rank, so that the smallest eigenvalue of $\mathbf{b}(v)$, $\lambda(v) \equiv \lambda_d(\mathbf{b}(v))$, satisfies

$$|\Omega_{\mathbf{b}}(\delta)| \lesssim \delta^{\alpha_{\mathbf{b}}} \quad \text{and} \quad \sup_{|\xi|=1}\ \sup_{v\in\Omega_{\mathbf{b}}(\delta)} |\langle \mathbf{b}'(v)\xi, \xi\rangle| \lesssim \delta^{1-\alpha_{\mathbf{b}}}, \tag{4.16}$$

where $\Omega_{\mathbf{b}}(\delta) := \{v \mid 0 \le \lambda_d(\mathbf{b}(v)) \le \delta\}$. In this case, $\omega_{\mathcal{L}}(J;\delta) \lesssim (\delta/J^2)^{\alpha_{\mathbf{b}}}$, which yields a gain of $W^{s,1}_{\mathrm{loc}}$-regularity of order $s < 2\alpha_{\mathbf{b}}/(2\alpha_{\mathbf{b}} + 1)$. If in addition the hyperbolic part of the equation has a nondegeneracy of order $\alpha_{\mathbf{a}}$, namely,

$$\sup_{\tau^2+|\xi|^2=1} |\Omega_{\mathbf{a}}((\tau,\xi),\delta)| \lesssim \delta^{\alpha_{\mathbf{a}}} \quad \text{and} \quad \sup_{|\xi|=1}\ \sup_{v\in\Omega_{\mathbf{a}}(\delta)} |\mathbf{a}'(v)\cdot\xi| \lesssim \delta^{1-\alpha_{\mathbf{a}}},$$

where $\Omega_{\mathbf{a}}((\tau,\xi);\delta) := \{v : |\tau + \mathbf{a}(v)\cdot\xi| \le \delta\}$; then we can argue along the lines of Corollary 4.3 to conclude that there is an overall $W^{s,1}_{\mathrm{loc}}$-regularity of order dictated by the relative size of $2\alpha_{\mathbf{b}}/(2\alpha_{\mathbf{b}} + 1)$ and $\alpha_{\mathbf{a}}/(2\alpha_{\mathbf{a}} + 1)$. As an example, we have the following:

COROLLARY 4.4 *Consider the two-dimensional convection-diffusion equation*

$$\frac{\partial}{\partial t}\rho(t,x)+\frac{\partial}{\partial x_1}\left\{\frac{1}{\ell+1}\rho^{\ell+1}(t,x)\right\}+\frac{\partial}{\partial x_2}\left\{\frac{1}{m+1}\rho^{m+1}(t,x)\right\} - \sum_{j,k=1}^{2}\frac{\partial^2}{\partial x_j\partial x_k}B_{jk}(\rho(t,x))=0, \tag{4.17}$$

with nondegenerate diffusion, $B'(v)\gtrsim |v|^n$. Then its renormalized kinetic solution admits a $W^{s,1}_{\mathrm{loc}}$-regularizing effect, $\rho_0\in L^\infty\mapsto \rho(t,\cdot)\in W^{s,1}_{\mathrm{loc}}(\mathbb{R}^2_x)$, of order

$$s<\begin{cases} s_{\ell,m}:=\min\left(\dfrac{1}{\ell+2},\dfrac{1}{m+2}\right) & \text{if } n\ge 2\max(\ell,m) \text{ and } \ell\neq m,\\ s_n:=\dfrac{2}{n+2} & \text{if } n\le \min(\ell,m) \text{ or } \ell=m,\\ s_{\ell,m,n}\in[s_{\ell,m},s_n] & \text{if } \min(\ell,m)<n<2\max(\ell,m).\end{cases}$$

Finally, we close this section with a third example of a fully degenerate equation

$$\frac{\partial}{\partial t}\rho(t,x)+\left(\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2}\right)A(\rho(t,x)) - \left(\frac{\partial^2}{\partial x_1^2}-2\frac{\partial^2}{\partial x_1\partial x_2}+\frac{\partial^2}{\partial x_2^2}\right)B(\rho(t,x))=0. \tag{4.18}$$

In this case, there is a stronger, rank-1 parabolic degeneracy with no regularizing effect from the purely diffusion part, since $\langle \mathbf{b}(v)\xi,\xi\rangle\equiv 0\ \ \forall \xi_1-\xi_2=0$, and no regularizing effect from the purely convection part where $\mathbf{a}(v)\cdot\xi\equiv 0\ \ \forall\xi_1+\xi_2=0$. Nevertheless, the combined *convection-diffusion* does have a regularizing effect as demonstrated in the following:

COROLLARY 4.5 *Consider the two-dimensional convection-diffusion equation*

$$\frac{\partial}{\partial t}\rho(t,x)+\left(\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2}\right)\left\{\frac{1}{\ell+1}\rho^{\ell+1}(t,x)\right\} - \left(\frac{\partial^2}{\partial x_1^2}-2\frac{\partial^2}{\partial x_1\partial x_2}+\frac{\partial^2}{\partial x_2^2}\right)\left\{\frac{1}{n+1}|\rho^n(t,x)|\rho(t,x)\right\}=0. \tag{4.19}$$

For $n\ge 2\ell$ it admits a regularizing effect, $\rho_0\in L^\infty\cap L^1\mapsto \rho(t,\cdot)\in W^{s,1}_{\mathrm{loc}}(\mathbb{R}^2_x)$, of order $s<6/(2+2n-\ell)$.

As before, the Sobolev exponent computed here is not necessarily sharp, and as a result no gain of regularity is stated for $n<2\ell$.

PROOF: We use Averaging Lemma 2.3 with the usual $(p,q)=(2,1)$, which yields a Sobolev regularity exponent of order $s=\beta\alpha(2-\mu-\lambda)/(\alpha+4-2\mu)$, where α, β, λ, and μ characterize the degeneracy of the symbol associated with equation (4.19),

$$\mathcal{L}((i\tau, i\xi), v) = Ji\tau + Jv^\ell i(\xi_1 + \xi_2) + J^2|v|^n|\xi_1 - \xi_2|^2, \quad \tau^2 + |\xi|^2 = 1.$$

We consider first those ξ's, $|\xi| = 1$, such that $|\xi_1 - \xi_2| \le 1/5$. Here we have (say) $|\xi_1 + \xi_2| \ge 1/10$, so that $\Omega_{\mathcal{L}}(J;\delta) \subset \Omega_{\mathbf{a}} = \{v : |v|^\ell \lesssim \delta/J\}$, and (2.19) holds with $(\alpha_a, \beta_a) = (1/\ell, 1)$; moreover, the growth of $\mathcal{L}_v$ sought in (2.20) is bounded by

$$\sup_{v\in\Omega_{\mathcal{L}}(J;\delta)} |\mathcal{L}_v\big((i\tau, i\xi), v\big)| \lesssim J\left(\frac{\delta}{J}\right)^{(\ell-1)/\ell} + J^2\left(\frac{\delta}{J}\right)^{(n-1)/\ell} \lesssim J^\lambda\delta^\mu,$$

where

$$\begin{cases} \lambda := 1/\ell \text{ and } \mu := 1 - 1/\ell & \text{if } n \ge 2\ell, \\ \lambda := 2 - (n-1)/\ell \text{ and } \mu := (n-1)/\ell & \text{if } n < 2\ell. \end{cases}$$

In particular, if $n < 2\ell$ then the Sobolev exponent vanishes since $2 - \mu - \lambda = 0$, and we cannot deduce any regularizing effect in this case. If $n \ge 2\ell$, however, we compute the Sobolev exponent as before, $s = s_\ell = 1/(2\ell + 3)$. Next, we consider the case when $|\xi_1 - \xi_2| \ge 1/5$ so that $\Omega_{\mathcal{L}}(J;\delta) \subset \Omega_{\mathbf{b}} = \{v : |v|^n \lesssim \delta/J^2\}$ and (2.19) holds with $(\alpha_b, \beta_b) = (1/n, 2)$. As for (2.20), we have

$$\sup_{v\in\Omega_{\mathcal{L}}(J;\delta)} |\mathcal{L}_v\big((i\tau, i\xi), v\big)| \lesssim J\left(\frac{\delta}{J^2}\right)^{(\ell-1)/n} + J^2\left(\frac{\delta}{J^2}\right)^{(n-1)/n} \lesssim J^{2\lambda}\delta^\mu,$$

where

$$\begin{cases} \lambda = 1/n \text{ and } \mu = 1 - 1/n & \text{if } n \le \ell, \\ \lambda = 1/2 - (\ell-1)/n \text{ and } \mu = (\ell-1)/n & \text{if } n > \ell. \end{cases}$$

In particular, if $n \ge 2\ell$ we compute in this case a smaller Sobolev exponent $s = 3/(3 - 2\ell + 4n) \le s_\ell$. The regularity result follows from the bootstrap argument we discussed earlier, which yields the final Sobolev exponent $s = 2\beta\alpha(2 - \mu - \lambda)/(\alpha + 2 - \mu)$. □

5 Nonlinear Degenerate Elliptic Equations

We consider the nonlinear, possibly degenerate elliptic equation

$$(5.1) \qquad -\sum_{j,k=1}^{d} \frac{\partial^2}{\partial x_j \partial x_k} B_{jk}(\rho(x)) = S(\rho(x)) \quad \text{in } \mathcal{D}'(\Gamma), \quad \mathbf{b}(\cdot) := \mathbf{B}'(\cdot) \ge 0,$$

augmented with proper boundary conditions along the $C^{1,1}$-boundary $\partial\Gamma$. We assume that the nonlinear source term $S(\rho)$ is further restricted so that blowup is avoided.

We begin with formal manipulations, multiplying (5.1) by $\eta'(\rho)$ and "differentiating by parts" to find

$$
\begin{aligned}
&-\sum_{j,k=1}^{d} \eta'(\rho)\frac{\partial^2}{\partial x_j \partial x_k} B_{jk}(\rho) - \eta'(\rho)S(\rho)\\
&\quad = -\sum_{j=1}^{d} \frac{\partial}{\partial x_j}\left(\eta'(\rho)b_{jk}(\rho)\frac{\partial \rho}{\partial x_k}\right) + \sum_{j,k=1}^{d} \eta''(\rho)b_{jk}(\rho)\frac{\partial \rho}{\partial x_j}\frac{\partial \rho}{\partial x_k} - \eta'(\rho)S(\rho)\\
&\quad = -\operatorname{tr}(\nabla_x \otimes \nabla_x \mathbf{B}^{\eta}(\rho)) + \eta''(\rho)\langle \mathbf{b}(\rho)\nabla_x \rho, \nabla_x \rho\rangle - \eta'(\rho)S(\rho).
\end{aligned}
$$

We arrive at the *entropy inequalities* associated with (5.1), stating that sufficiently smooth solutions of (5.1) satisfy, for all convex η's,

$$
-\operatorname{tr}(\nabla_x \otimes \nabla_x \mathbf{B}^{\eta}(\rho)) \le \eta'(\rho)S(\rho), \tag{5.2}
$$

where $\mathbf{B}^{\eta}(\rho) = \{B^{\eta}_{jk}(\rho)\}_{j,k=1}^{d}$ and $B^{\eta}_{jk}(\rho) := \int_0^{\rho} \eta'(s)b_{jk}(s)ds$. So far, we have not specified the notion of solutions for (5.1) since it seems that relatively little is known about a general stability theory for degenerate equations such as (5.1). The difficulty lies with the type of degeneracy that does not lend itself to standard elliptic regularity theory, because of the B_{jk}'s degenerate dependence on ρ, nor does it admit the regularity theory for viscosity solutions, e.g., [7, 13] because of their degenerate dependence on ρ rather than $\nabla_x \rho$. We refer to the works of Guan [27, 28], who shows that in certain cases, one is able to "lift" a $C^{1,1}$-regularity of ρ into a statement of C^{∞}-regularity. By using the existence of smooth viscosity solutions in the uniformly elliptic case where $\mathbf{b}(v) \ge \lambda > 0$, (5.2) can be justified by the "vanishing viscosity limit," forming a family of regularized solutions ρ^{λ} associated with $\mathbf{b}^{\lambda}(s) := \mathbf{b}(s) + \lambda I_{d\times d}$ and letting $\lambda \downarrow 0_+$. Next comes the kinetic formulation of (5.1), which takes the form

$$
-\nabla_x^{\top} \cdot \mathbf{b}(v)\nabla_x \chi_\rho(v) + S(v)\frac{\partial}{\partial v}\chi_\rho(v) = \frac{\partial}{\partial v} m(x,v) \quad \text{in } \mathcal{D}'(\Gamma \times \mathbb{R}_v) \tag{5.3}
$$

for some nonnegative $m \in \mathcal{M}^{+}$ that measures "entropy production." Indeed, for an arbitrary convex "entropy" η, the moments of (5.3) yield

$$
\begin{aligned}
0 &\ge -\int \eta''(v)m(x,v)dv\\
&= \int \eta'(v)\frac{\partial}{\partial v} m(x,v)dv\\
&= -\sum_{j,k=1}^{d}\int \eta'(v)b_{jk}(v)\frac{\partial^2}{\partial x_j \partial x_k}\chi_\rho(v)dv + \int \eta'(v)S(v)\frac{\partial}{\partial v}\chi_\rho(v)dv
\end{aligned}
$$

$$= -\sum_{j,k=1}^{d} \frac{\partial^2}{\partial x_j \partial x_k} B^{\eta}_{jk}(\rho) - \int (\eta'(v)S(v))' \chi_\rho(v)dv$$

$$= -\operatorname{tr}(\nabla_x \otimes \nabla_x \mathbf{B}^{\eta}(\rho)) - \eta'(\rho)S(\rho).$$

Thus, the kinetic formulation (5.3) is the dual statement for the entropy inequalities (5.2). We postulate that ρ is a *kinetic solution* of (5.1) if the corresponding distribution function $\chi_{\rho(x)}(v)$ satisfies (5.3), and we address the regularizing effect of such kinetic solutions.

To use the averaging lemma, we first extend (5.3) over the full $\mathbb{R}^d_x \times \mathbb{R}_v$–space. Let ψ be a $C_0^\infty(R^+)$-cutoff function, $\psi(s) \equiv 1$ for $s \geq \epsilon$, and let $\zeta(x)$ denote the smoothed distance function to the boundary, $\zeta(x) = \psi(\operatorname{dist}(x, \partial\Gamma))$; then $f(x, v) := \chi_{\rho(x)}(v)\zeta(x)$ satisfies

$$\begin{aligned}
&-\sum_{j,k=1}^{d} b_{jk}(v) \frac{\partial^2}{\partial x_j \partial x_k} f(x, v) \\
&\quad = \frac{\partial}{\partial v} \zeta(x) m(x, v) \\
&\qquad + \sum_{j,k=1}^{d} \left(\frac{\partial}{\partial x_j} (b_{jk}(v)\zeta_{x_k}(x)\chi_\rho(v)) + \frac{\partial}{\partial x_k} (b_{jk}(v)\zeta_{x_j}(x)\chi_\rho(v)) \right) \\
&\qquad + S(v)\zeta(x) \frac{\partial}{\partial v} \chi_\rho(v) - \sum_{j,k=1}^{d} b_{jk}(v)\zeta_{x_j x_k}(x)\chi_\rho(v) \\
&\quad =: \frac{\partial}{\partial v} g_1(x, v) + \Lambda^{\eta}_x g_2(x, v) + g_3(x, v) + g_4(x, v) \quad \text{in } \mathcal{D}'(\mathbb{R}^d_x \times \mathbb{R}_v).
\end{aligned} \tag{5.4}$$

Assume that (5.1) is nondegenerate in the sense that there exists an $\alpha \in (0, 1)$ such that

$$|\Omega_{\mathbf{b}}(\delta)| \lesssim \delta^{\alpha} \quad \text{and} \quad \sup_{|\xi|=1} \sup_{v \in \Omega_{\mathbf{b}}(\delta)} |\langle \mathbf{b}'(v)\xi, \xi \rangle| \lesssim \delta^{1-\alpha}, \tag{5.5}$$

where $\Omega_{\mathbf{b}}(\delta) := \{v \in I : \langle \mathbf{b}(v)\xi, \xi \rangle \leq \delta\}$.

We examine the contribution of each of the four terms on the right of (5.4) to the overall $W^{s,1}$-regularity of $\overline{f}$, appealing to the different averaging lemmas term by term. The first term on the right involves the bounded measure $g_1 = \zeta m$; Averaging Lemma 2.3, with the usual $(p, q) = (2, 1)$, then yields that the corresponding average $\overline{f}_1$ has a $W^{s_1,1}_{\text{loc}}$-regularity of order $s_1 < 2\theta_1$, $\theta_1 = \alpha/(3\alpha + 2)$. The second term on the right-hand side of (5.4) involves the gradient of the uniformly bounded term $g_2 = \sum b_{jk}(v)\zeta_{x_k}(x)\chi_\rho(v)$; here we can use Averaging Lemma 2.1 with $\eta = q = 1$ to conclude that the corresponding average $\overline{f}_2$ has $W^{s_2,1}_{\text{loc}}$-regularity of order $s_2 < \theta_2 = \alpha/(\alpha + 2)$ (in fact, with $q = 2$ one concludes a better $W^{s_2,2}$-regularity of order $s_2 < \alpha/2$). The remaining two terms on the right-hand side of (5.4) yield smoother averages and therefore do not affect the overall regularity dictated by the first two. Indeed, $\partial_v \chi_\rho(v)$ and hence $g_3(x,v) = S(v)\zeta(x)\partial_v\chi_\rho(v)$ is a bounded measure, and Averaging Lemma 2.1 with $\eta = N = 0$ implies that the corresponding average $\overline{f}_3$ belongs to the smaller Sobolev space $W^{s_3,1}_{\mathrm{loc}}$ of order $s_3 < 2\theta_3$, $\theta_3 = \alpha/(\alpha+2)$. Finally, the last term on the right of (5.4) consists of the bounded sum, $g_4 = -\sum b_{jk}(v)\zeta_{x_j x_k}(x)\chi_\rho$; with $\eta = N = 0$ and $q = 2$, the corresponding average $\overline{f}_4$ has a $W^{s_4,2}_{\mathrm{loc}}$-regularity of order $s_4 < 2\theta_4$, $\theta_4 = \alpha/2$.

Next, we iterate the bootstrap argument we mentioned earlier in the context of hyperbolic conservation laws. The first $W^{s_1,1}$-bound together with the L^∞-bound of f imply a $W^{\sigma_1,2}_{\mathrm{loc}}$-bound with $\sigma_1 = s_1/2$, which in turn yields the improved regularity of $\overline{f}_1 \in W^{s,1}_{\mathrm{loc}}$, $s < (1-\theta_1)\sigma_1 + 2\theta_1$. Thus, for the first term we can iterate the improved regularity, $s_1 \mapsto (1-\theta_1)s_1/2 + 2\theta_1$, converging to the same fixed point we had in the parabolic case before, $s_1 < 2\alpha/(2\alpha+1)$. The second term requires a more careful treatment: as we iterate the improved regularity of $\chi_\rho \in W^{s,1}_{\mathrm{loc}}$, we can express the term on the right of (5.4) as $\Lambda^{\eta_s} g_2$ with $\eta_s := 1 - s$ and with g_2 standing for the sum of L^1-bounded terms, $g_2 = \Lambda^s \sum b_{jk}(v)\zeta_{x_k}\chi_\rho(v)$. Consequently, Averaging Lemma 2.1 yields the fixed-point iterations $s_2 \mapsto (1 - \theta_2)s_2/2 + (2 - \eta_{s_2})\theta_2$ with limiting regularity of order $s_2 < \alpha$. The remaining two terms are smoother and do not affect the overall regularity: a similar argument for the third term yields the fixed-point iterations, $s_3 \mapsto (1-\theta_3)s_3/2 + 2\theta_3$ with a fixed point $s_3 < 2\alpha/(\alpha+1)$, while the fourth term remains in the smaller Sobolev space $W^{\alpha,2}$. We summarize with the following statement:

COROLLARY 5.1 *Let $\rho \in L^\infty$ be a kinetic solution of the nonlinear elliptic equation* (5.1) *and assume the nondegeneracy condition* (5.5) *holds. Then we have the interior regularity estimate for all $D \subset \Gamma$,*

$$\rho(x) \in W^{s,1}_{\mathrm{loc}}(D), \qquad s < \begin{cases} \alpha & \text{if } \alpha < \frac{1}{2}, \\ \dfrac{2\alpha}{2\alpha+1} & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

Acknowledgments. Part of this research was carried out while E.T. was visiting the Weizmann Institute for Science, and it is a pleasure to thank the faculty of mathematics and computer science for their hospitality. Research was supported in part by National Science Foundation Grant DMS #DMS04-07704 and ONR #N00014-91-J-1076 (E.T.) and by the David and Lucille Packard Foundation (T.T.).

Bibliography

[1] Agoshkov, V. I. Spaces of functions with differential-difference characteristics and the smoothness of solutions of the transport equation. *Dokl. Akad. Nauk SSSR* **276** (1984), no. 6, 1289–1293; translation in *Soviet Math. Dokl.* **29** (1984), no. 3, 662–666.

[2] Bénilan, P.; Crandall, M. G. The continuous dependence on φ of solutions of $u_t - \Delta\varphi(u) = 0$. *Indiana Univ. Math. J.* **30** (1981), no. 2, 161–177.

[3] Bénilan, P.; Crandall, M. G. Regularizing effects of homogeneous evolution equations. *Contributions to analysis and geometry (Baltimore, Md., 1980)*, 23–39. Johns Hopkins University Press, Baltimore, Md., 1981.

[4] Bézard, M. Régularité L^p précisée des moyennes dans les équations de transport. *Bull. Soc. Math. France* **122** (1994), no. 1, 29-76.

[5] Bouchut, F.; Desvillettes, L. Averaging lemmas without time Fourier transform and application to discretized kinetic equations. *Proc. Roy. Soc. Edinburgh Sect. A* **129** (1999), no. 1, 19–36.

[6] Brézis, H.; Crandall, M. G. Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$. *J. Math. Pures Appl. (9)* **58** (1979), no. 2, 153–163.

[7] Caffarelli, L. A.; Cabré, X. *Fully nonlinear elliptic equations*. American Mathematical Society Colloquium Publications, 43. American Mathematical Society, Providence, R.I., 1995.

[8] Carrillo, J. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal*. **147** (1999), no. 4, 269–361.

[9] Chen, G.-Q. Some recent methods for partial differential equations of divergence form. Dedicated to the 50th anniversary of IMPA. *Bull. Braz. Math. Soc. (N.S.)* **34** (2003), no. 1, 107–144.

[10] Chen, G.-Q.; Karlsen, K. H. L^1-framework for continuous dependence and error estimates for quasilinear anisotropic degenerate parabolic equations. *Trans. Amer. Math. Soc*. **358** (2006), no. 3, 937–963 (electronic).

[11] Chen, G.-Q.; Perthame, B. Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **20** (2003), no. 4, 645–668.

[12] Cheverry, C. Regularizing effects for multidimensional scalar conservation laws. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **17** (2000), no. 4, 413–472.

[13] Crandall, M. G.; Ishii, H.; Lions, P.-L. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), no. 1, 1–67.

[14] Crandall, M. G.; Kocan, M.; Lions, P. L.; Święch, A. Existence results for boundary problems for uniformly elliptic and parabolic fully nonlinear equations. *Electron. J. Differential Equations* **1999**, no. 24, 22 pp. (electronic).

[15] De Lellis, C.; Westdickenberg, M. On the optimality of velocity averaging lemmas. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **20** (2003), no. 6, 1075–1085.

[16] DeVore, R.; Petrova, G. The averaging lemma. *J. Amer. Math. Soc*. **14** (2001), no. 2, 279–296 (electronic).

[17] DiBenedetto, E. *Degenerate parabolic equations*. Universitext. Springer, New York, 1993.

[18] DiPerna, R. J.; Lions, P.-L. Global weak solutions of Vlasov-Maxwell systems. *Comm. Pure Appl. Math*. **42** (1989), no. 6, 729–757.

[19] DiPerna, R. J.; Lions, P.-L. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)* **130** (1989), no. 2, 321–366.

[20] DiPerna, R. J.; Lions, P.-L.; Meyer, Y. L^p regularity of velocity averages. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **8** (1991), no. 3-4, 271–287.

[21] Engquist, B.; E, W. Large time behavior and homogenization of solutions of two-dimensional conservation laws. *Comm. Pure Appl. Math*. **46** (1993), no. 1, 1–26.

[22] Gérard, P. Moyennisation et régularité deux-microlocale. *Ann. Sci. École Norm. Sup. (4)* **23** (1990), no. 1, 89–121.

[23] Gérard, P.; Golse, F. Averaging regularity results for PDEs under transversality assumptions. *Comm. Pure Appl. Math*. **45** (1992), no. 1, 1–26.

[24] Golse, F.; Lions, P.-L.; Perthame, B.; Sentis, R. Regularity of the moments of the solution of a transport equation. *J. Funct. Anal*. **76** (1988), no. 1, 110–125.

[25] Golse, F.; Perthame, B.; Sentis, R. Un résultat de compacité pour les équations de transport et application au calcul de la limite de la valeur propre principale d'un opérateur de transport. *C. R. Acad. Sci. Paris Sér. I Math*. **301** (1985), no. 7, 341–344.

[26] Golse, F.; Saint-Raymond, L. Velocity averaging in L^1 for the transport equation. *C. R. Math. Acad. Sci. Paris* **334** (2002), no. 7, 557–562.

[27] Guan, P. Regularity of a class of quasilinear degenerate elliptic equations. *Adv. Math*. **132** (1997), no. 1, 24–45.

[28] Guan, P. Nonlinear degenerate elliptic equations. *Proceedings of ICCM2001 (Taiwan, 2001)*, 257–266. New Studies in Advanced Mathematics, 4. International, Somerville, Mass., 2004. Available online at `http://www.math.mcgill.ca/guan/iccm.dvi`

[29] Jabin, P.-E.; Perthame, B. Regularity in kinetic formulations via averaging lemmas. A tribute to J. L. Lions. *ESAIM Control Optim. Calc. Var*. **8** (2002), 761–774 (electronic).

[30] Jabin, P.-E.; Vega, L. Averaging lemmas and the X-ray transform. *C. R. Math. Acad. Sci. Paris* **337** (2003), no. 8, 505–510.

[31] Jabin, P.-E.; Vega, L. A real space method for averaging lemmas. *J. Math. Pures Appl. (9)* **83** (2004), no. 11, 1309–1351.

[32] Kružkov, S. N. First order quasilinear equations in several independent variables. *Mat. Sb. (N.S.)* **81** (1970), no. 123, 228–255.

[33] Lax, P. D. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, 11. Society for Industrial and Applied Mathematics, Philadelphia, 1973.

[34] Lions, P.-L. Régularité optimale des moyennes en vitesses. *C. R. Acad. Sci. Paris Sér. I Math*. **320** (1995), no. 8, 911–915.

[35] Lions, P.-L.; Perthame, B.; Souganidis, P. E. Existence and stability of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates. *Comm. Pure Appl. Math*. **49** (1996), no. 6, 599–638.

[36] Lions, P.-L.; Perthame, B.; Tadmor, E. A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Amer. Math. Soc*. **7** (1994), no. 1, 169–191.

[37] Lions, P.-L.; Perthame, B.; Tadmor, E. Kinetic formulation of the isentropic gas dynamics and p-systems. *Comm. Math. Phys*. **163** (1994), no. 2, 415–431.

[38] Oleĭnik, O. A. Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl. (2)* **26** (1963), 95–172.

[39] Perthame, B. *Kinetic formulation of conservation laws*. Oxford Lecture Series in Mathematics and Its Applications, 21. Oxford University Press, Oxford, 2002.

[40] Perthame, B. Mathematical tools for kinetic equations. *Bull. Amer. Math. Soc. (N.S.)* **41** (2004), no. 2, 205–244.

[41] Perthame, B.; Souganidis, P. E. A limiting case for velocity averaging. *Ann. Sci. École Norm. Sup. (4)* **31** (1998), no. 4, 591–598.

[42] Perthame, B.; Souganidis, P. E Dissipative and entropy solutions to non-isotropic degenerate parabolic balance laws. *Arch. Ration. Mech. Anal*. **170** (2003), no. 4, 359–370.

[43] Perthame, B.; Tadmor, E. A kinetic equation with kinetic entropy functions for scalar conservation laws. *Comm. Math. Phys*. **136** (1991), no. 3, 501–517.

[44] Stein, E. M. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, 3. Princeton University Press, Princeton, N.J., 1993.

[45] Tadmor, E.; Rascle, M.; Bagnerini, P. Compensated compactness for 2D conservation laws. *J. Hyperbolic Differ. Equ*. **2** (2005), no. 3, 697–712.

[46] Tartar, L. Compensated compactness and applications to partial differential equations. *Nonlinear analysis and mechanics: Heriot-Watt Symposium*, vol. 4, 136–212. Research Notes in Mathematics, 39. Pitman, Boston–London, 1979.

[47] Tartar, L. The compensated compactness method for a scalar hyperbolic equation. Lecture notes, 87-20. Carnegie-Mellon University, Pittsburgh, Pa., 1987.

[48] Tartar, L. H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A* **115** (1990), no. 3-4, 193–230.

[49] Tartar, L. An introduction to Sobolev spaces and interpolation spaces. Available online at `http://www.math.cmu.edu/cna/publications.html/SOB+Int.pdf`

[50] Tassa, T. Regularity of weak solutions of the nonlinear Fokker-Planck equation. *Math. Res. Lett*. **3** (1996), no. 4, 475–490.

[51] Tassa, T. Uniqueness of piecewise smooth weak solutions of multidimensional degenerate parabolic equations. *J. Math. Anal. Appl*. **210** (1997), no. 2, 598–608.

[52] Vasseur, A. Kinetic semidiscretization of scalar conservation laws and convergence by using averaging lemmas. *SIAM J. Numer. Anal*. **36** (1999), no. 2, 465–474 (electronic).

[53] Vol′pert, A. I.; Hudjaev, S. I. The Cauchy problem for second order quasilinear degenerate parabolic equations. *Mat. Sb. (N.S.)* **78** (120) (1969), 374–396; translation in *Math. USSR-Sb*. **7** (1969), no. 3, 365–387.

[54] Westdickenberg, M. Some new velocity averaging results. *SIAM J. Math. Anal*. **33** (2002), no. 5, 1007–1032 (electronic).

EITAN TADMOR
University of Maryland
Department of Mathematics
CSCAMM, 4149 CSIC Building #406
Paint Branch Drive
College Park, MD 20742-2389
E-mail: `tadmor@cscamm.umd.edu`

TERENCE TAO
UCLA
Department of Mathematics
Box 951555
Los Angeles, CA 90095-1555
E-mail: `tao@math.ucla.edu`

Received November 2005.