A variational weak weighted derivative: Sobolev spaces and degenerate elliptic equations*

Kristian Bredies†

Abstract

A new class of weak weighted derivatives and its associated Sobolev spaces is introduced and studied. The proposed notion uses a variational formulation in its definition which generalizes the usual weighting of the classical weak derivative. Such a construction naturally leads to Sobolev spaces containing classes of discontinuous functions. Weak closedness with respect to both varying functions and weights are obtained as well as density results and the validity of certain calculus rules in the respective spaces. Moreover, the local properties of functions whose weak weighted derivative exists are examined. Connections to the classical partial weak differentiation are established. In order to demonstrate the applicability, a quasi-linear degenerate elliptic equation modeling an edge-preserving denoising procedure in image processing is analyzed. It turns out that the existence of potentially discontinuous weak solutions for such problems can be ensured, utilizing the closedness and density results for the weighted Sobolev spaces.

Keywords: Weighted Sobolev spaces, closedness of differential operators, density theorems, degenerate elliptic equations, discontinuous solutions, image denoising.

AMS subject classifications: 47F05, 46E35, 35J70, 68U10.

1 Introduction

This article is concerned with a special class of weak weighted derivatives and its associated Sobolev spaces. Differently from the common approach of measuring the weak derivative by multiplication with a spatially dependent positive function, we realize a weighting of the weak derivative through a variational formulation. This can be motivated by the investigation of how linear degenerate elliptic partial differential equations and their solutions behave when the degeneracies are varied. For a fixed, possibly degenerate diffusion tensor field, it usually suffices to construct a suitable completion of a classical Sobolev space with respect to a weaker norm in order to obtain existence and uniqueness [22, 6]. Unfortunately, the abstract device of completing a normed space, i.e. going to equivalence classes of Cauchy sequences or to a suitable subspace of the bidual space, does not reveal the structure of the functions which are contained. The situation becomes even more difficult when one, e.g., considers a sequence of coefficients and data such that each solution is contained in a different space for which in general no inclusion relation holds. This

*The author acknowledges support by the "Center for Industrial Mathematics" (Fachbereich 3, University of Bremen, Germany) during the preparation of this work.

†`kristian.bredies@uni-graz.at`, Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstraße 36, A-8010 Graz, Austria.

problem is addressed by introducing a variational formulation of the weak weighted derivative allowing to define respective Sobolev spaces which possess desireable properties. Furthermore, it is shown how such an approach can be applied to solve certain regularized quasi-linear degenerate elliptic equations arising from a denoising problem in image processing.

In the existing literature, weighted Sobolev spaces are often introduced in the context of degenerate partial differential equations. Usually, the weight is not a full tensor field. Instead, w denotes a non-negative scalar weight on the domain Ω which is assumed to vanish only at the boundary, often monotonically in some sense [17] or with some integrability condition fullfilled such as being in the A_r-class of Muckenhoupt weights [21, 20], i.e. $w^{1/(r-1)} \in L^1_{\text{loc}}(\Omega)$ as well as

$$\sup_{Q\subset\Omega} \Big(\frac{1}{|Q|}\int_Q w \,\mathrm{d}x\Big)\Big(\frac{1}{|Q|}\int_\Omega w^{1/(1-r)} \,\mathrm{d}x\Big)^{r-1} < \infty$$

where the supremum over all cubes $Q \subset \Omega$ is taken.

The associated weighted Sobolev spaces with weight $w \geq 0$ are usually defined with the classical weak gradient involved and with the requirement that

$$\int_\Omega \big(|y|^r + |\nabla y|^r\big) w \,\mathrm{d}x < \infty \ ,$$

holds. This construction is well-analyzed and widely known in the literature (see [24, 15, 25], for example). In particular, it is known that for general measurable weights, approximation with smooth functions fails ($H \neq W$, see [31]). But this is an important property since it is closely connected with the closedness of the gradient operator in the weighted L^r-space and the existence of Poincaré-Wirtinger inequalities [16]. Moreover, varying w in the classical construction causes problems. To the best knowledge of the author, there is no analytic structure which reflects the conditions on the weights properly and simultaneously introduces notions of convergence for the weights and maintains properties like closedness. Another issue is that the conditions on the weights are sometimes too restrictive. On can think of, e.g. weights which are not monotone or where local integrability of $w^{1/(1-r)}$ may fail due to w vanishing "too fast". It will turn out that with the notion of the variational weak weighted derivative, it is possible to overcome these difficulties. This is essentially due to smoothness conditions on the weights which are implicit in the definition. Finally, it is worth mentioning that all results apply to general tensor fields as weights and therefore cover, in particular, the case of anisotropic spaces.

The plan of the paper is as follows: First, in Section 2 we investigate the spaces of tensor fields we allow as weights for the weak weighted derivative. A dual space construction is utilized which gives us a notion of weak* convergence as well as weak* sequential compactness. The weak weighted derivative and the associated Sobolev spaces are introduced and investigated in Section 3. We show that these spaces are Banach spaces in which smooth functions are dense (meaning that $H = W$ holds, just as in the unweighted case [19]) and derive the validity of some common calculus rules. In particular, a closedness result is obtained which allows to make statements about the weak limit of a sequence of functions in which the weights are also varied. Section 4 deals with the investigation of special coefficient matrix fields D and establishes connections with the classical weak derivative. Specifically, the cases where D is a scalar multiple of the identity as well as D implementing the scalar product with a vector field are studied. For the former, it turns out that the weak weighted derivative coincides with the classical notion where the scalar weight does not vanish. Such an observation also leads to more refined closedness properties. The latter type of coefficients gives a notion of a weak directional derivative. Functions which posssess such kind of derivative can locally be regarded as the image of a function being weakly differentiable with respect to one space variable under the coordinate transform induced by the flow with respect to the corresponding vector field. Eventually, in Section 5, these results are applied to derive existence results for a non-linear degenerate elliptic equation arising, e.g., in image processing:

$$\lambda y - \operatorname{div}\Big(\Big(I - \tau(|\nabla y_\sigma|)\frac{\nabla y_\sigma^{\mathrm{T}}}{|\nabla y_\sigma^{\mathrm{T}}|} \otimes \frac{\nabla y_\sigma^{\mathrm{T}}}{|\nabla y_\sigma^{\mathrm{T}}|}\Big)\nabla y^{\mathrm{T}}\Big) = u$$

with corresponding homogeneous Neumann boundary conditions where ∇y_σ denotes a smoothed version of ∇y and τ represents a weighting function which allows for "true" degeneracies, i.e. is equal to one on some interval $[t_0, \infty[$. The solutions of such an regularized anisotropic degenerate quasi-linear equation can model denoised versions of a noisy image u. It is motivated by the famous, generally ill-posed Perona-Malik model [23] which can be regularized in a similar manner [7] (and [27] for the anisotropic case). But, in contrast to the existing regularizations, the new model also allows for "true" discontinuities across hypersurfaces. The article finally concludes with some remarks in Section 6.

2 Preliminaries

In what follows, we will assume that the space dimension $d \geq 1$ is fixed, $\Omega \subset \mathrm{I\!R}^d$ is a bounded domain and we have moreover a fixed $M_x > 0$. The following notation will be commonly used within the article:

$$(\nabla y)_{i,j,k} = \frac{\partial y_{i,j}}{\partial x_k} \quad , \quad (\operatorname{div} y)_{i,j} = \sum_{k=1}^{d} \frac{\partial y_{i,j,k}}{\partial x_k} \quad ,$$
$$y \cdot z = \sum_{i,j,k=1}^{d} y_{i,j,k} z_{i,j,k} \quad , \quad |y| = \Big(\sum_{i,j,k=1}^{d} |y_{i,j,k}|^2\Big)^{1/2}$$

and the analogs in which summation over i and i,j is dropped, respectively, for vector-/matrix-valued and scalar-/vector-valued functions. The matrix/tensor norms according to the above are in particular utilized to compute the associated Lebesgue and Sobolev norms.

To formulate a variational definition of weak weighted derivatives, we first introduce the function space we consider for the coefficient tensor fields D and examine its topology. This will essentially be $W^{1,\infty}(\Omega, \mathrm{I\!R}^{d\times d})$ interpreted as a dual space.

Definition 2.1. Let Ω be a bounded domain and $M_x > 0$. We define the space $Y = L^1(\Omega, \mathrm{I\!R}^{d\times d})$ equipped with the norm

$$\|v\|_Y = \inf\ \{\|v_1\|_1 + \|v_2\|_1 \ \big|\ (v_1, v_2) \in \Sigma(v)\} \tag{1a}$$

where

$$\Sigma(v) = \{(v_1, v_2) \in L^1(\Omega, \mathrm{I\!R}^{d\times d}) \times W_0^{1,1}(\Omega, \mathrm{I\!R}^{d\times d\times d}) \ \big|\ v = v_1 - M_x \operatorname{div} v_2\} \ . \tag{1b}$$

Likewise, define the space $Y_{\mathrm{div}} = Y$, but equipped with the stronger norm

$$\|v\|_{Y_{\mathrm{div}}} = \inf\{\|v_1\|_1 + \|v_2\|_1 \ \big|\ (v_1, v_2) \in \Sigma_{\mathrm{div}}(v)\} \ , \tag{2a}$$
$$\Sigma_{\mathrm{div}}(v) = \{(v_1, v_2) \in L^1(\Omega, \mathrm{I\!R}^{d\times d}) \times W_0^{1,1}(\Omega, \mathrm{I\!R}^{d}) \ \big|\ v = v_1 - M_x \nabla v_2\} \ . \tag{2b}$$

Proposition 2.2. *The spaces Y, Y_{div} are separable normed spaces. Their duals are given by*

$$Y^* = W^{1,\infty}(\Omega, \mathrm{IR}^{d\times d}) \quad , \quad \|w\|_{Y^*} = \max\left\{\|w\|_\infty, M_x\|\nabla w\|_\infty\right\} , \tag{3a}$$

$$\begin{cases} Y_{\mathrm{div}}^* = \{w \in L^\infty(\Omega, \mathrm{IR}^{d\times d}) \mid \operatorname{div} w \in L^\infty(\Omega, \mathrm{IR}^d)\} , \\ \|w\|_{Y_{\mathrm{div}}^*} = \max\left\{\|w\|_\infty, M_x\|\operatorname{div} w\|_\infty\right\} . \end{cases} \tag{3b}$$

Proof. We will show the statements for Y^* as the argumentation is completely analogous for Y_{div}^*. First, $\Sigma(v) \neq \emptyset$ for each $v \in L^1(\Omega, \mathrm{IR}^{d\times d})$, hence the norm is well-defined as a function on $L^1(\Omega, \mathrm{IR}^{d\times d})$. The positive homogeneity and the triangle inequality can be directly obtained from the definition, while the positive definiteness follows easily from the closedness of ∇ and div in the respective L^1-spaces.

Next, by definition, the Y-norm is weaker than the L^1-norm, so Y is also separable. Moreover, we have the embedding $Y^* \hookrightarrow L^\infty(\Omega, \mathrm{IR}^{d\times d})$. Hence, an element $w \in Y^*$ can be interpreted as a L^∞-function which acts on $v \in L^1(\Omega, \mathrm{IR}^{d\times d})$ via

$$v \mapsto \int_\Omega w \cdot v \, \mathrm{d}x .$$

Additionally, for each $v \in \mathcal{C}_0^\infty(\Omega, \mathrm{IR}^{d\times d\times d})$,

$$\left|\int_\Omega w \cdot \operatorname{div} v \, \mathrm{d}x\right| \leq M_x^{-1}\|w\|_{Y^*}\|v\|_1$$

holds, meaning that the distributional derivative of w is in $L^\infty(\Omega, \mathrm{IR}^{d\times d})$. Consequently, $w \in W^{1,\infty}(\Omega, \mathrm{IR}^{d\times d})$. Conversely, if $w \in W^{1,\infty}(\Omega, \mathrm{IR}^{d\times d})$, we have for each $(v_1, v_2) \in L^1(\Omega, \mathrm{IR}^{d\times d}) \times W_0^{1,1}(\Omega, \mathrm{IR}^{d\times d\times d})$

$$\begin{aligned} \left|\int_\Omega w \cdot v_1 - M_x w \cdot \operatorname{div} v_2 \, \mathrm{d}x\right| &\leq \left|\int_\Omega w \cdot v_1 + M_x \nabla w \cdot v_2 \, \mathrm{d}x\right| \\ &\leq \max\left\{\|w\|_\infty, M_x\|\nabla w\|_\infty\right\}(\|v_1\|_1 + \|v_2\|_1) , \end{aligned}$$

meaning that $W^{1,\infty}(\Omega, \mathrm{IR}^{d\times d}) \hookrightarrow Y^*$. The claimed identity for the norm in Y^* finally follows from the observation that

$$\begin{aligned} \sup_{v\neq 0} \frac{\left|\int_\Omega v \cdot w \, \mathrm{d}x\right|}{\|v\|_Y} &= \sup_{v\neq 0} \sup_{(v_1,v_2)\in\Sigma(v)} \frac{\left|\int_\Omega v \cdot w \, \mathrm{d}x\right|}{\|v_1\|_1 + \|v_2\|_1} \\ &= \sup_{v_1 - M_x \operatorname{div} v_2 \neq 0} \frac{\left|\int_\Omega w \cdot v_1 + M_x \nabla w \cdot v_2 \, \mathrm{d}x\right|}{\|v_1\|_1 + \|v_2\|_1} \\ &= \sup_{(v_1,v_2)\neq 0} \frac{\left|\int_\Omega w \cdot v_1 + M_x \nabla w \cdot v_2 \, \mathrm{d}x\right|}{\|v_1\|_1 + \|v_2\|_1} , \end{aligned}$$

the latter since the expression is zero where $v_1 - M_x \operatorname{div} v_2 = 0$ and $(v_1, v_2) \neq 0$. The well-known L^1-L^∞ duality as well as density yields the identity $\|w\|_{Y^*} = \max\left\{\|w\|_\infty, M_x\|\nabla w\|_\infty\right\}$. □

Remark 1. The are some subspaces of Y^* which will be of interest later:

$$Y_0^* = \{w \in Y^* \mid w = vI \text{ almost everywhere}\} \tag{4a}$$

$$Y_1^* = \{w \in Y^* \mid w = \mathbf{1} \otimes q \text{ almost everywhere}\} . \tag{4b}$$

We denote by $Y^*_{\mathrm{div},0}$ and $Y^*_{\mathrm{div},1}$ the analog versions for Y^*_{div}. Note that one can identify these subspaces with

$$Y_0^* = W^{1,\infty}(\Omega) \quad , \quad \|w\|_{Y_0^*} = \sqrt{d}\max\{\|w\|_\infty, M_x\|\nabla w\|_\infty\}\ , \tag{5a}$$

$$Y_1^* = W^{1,\infty}(\Omega, \mathbb{R}^d) \quad , \quad \|q\|_{Y_1^*} = \sqrt{d}\max\{\|q\|_\infty, M_x\|\nabla q\|_\infty\}\ , \tag{5b}$$

$$Y^*_{\mathrm{div},0} = W^{1,\infty}(\Omega) \quad , \quad \|w\|_{Y^*_{\mathrm{div},0}} = \max\{\sqrt{d}\|w\|_\infty, M_x\|\nabla w\|_\infty\}\ , \tag{5c}$$

$$\begin{cases} Y^*_{\mathrm{div},1} = \{q \in L^\infty(\Omega, \mathbb{R}^d) \mid \operatorname{div} q \in L^\infty(\Omega)\}\ , \\ \|q\|_{Y^*_{\mathrm{div},1}} = \sqrt{d}\max\{\|q\|_\infty, M_x\|\operatorname{div} q\|_\infty\}\ . \end{cases} \tag{5d}$$

The space Y^* (as well as the subspaces introduced above) is intrinsically connected with Lipschitz continuous functions.

Proposition 2.3.

1. *A matrix field $w : \Omega \to \mathbb{R}^{d\times d}$ belongs to $W^{1,\infty}(\Omega, \mathbb{R}^{d\times d}) = Y^*$ if and only if w is bounded and locally Lipschitz continuous with Lipschitz constant bound C not depending on Ω' for each convex $\Omega' \subset\subset \Omega$.*

2. *Each $w \in Y^*$ is differentiable almost everywhere and the gradient satisfies $\|\nabla w\|_\infty \le C'C$ with some $C' > 0$ and C the above Lipschitz constant.*

3. *For each $\Omega' \subset\subset \Omega$ there is a constant $C_{\Omega'} \ge 1$ ($C_{\Omega'} = 1$ for convex Ω') such that each $w \in Y^*$ is Lipschitz continuous with constant not greater that $C_{\Omega'}\|\nabla w\|_\infty$.*

Proof. The first statement is an extension of Theorem 4.2.5 in [11] which states that w being locally Lipschitz is equivalent to $w \in W^{1,\infty}_{\mathrm{loc}}(\Omega, \mathbb{R}^{d\times d})$.

To prove the only if part, it remains to show the independence of the Lipschitz constant on convex sets $\Omega' \subset\subset \Omega$. With the standard mollifier G and its dilated versions G_ε, we define $w_\varepsilon = w * G_\varepsilon$ (componentwise) on Ω' for sufficiently small $\varepsilon > 0$. We have

$$w_\varepsilon(x) - w_\varepsilon(\xi) = \int_0^1 \nabla w_\varepsilon\big(\xi + s(x-\xi)\big)(x-\xi)\ \mathrm{d}x$$

since Ω' is convex and w_ε is smooth, hence

$$|w(x) - w(\xi)| = \lim_{\varepsilon\to 0} |w_\varepsilon(x) - w_\varepsilon(\xi)| \le \|\nabla w_\varepsilon\|_\infty |x-\xi| \le \|\nabla w\|_\infty |x-\xi|\ ,$$

using that w_ε converges uniformly in Ω' as well as $\nabla w_\varepsilon = \nabla w * G_\varepsilon$ (see Theorem 4.2.1 in [11]).

Conversely, we only have to show that the weak gradient ∇w belongs to the space $L^\infty(\Omega, \mathbb{R}^{d\times d\times d})$. It is known that the L^∞-norm of the gradient in a $\Omega' \subset \Omega$ can be estimated by the Lipschitz constant in Ω'. In particular, denoting by C' the constant which arises from the change from operator tensor norm to the Frobenius norm (i.e. $|A| \le C' \sup_{|u|\le 1} |Au|$),

$$\operatorname*{ess\,sup}_{\xi\in B_\varepsilon(x)} |\nabla w(\xi)| \le C'C \quad \text{for all} \quad \overline{B_\varepsilon(x)} \subset \Omega$$

by assumption, hence $\nabla w \in L^\infty(\Omega, \mathbb{R}^{d\times d\times d})$ since Ω is open.

The second statement follows from the fact that one can choose a sequence Ω'_k of open convex sets satisfying

$$\bigcup_k \Omega'_k = \Omega \quad , \quad \overline{\Omega'_k} \subset\subset \Omega \text{ for } k = 1, 2, \ldots$$

and applying Rademacher's theorem to each $\overline{\Omega_k'}$ thus reducing the set where w is not differentiable to a countable union of null-sets which is also a null-set.

Finally, the last statement is well-known, we only have to derive the Lipschitz constant estimate. Suppose that Ω' is connected. This is not less general since each $\Omega' \subset\subset \Omega$ is subset of a connected compact set in Ω. Choose an $\varepsilon > 0$ such that $\Omega_\varepsilon = \Omega' + B_\varepsilon(0)$ satisfies $\overline{\Omega_\varepsilon} \subset\subset \Omega$ and define for each $x, \xi \in \Omega'$ the usual distance

$$\rho(x, \xi) = \inf \left\{ \int_0^1 |\gamma'(s)| \, \mathrm{d}s \;\middle|\; \gamma \in \mathcal{C}^1([0,1], \overline{\Omega_\varepsilon}), \gamma(0) = x, \gamma(1) = \xi \right\} .$$

One can easily deduce the triangle inequality for ρ as well as $\rho(x, x) = 0$ for $x \in \Omega'$. It moreover follows that $\rho(x, \xi) = |x - \xi|$ if $|x - \xi| < \varepsilon$, hence

$$|\rho(\tilde{x}, \tilde{\xi}) - \rho(x, \xi)| \leq |\tilde{x} - x| + |\tilde{\xi} - \xi| < \delta$$

if $|\tilde{x} - x| < \delta/2 < \varepsilon$, $|\tilde{\xi} - \xi| < \delta/2 < \varepsilon$, yielding the continuity of ρ on $\Omega' \times \Omega'$. In particular, $\rho(x, \xi)/|x - \xi|$ is continuous on $(\Omega' \times \Omega') \backslash \{|x - \xi| < \varepsilon\}$, admitting a maximum $C_{\Omega'} \geq 1$. On $(\Omega' \times \Omega') \cap \{0 < |x - \xi| < \varepsilon\}$ we have $\rho(x, \xi)/|x - \xi| = 1$, hence the function is bounded on $\Omega' \times \Omega'$.

Recalling from the proof of the first statement we deduce for each $w \in Y^*$ that

$$\begin{aligned} |w(x) - w(\xi)| \leq \rho(x, \xi) \|\nabla w\|_\infty &\leq \sup_{\tilde{x}, \tilde{\xi} \in \Omega', \tilde{x} \neq \tilde{\xi}} \frac{\rho(\tilde{x}, \tilde{\xi})}{|\tilde{x} - \tilde{\xi}|} \|\nabla w\|_\infty |x - \xi| \\ &\leq C_{\Omega'} \|\nabla w\|_\infty |x - \xi| \ . \end{aligned}$$

It follows that the Lipschitz constant of w on the compact subset Ω' can be estimated by $C_{\Omega'} \|\nabla w\|_\infty$ where $C_{\Omega'}$ only depends on Ω'. In particular, $C_{\Omega'} = 1$ if Ω' is convex. □

Finally, note some topological properties of the space Y^*.

Proposition 2.4.

1. *Each bounded sequence $\{w_k\} \subset Y^*$ admits a weakly* convergent subsequence.*

2. *A sequence $\{w_k\} \subset Y^*$ converges to $w \in Y^*$ weakly* if and only if $w_k \overset{*}{\rightharpoonup} w$ in $L^\infty(\Omega, \mathrm{IR}^{d \times d})$ and $\nabla w_k \overset{*}{\rightharpoonup} \nabla w$ in $L^\infty(\Omega, \mathrm{IR}^{d \times d \times d})$.*

3. *Y^* is continously embedded in $\mathcal{C}(\Omega, \mathrm{IR}^{d \times d})$ such that $w_k \overset{*}{\rightharpoonup} w$ in Y^* implies $w_k \to w$ in $\mathcal{C}(\Omega, \mathrm{IR}^{d \times d})$.*

Proof. Regarding the first statement, by Proposition 2.2, we know that the predual space Y is separable, hence bounded sequences in Y^* admit weakly* convergent subsequences (confer [28], for example).

To prove the second statement, assume that $w_k \overset{*}{\rightharpoonup} w$ in Y^*. So $\|w_k\|_\infty$ and $\|\nabla w_k\|_\infty$ are bounded, thus $w_k \overset{*}{\rightharpoonup} w$ and $\nabla w_k \overset{*}{\rightharpoonup} \nabla w$ for each subsequence in the respective L^∞-spaces, and hence for the whole sequence. Conversely, if $w_k \overset{*}{\rightharpoonup} w$ and $\nabla w_k \overset{*}{\rightharpoonup} w$ in the respective L^∞-spaces, we get a bounded sequence in Y^* for which each subsequence converges weakly* to w and consequently the whole sequence.

For the compact embedding property in the third statement, suppose that we are given a $\{w_k\} \subset Y^*$ which converges in the weak* sense to a $w \in Y^*$. Recall that each w_k is Lipschitz continuous in $\Omega' \subset\subset \Omega$ with a constant $C_{\Omega'}$ which is only dependent on Ω' and on the bound for $\|\nabla w_k\|_\infty$, see Proposition 2.3. Since this holds for all such Ω', $w_k \in \mathcal{C}(\Omega, \mathrm{IR}^k)$.

Now pick a $\Omega' \subset\subset \Omega$ with $\Omega' = \overline{\mathrm{int}(\Omega')}$. We want to show that $\{w_k\}$ converges in $\mathcal{C}(\Omega', \mathrm{IR}^{d \times d})$ by applying the theorem of Arzelà and Ascoli. Since $\{w_k\}$ is equicontinuous, there exists a subsequence of $\{w_k\}$ which converges in $\mathcal{C}(\Omega', \mathrm{IR}^{d \times d})$ to a

$\tilde{w}$. For the subsequence there also holds $\nabla w_k \overset{*}{\rightharpoonup} \nabla w$ in $L^\infty(\Omega', \mathrm{IR}^{d\times d\times d})$, hence the closedness of the gradient yields that $\nabla\tilde{w} = \nabla w$. But this means that on each connected component $\Omega'' \subset \Omega'$ with non-empty interior we have $\tilde{w} = w + c_{\Omega''}$ with a $c_{\Omega''} \in \mathrm{IR}^{d\times d}$. Testing $w_k \in L^\infty(\Omega, \mathrm{IR}^{d\times d})$ with $\chi_{\Omega''} e_{i,j}$, leads to

$$\int_{\Omega''} \tilde{w} \cdot e_{i,j} \, \mathrm{d}x = \lim_{k\to\infty} \int_{\Omega''} w_k \cdot e_{i,j} \, \mathrm{d}x = \int_{\Omega''} w \cdot e_{i,j} \, \mathrm{d}x$$

which means that actually $c_{\Omega''} = 0$. Since this holds for each connected component of Ω' we conclude $\tilde{w} = w$. Additionally, this is true for each subsequence, so the whole sequence converges to w in $\mathcal{C}(\Omega', \mathrm{IR}^{d\times d})$.

Now choose for $\varepsilon > 0$ the compact set $\Omega'_\varepsilon = \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) \geq \varepsilon\}$ which satisfies $\Omega'_\varepsilon = \overline{\operatorname{int}(\Omega'_\varepsilon)}$ for ε small enough. By the above arguments, $w_k \to w$ in $\mathcal{C}(\Omega'_\varepsilon, \mathrm{IR}^{d\times d})$, hence for $k \geq k_0$ we have $|w_k(x) - w(x)| < \varepsilon$ for each $x \in \Omega'_\varepsilon$. Moreover, for each $\xi \in \Omega$ an $x \in \Omega'_\varepsilon$ with $|x - \xi| < \varepsilon_0 < \varepsilon$ can be found and by local Lipschitz continuity on $\overline{B_{\varepsilon_0}(x)} \subset\subset \Omega$ we have

$$|w_l(\xi) - w(\xi)| \leq |w_l(\xi) - w_l(x)| + |w_l(x) - w(x)| + |w(x) - w(\xi)| \leq (2C + 1)\varepsilon$$

where C is a bound on $\{\|\nabla w_k\|_\infty\}$. Consequently, $\|w_k - w\|_\infty \to 0$ on Ω which proves the desired statement. In particular, each $w \in Y^*$ belongs to $\mathcal{C}(\Omega, \mathrm{IR}^{d\times d})$. □

3 Weak weighted derivatives and weighted Sobolev spaces

In the following, we introduce and study the notion of variational weak weighted derivatives. Taking the weights in Y^*_{div}, we are able to state the definition for which we can formulate associated Sobolev spaces. It turns out that to derive density results for these spaces as well as finer properties of the weak weighted gradients, we need matrix fields in the smaller space Y^*. We start with some basic considerations.

3.1 Definition and fundamental properties

Definition 3.1. Let $y : \Omega \to \mathrm{IR}$ be locally integrable and $D \in Y^*_{\mathrm{div}}$. Then the *weak weighted gradient* $v = D\nabla y^{\mathrm{T}}$ is defined as the locally integrable function $v : \Omega \to \mathrm{IR}^d$ fulfilling

$$\int_\Omega v \cdot z \, \mathrm{d}x = -\int_\Omega y(D \cdot \nabla z + \operatorname{div} D \cdot z) \, \mathrm{d}x \quad \text{for all } z \in \mathcal{C}_0^\infty(\Omega, \mathrm{IR}^d) \; . \tag{6}$$

Remark 2. (i) By the results from Section 2, it is immediate that the definition makes sense and that $D\nabla y^{\mathrm{T}}$ has to be unique, since it is tested against all functions $z \in \mathcal{C}_0^\infty(\Omega, \mathrm{IR}^d)$.

(ii) Note that this definition allows for functions which are of lower regularity than weakly differentiable functions: Since D can have arbitrary eigenvalues on non-null sets in Ω, the weak weighted gradient may exist in $L^r(\Omega)$ even if y is not weakly differentiable.

(iii) For $y \in W^{1,r}(\Omega)$, it is easily verified that $(D\nabla y^{\mathrm{T}}) = D(\nabla y^{\mathrm{T}})$, i.e. the weak weighted derivative is the pointwise application of D to ∇y^{T}. The operation $(y, D) \to D\nabla y^{\mathrm{T}}$ mapping $W^{1,r}(\Omega) \times L^\infty(\Omega) \to L^r(\Omega, \mathrm{IR}^d)$ is of course bilinear and continuous since

$$\|D\nabla y^{\mathrm{T}}\|_r \leq \|D\|_\infty \big(\|y\|_r^r + \|\nabla y\|_r^r\big)^{1/r} \; .$$

This also applies if the space Y^*_{div} or Y^* is chosen for D.

Definition 3.2. Let $1 \le r \le \infty$ and $D \in Y_{\mathrm{div}}^*$. Then we define the *weighted Sobolev space* $W_D^{1,r}$ as the vector space of functions satisfying

$$y \in L^r(\Omega) \quad , \quad D\nabla y^{\mathrm{T}} \in L^r(\Omega, \mathrm{I\!R}^d)$$

with norm

$$\|y\|_{W_D^{1,r}} = \Big(\int_\Omega |y|^r + |D\nabla y^{\mathrm{T}}|^r \,\mathrm{d}x\Big)^{1/r} \quad \text{if } 1 \le r < \infty \;, \tag{7a}$$

$$\|y\|_{W_D^{1,\infty}} = \max\,\{\|y\|_\infty, \|D\nabla y^{\mathrm{T}}\|_\infty\}\;. \tag{7b}$$

In order to examine the Banach space properties of $W_D^{1,r}$, we need statements about the closedness of the underlying differential operators. In this context, we will also address preferably weak closedness properties with respect to varying coefficients, one of the main points of this article. It turns out that the variational definition of the weak gradient is naturally connected with such properties.

Proposition 3.3. *Let $1 < r < \infty$ and the sequences $\{y_k\} \subset L^r(\Omega)$ as well as $\{D_k\} \subset Y_{\mathrm{div}}^*$ be given such that*

$$y_k \rightharpoonup y \text{ in } L^r(\Omega) \quad , \quad D_k \overset{*}{\rightharpoonup} D \text{ in } Y_{\mathrm{div}}^* \quad \text{and} \quad D_k \nabla y_k^{\mathrm{T}} \rightharpoonup v \text{ in } L^r(\Omega, \mathrm{I\!R}^d)\;.$$

If, additionally, $y_k \to y$ or $D_k \to D$ together with $\operatorname{div} D_k \to \operatorname{div} D$ pointwise almost everywhere in Ω, then $v = D\nabla y^{\mathrm{T}}$.

Proof. Note that for a fixed $z \in \mathcal{C}_0^\infty(\Omega, \mathrm{I\!R}^d)$ one can define, on the dense subset $\mathcal{C}^\infty(\overline{\Omega})$, the mapping $J_z : v \mapsto \nabla v \otimes z$ from $L^1(\Omega) \to Y_{\mathrm{div}}$. We can write $z \otimes \nabla v = v\nabla z - \nabla(vz)$ which leads, by definition of the norm in Y_{div}, to the continuity estimate

$$\|z \otimes \nabla v\|_{Y_{\mathrm{div}}} \le \|v\nabla z\|_1 + M_x^{-1}\|vz\|_1 \le (\|\nabla z\|_\infty + M_x^{-1}\|z\|_\infty)\|v\|_1\;.$$

This makes J^* weak*-convergence preserving, hence $J_z^* D_k = -(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z) \overset{*}{\rightharpoonup} J_z^* D = -(D \cdot \nabla z + \operatorname{div} D \cdot z)$ in $L^\infty(\Omega)$.

Our aim now is to show convergence of the defining integrals in (6). First suppose that $y_k \to y$ pointwise almost everywhere. By Egorov's Theorem one can choose for each $\varepsilon > 0$ a measurable $\Omega' \subset \Omega$ with $|\Omega \backslash \Omega'| < \varepsilon$ such that $y_k \to y$ uniformly on Ω'. In particular, $y_k \to y$ in $L^r(\Omega')$. Hence, splitting the defining integral as follows

$$\int_\Omega y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x = \int_{\Omega'} y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x + \int_{\Omega\backslash\Omega'} y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x$$

leads to

$$\lim_{k\to\infty} \int_{\Omega'} y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x = \int_{\Omega'} y(D \cdot \nabla z + \operatorname{div} D \cdot z)\,\mathrm{d}x$$

as well as

$$\Big|\int_{\Omega\backslash\Omega'} y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x\Big| \le (\|\nabla z\|_\infty + M_x^{-1}\|z\|_\infty)\|D_k\|_{Y_{\mathrm{div}}^*} \int_{\Omega\backslash\Omega'} |y_k|\,\mathrm{d}x \le C_1|\Omega\backslash\Omega'|^{1/r'}\|y_k\|_r \le C_2\varepsilon^{1/r'}$$

where $\frac{1}{r} + \frac{1}{r'} = 1$ and $C_1, C_2 > 0$ are suitable constants which can be chosen independently of k since $\{y_k\}$ and $\{D_k\}$ are bounded sequences. Letting $\varepsilon \to 0$ then implies that

$$\lim_{k\to\infty} \int_\Omega y_k(D_k \cdot \nabla z + \operatorname{div} D_k \cdot z)\,\mathrm{d}x = \int_\Omega y(D \cdot \nabla z + \operatorname{div} D \cdot z)\,\mathrm{d}x\;. \tag{8}$$

Now suppose that $D_k \to D, \operatorname{div} D_k \to \operatorname{div} D$ pointwise almost everywhere in Ω. From the above considerations we have that $\|D_k \cdot \nabla z + \operatorname{div} D_k \cdot z\|_\infty$ is bounded, thus, by Lebesgue's convergence theorem, $D_k \cdot \nabla z + \operatorname{div} D_k \cdot z \to D \cdot \nabla z + \operatorname{div} D \cdot z$ in $L^{r'}(\Omega)$ from which it immediately follows that (8) is also satisfied in this case.

So, convergence of both sides of the defining integral is guaranteed and one has

$$\int_\Omega v \cdot z \, \mathrm{d}x = - \int_\Omega y(D \cdot \nabla z + \operatorname{div} D \cdot z) \, \mathrm{d}x$$

consequently leading to $v = D\nabla y^{\mathrm{T}}$. □

Remark 3. (i) It can be easily seen that in the case of strong convergence $y_k \to y$ in $L^r(\Omega)$, the statement is also valid for $1 \leq r \leq \infty$ (with weak*-convergence of $\{D_k \nabla y_k\}$ for $r = \infty$).

(ii) If each $D_k \in Y^*$ and $D_k \overset{*}{\rightharpoonup} D$ there, then it suffices that $\operatorname{div} D_k^{\mathrm{T}} \to \operatorname{div} D^{\mathrm{T}}$ pointwise almost everywhere since $D_k \to D$ uniformly in Ω, due to Proposition 2.4.

An immediate consequence of the strong closedness property is the completeness of the spaces $W_D^{1,r}$.

Proposition 3.4. *The weighted Sobolev space $W_D^{1,r}$ according to Definition 3.2 and associated with a $D \in Y^*_{\operatorname{div}}$ is a Banach space which is reflexive for $1 < r < \infty$.*

Proof. The normed space $W_D^{1,r}$ can be interpreted as a subspace, denoted by $\tilde{W}_D^{1,r}$, of a Lebesgue space via

$$(y, v) \in \tilde{W}_D^{1,r} \subset L^r(\Omega) \times L^r(\Omega, \mathrm{IR}^d) \qquad \Leftrightarrow \qquad v = D\nabla y^{\mathrm{T}} \ .$$

All statements follow from the closedness of $\tilde{W}_D^{1,r}$ since closed subspaces of (reflexive) Banach spaces are (reflexive) Banach spaces as well. But this corresponds to the closedness of the differential operator $y \mapsto D\nabla y^{\mathrm{T}}$ which has been established in Remark 3. □

Remark 4. It is easy to check that $W_D^{1,2}$ is also a real Hilbert space.

3.2 Density theorems

In the following, we will be interested in density statements for the weighted spaces $W_D^{1,r}$, in particular in the density of $\mathcal{C}^\infty(\Omega) \cap W_D^{1,r}$ which is known for classical Sobolev spaces ("$H = W$", see [19] and also [9]) but fails in general for weighted Sobolev spaces [31]. Also, we want to establish the density of $\mathcal{C}^\infty(\overline{\Omega})$ which is known to be true, in case classical Sobolev spaces, whenever the boundary satisfies certain regularity assumptions.

Following the lines of proof which can be found in many textbooks (e.g. [1]), the density of $\mathcal{C}^\infty(\Omega)$ can be proven under abstract conditions on the differential operator.

Proposition 3.5. *Let $l \geq 1$, $1 < r < \infty$ and $\Lambda : \mathcal{D}(\Lambda) \subset L^r_{\mathrm{loc}}(\Omega) \to L^r_{\mathrm{loc}}(\Omega, \mathrm{IR}^l)$ with $\mathcal{C}_0^\infty(\Omega) \subset \mathcal{D}(\Lambda)$ a closed linear operator with the property that*

1. *$y \in \mathcal{D}(\Lambda)$ if and only if for each $\zeta \in \mathcal{C}_0^\infty(\Omega)$ follows $\zeta y \in \mathcal{D}(\Lambda)$,*
2. *for each $y \in \mathcal{D}(\Lambda)$ and $\zeta \in \mathcal{C}_0^\infty(\Omega)$ follows $\operatorname{supp} \Lambda(\zeta y) \subset \operatorname{supp} \zeta$,*
3. *for each $y \in \mathcal{D}(\Lambda)$ with compact support in Ω, there exists a $\delta_0 > 0$ such that the sequence of mollified $y_\delta = y * G_\delta$ satisfies $\|\Lambda y_\delta\|_r \leq C$ for every $0 < \delta < \delta_0$ and some $C > 0$.*

Then, for each $\varepsilon > 0$ and $y \in L^r(\Omega)$ with $\Lambda y \in L^r(\Omega, \mathrm{IR}^l)$, there exists a $\bar{y} \in \mathcal{C}^\infty(\Omega)$ such that

$$\Big(\int_\Omega |\bar{y}(x) - y(x)|^r + |(\Lambda\bar{y})(x) - (\Lambda y)(x)|^r \,\mathrm{d}x\Big)^{1/r} < \varepsilon \ . \tag{9}$$

Proof. We introduce the sets $\Omega_k = \{x \in \Omega \mid \mathrm{dist}(x, \partial\Omega) > \frac{1}{k}\}$, $\Omega_0 = \emptyset$ and $\Omega'_k = \Omega_{k+1}\backslash\overline{\Omega_{k-1}}$ for $k = 1, 2, \ldots$. Let $\varepsilon > 0$ be given and choose $\zeta_k \in \mathcal{C}_0^\infty(\Omega'_k)$ to form the usual partition of unity subordinate to Ω_k, i.e.

$$0 \leq \zeta_k \leq 1 \quad \text{and} \quad \sum_{k'=1}^{k} \zeta_{k'} = 1 \quad \text{in } \Omega_k \ .$$

According to the third assumption on Λ, for each $\zeta_k y$, the mollified versions satisfy $\|\Lambda\big((\zeta_k y) * G_\delta\big)\|_r \leq C_k$ for every $0 < \delta < \min\,\{\mathrm{dist}(\mathrm{supp}\,\zeta_k, \partial\Omega'_k), \delta_{0,k}\}$, yielding a weakly converging subsequence $\Lambda\big((\zeta_k y) * G_{\delta_{k'}}\big)$ with limit θ_k as $\delta_{k'} \to 0$. By the theorem of Banach-Saks, there is a subsequence (not relabeled), such that

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k'=1}^{m} \Lambda\big((\zeta_k y) * G_{\delta_{k'}}\big) = \theta_k \quad , \quad \lim_{m\to\infty} \frac{1}{m} \sum_{k'=1}^{m} (\zeta_k y) * G_{\delta_{k'}} = \zeta_k y$$

the latter being a consequence of $(\zeta_k y) * G_{\delta_{k'}} \to \zeta_k y$ as $k' \to \infty$. Due to the closedness of Λ, there has to be $\theta_k = \Lambda(\zeta_k y)$. Choosing m_k large enough and writing

$$y_k = \frac{1}{m_k} \sum_{k'=1}^{m_k} (\zeta_k y) * G_{\delta_{k'}}$$

we can achieve that, exploiting the linearity of Λ,

$$\|y_k - \zeta_k y\|_r < \frac{\varepsilon}{2^{k+1}} \qquad , \qquad \|\Lambda y_k - \Lambda(\zeta_k y)\|_r < \frac{\varepsilon}{2^{k+1}} \ .$$

Note that by construction $y_k \in \mathcal{C}_0^\infty(\Omega'_k)$ since δ is chosen small enough. Moreover, the sets $\Omega'_1, \Omega'_2, \ldots$ are locally finite, hence the summation $\bar{y} = \sum_{k'=1}^\infty y_{k'}$ is finite in a neighborhood of each $x \in \Omega$, yielding that $\bar{y} \in \mathcal{C}^\infty(\Omega)$. Furthermore, there has to be $\bar{y} \in \mathcal{D}(\Lambda)$ meaning that $\Lambda\bar{y} \in L^r_{\mathrm{loc}}(\Omega, \mathrm{IR}^l)$: This follows from the first assumption on Λ since for each $\zeta \in \mathcal{C}_0^\infty(\Omega)$ there has to be a k such that $\mathrm{supp}\,\zeta \subset \Omega_{k-1}$ and consequently

$$\zeta\bar{y} = \sum_{k'=1}^{k} \zeta y_{k'} \quad \text{as well as} \quad \Lambda(\zeta\bar{y}) = \sum_{k'=1}^{k} \Lambda(\zeta y_{k'}) \ ,$$

taking the second assumption on Λ into account. In particular, $\zeta y \in \mathcal{D}(\Lambda)$.

Also, observe that one can deduce similarly that

$$\chi_{\Omega_k}(\bar{y} - y) = \chi_{\Omega_k} \sum_{k'=1}^{k+1} (y_{k'} - \zeta_{k'} y) \ , \ \chi_{\Omega_k}(\Lambda\bar{y} - \Lambda y) = \chi_{\Omega_k} \sum_{k'=1}^{k+1} \big(\Lambda y_{k'} - \Lambda(\zeta_{k'} y)\big)$$

leading to the estimate

$$\int_{\Omega_k} |\bar{y}(x) - y(x)|^r + |(\Lambda\bar{y})(x) - (\Lambda y)(x)|^r \,\mathrm{d}x$$
$$\leq \Big(\sum_{k'=1}^{k} \|y_{k'} - \zeta_k y\|_r\Big)^r + \Big(\sum_{k'=1}^{k} \|\Lambda y_{k'} - \Lambda(\zeta_{k'} y)\|_r\Big)^r < \frac{(1 - 2^{-k})^r}{2^{r-1}} \varepsilon^r \ .$$

The fact that $\Lambda\bar{y} \in L^r(\Omega, \mathrm{IR}^l)$ and, consequently, the claimed statement (9) then follows from the application of Levi's monotone-convergence theorem on the sequence $\{\chi_{\Omega_k}\big(|\bar{y} - y|^r + |\Lambda\bar{y} - \Lambda y|^r\big)\}$. $\square$

The next lemmas are concerned with providing the necessary arguments for the verification the prerequisites of this abstract approximation result. The difficulty here lies in establishing the boundedness of the $\{D\nabla(G_\delta * y)\}$ as δ approaches zero. Therefore, we will eventually assume the stronger regularity condition $D \in Y^*$.

Lemma 3.6. *If $D\nabla y^{\mathrm{T}}$ exists for a locally integrable y and a $D \in Y^*_{\mathrm{div}}$, then* (6) *also holds for vector fields $z \in W^{1,\infty}(\Omega, \mathbb{R}^d)$ with compact support in Ω.*

Proof. This can be seen by using the standard mollifier G and its dilated versions G_ε. Denoting $z_\varepsilon = z * G_\varepsilon$ with $\varepsilon > 0$ small enough such that $z_\varepsilon \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d)$, it is clear that the left-hand side of

$$\int_\Omega D\nabla y^{\mathrm{T}} \cdot z_\varepsilon \, \mathrm{d}x = -\int_\Omega y(D \cdot \nabla z_\varepsilon + \operatorname{div} D \cdot z_\varepsilon) \, \mathrm{d}x$$

converges to $\int_\Omega D\nabla y^{\mathrm{T}} \cdot z \, \mathrm{d}x$ as $\varepsilon \to 0$ since z is Lipschitz continuous in $\operatorname{supp} z$ (by Proposition 2.3 and the identification (5b)) and hence the convergence is uniform. Moreover, the right-hand side converges pointwise a.e. and can be bounded by an integrable function, thus it also converges by virtue of Lebesgue's dominated-convergence theorem. The latter can be seen by

$$\big|y(x)(D \cdot \nabla z_\varepsilon + \operatorname{div} D \cdot z_\varepsilon)(x)\big| \leq (\|D\|_\infty \|\nabla z\|_\infty + \|\operatorname{div} D\|_\infty \|z\|_\infty)|y(x)|$$

considered a.e. in a subset $\Omega' \subset\subset \Omega$ which contains the supports of $\{z_\varepsilon\}$. □

Lemma 3.7. *Let $1 \leq r \leq \infty$, $D \in Y^*_{\mathrm{div}}$ and $\zeta \in W^{1,\infty}(\Omega)$. Then, from $y \in W^{1,r}_D$ follows $\zeta y \in W^{1,r}_D$ with the identity*

$$D\nabla(\zeta y)^{\mathrm{T}} = \zeta D\nabla y^{\mathrm{T}} + y D\nabla\zeta^{\mathrm{T}} \ . \tag{10}$$

Moreover, the multiplication with ζ maps $W^{1,r}_D \to W^{1,r}_D$ continuously.

Proof. It is clear that for each $z \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d)$ we have

$$\zeta(D \cdot \nabla z + \operatorname{div} D \cdot z) = \zeta \operatorname{div}(D^{\mathrm{T}} z) = \operatorname{div}(D\zeta z) - D\nabla\zeta^{\mathrm{T}} \cdot z$$

almost everywhere in Ω and since ζz can be used as a test function (see Lemma 3.6), it follows

$$\begin{aligned} -\int_\Omega \zeta y(D \cdot \nabla z + \operatorname{div} D \cdot z) \, \mathrm{d}x &= -\int_\Omega y\big(\operatorname{div}(D\zeta z) - D\nabla\zeta^{\mathrm{T}} \cdot z\big) \, \mathrm{d}x \\ &= \int_\Omega (\zeta D\nabla y^{\mathrm{T}} + y D\nabla\zeta^{\mathrm{T}}) \cdot z \, \mathrm{d}x \ . \end{aligned}$$

Moreover, the norm can be estimated by

$$\|D\nabla(\zeta y)^{\mathrm{T}}\|_r \leq \|\zeta\|_\infty \|D\nabla y^{\mathrm{T}}\|_r + \|D\|_\infty \|\nabla\zeta\|_\infty \|y\|_r \ ,$$

yielding the continuity. □

Lemma 3.8. *Let $1 \leq r < \infty$ and $D \in Y^*$. For each $y \in W^{1,r}_D(\Omega)$ with $\operatorname{supp} y \subset \Omega' \subset\subset \Omega$ there exists an $\varepsilon_0 > 0$ and a C which only depends on r, $\|D\|_{Y^*}$ and the standard mollifier G with $\operatorname{supp} G \subset B_1(0)$ such that for each $0 < \varepsilon < \varepsilon_0$ the estimate*

$$\|y_\varepsilon\|_{W^{1,r}_D} \leq C\|y\|_{W^{1,r}_D} \qquad \text{where} \qquad y_\varepsilon = y * G_\varepsilon$$

holds with G_ε being the dilated versions of G.

Proof. Choose $0 < \varepsilon_0 < \operatorname{dist}(\operatorname{supp} y, \partial\Omega)$, so all $y_\varepsilon \in \mathcal{C}_0^\infty(\Omega)$. Note that $D \in Y^*$ means that D is bounded and Lipschitz continuous in the sense that

$$|D(x) - D(\xi)| \leq \|\nabla D\|_\infty \varepsilon \quad \text{for all } \xi \in \operatorname{supp} y\ , |x - \xi| \leq \varepsilon\ , \tag{11}$$

by Proposition 2.3. It is moreover clear that $\|y_\varepsilon\|_r \leq \|G\|_1 \|y\|_r$. To estimate the approximation of the weak weighted gradient consider $D\nabla y_\varepsilon^{\mathrm{T}}$ which can, because of the smoothness of y_ε, be written as

$$\begin{aligned}(D\nabla y_\varepsilon^{\mathrm{T}})(x) &= \int_\Omega y(\xi) D(x) \nabla G_\varepsilon(x-\xi)^{\mathrm{T}} \ \mathrm{d}\xi \\ &= \int_\Omega y(\xi)\big(D(x) - D(\xi)\big)\nabla G_\varepsilon(x-\xi)^{\mathrm{T}} \ \mathrm{d}\xi \\ &\quad + \int_\Omega y(\xi) D(\xi) \nabla G_\varepsilon(x-\xi)^{\mathrm{T}} \ \mathrm{d}\xi\end{aligned} \tag{12}$$

but $D(\xi)\nabla G_\varepsilon(x-\xi)^{\mathrm{T}} = -D(\xi)\nabla G_{\varepsilon,x}(\xi)^{\mathrm{T}}$ where $G_{\varepsilon,x}(\xi) = G_\varepsilon(x-\xi)$, so, introducing a $z \in \mathbb{R}^d$, one is able to reformulate the latter term to

$$\begin{aligned}-\Big(\int_\Omega yD\nabla G_{\varepsilon,x}^{\mathrm{T}} \ \mathrm{d}\xi\Big) \cdot z &= -\int_\Omega yD \cdot \nabla(zG_{\varepsilon,x}) \ \mathrm{d}\xi \\ &= -\int_\Omega y\big(D\cdot\nabla(zG_{\varepsilon,x}) + \operatorname{div} D \cdot zG_{\varepsilon,x}\big) \ \mathrm{d}\xi \\ &\quad + \int_\Omega y \operatorname{div} D \cdot zG_{\varepsilon,x} \ \mathrm{d}\xi \\ &= \Big(\int_\Omega \big(D\nabla y^{\mathrm{T}} + y \operatorname{div} D\big)(\xi) G_\varepsilon(x-\xi) \ \mathrm{d}\xi\Big) \cdot z\end{aligned}$$

with the help of the definition of the weak weighted derivative (6). This leads to

$$\|(D\nabla y^{\mathrm{T}} + y\operatorname{div} D) * G_\varepsilon\|_r \leq \|G\|_1(\|D\nabla y^{\mathrm{T}}\|_r + \sqrt{d}\|\nabla D\|_\infty \|y\|_r)$$

so let us estimate, using (11), the other term in (12) according to

$$\begin{aligned}&\Big(\int_\Omega\Big|\int_\Omega y(\xi)\big(D(x) - D(\xi)\big)\nabla G_\varepsilon(x-\xi)^{\mathrm{T}} \ \mathrm{d}\xi\Big|^r \ \mathrm{d}x\Big)^{1/r} \\ &\qquad \leq \|\nabla D\|_\infty \Big(\int_\Omega\Big|\int_\Omega \varepsilon^{-d}|y(\xi)|\Big|\nabla G\Big(\frac{x-\xi}{\varepsilon}\Big)\Big| \ \mathrm{d}\xi\Big|^r \ \mathrm{d}x\Big)^{1/r} \\ &\qquad \leq \|\nabla D\|_\infty \|\nabla G\|_1 \|y\|_r\ .\end{aligned}$$

Altogether, one gets

$$\begin{aligned}\|D\nabla y_\varepsilon^{\mathrm{T}}\|_r &\leq (\|\nabla G\|_1 + \sqrt{d}\|G\|_1)\|\nabla D\|_\infty \|y\|_r + \|G\|_1 \|D\nabla y^{\mathrm{T}}\|_r \\ &\leq \big(\|\nabla D\|_\infty^{r'}(\|\nabla G\|_1 + \sqrt{d}\|G\|_1)^{r'} + \|G\|_1^{r'}\big)^{1/r'} \|y\|_{W_D^{1,r}}\end{aligned}$$

again, with the dual exponent r' such that $\frac{1}{r} + \frac{1}{r'} = 1$. This leads to the desired result if C is chosen as follows

$$C = \Big(\big(\|\nabla D\|_\infty^{r'}(\|\nabla G\|_1 + \sqrt{d}\|G\|_1)^{r'} + \|G\|_1^{r'}\big)^{r-1} + \|G\|_1^r\Big)^{1/r}\ . \qquad \square$$

Collecting the results, the first main result of this section can be proven.

Theorem 3.9. *Let $1 < r < \infty$ and $D \in Y^*$. Then $\mathcal{C}^\infty(\Omega)$ is dense in $W_D^{1,r}$. In particular, $W_D^{1,r}$ can be regarded as the closure of $\mathcal{C}^\infty(\Omega)$ under the norm* (7a).

Proof. To show the density, we want to apply Proposition 3.5 to the closed differential operator $\Lambda : y \mapsto D\nabla y^{\mathrm{T}}$ (see Remark 3) whose domain contains $\mathcal{C}_0^\infty(\Omega)$ and which can be interpreted in $\mathcal{D}(\Lambda) \subset L^r_{\mathrm{loc}}(\Omega) \to L^r_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$. By virtue of Lemma 3.7, Λ meets the first requirement of the proposition (the statement that $\zeta y \in \mathcal{D}(\Lambda)$ for each $\zeta \in \mathcal{C}_0^\infty(\Omega)$ implies $y \in \mathcal{D}(\Lambda)$ follows directly from the definition of the weak weighted derivative). The second is easily obtained from the product formula (10) for the weak weighted gradient, while the third is a direct consequence of Lemma 3.8. Hence, the asserted statement follows from the application of Proposition 3.5. □

Note that we used $\mathcal{C}^\infty(\Omega) \subset W_D^{1,r}$ as a short-hand notation for $\mathcal{C}^\infty(\Omega) \cap W_D^{1,r}$ being a subset of $W_D^{1,r}$, a slight abuse of notation which seems to be common in the literature.

In addition to the density of $\mathcal{C}^\infty(\Omega)$, which allows to approximate with smooth functions, it is also of interest to know about the approximation properties of $\mathcal{C}^\infty(\overline{\Omega})$. Since this set is often dense in Sobolev spaces, such density results allow to approximate with smooth functions which are moreover contained spaces with "higher" integrability, i.e. $W^{1,r}(\Omega)$. In the following, we will see that, provided that $\partial\Omega$ is Lipschitz, $\mathcal{C}^\infty(\overline{\Omega})$ is dense in each $W_D^{1,r}$ and, hence, each $W_D^{1,r}$ can be interpreted as the closure of $W^{1,r}(\Omega)$ under the norm (7).

We establish an abstract approximation result for $\mathcal{C}^\infty(\overline{\Omega})$ which is analogous to Proposition 3.5 in the first step and prove the necessary boundedness of the approximizing sequence subsequently.

Proposition 3.10. *Let $l \geq 1$, $1 < r < \infty$ and $\Lambda : \mathcal{D}(\Lambda) \subset L^r(\Omega) \to L^r(\Omega, \mathbb{R}^l)$ with $\mathcal{C}^\infty(\overline{\Omega}) \subset \mathcal{D}(\Lambda)$ a closed linear operator with the property that for each $y \in \mathcal{D}(\Lambda)$, there exists a sequence $\{y_\delta\} \subset \mathcal{C}^\infty(\overline{\Omega})$ for $0 < \delta < \delta_0$ such that $y_\delta \to y$ in $L^r(\Omega)$ as $\delta \to 0$ and $\|\Lambda y_\delta\|_r \leq C$ with some $C > 0$.*

Then, for each $\varepsilon > 0$ and $y \in \mathcal{D}(\Lambda)$, there exists a $\bar{y} \in \mathcal{C}^\infty(\overline{\Omega})$, which is a finite linear combination of some y_δ, such that the approximation property (9) *holds true.*

Proof. Choose a $y \in \mathcal{D}(\Lambda)$ and δ_0 according to the assumptions. Then, the sequence $\{\Lambda y_{\delta_k}\}$ remains bounded in $L^r(\Omega, \mathbb{R}^l)$ as $\delta_k \to 0$, hence there exists a weakly convergent subsequence (still labeled with k) such that $y_{\delta_k} \rightharpoonup \theta$. Again, by the theorem of Banach-Saks,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \Lambda y_{\delta_k} = \theta \quad , \quad \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} y_{\delta_k} = y$$

for a subsequence, where the latter follows from $y_{\delta_k} \to y$ in $L^r(\Omega)$ as $k \to \infty$. Moreover, $\theta = \Lambda y$ since the operator is closed. Choosing m large enough and exploiting the linearity of Λ finally gives a $\bar{y}$ according to

$$\bar{y} = \frac{1}{m} \sum_{k=1}^{m} y_{\delta_k}$$

such that (9) holds true. From the construction then follows $\bar{y} \in \mathcal{C}^\infty(\overline{\Omega})$. □

Lemma 3.11. *Let Ω be a bounded Lipschitz domain, $1 < r < \infty$ and $D \in Y^*$. Then, there exists a family of linear operators $\mathcal{M}_\varepsilon$ for $0 < \varepsilon < \tilde{\varepsilon}$ which only depends on Ω, with the following property:*

For each $y \in W_D^{1,r}$, the functions $y_\varepsilon = \mathcal{M}_\varepsilon y$ are in $\mathcal{C}^\infty(\overline{\Omega})$, $y_\varepsilon \to y$ in $L^r(\Omega)$ as $\varepsilon \to 0$ and

$$\|D\nabla y_\varepsilon^{\mathrm{T}}\|_r \leq C\|y\|_{W_D^{1,r}}$$

for some $C > 0$ which is independent of y.

Proof. The proof is very technical and similar to the proof of Lemma 3.8 and therefore presented in Appendix A.1. □

Theorem 3.12. *Let Ω be a bounded Lipschitz domain, $1 < r < \infty$ and $D \in Y^*$. Then $\mathcal{C}^\infty(\overline{\Omega})$ is dense in $W^{1,r}_D$.*

Proof. The result follows in analogy to the proof of Theorem 3.9. The essential step here is the combination of Lemma 3.11 with the approximation result of Proposition 3.10. □

3.3 Calculus rules

The approximation results in Theorems 3.9 and 3.12 can be applied to establish the usual calculus rules for the spaces $W^{1,r}_D$, namely the product rule (which was already established in Lemma 3.7 but can now be extended to $W^{1,\infty}_D$) as well as the chain rule.

Proposition 3.13. *Let $1 < r < \infty$ and $D \in Y^*$. Then, for each $y \in W^{1,r}_D$ and $\zeta \in W^{1,\infty}_D$ follows $\zeta y \in W^{1,r}_D$ with the identity* (10) *remaining valid. Moreover, $y \mapsto \zeta y$ is a continuous operation mapping $W^{1,r}_D$ into itself.*

Proof. The idea is to choose $y \in \mathcal{C}^\infty(\Omega) \cap W^{1,r}_D$ and a $\Omega' \subset\subset \Omega$ such that $y \in W^{1,\infty}(\Omega')$ in order to apply Lemma 3.7 with y and ζ interchanged. This yields the existence of $D\nabla(\zeta y)^{\mathrm{T}}$ and the asserted identity in $L^r_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$.

Next, we find the continuity estimate

$$\|D\nabla(\zeta y)^{\mathrm{T}}\|_{r,\Omega'} \leq \|D\nabla\zeta^{\mathrm{T}}\|_\infty \|y\|_r + \|\zeta\|_\infty \|D\nabla y^{\mathrm{T}}\|_r$$

which is uniform for all $\Omega' \subset\subset \Omega$, allowing to extend the above to $L^r(\Omega, \mathbb{R}^d)$. Finally,

$$\|\zeta y\|_{W^{1,r}_D} \leq C\|y\|_{W^{1,r}_D}$$

with a C which only depends on $\|\zeta\|_{W^{1,\infty}_D}$. Hence, $y \mapsto \zeta y$ can be extended continuously to the whole space by density (Theorem 3.9). □

Proposition 3.14. *Let $1 < r < \infty$, $D \in Y^*$ and $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function with bounded derivative. Then, from $y \in W^{1,r}_D$ follows $\varphi \circ y \in W^{1,r}_D$ and*

$$D\nabla(\varphi \circ y)^{\mathrm{T}} = (\varphi' \circ y)(D\nabla y^{\mathrm{T}}) \ .$$

The statement is still valid for $\varphi(s) = \max\{0, s\}$, $\varphi(s) = \min\{0, s\}$ as well as $\varphi(s) = |s|$ (with $\varphi'(0) = 0$). In particular, $y(x) = 0$ implies $(D\nabla y^{\mathrm{T}})(x) = 0$ almost everywhere.

Moreover, the superposition $y \mapsto \varphi \circ y$ is continuous mapping $W^{1,r}_D \to W^{1,r}_D$.

Proof. The proof is an adaptation of the chain rule for classical Sobolev functions to the weak weighted and directional derivatives (confer [13], for example). Let $y \in \mathcal{C}^\infty(\Omega)$ so that $\varphi \circ y \in \mathcal{C}^1(\Omega)$ by the usual chain rule. The asserted identities in this case follow easily from $\nabla(\varphi \circ y) = (\varphi' \circ y)\nabla y$.

Now approximate $y \in W^{1,r}_D$ with a sequence of $\{y_\varepsilon\} \subset \mathcal{C}^\infty(\Omega)$ with $\varepsilon > 0$. In particular, one can find a subsequence such that we have pointwise a.e. convergence in Ω in addition to $y_\varepsilon \to y$ in $L^r(\Omega)$ and $D\nabla y_\varepsilon^{\mathrm{T}} \to D\nabla y^{\mathrm{T}}$ in $L^r(\Omega, \mathbb{R}^d)$. It is well-known that $\varphi \circ y_\varepsilon \to \varphi \circ y$ in $L^r(\Omega)$. Observe that $\varphi' \circ y_\varepsilon$ converges pointwise a.e. and is bounded almost everywhere. Consequently, Lebesgue's dominated-convergence theorem gives

$$(\varphi' \circ y_\varepsilon) D\nabla y_\varepsilon^{\mathrm{T}} \to (\varphi' \circ y) D\nabla y^{\mathrm{T}}$$

in $L^r(\Omega, \mathrm{I\!R}^d)$ leading to the asserted identity for $D\nabla(\varphi \circ y)^{\mathrm{T}}$ by the closedness of the differential operators (see Remark 3). This proves the chain rule in $W_D^{1,r}$.

For the remainder, it is sufficient to consider $\varphi(s) = \max\{0, s\}$ since $\min\{0, s\} = -\max\{0, -s\}$ and $|s| = \max\{0, s\} - \min\{0, s\}$. The usual way is to approximate φ by

$$\varphi_\varepsilon(s) = \chi_{]0,\infty[}(s)\sqrt{s^2 + \varepsilon^2} - \varepsilon \quad , \quad \varphi_\varepsilon'(s) = \chi_{]0,\infty[}(s)\frac{s}{\sqrt{s^2 + \varepsilon^2}}$$

and to see that $\varphi_\varepsilon \circ y \to \varphi \circ y$ as well as $(\varphi_\varepsilon' \circ y)D\nabla y^{\mathrm{T}} \to (\varphi' \circ y)D\nabla y^{\mathrm{T}}$, in $L^r(\Omega, \mathrm{I\!R}^d)$, again by Lebesgue's dominated-convergence theorem. Note that from $y = \max\{0, y\} + \min\{0, y\}$ then follows that $D\nabla y^{\mathrm{T}} = 0$ almost everywhere where $y = 0$.

The continuity of $y \mapsto \varphi \circ y$ can be proven as follows. Suppose $y_l \to y$ in $W_D^{1,r}$. First, $\varphi \circ y_l \to \varphi \circ y$ in $L^r(\Omega)$ by classical results about superposition operators (see [4], for example). One can moreover assume that $D\nabla y_l^{\mathrm{T}} \to D\nabla y^{\mathrm{T}}$ pointwise a.e. and in $L^r(\Omega, \mathrm{I\!R}^d)$. Consequently, we have

$$(\varphi' \circ y_l)(D\nabla y_l^{\mathrm{T}}) \to (\varphi' \circ y)(D\nabla y^{\mathrm{T}})$$

almost everywhere in Ω. This follows from the pointwise a.e. convergence of $\varphi' \circ y$ in the case where φ' is continuous. In the case where $\varphi(s) = \max\{0, s\}$, φ' is discontinuous in 0, so one also has to investigate the behavior on the set $\{y = 0\}$. The pointwise a.e. convergence is then derived from $|D\nabla y_l^{\mathrm{T}}| \to |D\nabla y^{\mathrm{T}}| = 0$ a.e. where $y = 0$, since $|(\varphi' \circ y_l)(D\nabla y_l^{\mathrm{T}})| \leq |D\nabla y_l^{\mathrm{T}}| \to 0$ there. Similar arguments show that this is also true for $\varphi(s) = \min\{0, s\}$ as well as $\varphi(s) = |s|$.

Returning to general φ, by Vitali's convergence theorem (cf. [2], for example), we have

$$\sup_{l \in \mathrm{I\!N}} \int_{\Omega'} |D\nabla y_l^{\mathrm{T}}|^r \,\mathrm{d}x \to 0 \quad , \quad \text{if } |\Omega'| \to 0$$

and for each $\varepsilon > 0$ the existence of an Ω_ε such that

$$\sup_{l \in \mathrm{I\!N}} \int_{\Omega_\varepsilon} |D\nabla y_l^{\mathrm{T}}|^r \,\mathrm{d}x \leq \varepsilon \ .$$

Since φ' is bounded by, say, $C > 0$, it follows that

$$\sup_{l \in \mathrm{I\!N}} \int_{\Omega'} |(\varphi' \circ y_l)(D\nabla y_l^{\mathrm{T}})|^r \,\mathrm{d}x \to 0 \quad , \quad \text{if } |\Omega'| \to 0$$

as well as the fact that for each $\varepsilon > 0$ we have for $\delta = C^{-r}\varepsilon$ that

$$\sup_{l \in \mathrm{I\!N}} \int_{\Omega_\delta} |(\varphi' \circ y_l)(D\nabla y_l^{\mathrm{T}})|^r \,\mathrm{d}x \leq \varepsilon$$

with the above Ω_δ. This, in turn implies $(\varphi' \circ y_l)(D\nabla y_l^{\mathrm{T}}) \to (\varphi' \circ y)(D\nabla y^{\mathrm{T}})$ in $L^r(\Omega, \mathrm{I\!R}^d)$, again by using Vitali's convergence theorem. Consequently, $\varphi \circ y_l \to \varphi \circ y$ in $W_D^{1,r}$. □

4 Scalar weights and directional derivatives

As it was already mentioned in Remark 1, there are some special subspaces of coefficients $D \in Y^*$ for which more about the structure of the weak weighted derivative as well as the associated $W_D^{1,r}$ can be said. In the following, we investigate the case $D \in Y_0^* = Y_{\mathrm{div},0}^*$ (confer Remark 1 for the definition) which corresponds to scalar weighting functions. Furthermore, the weak weighted gradient for coefficients $D \in Y_1^*$ are examined, corresponding to weak directional derivatives. Besides the connections to the classical weak derivative, will see that for scalar coefficients, a stronger closedness result holds, assuming only weak*-convergence for $D = wI$ in Y^*_{div} and no additional convergence for ∇w, see Proposition 3.3. Furthermore, if $D \in Y^*_1$, one can locally associate coordinate transforms such that the y for which $D\nabla y^{\mathrm{T}}$ exists can be expressed locally, under the coordinate transform, by the functions which are weakly differentiable with respect to one distinguished space variable.

4.1 Scalar weights

Suppose that $D \in Y^*_0 = Y^*_{\mathrm{div},0}$ according to (4a) meaning that $D = wI$ for a function w, which has to be in $W^{1,\infty}(\Omega)$, see Remark 1. For weak weighted gradients associated with $D \in Y^*_0$ we write $D\nabla y^{\mathrm{T}} = w\nabla y^{\mathrm{T}}$ and, as long as it creates no confusion, we also denote by $w \in Y^*_0$ a function $w \in W^{1,\infty}(\Omega)$ for which $wI \in Y^*_0$. First note that Definition 3.1 can equivalently be expressed by the following.

Remark 5. Let $w \in Y^*_0$ and $y \in L^1_{\mathrm{loc}}(\Omega)$, $v \in L^1_{\mathrm{loc}}(\Omega, \mathrm{IR}^d)$. Then, $v = w\nabla y^{\mathrm{T}}$ according to Definition 3.1 if and only if

$$\int_\Omega v \cdot z \, \mathrm{d}x = -\int_\Omega y(w \operatorname{div} z + \nabla w \cdot z) \, \mathrm{d}x \quad \text{for all } z \in \mathcal{C}^\infty_0(\Omega, \mathrm{IR}^d) \, . \tag{13}$$

Subsequently, we use the more concise version (13) instead of (6). The first result addresses the pointwise properties of $w\nabla y^{\mathrm{T}}$.

Proposition 4.1. *Let $w \in Y^*_0$ and y be locally integrable on Ω such that the weak weighted gradient $w\nabla y^{\mathrm{T}} \in L^1_{\mathrm{loc}}(\Omega, \mathrm{IR}^d)$ exists. Then*

1. *∇y exists where $w \neq 0$ and there we have $\nabla y^{\mathrm{T}} = w^{-1}(w\nabla y^{\mathrm{T}})$ almost everywhere and*

2. *$w\nabla y^{\mathrm{T}} = 0$ almost everywhere where $w = 0$.*

Proof. We know from Proposition 2.4 that w is continuous in Ω, hence the subset $\Omega_+ = \{x \in \Omega \mid w(x) \neq 0\}$ is relatively open. Moreover, for each $z \in \mathcal{C}^\infty_0(\Omega_+, \mathrm{IR}^d)$ there exists a $\delta > 0$ such that $|w(x)| \geq \delta$ for each $x \in \operatorname{supp} z$. Hence $\bar{z} = w^{-1} z$ is still in $W^{1,\infty}(\Omega_+, \mathrm{IR}^d)$ with compact support in Ω_+ and can be used as a test function (see Lemma 3.6). This yields

$$-\int_{\Omega_+} y \operatorname{div} z \, \mathrm{d}x = -\int_{\Omega_+} y(w \operatorname{div} \bar{z} + \nabla w \cdot \bar{z}) \, \mathrm{d}x = \int_{\Omega_+} \frac{w\nabla y^{\mathrm{T}}}{w} \cdot z \, \mathrm{d}x$$

so $\nabla y^{\mathrm{T}} = w^{-1}(w\nabla y^{\mathrm{T}})$ exists in the weak sense (13) in Ω_+. Especially, the asserted identity holds for almost every $x \in \Omega_+$.

Now fix an $1 \leq i \leq d$ and a compact $\Omega' \subset \Omega_0 = \{x \in \Omega \mid w(x) = 0\}$ and let $z_\varepsilon = \chi_{\Omega'} * G_\varepsilon$ where G denotes again the standard mollifier with $\operatorname{supp} G \subset B_1(0)$ and G_ε its dilated versions. We choose $\varepsilon > 0$ sufficiently small, i.e. such that $z_\varepsilon \in \mathcal{C}^\infty_0(\Omega)$. It is clear that $|\frac{\partial z_\varepsilon}{\partial x_i}(x)| \leq \|\chi_{\Omega'}\|_\infty \|\frac{\partial G_\varepsilon}{\partial x_i}\|_1 \leq \varepsilon^{-1} C_1$ with a suitable $C_1 > 0$. Since $w \in Y^*_0$, we moreover know that $\max\{\|w\|_\infty, \|\nabla w\|_\infty\} \leq C_2$ for some $C_2 > 0$. Hence, if $x^* \in \Omega'$, then for each $x \in \Omega$ with $|x - x^*| \leq \varepsilon$ follows $|w(x)| \leq C_2 C_{\Omega'} \varepsilon$ (with a $C_{\Omega'} > 0$ given by Proposition 2.3). Furthermore, $w = 0$ implies $\nabla w = 0$ almost everywhere on the respective subset. Now consider the estimate

$$\Big|\int_\Omega y\big(w \frac{\partial z_\varepsilon}{\partial x_i} + \frac{\partial w}{\partial x_i} z_\varepsilon\big) \, \mathrm{d}x\Big| \leq C_2(C_1 C_{\Omega'} + 1) \int_{\operatorname{supp} z_\varepsilon \setminus \Omega'} |y| \, \mathrm{d}x$$

which tends to zero as $\varepsilon \to 0$ since from the compactness of Ω' follows $|\mathrm{supp}\, z_\varepsilon \backslash \Omega'| \to 0$. The weak weighted derivative exists, hence we have, by Lebesgue's dominated-convergence theorem that $\int_{\Omega'} (w\nabla y^{\mathrm{T}})_i \, \mathrm{d}x = 0$ for all compact $\Omega' \subset \Omega_0$ and $1 \leq i \leq d$. This yields $w\nabla y^{\mathrm{T}} = 0$ almost everywhere in Ω_0, the desired result. □

The following results states the closedness of the weak weighted derivative considered in $L^r(\Omega) \times Y_0^* \to L^r(\Omega, \mathbb{R}^d)$ with weak and weak* convergence. It will turn out that the sequence of weak weighted gradients converges weakly to the weak weighted gradient of the limit only on the set where the limit weight does not vanish. The main argument is utilizing the compact embedding $W^{1,r}(\Omega) \hookrightarrow L^r(\Omega)$ which is done in the following preparatory lemma.

Lemma 4.2. *Let Ω fulfill the cone condition, $\{w_k\} \subset Y_0^*$, $w \in Y_0^*$ be given with $w_k \overset{*}{\rightharpoonup} w$. Moreover, let $1 < r < \infty$, $\{y_k\} \subset L^r(\Omega)$, $y \in L^r(\Omega)$ with $y_k \rightharpoonup y$ as well as $\{w_k\nabla y_k\} \subset L^r(\Omega, \mathbb{R}^d)$, $\theta \in L^r(\Omega, \mathbb{R}^d)$ with $w_k \nabla y_k \rightharpoonup \theta$. For $\varepsilon > 0$ denote by $\Omega_\varepsilon = \{x \in \Omega \mid |w(x)| > \varepsilon\}$.*

Then, for each $\varepsilon > 0$, we have $y_k \to y$ in $L^r(\Omega_\varepsilon)$ and $\theta = w\nabla y$ in Ω_ε.

Proof. First note that $w_k \to w$ in $\mathcal{C}(\Omega)$ (a consequence of Proposition 2.4 and (5a)), so for sufficiently large k we have $|w_k| > \varepsilon/2$ on Ω_ε and thus ∇y_k as well as ∇y exist there according to Proposition 4.1. Moreover, we know that $\nabla y_k^{\mathrm{T}} = w_k^{-1}(w_k \nabla y_k^{\mathrm{T}})$ on Ω_ε, so

$$\|\nabla y_k\|_r \leq 2\varepsilon^{-1} \|w_k \nabla y_k^{\mathrm{T}}\|_r \leq C$$

by assumption. Hence, we can extract a subsequence, also denoted by $\{\nabla y_k\}$, which converges weakly in $L^r(\Omega_\varepsilon, \mathbb{R}^d)$ with limit ∇y since taking the gradient is closed. Now $W^{1,r}(\Omega_\varepsilon)$ is compactly embedded in $L^r(\Omega_\varepsilon)$ (since Ω_ε also satisfies the cone condition, see [1]) so actually $y_k \to y$ in $L^r(\Omega_\varepsilon)$. Finally, one can easily deduce in analogy to Proposition 3.3 that $w_k \operatorname{div} z + \nabla w_k \cdot z \overset{*}{\rightharpoonup} w \operatorname{div} z + \nabla w \cdot z$ for $z \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d)$ with $\mathrm{supp}\, z \subset\subset \Omega_\varepsilon$. We have

$$\int_\Omega y_k (w_k \operatorname{div} z + \nabla w_k \cdot z) \, \mathrm{d}x = -\int_\Omega w_k \nabla y_k^{\mathrm{T}} \cdot z \, \mathrm{d}x$$

for each $z \in \mathcal{C}_0^\infty(\Omega_\varepsilon, \mathbb{R}^d)$, so taking limits yields

$$\int_\Omega y (w \operatorname{div} z + \nabla w \cdot z) \, \mathrm{d}x = -\int_\Omega \theta \cdot z \, \mathrm{d}x$$

and thus, $\theta = w\nabla y^{\mathrm{T}}$ in Ω_ε. □

Now, the weak closedness result follows immediately.

Proposition 4.3. *Suppose that Ω satisfies the cone condition, $1 < r < \infty$ and let the sequences $\{y_k\} \subset L^r(\Omega)$ and $\{w_k\} \subset Y_0^*$ with limits $y_k \rightharpoonup y$ and $w_k \overset{*}{\rightharpoonup} w$ in the respective spaces be given. If $w_k \nabla y_k^{\mathrm{T}} \rightharpoonup \theta$ in $L^r(\Omega, \mathbb{R}^d)$, then $w\nabla y^{\mathrm{T}} = \chi_{\{w \neq 0\}} \theta$.*

Proof. Choose an arbitrary $z \in \mathcal{C}_0^\infty(\Omega, \mathbb{R}^d)$ and $C_1 > 0$ such that C_1 is an estimate of the Lipschitz constant for all w_k on $\mathrm{supp}\, z$ (again, see Proposition 2.3). For each $0 < \varepsilon < \mathrm{dist}(\mathrm{supp}\, z, \partial\Omega)$, we define $\Omega_\varepsilon = \{|w| > \varepsilon\}$ and observe that for the dilated standard mollifier G_ε the function $z_\varepsilon^1 = (G_{\varepsilon/(2C_1)} * \chi_{\Omega_\varepsilon}) z$ belongs to $\mathcal{C}_0^\infty(\Omega_{\varepsilon/2}, \mathbb{R}^d)$ since

$$|w(\xi)| \geq |w(x)| - |w(\xi) - w(x)| > \varepsilon - C_1|\xi - x| > \varepsilon/2$$

whenever $x \in \Omega_\varepsilon$ and $\xi \in B_{\varepsilon/(2C_1)}(x)$. As we have proven in Lemma 4.2, $y_k \to y$ in $L^r(\Omega_{\varepsilon/2})$ which implies

$$\int_\Omega y (w \operatorname{div} z_\varepsilon^1 + \nabla w \cdot z_\varepsilon^1) \, \mathrm{d}x = \lim_{k\to\infty} \int_\Omega y_k (w_k \operatorname{div} z_\varepsilon^1 + \nabla w_k \cdot z_\varepsilon^1) \, \mathrm{d}x \ .$$

Setting $z_\varepsilon^2 = z - z_\varepsilon^1$, we can moreover observe that $\operatorname{supp} z_\varepsilon^2 \subset \{|w| \leq 3\varepsilon/2\}$ since

$$|w(\xi)| \geq |w(x)| - |w(\xi) - w(x)| > 3\varepsilon/2 - C_1|\xi - x| \geq \varepsilon$$

whenever $|w(x)| > 3\varepsilon/2$ and $|\xi - x| \leq \varepsilon/(2C_1)$. Hence, we deduce

$$\Big|\int_\Omega y(w \operatorname{div} z_\varepsilon^2 + \nabla w \cdot z_\varepsilon^2)\ \mathrm{d}x\Big| \leq \operatorname*{ess\,sup}_{\{|w|\leq 3\varepsilon/2\}} |w \operatorname{div} z_\varepsilon^2 + \nabla w \cdot z_\varepsilon^2| \int\limits_{\{0<|w|\leq 3\varepsilon/2\}} |y|\ \mathrm{d}x$$

and estimate

$$\begin{aligned} |w \operatorname{div} z_\varepsilon^2| &\leq 3\varepsilon/2 \big\|(1 - G_{\varepsilon/(2C_1)} * \chi_{\Omega_\varepsilon}) \operatorname{div} z - (\nabla G_{\varepsilon/(2C_1)} * \chi_{\Omega_\varepsilon}) \cdot z\big\|_\infty \\ &\leq 3\varepsilon/2\big(\|\operatorname{div} z\|_\infty + 2C_1/\varepsilon \|\nabla G\|_1 \|z\|_\infty\big) \leq C_2 \end{aligned}$$

as well as

$$\|\nabla w \cdot z_\varepsilon^2\|_\infty \leq \|\nabla w\|_\infty \|z\|_\infty \leq C_3$$

with suitable $C_2, C_3 > 0$ which can be chosen to be independent of ε. Additionally, $\int_{\{0<|w|\leq 3\varepsilon/2\}} |y|\ \mathrm{d}x$ tends to zero as $\varepsilon \to 0$, which allows to choose $\varepsilon > 0$ such that

$$\Big|\int_\Omega y(w \operatorname{div} z_\varepsilon^2 + \nabla w \cdot z_\varepsilon^2)\ \mathrm{d}x\Big| < \delta$$

with $\delta > 0$ arbitrarily small. Together with the above, we have

$$\begin{aligned} \int_\Omega y(w \operatorname{div} z + \nabla w \cdot z)\ \mathrm{d}x &= \lim_{k\to\infty} \int_\Omega y_k(w_k \operatorname{div} z_\varepsilon^1 + \nabla w_k \cdot z_\varepsilon^1)\ \mathrm{d}x \\ &\quad + \int_\Omega y(w \operatorname{div} z_\varepsilon^2 + \nabla w \cdot z_\varepsilon^2)\ \mathrm{d}x \\ &= -\lim_{k\to\infty} \int_\Omega w_k \nabla y_k^{\mathrm{T}} \cdot z_\varepsilon^1\ \mathrm{d}x + \int_\Omega y(w \operatorname{div} z_\varepsilon^2 + \nabla w \cdot z_\varepsilon^2)\ \mathrm{d}x \end{aligned}$$

so we can achieve

$$\Big|\int_\Omega y(w \operatorname{div} z + \nabla w \cdot z)\ \mathrm{d}x + \int_\Omega \theta \cdot z_\varepsilon^1\ \mathrm{d}x\Big| < \delta$$

for each $\delta > 0$. Finally, letting $\delta \to 0$ and consequently $\varepsilon \to 0$ yields

$$\int_\Omega y(w \operatorname{div} z + \nabla w \cdot z)\ \mathrm{d}x = -\int_\Omega \chi_{\{w\neq 0\}} \theta \cdot z\ \mathrm{d}x$$

which implies the assertion. □

4.2 Directional derivatives

A different class of weights which can be examined further are described by $Y^*_{\mathrm{div},1}$ and Y_1^*, once again, see Remark 1. For $D = \mathbf{1} \otimes q$ and smooth functions y, the weighted gradient $D\nabla y^{\mathrm{T}}$ consists of d copies of the directional derivative $q \cdot \nabla y$. On the other hand, using pointwise singular value decomposition of a given matrix field D, one is able to obtain vector fields $q_1, \ldots, q_d$ as well as $p_1, \ldots, p_d$ such that

$$(D\nabla y^{\mathrm{T}})(x) = \sum_{i=1}^{d} \frac{1}{d}\big(p_i(x) \otimes \mathbf{1}\big)\big(\mathbf{1} \otimes q_i(x)\big)\nabla y(x)^{\mathrm{T}} \tag{14}$$

for example (it is not necessary to choose $\mathbf{1}$ as the factors for the outer products, two non-orthogonal vectors in $\mathbb{R}^d$ will also work). From (14) we see that $D\nabla y^{\mathrm{T}}$ can always be decomposed into terms involving directional derivatives.

In the following, the weak analogon to the directional derivatives is examined, i.e. $D\nabla y^{\mathrm{T}}$ in the sense of (6) such that $D = \mathbf{1} \otimes q$ with $q \in W^{1,\infty}(\Omega, \mathbb{R}^d)$ or, equivalently, $D \in Y_1^*$. Again, first remark that there is an equivalent scalar variational definition.

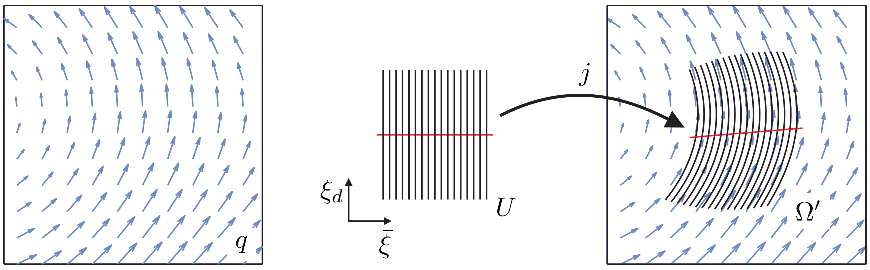

Fig. 1: Illustration of a local coordinate transformation j associated with a vector field q. On the left-hand side, a sample vector field q is depicted, while in the middle and the right, the coordinate transform j defined on V is illustrated. Note that the mapping takes lines along ξ_d to integral lines with respect to q.

Remark 6. Let $D = 1 \otimes q \in Y^*_{\mathrm{div},1}$. Then, for a locally integrable u, we have $u = D\nabla y^{\mathrm{T}}$ in the weak sense if and only if each $u_i = v$ with v being locally integrable and satisfying

$$\int_\Omega vz \,\mathrm{d}x = -\int_\Omega y(z \operatorname{div} q + \nabla z \cdot q) \,\mathrm{d}x \qquad \text{for all } z \in \mathcal{C}_0^\infty(\Omega) \ . \tag{15}$$

This follows from plugging $1 \otimes q$ into the variational equation (6) and noting that it is equivalent to test against all $z \in \mathcal{C}_0^\infty(\Omega)$ instead of all $\sum_{i=1}^d z_i$ with $z \in \mathcal{C}_0^\infty(\Omega, \mathrm{IR}^d)$.

As an abbreviation, will denote by $\partial_q y = v$ when referring to the above definition.

The main result about weak directional derivatives will be that we can express $D\nabla y^{\mathrm{T}}$ locally as a function being weakly differentiable in one direction in the classical sense under a Lipschitz coordinate transformation. For that, we need three preparatory lemmas whose proofs can be found in the appendix. The first two deal with the integral curves associated with a vector field q.

Lemma 4.4. *Let $q \in W^{1,\infty}(\Omega, \mathrm{IR}^d)$ and a point $x \in \Omega$ be given such that $q(x) \neq 0$. Then there exists an open neighborhood U of 0 in IR^d and a Lipschitz mapping j on $U \to \Omega' \subset \Omega$ such that Ω' is an open neighborhood of x with*

$$q(x) \cdot q(x') \geq \tfrac{1}{2}|q(x)||q(x')| \quad \text{for all } x' \in \Omega' \ , \tag{16}$$

and j has a Lipschitz-continuous inverse $j^{-1} : \Omega' \to U$. Moreover, j is partially differentiable in U with respect to the d-th component (denoted by $\xi = (\bar{\xi}, \xi_d)$) with derivative

$$\frac{\partial j}{\partial \xi_d}(\bar{\xi}, \xi_d) = q\big(j(\bar{\xi}, \xi_d)\big) \ . \tag{17}$$

In Figure 1, an example for the j constructed in Lemma 4.4 is depicted.

Lemma 4.5. *Let the situation of Lemma 4.4 be given. Then, with j according to Lemma 4.4, the function $J(\bar{\xi}, \xi_d) = \det \nabla j(\bar{\xi}, \xi_d)$ fulfills*

$$\frac{\partial J}{\partial \xi_d}(\bar{\xi}, \xi_d) = \operatorname{div} q\big(j(\bar{\xi}, \xi_d)\big) J(\bar{\xi}, \xi_d) \quad , \quad J(\bar{\xi}, 0) = \frac{q(x)}{|q(x)|} \cdot q\big(j(\bar{\xi}, 0)\big) \tag{18}$$

for almost every $(\bar{\xi}, \xi_d) \in U$. In particular, J is given a.e. by

$$J(\bar{\xi}, \xi_d) = \frac{q(x)}{|q(x)|} \cdot q\big(j(\bar{\xi}, 0)\big) \mathrm{e}^{\int_0^{\xi_d} \operatorname{div} q\left(j(\bar{\xi}, s)\right) \,\mathrm{d}s} \tag{19}$$

The third lemma notes which class of functions can be used alternatively to test weak derivatives with respect to the d-th direction.

Lemma 4.6. *In the situation of Lemmas 4.4 and 4.5, a $y \in L^1_{\mathrm{loc}}(U)$ possesses a weak derivative $\frac{\partial y}{\partial \xi_d} \in L^1_{\mathrm{loc}}(U)$ if and only if*

$$\int_U y \frac{\partial \tilde{z}}{\partial \xi_d} \,\mathrm{d}\xi = -\int_U \frac{\partial y}{\partial \xi_d} \tilde{z} \,\mathrm{d}\xi \tag{20}$$

for each $\tilde{z}$ given by $\tilde{z}(\xi) = J(\xi) z\big(j(\xi)\big)$ where $z \in \mathcal{C}^\infty_0(j(U))$.

Now, we are able to give a characterization of $D\nabla y^{\mathrm{T}}$ with $D \in Y_1^*$.

Proposition 4.7. *Let $D = \mathbf{1} \otimes q \in Y_1^*$ be given. Then there exists an open locally finite covering Ω_k, $k = 1, 2, \ldots$ of $\Omega_+ = \{x \in \Omega \mid q(x) \neq 0\}$, Lipeomorphisms $j_k : U_k \to \Omega_k$ (with associated $J_k = \det \nabla j_k$) and a partition of unity ζ_k subordinate to this covering such that the following characterization holds:*

A locally integrable y possesses a weak weighted derivative $D\nabla y^{\mathrm{T}}$ if and only if each $\bar{y}_k = y \circ j_k$ has a weak derivative with respect to ξ_d in U_k with

$$\sum_{\{k' \mid \Omega_{k'} \cap \Omega' \neq \emptyset\}} \int_{U_{k'}} J_{k'}(\zeta_{k'} \circ j_{k'}) \Big| \frac{\partial \bar{y}_{k'}}{\partial \xi_d} \Big| \,\mathrm{d}\xi < \infty \tag{21}$$

for each $\Omega' \subset\subset \Omega$. In the case $D\nabla y^{\mathrm{T}}$ exists, the identity

$$(D\nabla y^{\mathrm{T}})_i = \partial_q y = \sum_{k=1}^{\infty} \zeta_k \Big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1} \Big) \tag{22}$$

holds almost everywhere in Ω for each $1 \leq i \leq d$.

In particular, $D\nabla y^{\mathrm{T}} = 0$ almost everywhere where $q = 0$.

Proof. Again, note that $D = \mathbf{1} \otimes q$ with a $q \in W^{1,\infty}(\Omega, \mathbb{R}^d)$, see Remark 1. We first construct the open covering. Set

$$\Omega'_k = \{x \in \Omega \mid |q(x)| \geq \tfrac{1}{k} \wedge \operatorname{dist}(x, \partial\Omega) \geq \tfrac{1}{k}\} \quad \text{for } k = 1, 2, \ldots$$

which is family of compact sets in Ω. For $k < 1$, let $\Omega'_k = \emptyset$. Now, cover the compact sets $\Omega'_k \setminus \operatorname{int}(\Omega'_{k-1})$ with finitely many open $\Omega'_{k,l}$ associated with the anchor points $x_{k,l} \in \Omega'_k$ and Lipeomorphisms $j_{k,l} : U_{k,l} \to \Omega'_{k,l}$ according to Lemma 4.4. Note that due to the construction, each $\Omega'_{k,l}$ has to be contained in $\Omega_+ = \{x \in \Omega \mid q(x) \neq 0\}$. Without loss of generality one can also assume that

$$\Omega'_{k,l} \subset \operatorname{int}(\Omega'_{k+1}) \setminus \Omega'_{k-2} \ .$$

for each k and l. This covering has to be locally finite: Each $x \in \Omega_+$ has to be contained in some open $\Omega'_{k,l}$ which has empty intersection with each $\Omega'_{k',l'}$ for $k' > k + 2$, leaving only finitely many possibilities for non-empty intersection. See Figure 2 for an illustration of this construction.

Now choose for each k a smooth partition of unity $\eta_{k,l}$ of $\Omega'_k \setminus \operatorname{int}(\Omega'_{k-1})$ subordinate to $\Omega'_{k,l}$. Summing all up $\eta = \sum_{k,l} \eta_{k,l}$ yields a function which is finite on Ω_+ and not less than 1. Thus one can set $\zeta_{k,l} = \eta_{k,l}/\eta$ and gets a partition of unity with $\zeta_{k,l} \in \mathcal{C}^\infty_0(\Omega_+)$. By renumbering the sets $U_{k,l}$, $\Omega'_{k,l}$ as well as the mappings $j_{k,l}$ and functions $\zeta_{k,l}$, we can drop the index l and change the notation to U_k, Ω_k, j_k and ζ_k, respectively.

Let y be locally integrable with weak weighted derivative $D\nabla y^{\mathrm{T}}$ which corresponds to the existence of $\partial_q y$, see Remark 6. With $\bar{y}_k = y \circ j_k$, the function y

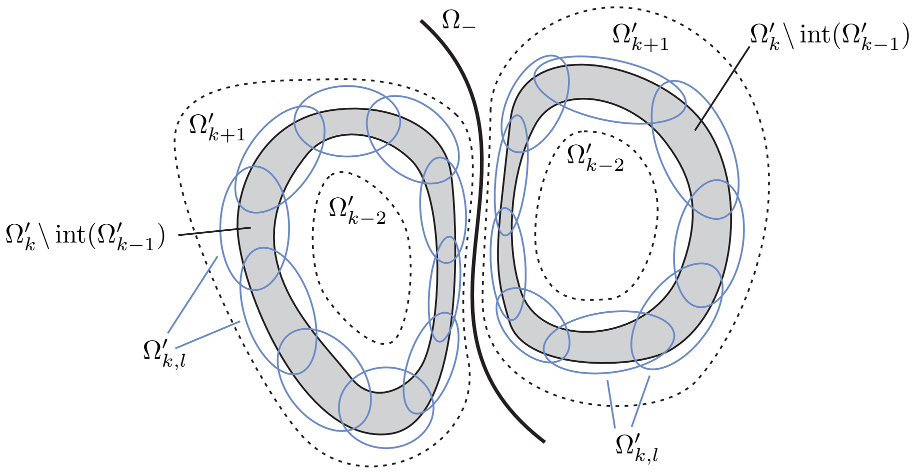

Fig. 2: Illustration of the covering constructed in the proof of Proposition 4.7. For a fixed k, the nested sets $\Omega'_{k-2}, \dots, \Omega'_{k+1}$ are depicted as well as the open $\Omega'_{k,l}$ (indicated blue) which cover $\Omega'_k \setminus \mathrm{int}(\Omega'_{k-1})$ (shaded regions). Note that all $\Omega'_{k,l}$ are contained in $\mathrm{int}(\Omega'_{k+1}) \setminus \Omega'_{k-2}$ (the regions within the dashed lines). Additionally, the set $\Omega_- = \{x \in \Omega \mid q(x) = 0\}$ is shown.

admits the representation $y = \sum_{k=1}^{\infty} \zeta_k \bar{y}_k \circ j_k^{-1}$ in $L^1_{\mathrm{loc}}(\Omega_+)$. Choose a $\tilde{z} = J_k(z \circ j_k)$ with $z \in \mathcal{C}_0^\infty(\Omega_k)$ as a test function. Applying the change of variables theorem in [11] as well as (17), (18) gives

$$
\begin{aligned}
-\int_{U_k} \bar{y}_k \frac{\partial \tilde{z}}{\partial \xi_d} \, \mathrm{d}\xi &= -\int_{U_k} \bar{y}_k \frac{\partial}{\partial \xi_d} \big(J_k(z \circ j_k)\big) \, \mathrm{d}\xi \\
&= -\int_{U_k} \bar{y}_k \Big((z \circ j_k) \frac{\partial J_k}{\partial \xi_d} + \frac{\partial (z \circ j_k)}{\partial \xi_d} J_k \Big) \, \mathrm{d}\xi \\
&= -\int_{U_k} \bar{y}_k \Big((z \circ j_k)\big((\operatorname{div} q) \circ j_k\big) + \frac{\partial (z \circ j_k)}{\partial \xi_d} \Big) J_k \, \mathrm{d}\xi \\
&= -\int_{\Omega_k} y \big(z \operatorname{div} q + \nabla z \cdot q \big) \, \mathrm{d}x
\end{aligned}
$$

which becomes, using the equivalent defining property of the weak directional derivative (15) and the notation introduced in Remark 6,

$$
\begin{aligned}
-\int_{\Omega_k} y \big(z \operatorname{div} q + \nabla z \cdot q \big) \, \mathrm{d}x &= \int_{\Omega_k} (\partial_q y) z \, \mathrm{d}x \\
&= \int_{U_k} (\partial_q y \circ j_k)(z \circ j_k) J_k \, \mathrm{d}\xi = \int_{U_k} (\partial_q y \circ j_k) \tilde{z} \, \mathrm{d}\xi \; .
\end{aligned}
$$

Hence, by Lemma 4.6, $\bar{y}_k$ is weakly differentiable with respect to ξ_d with derivative $\frac{\partial \bar{y}_k}{\partial \xi_d} = \partial_q y \circ j_k$. Moreover, for $\Omega' \subset\subset \Omega$

$$
\begin{aligned}
&\sum_{\{k' \mid \Omega_{k'} \cap \Omega' \neq \emptyset\}} \int_{U_{k'}} J_{k'} (\zeta_{k'} \circ j_{k'}) \Big| \frac{\partial \bar{y}_{k'}}{\partial \xi_d} \Big| \, \mathrm{d}\xi \\
&\qquad = \sum_{\{k' \mid \Omega_{k'} \cap \Omega' \neq \emptyset\}} \int_{\Omega_{k'}} \zeta_{k'} |\partial_q y| \, \mathrm{d}x = \int_{\Omega''} |\partial_q y| \, \mathrm{d}x < \infty
\end{aligned}
$$

where $\Omega'' = \bigcup_{\{k' | \Omega_{k'} \cap \Omega' \neq \emptyset\}} \Omega_{k'}$ whose closure is also compact in Ω due to construction. This establishes (21) and therefore the "only if" part.

Before proceeding to the "if" part, let us note that, using the partition of unity, one obtains (22) in $L^1_{\mathrm{loc}}(\Omega_+)$. Moreover, we first verify that $\partial_q y = 0$ almost everywhere on the set $\Omega_- = \{x \in \Omega \mid q(x) = 0\}$. Choose a $\Omega'_- \subset \Omega_-$ which is compact in Ω and enlarge it to $\Omega'_{-,\varepsilon} = \overline{\Omega'_- + B_\varepsilon(0)}$ for small $\varepsilon > 0$. If ε is chosen small enough, the test functions $z_\varepsilon = \chi_{\Omega'} * G_\varepsilon$, where G_ε denote the dilated versions of the standard mollifier G, are still in $\mathcal{C}_0^\infty(\Omega)$. Let $C\|\nabla q\|_\infty$ be a Lipschitz constant for q which is common to all considered $\operatorname{supp} z_\varepsilon \subset\subset \Omega$.

Hence, we can estimate $|q(x')| \leq \varepsilon C \|\nabla q\|_\infty$ whenever $x' \in B_\varepsilon(x)$ with $x \in \Omega'_-$. Moreover $|\operatorname{div} q| \leq C\sqrt{d}\|\nabla q\|_\infty$ as well as $\operatorname{div} q = 0$ almost everywhere in Ω_-, so the defining integral can be estimated by

$$\begin{aligned} \Big| \int_\Omega \partial_q y z_\varepsilon \, \mathrm{d}x \Big| &= \Big| \int_\Omega y (z_\varepsilon \operatorname{div} q + \nabla z_\varepsilon \cdot q) \, \mathrm{d}x \Big| \\ &\leq \big(C\sqrt{d} \|\nabla q\|_\infty \|G\|_1 + C \|\nabla q\|_\infty \|\nabla G\|_1 \big) \int_{\operatorname{supp} z_\varepsilon \setminus \Omega_-} |y| \, \mathrm{d}x \ , \end{aligned} \tag{23}$$

which vanishes as $\varepsilon \to 0$ since, due to the compactness of Ω'_- as well as $\Omega'_- \subset \Omega_-$, the Lebesgue measure of $\operatorname{supp} z_\varepsilon \setminus \Omega_-$ tends to zero. But since $\partial_q y$ is locally integrable, we can conclude $\int_{\Omega'_-} \partial_q y \, \mathrm{d}x = \lim_{\varepsilon \to 0} \int_\Omega \partial_q y z_\varepsilon \, \mathrm{d}x = 0$. Consequently, $\partial_q y = 0$ almost everywhere in Ω_-.

It remains to prove the converse direction of the equivalence statement. Suppose that each $\bar{y}_k$ is weakly differentiable in U_k with respect to ξ_d and (21) holds. Our aim is to show that $\partial_q y$ exists and is given by (22). First, take a test function $z \in \mathcal{C}_0^\infty(\Omega_+)$. Hence, testing y with $z \operatorname{div} q + \nabla z \cdot q$ leads to

$$\int_\Omega \sum_{k=1}^\infty \zeta_k y (z \operatorname{div} q + \nabla z \cdot q) \, \mathrm{d}x = \sum_{k=1}^K \int_{\Omega_k} \zeta_k y (z \operatorname{div} q + \nabla z \cdot q) \, \mathrm{d}x \ ,$$

where K is large enough such that $\Omega_{k'} \cap \operatorname{supp} z = \emptyset$ for $k' > K$. Each summand can be transformed to U_k which gives

$$\begin{aligned} \int_{\Omega_k} \zeta_k y (z \operatorname{div} q + \nabla z \cdot q) \, \mathrm{d}x &= \int_{U_k} \bar{\zeta}_k \bar{y}_k \Big(\frac{\partial J_k}{\partial \xi_d} (z_k \circ j_k) + J_k \frac{\partial (z_k \circ j_k)}{\partial \xi_d} \Big) \, \mathrm{d}\xi \\ &= - \int_{U_k} \frac{\partial}{\partial \xi_d} (\bar{\zeta}_k \bar{y}_k) J_k \tilde{z}_k \, \mathrm{d}\xi \end{aligned}$$

denoting $\bar{\zeta}_k = \zeta_k \circ j_k$, $\tilde{z}_k = J_k(\bar{z}_k \circ j_k)$ with a $\bar{z}_k \in \mathcal{C}_0^\infty(\Omega_k)$ and using (17) with the chain rule as well as (18). Furthermore,

$$\begin{aligned} - \int_{U_k} \frac{\partial}{\partial \xi_d} (\bar{\zeta}_k \bar{y}_k) \tilde{z}_k \, \mathrm{d}\xi &= - \int_{U_k} \Big(\frac{\partial \bar{\zeta}_k}{\partial \xi_d} \bar{y}_k + \bar{\zeta}_k \frac{\partial \bar{y}_k}{\partial \xi_d} \Big) \tilde{z}_k \, \mathrm{d}\xi \\ &= - \int_{\Omega_k} \Big((\nabla \zeta_k \cdot q) y + \zeta_k \big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1} \big) \Big) z \, \mathrm{d}x \end{aligned}$$

leading to, summed up,

$$\int_\Omega \Big(\sum_{k=1}^\infty \zeta_k (\bar{y}_k \circ j_k^{-1}) \Big) (z \operatorname{div} q + \nabla z \cdot q) \, \mathrm{d}x = - \int_\Omega \Big(\sum_{k=1}^\infty \zeta_k \big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1} \big) \Big) z \, \mathrm{d}x$$

since $\sum_{k=1}^K \nabla \zeta_k \cdot q = \nabla \big(\sum_{k=1}^K \zeta_k \big) \cdot q = 0$ on $\operatorname{supp} z$. This proves the weak directional differentiability on Ω_+.

Finally, for the case where $z \in \mathcal{C}_0^\infty(\Omega)$ we split z into a part with compact support in Ω_+ and a part for which the associated integral vanishes. This can be achieved by $z = \tilde{z}_\varepsilon + z_\varepsilon$ where z_ε is constructed analog to the above with $\Omega'_- = \operatorname{supp} z \cap \Omega_-$ and using $z\chi_{\Omega'_{-,\varepsilon}}$ instead of $\chi_{\Omega'_-}$, i.e. $z_\varepsilon = (z\chi_{\Omega'_{-,\varepsilon}}) * G_\varepsilon$. Set $\tilde{z}_\varepsilon = z - z_\varepsilon$. Testing then gives

$$\int_\Omega y(z \operatorname{div} q + \nabla z \cdot q) \, \mathrm{d}x = \int_\Omega y\big((\tilde{z}_\varepsilon + z_\varepsilon) \operatorname{div} q + (\nabla \tilde{z}_\varepsilon + \nabla z_\varepsilon) \cdot q\big) \, \mathrm{d}x$$
$$= -\int_\Omega \sum_{k=1}^\infty \zeta_k \Big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1}\Big) \tilde{z}_\varepsilon \, \mathrm{d}x + \int_\Omega y(z_\varepsilon \operatorname{div} q + \nabla z_\varepsilon \cdot q) \, \mathrm{d}x \, . \quad (24)$$

Consider the limits of both terms on the right-hand side as $\varepsilon \to 0$. The functions in the first term converge pointwise almost everywhere and can be estimated by $2\|z\|_\infty \sum_{k=1}^\infty \zeta_k \big|\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1}\big|$ which is integrable on $\overline{\operatorname{supp} z \backslash \Omega_-}$ due to

$$\int_{\operatorname{supp} z} \sum_{k=1}^\infty \zeta_k \Big|\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1}\Big| \, \mathrm{d}x \leq \sum_{\{k' | \Omega_{k'} \cap \operatorname{supp} z \neq \emptyset\}} \int_{\Omega_{k'}} \zeta_{k'} \Big|\frac{\partial \bar{y}_{k'}}{\partial \xi_d} \circ j_{k'}^{-1}\Big| \, \mathrm{d}x$$
$$= \sum_{\{k' | \Omega_{k'} \cap \operatorname{supp} z \neq \emptyset\}} \int_{U_{k'}} J_{k'}(\zeta_{k'} \circ j_{k'}) \Big|\frac{\partial \bar{y}_{k'}}{\partial \xi_d}\Big| \, \mathrm{d}\xi < \infty \, .$$

By virtue of Lebesgue's theorem, the first integral in (24) converges, as $\varepsilon \to 0$, to

$$-\int_\Omega \sum_{k=1}^\infty \zeta_k \Big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1}\Big) \tilde{z}_\varepsilon \, \mathrm{d}x \to -\int_{\operatorname{supp} z \cap \Omega_+} \sum_{k=1}^\infty \zeta_k \big(\frac{\partial \bar{y}_k}{\partial \xi_d} \circ j_k^{-1}\big) z \, \mathrm{d}x \, .$$

The second term in (24) tends to zero by the same argument utilized for (23). This establishes the weak directional differentiability on Ω. □

Remark 7. The result in Proposition 4.7 essentially tells that in case the weight D has rank 1, then weak weighted derivatives are locally respresented by a function being weakly differentiable in one direction. In the other extreme, from Proposition 4.1 follows that when D has full rank, the weak weighted differentiable functions are the usual weak differentiable functions (which are transformed by the identity mapping).

One could be tempted with extending these results to the intermediate cases where the rank r of D is locally fixed but ranges from $2, \ldots, d-1$, e.g. $D(x) = \sum_{i=1}^r p_i \otimes q_i(x)$ where $q_1, \ldots, q_r$ are pointwise linearly independent vector fields and $p_1, \ldots, p_r$ some linearly independent fixed vectors. When trying to establish an analogon of (22) with $\bar{y}_k$ being functions which are now weakly differentiable in the $d-r, \ldots, d$ unit directions, we have to adapt Lemma 4.4, which means to construct locally an r-dimensional foliation such that the tangential bundles of the leaves correspond to the bundle spanned by the vector fields $q_1, \ldots, q_r$. The Frobenius theorem states, at least for sufficiently smooth vector fields, necessary and sufficient conditions for existence of such foliations [18, 12]: Essentially, the Lie brackets $[q_i, q_j]$ have to be pointwise linear combinations of the q_k. Unfortunately, such a condition can be violated for arbitrary vector fields, so an r-dimensional analogon of Lemma 4.4 can only be established in special situations.

Subsequently, one has to determine how functions behave under the coordinate transformations associated with the local foliations in order to obtain an analogon of (22). A detailed and rigorous investigation of what is only roughly sketched here, however, is beyond the scope of this paper and may be subject to further studies.

5 Application to degenerate elliptic equations

As it has already been indicated in [5, 6], degenerate equations are suitable to model the smoothing of images while preserving relevant edges. There, a degenerate parabolic equation with fixed degeneracies has been proposed for image denoising. The coefficients were obtained from the noisy image and updated from time to time. Here, our intention is to study a full non-linear coupling between the coefficients and the solution for the stationary case.

Therefore, the following regularized model for denoising an image u is considered:

$$\text{solve} \quad \begin{cases} \lambda y - \operatorname{div}\bigl(D^2(\nabla y_\sigma)\nabla y^{\mathrm{T}}\bigr) = u & \text{in } \Omega \\ \dfrac{\partial y}{\partial_{D^2(\nabla y_\sigma)}\nu} = 0 & \text{on } \partial\Omega \end{cases} \tag{25a}$$

$$\nabla y_\sigma^{\mathrm{T}} = \nabla(y * G_\sigma)^{\mathrm{T}} \quad , \quad D^2(\nabla y_\sigma) = \Bigl(I - \tau(|\nabla y_\sigma|^2)\frac{\nabla y_\sigma^{\mathrm{T}}}{|\nabla y_\sigma|} \otimes \frac{\nabla y_\sigma^{\mathrm{T}}}{|\nabla y_\sigma|}\Bigr) \tag{25b}$$

where $\lambda > 0$ is “small”, G_σ respresents a Gaussian kernel with variance $\sigma > 0$ which has been normalized to area 1 and $\tau : [0, \infty[\to [0, 1]$ is suitable, monotone weighting function with $\tau(0) = 0$ and, with a $t_0 > 0$, $\tau(t) = 1$ for all $t \geq t_0$. Such a degenerate equation is suitable for image denoising since on the one hand, on areas where y is flat, $D^2(\nabla y_\sigma)$ is close to the identity and consequently, y is the result of an second-order smoothing of the data u. On the other, in regions where y has edges, ∇y_σ has an large absolute value and is approximately pointing in the direction normal to the edge, hence $D^2(\nabla y_\sigma)$ completely blocks smoothing into that direction and y corresponds approximately to a second-order smoothing of u only in the directions which are tangential to the edge.

We will see that under certain assumptions, (25) has a weak solution y in a weighted Sobolev space associated with $D(\nabla y_\sigma)$, the pointwise square root of the diffusion tensor in (25). Such a solution typically leads the partial differential equation to degenerate on arbitrary parts of the domain and is therefore discontinuous in general, a feature which is desired when dealing with images (see [6] for an argumentation why $W^{1,2}_{D(\nabla y_\sigma)}$ is a suitable model for images). The proof of the existence result is inspired by [7] and also utilizes the Schauder fixed point theorem on the mapping which assigns y the weak solution of the linear equation:

$$\lambda z - \operatorname{div}\bigl(D^2(\nabla y_\sigma)\nabla z^{\mathrm{T}}\bigr) = u \quad \text{in } \Omega \quad , \quad \frac{\partial z}{\partial_{D^2(\nabla y_\sigma)}\nu} = 0 \quad \text{on } \partial\Omega \, . \tag{26}$$

As the underlying equation is supposed to degenerate, the proof in [7] cannot be adapted in a straightforward manner since special care with respect to the function spaces has to be taken causing that major parts of the proof rely on the results presented in the previous sections.

Turning towards showing the existence, first define the class of weighting functions τ we like to consider as well as appropriate notions of a weak solutions.

Definition 5.1. A continuous monotone increasing $\tau : [0, \infty[\to [0, 1]$ is an *admissible weighting function* if the function ϑ given by $\vartheta(t) = (1 - \sqrt{1 - \tau(t)})/t$ satisfies the following conditions:

1. There exists a $t_0 > 0$ such that $\vartheta(t) = 1/t$ for all $t \geq t_0$ and $\lim_{t\to 0} \vartheta(t)t^{1/2} = 0$,
2. ϑ is continuously differentiable on $]0, \infty[$ and we have $\lim_{t\to 0} t^{3/2}\vartheta'(t) = 0$.

For admissible τ and $y \in L^2(\Omega)$, we define

$$D(\nabla y_\sigma) = \bigl(I - \vartheta(|\nabla y_\sigma|^2)\nabla y_\sigma^{\mathrm{T}} \otimes \nabla y_\sigma^{\mathrm{T}}\bigr) \tag{27}$$

where $\nabla y_\sigma^{\mathrm{T}} = \nabla(y * G_\sigma)^{\mathrm{T}}$ with zero extension of y outside of Ω.

One can, for example, choose τ according to

$$\tau(t) = \begin{cases} (t/t_0)^{3/2}\big(6t/t_0 - 15\sqrt{t/t_0} + 10\big) & \text{if } t \leq t_0 \\ 1 & \text{else} \end{cases}$$

for some $t_0 > 0$ in order to get an admissible weighting function.

Note that $D(\nabla y_\sigma)$ is the positive semi-definite pointwise square root of $D^2(\nabla y_\sigma)$ defined in (25b). We will immediately see how the conditions on τ imply that the superposition operator $y \mapsto D(\nabla y_\sigma)$ is sufficiently continuous. But before, we give the following definitions of weak solutions for (25) as well as (26).

Definition 5.2. Let $\lambda > 0$, τ be an admissible weighting function and $u \in L^2(\Omega)$ be given. Then, $y \in L^2(\Omega)$ is a weak solution of (25) if for $D = D(\nabla y_\sigma)$ according to (27) we have $y \in W_D^{1,2}$ and for each $z \in W_D^{1,2}$ it holds that

$$\lambda\langle y, z\rangle_2 + \langle D\nabla y^{\mathrm{T}}, D\nabla z^{\mathrm{T}}\rangle_2 = \langle u, z\rangle_2 \ . \tag{28}$$

Likewise, for arbitrary $y \in L^2(\Omega)$ and $D = D(\nabla y_\sigma)$, $z \in W_D^{1,2}$ is said to be a weak solution of (26) if for each $\zeta \in W_D^{1,2}$ the variational equation

$$\lambda\langle z, \zeta\rangle_2 + \langle D\nabla z^{\mathrm{T}}, D\nabla \zeta^{\mathrm{T}}\rangle_2 = \langle u, \zeta\rangle_2 \tag{29}$$

is satisfied.

Lemma 5.3. *Let τ be an admissible weighting function. Then $\varphi : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ defined by*

$$\varphi(\xi) = I - \vartheta(|\xi|^2)\xi \otimes \xi$$

is a continuously differentiable mapping.

The operator $y \mapsto D(\nabla y_\sigma)$ according to (27) is well-defined as $L^2(\Omega) \to Y^*$ and possesses the following continuity property:

$$y_k \rightharpoonup y \text{ in } L^2(\Omega) \quad \Rightarrow \quad \begin{cases} D\big(\nabla y_{k,\sigma}\big) \overset{*}{\rightharpoonup} D(\nabla y_\sigma) \text{ in } Y^* \\ \nabla D\big(\nabla y_{k,\sigma}\big) \to \nabla D(\nabla y_\sigma) \text{ pointwise a.e.} \end{cases}$$

Proof. It is easy to verify the continuous differentiability of φ on $\mathbb{R}^d\backslash\{0\}$ since the partial derivatives are given by

$$\frac{\partial\varphi}{\partial\xi_i}(\xi) = -2\xi_i\vartheta'(|\xi|^2)\xi\otimes\xi - \vartheta(|\xi|^2)e_i\otimes\xi - \vartheta(|\xi|^2)\xi\otimes e_i$$

which are continuous functions. Moreover, as $\xi \to 0$ it holds that

$$\lim_{\xi\to 0}|\nabla\varphi(\xi)| \leq 2\big(\lim_{\xi\to 0}|\vartheta'(|\xi|^2)||\xi|^3 + d\vartheta(|\xi|^2)|\xi|\big) = 0$$

since ϑ satisfies the conditions stated in Definition 5.1 making $\nabla\vartheta$ continuous on the whole space. Likewise, since ϑ is continuous, we have that

$$\lim_{\xi\to 0}\frac{|\varphi(\xi) - \varphi(0)|}{|\xi|} \leq \lim_{\xi\to 0}\vartheta(|\xi|^2)|\xi| = 0 \qquad \Rightarrow \qquad \nabla\varphi(0) = 0 \ .$$

For the continuity statement on the associated regularized superposition operator, first observe that $y \mapsto \nabla y_\sigma$ maps such that

$$\|\nabla y_\sigma\|_\infty \leq \|\nabla G_\sigma\|_2\|y\|_2 \qquad , \qquad \|\nabla^2 y_\sigma\|_\infty \leq \|\nabla^2 G_\sigma\|_2\|y\|_2$$

meaning that ∇y_σ is in $W^{1,\infty}(\Omega, \mathrm{I\!R}^d)$. Moreover, whenever $y_k \rightharpoonup y$ in $L^2(\Omega)$, for each $x \in \Omega$ holds

$$\nabla y_{k,\sigma}(x) \to \nabla y_\sigma(x) \qquad , \qquad \nabla^2 y_{k,\sigma}(x) \to \nabla^2 y_\sigma(x)$$

since the convolution evaluated in a point x is an L^2 dual pairing and one is allowed to interchange integration and differentiation. Now, written as superposition operator $D(\nabla y_\sigma)(x) = \varphi\big(\nabla y_\sigma(x)\big)$ (and the analog for $D(\nabla y_{k,\sigma})$) with φ being a continuously differentiable mapping, hence, for each $x \in \Omega$,

$$\varphi\big(\nabla y_{k,\sigma}(x)\big) \to \varphi\big(\nabla y_\sigma(x)\big) \ ,$$
$$\begin{cases} \nabla\big(\varphi(\nabla y_{k,\sigma})\big)(x) = \nabla\varphi\big(\nabla y_{k,\sigma}(x)\big)\nabla^2 y_{k,\sigma}(x) \\ \qquad\qquad \to \nabla\varphi\big(\nabla y_\sigma(x)\big)\nabla^2 y_\sigma(x) = \nabla\big(\varphi(\nabla y_\sigma)\big)(x) \ . \end{cases}$$

Pointwise convergence means in particular weak*-convergence in the respective L^∞-spaces, hence, by Proposition 2.4, we additionally have $D(\nabla y_{k,\sigma}) \overset{*}{\rightharpoonup} D(\nabla y_\sigma)$ in Y^* what was to show. □

Proposition 5.4. *Let Ω be a bounded Lipschitz domain, $u \in L^2(\Omega)$ be fixed and τ an admissible weighting function. The operator $\mathcal{T} : L^2(\Omega) \to L^2(\Omega)$ which maps $y \mapsto z$ according to* (29) *is well-defined and possesses the following properties:*

1. *each $\mathcal{T}(y)$ belongs to the corresponding $W^{1,2}_{D(\nabla y_\sigma)}$,*
2. *$\mathcal{T}$ maps $C = \{y \in L^2(\Omega) \mid \|y\|_2 \le \lambda^{-1}\|u\|_2\}$ into itself,*
3. *it is continuous with respect to the induced weak topology in C.*

Proof. First of all, examine the existence and uniqueness for solutions of (29): Note that for each $y \in L^2(\Omega)$, Lemma 5.3 implies $D = D(\nabla y_\sigma) \in Y^*$, hence we can consider the associated weighted Sobolev space $W^{1,2}_D$ and define the bilinear form as well as the linear form

$$a(z,\zeta) = \lambda\langle z,\, \zeta\rangle_2 + \langle D\nabla z^{\mathrm{T}},\, D\nabla\zeta^{\mathrm{T}}\rangle_2 \qquad , \qquad f(\zeta) = \langle u,\, \zeta\rangle_2 \ .$$

Of course, (29) is equivalent to finding a $z \in W^{1,2}_D$ for which $a(z,\zeta) = f(\zeta)$ for all $\zeta \in W^{1,2}_D$. The continuity and coercivity estimates

$$|a(z,\zeta)| \le \max(\lambda,1)\|z\|_{W^{1,2}_D}\|\zeta\|_{W^{1,2}_D} \quad , \quad \min(\lambda,1)\|z\|^2_{W^{1,2}_D} \le a(z,z) \ ,$$

as well as $|f(\zeta)| \le \|u\|_2\|\zeta\|_{W^{1,2}_D}$ can easily be derived with standard arguments, consequently, the Lax-Milgram theorem ensures the existence of a unique solution $z = \mathcal{T}(y)$ in $W^{1,2}_D$. Thus, $\mathcal{T}$ is also well-defined as $\mathcal{T} : L^2(\Omega) \to L^2(\Omega)$ by the embedding $W^{1,2}_D \hookrightarrow L^2(\Omega)$. Now, if we test (29) with $\zeta = z$,

$$\lambda\|z\|_2^2 + \|D\nabla z^{\mathrm{T}}\|_2^2 = \langle u,\, z\rangle_2 \le \|u\|_2\|z\|_2 \quad \Rightarrow \quad \begin{cases} \|z\|_2 \le \lambda^{-1}\|u\|_2 \\ \|D\nabla z^{\mathrm{T}}\|_2^2 \le \lambda^{-1}\|u\|_2^2 \ , \end{cases} \tag{30}$$

so $\mathcal{T}$ maps $L^2(\Omega)$ into C and in particular, the set C into itself.

Let us finally prove the continuity statement. First, note that since C is bounded and weakly closed in the separable space $L^2(\Omega)$, the weak continuity can equivalently be described by weak sequential continuity [10]. Hence, we have to show that $y_k \rightharpoonup y^*$ with $\{y_k\} \subset C$ and $y^* \in C$ implies the convergence $z_k \rightharpoonup z^*$ for $z_k = \mathcal{T}(y_k)$ and $z^* = \mathcal{T}(y^*)$. Denote by $D_k = D(\nabla y_{k,\sigma})$ as well as $D = D(\nabla y^*_\sigma)$ for which $D_k \overset{*}{\rightharpoonup} D$ in Y^* holds as well as $\nabla D_k \to \nabla D$ almost everywhere in Ω due to Lemma 5.3. Looking at the a-priori estimates (30), we see that $\{z_k\}$ and

$\{D_k \nabla z_k^{\mathrm{T}}\}$ are bounded in the respective L^2-spaces, so going to subsequences (without relabeling), the existence of $z \in L^2(\Omega)$ and $v \in L^2(\Omega, \mathrm{IR}^d)$ such that $z_k \rightharpoonup z$ and $D_k \nabla z_k^{\mathrm{T}} \rightharpoonup v$ can be ensured. All prerequisites of Proposition 3.3 are satisfied, so one concludes $v = D\nabla z^{\mathrm{T}}$ meaning that $z \in W_D^{1,2}$.

It remains to show that actually $z = z^*$, which amounts to proving that z satisfies (29) since its solution is unique in $W_D^{1,2}$. Choose a test function $\zeta \in W^1(\Omega)$ and observe that $D_k \nabla \zeta^{\mathrm{T}} \to D\nabla\zeta^{\mathrm{T}}$ in $L^2(\Omega)$ since $\{D_k\}$ converges in Y^* in the weak* sense and consequently uniformly in $\mathcal{C}(\Omega, \mathrm{IR}^{d\times d})$ (see Proposition 2.4). Each z_k solves (29) with the corresponding y_k, i.e.

$$\lambda\langle z_k, \zeta\rangle_2 + \langle D_k \nabla z_k^{\mathrm{T}}, D_k \nabla \zeta^{\mathrm{T}}\rangle_2 = \langle u, \zeta\rangle_2 \ .$$

Here, only dual pairing of weakly and strongly convergent sequences are involved, so passage to the limit yields

$$\lambda\langle z, \zeta\rangle_2 + \langle D \nabla z^{\mathrm{T}}, D \nabla \zeta^{\mathrm{T}}\rangle_2 = \langle u, \zeta\rangle_2 \ .$$

Since $W^1(\Omega) \hookrightarrow W_D^{1,2}$ densely as an immediate consequence of Theorem 3.12, the variational equation holds for sufficiently many test functions, hence z is the solution of (29) with the corresponding y^* meaning that $z = z^*$. This holds for each subsequence for which $\{z_k\}$ and $\{D_k \nabla z_k^{\mathrm{T}}\}$ converge weakly, so the whole sequence must satisfy $z_k \rightharpoonup z^*$ what was to show. □

Proposition 5.4 provides us with everything we need for the application of Schauder's fixed point theorem.

Theorem 5.5. *Let Ω be a bounded Lipschitz domain, $u \in L^2(\Omega)$ and τ an admissible weighting function. Then, there exists a weak solution of the degenerate non-linear equation* (25) *in the sense of Definition 5.2.*

Proof. The fixed point operator $\mathcal{T}$ maps the compact and convex subset $C = \{y \in L^2(\Omega) \mid \|y\|_2 \leq \lambda^{-1}\|u\|_2\}$ endowed with the induced weak topology continuously into itself, see Proposition 5.4. Application of the Schauder fixed-point theorem [30] yields the existence of a $y \in C$ with $y = \mathcal{T}(y)$ meaning that with $D = D(\nabla y_\sigma)$ we have on the one hand that $y = \mathcal{T}(y) \in W_D^{1,2}$ as well as (29) with $z = y$ on the other hand, eventually leading to (28). □

Remark 8. (i) In the existence proof, some ingredients played an important role: First, we had to find appropriate solution spaces for the degenerate elliptic equations. Usually, this is done by considering weights which allow the utilization of the classical weak gradient leading to the classical weighted Sobolev spaces [29, 14] or by completing a classical function space with respect to a weaker norm [22]. Here, we utilized the variational notion of the weak weighted derivative and associated Sobolev spaces, serving as a replacement for the classical gradient which does not necessarily exist in the degenerate case. In order to deal with the solution space varying with y, weak closedness properties became important. In contrast to the approach via completion, such closedness results are easily accessible in weighted Sobolev spaces. They give, roughly speaking, the property that a weak limit of solutions belongs to the corresponding limit solution space even if the sequence itself is not contained in the latter. Finally, in order to obtain the limit being a solution, we are utilizing the density of $W^1(\Omega)$, which is independent of D, in all $W_D^{1,2}$. Without having the density result, such a conclusion would not have been possible.

(ii) Unfortunately, not much can be said about uniqueness of a solution. We give some (mathematically not rigorous) indications for what the reasons might be. First, as the "forward operator" $\mathcal{A} : L^2(\Omega) \to \mathcal{L}\big(W^1(\Omega), W^1(\Omega)^*\big)$ defined by

$$\mathcal{A}(y)(z) = \lambda\langle y, z\rangle_2 + \langle D(\nabla y_\sigma)\nabla y^{\mathrm{T}}, D(\nabla y_\sigma)\nabla z^{\mathrm{T}}\rangle$$

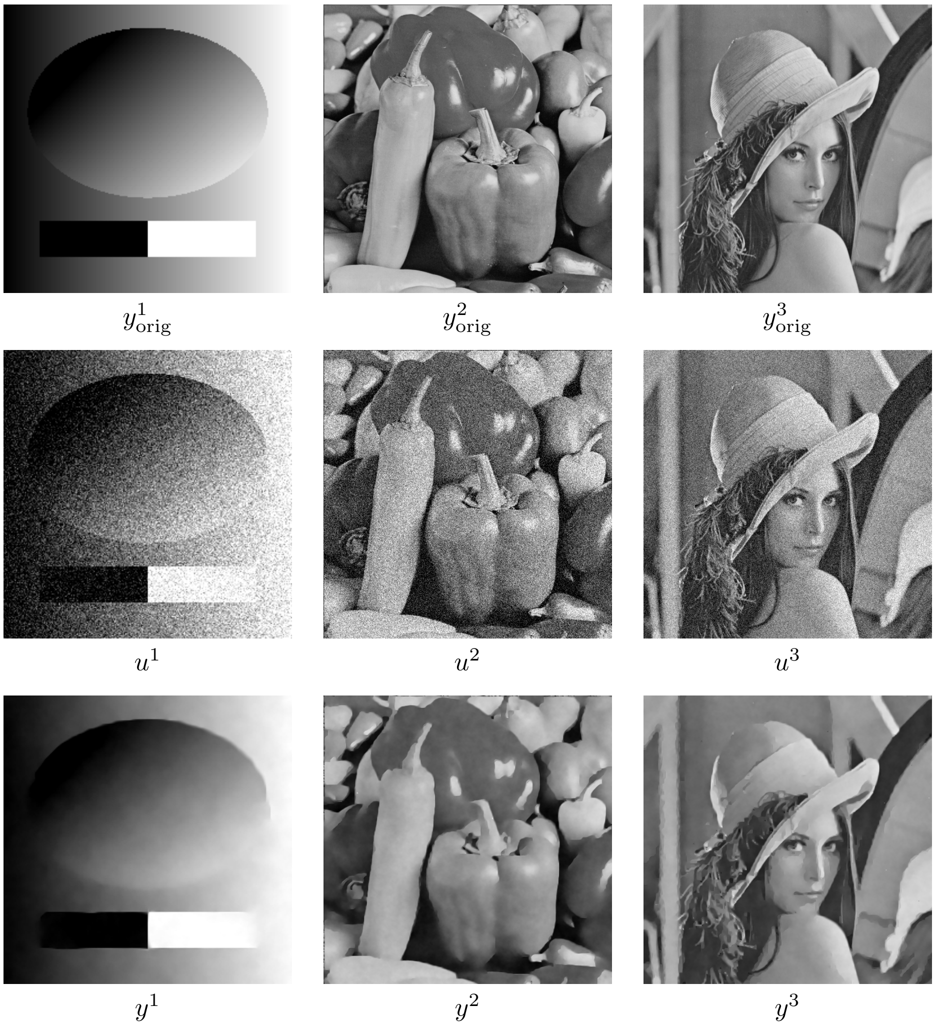

Fig. 3: Illustration of image denoising by solving (25). The top row depicts some sample images which are distorted by additive Gaussian noise (middle row). These noisy images were plugged into the right-hand side of (25), and a numerical approximation of a solution has been computed. In the bottom row, the outcome of this algorithm is depicted (up to a normalizing factor).

is not monotone, so one cannot use monotonicity arguments to obtain (global) uniqueness. One still might ask whether a solution is locally unique, meaning that possible solutions are discrete with respect to some norm possibly different from the L^2-norm. An approach then would be to examine the derivative of $\mathcal{T} - I$. However, variations in y may cause discontinuities to move (since $W^{1,2}_{D(\nabla y_\sigma)}$ is varying) suggesting that $\mathcal{T} - I$ is not differentiable in reasonable function spaces, thus, it is also not immediate how show local uniqueness.

To conclude this section, we show with some basic numerical examples that equation (25) is suitable for image denoising. We do not want to study the associated numerical analysis as this is beyond the scope of this paper, but it is worth mentioning that the naive approach of performing a fixed-point iteration with $\mathcal{T}$ usually turns out to be a successful strategy for computing solutions of (25). The denoising results for some well-known test images are depicted in Figure 3. As the image show, the partial differential equation is indeed able to preserve the edges while smoothing the noisy flat parts of the image. However, staircasing might occur depending on the noise level and amount of smoothing (adjusted with λ) as well as the edge sensitivity (adjusted with t_0). Also, as texture is not taken into account, fine structures may vanish. The method nevertheless performs pretty good on cartoon-like images.

6 Summary and conclusion

The notion of weighted gradients which is based on a variational formulation instead of an algebraic operation leads to a new class of weighted Sobolev spaces for which it is possible to establish well-known and desired properties such as closedness, density of smooth functions as well as the usual calculus rules. These spaces moreover model discontinuous functions with different "degree" of discontinuity, depending on the dimension of the kernel of the respective weight D in the neighborhood of a certain point. As we have seen, functions in $W_D^{1,r}$ can have no degree of smoothness where D vanishes and have to possess a classical gradient where D is regular. Likewise, if the weight is only of rank one, one can say that functions whose weak weighted derivative exist are smooth in one (possibly varying) direction and possess, in general, discontinuites with respect to each perpendicular direction. Another important aspect of the variational construction of weak weighted gradients is their natural connection to weak closedness with respect to varying functions and weights.

Such results can be applied to solve certain types of degenerate elliptic equations, e.g. to solve a denoising problem in image processing. For the latter, modeling the denoising procedure as a full anisotropic degenerate equation instead of the common way of regularizing it to become a uniformly elliptic problem, yields functions with "true" discontinuities across hypersurfaces, i.e. edges, as it is also the case for functions with bounded variation, a popular model in mathematical image processing. Moreover, the utilized notions and techniques are not limited to this topic; an adaptation to other problems also seems to be possible.

As the variational weak weighted derivative possesses many desireable properties, there is, however, the limitation that one usually has to assume a certain degree of smoothness of the weights which, roughly speaking, corresponds to Lipschitz continuity. Hence, the results in this paper do not include the existing theory of weighted Sobolev spaces but rather extend it to another class of weights. One can summarize that when analyzing weighted Sobolev spaces and degenerate elliptic equations, one approach is to impose integrability conditions (as for the Muckenhoupt classes) where restrictions on the decay behavior of the weights apply, while the approach introduced in this article bases on smoothness (which is also a restriction) but allows, among other things, for a more general decay behavior.

A Additional proofs

A.1 Proof of Lemma 3.11

We first construct the operators $\mathcal{M}_\varepsilon$. By assumption, Ω is a bounded Lipschitz domain, so according to that, there exist finitely many $U_l \subset \mathrm{I\!R}^{d-1}$ which are neighborhoods of zero (for $1 \leq l \leq L$), Lipschitz continuous mappings $\varphi_l : U_l \to \mathrm{I\!R}$, linear isometries $S_l : \mathrm{I\!R}^d \to \mathrm{I\!R}^d$ and $\bar{x}_l \in \partial\Omega$ such that

$$\begin{aligned}
S_l\big(\bar{\xi}, \varphi_l(\bar{\xi})\big) + \bar{x}_l &\in \partial\Omega && \text{for } \bar{\xi} \in U_l \\
S_l\big(\bar{\xi}, \varphi_l(\bar{\xi}) + \xi_d\big) + \bar{x}_l &\in \Omega && \text{for } \bar{\xi} \in U_l \ , \ \xi_d \in \,]0, \varepsilon_l[\\
S_l\big(\bar{\xi}, \varphi_l(\bar{\xi}) + \xi_d\big) + \bar{x}_l &\in \mathrm{I\!R}^d \backslash \overline{\Omega} && \text{for } \bar{\xi} \in U_l \ , \ \xi_d \in \,]{-\varepsilon_l}, 0[
\end{aligned}$$

with suitable $\varepsilon_l > 0$. One can moreover assume that the

$$\tilde{\Omega}_l = \{x \in \mathbb{R}^d \mid x = S_l\big(\bar{\xi}, \varphi_l(\bar{\xi}) + \xi_d\big) + \bar{x}_l \ , \bar{\xi} \in U_l \ , \ |\xi_d| < \varepsilon_l\}$$

are open and that $\partial\Omega \subset \bigcup_{1\leq l\leq L} \tilde{\Omega}_l$. Additionally, there has to be an $\varepsilon_0 > 0$ such that the sets

$$\Omega_{l,\varepsilon_0} = \{x \in \tilde{\Omega}_l \mid \operatorname{dist}(x, \partial\tilde{\Omega}_l) > \varepsilon_0\}$$

still cover $\partial\Omega$: Otherwise, there would be a sequence $\{x_\varepsilon\} \subset \partial\Omega$ such that for each $\varepsilon > 0$ the corresponding $x_\varepsilon \notin \Omega_{l,\varepsilon}$ for each $1 \leq l \leq L$. But $\partial\Omega$ is compact, so there exists a subsequence which converges to some $x \in \partial\Omega$. Now there is an l such that $x \in \tilde{\Omega}_l$ and since $\tilde{\Omega}_l$ is open, we have $\varepsilon_0 = \operatorname{dist}(x, \partial\tilde{\Omega}_l)/2 > 0$. Consequently, x is in Ω_{l,ε_0} as well as almost every member of the above subsequence, which results in a contradiction. Thus, there is an $\varepsilon_0 > 0$ such that $\Omega_l = \Omega_{l,\varepsilon_0}$ cover $\partial\Omega$. Moreover, by standard arguments, an open Ω_0 with $\overline{\Omega_0} \subset\subset \Omega$ can be found such that $\overline{\Omega} \subset \bigcup_{0\leq l\leq L} \Omega_l$.

Denote by η_l the image of e_d under S_l, i.e. the unit direction which locally points from $\partial\Omega$ into Ω. Introduce $t_l = 2(C_l + 1)$ with C_l being the Lipschitz constant for the boundary mapping φ_l and consider the mollifiers

$$G_{l,\varepsilon,x}(\xi) = G_\varepsilon(x + \varepsilon t_l \eta_l - \xi) \quad \text{for} \quad x \in \overline{\Omega} \cap \Omega_l \ , \ \xi \in \mathbb{R}^d \ .$$

For suitably small $\varepsilon > 0$ we want to obtain that $\operatorname{supp} G_{l,\varepsilon,x}$ is compact in Ω. First, by choosing $0 < \varepsilon < (1 + t_l)^{-1}\varepsilon_0$, we can ensure that $\operatorname{supp} G_{l,\varepsilon,x} \subset\subset \tilde{\Omega}_l$: If $\xi \in \operatorname{supp} G_{l,\varepsilon,x}$, then

$$|x + \varepsilon t_l \eta_l - \xi| \leq \varepsilon \quad \Rightarrow \quad |x - \xi| \leq |x + \varepsilon t_l \eta_l - \xi| + \varepsilon t_l \leq (1 + t_l)\varepsilon < \varepsilon_0$$

meaning that $\xi \in \Omega_l + B_{\varepsilon_0}(0) \subset \tilde{\Omega}_l$. Moreover, for $\xi \in \operatorname{supp} G_{l,\varepsilon,x}$ we can estimate the distance to the boundary part $\partial\Omega \cap \tilde{\Omega}_l$ from below as follows. Since $x \in \tilde{\Omega}_l$, there has to be a $\bar{\xi}_1 \in U_l$ and a $\xi_d \in \,]0, \varepsilon_l[$ such that $x = S_l\big(\bar{\xi}_1, \varphi_l(\bar{\xi}_1) + \xi_d\big) + \bar{x}_l$. Now any point $\bar{x} \in \partial\Omega \cap \tilde{\Omega}_l$ can be written as $\bar{x} = S_l\big(\bar{\xi}_2, \varphi(\bar{\xi}_2)\big) + \bar{x}_l$ for $\bar{\xi}_2 \in U_l$. Thus, we know that

$$\begin{aligned} |x + \varepsilon t_l \eta_l - \bar{x}| &= \big|S_l\big(\bar{\xi}_1, \varphi_l(\bar{\xi}_1) + \xi_d + \varepsilon t_l\big) - S_l\big(\bar{\xi}_2, \varphi_l(\bar{\xi}_2)\big)\big| \\ &= \big|\big(\bar{\xi}_1 - \bar{\xi}_2, \varphi_l(\bar{\xi}_1) - \varphi_l(\bar{\xi}_2) + \xi_d + \varepsilon t_l\big)\big| \\ &\geq \max\{|\bar{\xi}_1 - \bar{\xi}_2|, \varepsilon t_l - C_l|\bar{\xi}_1 - \bar{\xi}_2|\} \\ &\geq 2\varepsilon \end{aligned}$$

since there is only the possibility that $|\bar{\xi}_1 - \bar{\xi}_2| \geq 2\varepsilon$ or $|\bar{\xi}_1 - \bar{\xi}_2| < 2\varepsilon$. Consequently,

$$|\bar{x} - \xi| \geq |x + \varepsilon t_l \eta_l - \bar{x}| - |x + \varepsilon t_l \eta_l - \xi| \geq 2\varepsilon - \varepsilon = \varepsilon \ ,$$

so $\xi \in \Omega$ and in particular $\operatorname{supp} G_{l,\varepsilon,x} \subset\subset \Omega$.

In the following, choose a $\tilde{\varepsilon} > 0$ such that $\tilde{\varepsilon} < (1 + t_l)^{-1}\varepsilon_0$ for $1 \leq l \leq L$ as well as $\Omega_0 + B_{\tilde{\varepsilon}}(0) \subset\subset \Omega$ and let $0 < \varepsilon < \tilde{\varepsilon}$. Note that $\tilde{\varepsilon}$ then only depends on the domain Ω. The next step is to define $G_{l,\varepsilon}(x) = G_\varepsilon(x + \varepsilon t_l \eta_l)$ such that, for each $y \in W_D^{1,r}$,

$$y_{l,\varepsilon}(x) = (y * G_{l,\varepsilon})(x) = \langle y, G_{l,\varepsilon,x}\rangle$$

for $0 < \varepsilon < \tilde{\varepsilon}$ on each $\Omega \cap \Omega_l$. To take Ω_0 into account, let $t_0 = 0$, $\eta_0 = 0$ and set $y_{0,\varepsilon} = y * G_{0,\varepsilon} = y * G_\varepsilon$. Now, all $y_{l,\varepsilon}$ have to be assembled to an appropriate y_ε. This is of course done with a smooth partition of unity ζ_l on $\overline{\Omega}$ subordinate to the sets $\Omega_0, \ldots, \Omega_L$. We then define, as usual, the operators

$$\mathcal{M}_\varepsilon y = y_\varepsilon = \sum_{l=0}^{L} \zeta_l y_{l,\varepsilon} \ .$$

As one can easily see, this construction yields linear operators which only depend on Ω and whose images are always smooth functions, since by construction, each $\zeta_l y_{l,\varepsilon}$ is an element of $\mathcal{C}^\infty(\overline{\Omega})$ meaning that $y_\varepsilon \in \mathcal{C}^\infty(\overline{\Omega})$.

Before we prove that the boundedness statement $\|D\nabla y_\varepsilon^{\mathrm{T}}\|_r \leq C$ holds for these y_ε, we like to obtain estimates for the L^r-norm of $D\nabla y_{l,\varepsilon}^{\mathrm{T}}$ restricted to the sets $\Omega \cap \Omega_l$ for $1 \leq l \leq L$. Most of the arguments from the proof of Lemma 3.8 can be reused here, we will therefore give a shortened argumentation. Consider $D\nabla y_{l,\varepsilon}^{\mathrm{T}}$ which can be written as

$$\begin{aligned} D(x)\nabla y_{l,\varepsilon}(x)^{\mathrm{T}} = &\int_\Omega y(\xi)\big(D(x) - D(\xi)\big)\nabla G_\varepsilon(x + \varepsilon t_l \eta_l - \xi)^{\mathrm{T}} \,\mathrm{d}\xi \\ &+ \int_\Omega \big(D\nabla y^{\mathrm{T}} + y \operatorname{div} D\big)(\xi) G_\varepsilon(x + \varepsilon t_l \eta_l - \xi) \,\mathrm{d}\xi \ . \end{aligned}$$

The crucial part is showing that the L^r-norm of the first term restricted to $\Omega \cap \Omega_l$ is bounded independently of ε, the second term can be estimated by standard arguments (see Lemma 3.8). Here, observe that due to Proposition 2.3 as well as the above considerations, we have the estimates

$$\begin{aligned} |D(x + \varepsilon t_l \eta_l) - D(\xi)| &\leq \|\nabla D\|_\infty \varepsilon \qquad \text{for } x \in \Omega \cap \Omega_l \ , \ |x + \varepsilon t_l \eta_l - \xi| \leq \varepsilon \\ |D(x) - D(x + \varepsilon t_l \eta_l)| &\leq \|\nabla D\|_\infty t_l \varepsilon \qquad \text{for } x \in \Omega \cap \Omega_l \end{aligned}$$

since $B_\varepsilon(x + \varepsilon t_l \eta_l)$ is convex and compact in Ω as well as the line connecting x and $x + \varepsilon t_l \eta_l$. Thus,

$$\begin{aligned} &\Big(\int_{\Omega\cap\Omega_l} \Big|\int_\Omega y(\xi)\big(D(x) - D(\xi)\big)\nabla G_\varepsilon(x + \varepsilon t_l \eta_l - \xi)^{\mathrm{T}} \,\mathrm{d}\xi\Big|^r \,\mathrm{d}x\Big)^{1/r} \\ &\qquad \leq (1 + t_l)\|\nabla D\|_\infty \Big(\int_{\Omega\cap\Omega_l} \Big|\int_\Omega \varepsilon^{-d}|y(\xi)|\Big|\nabla G\Big(\frac{x - \xi}{\varepsilon} + t_l \eta_l\Big)\Big| \,\mathrm{d}\xi\Big|^r \,\mathrm{d}x\Big)^{1/r} \\ &\qquad \leq (1 + t_l)\|\nabla D\|_\infty \|\nabla G\|_1 \|y\|_r \end{aligned}$$

from which follows that

$$\begin{aligned} &\Big(\int_{\Omega\cap\Omega_l} |D(x)\nabla y_{l,\varepsilon}(x)^{\mathrm{T}}|^r \,\mathrm{d}x\Big)^{1/r} \\ &\qquad \leq \big((1 + t_l)\|\nabla G\|_1 + \sqrt{d}\|G\|_1\big)\|\nabla D\|_\infty \|y\|_r + \|G\|_1 \|D\nabla y^{\mathrm{T}}\|_r \ . \qquad (31) \end{aligned}$$

Note that the estimate also applies to $y_{0,\varepsilon}$.

We now turn to proving the claimed statements and verify that $y_\varepsilon \to y$ in $L^r(\Omega)$ as $\varepsilon \to 0$. It is clear that $y_{l,\varepsilon} \to y$ in $L^r(\Omega \cap \Omega_l)$ since $y_{l,\varepsilon}$ emerges from the convolution of y with the dilated versions of the translated mollifier $G(\,\cdot\, + t_l \eta_l)$. Hence,

$$\|y - y_\varepsilon\|_r \leq \sum_{l=0}^L \Big(\int_{\Omega\cap\Omega_l} |y(x) - y_{l,\varepsilon}(x)|^r \,\mathrm{d}x\Big)^{1/r} \to 0$$

as $\varepsilon \to 0$. Regarding the weak weighted and directional derivatives, the estimate (31) finally yields

$$\begin{aligned} \|D\nabla y_\varepsilon^{\mathrm{T}}\|_r &\leq \sum_{l=0}^L \|\nabla\zeta_l\|_\infty \|D\|_\infty \Big(\int_{\Omega\cap\Omega_l} |y_{l,\varepsilon}(x)|^r \,\mathrm{d}x\Big)^{\frac{1}{r}} + \Big(\int_{\Omega\cap\Omega_l} |(D\nabla y_{l,\varepsilon}^{\mathrm{T}})(x)|^r \,\mathrm{d}x\Big)^{\frac{1}{r}} \\ &\leq \Big(\sum_{l=0}^L \|\nabla\zeta_l\|_\infty \|G\|_1 \|D\|_\infty + \big((1 + t_l)\|\nabla G\|_1 + \sqrt{d}\|G\|_1\big)\|\nabla D\|_\infty\Big)\|y\|_r \\ &\quad + (L+1)\|G\|_1 \|D\nabla y^{\mathrm{T}}\|_r \\ &\leq \tilde{C}_1\big(\|D\|_{Y^*}\big)\|y\|_r + \tilde{C}_2 \|D\nabla y^{\mathrm{T}}\|_r \ , \end{aligned}$$

with suitable $\tilde{C}_1\big(\|D\|_{Y^*}\big)$ and $\tilde{C}_2$. Note that $\tilde{C}_1$ remains bounded whenever D is bounded in Y^*. The desired result $\|D\nabla y_\varepsilon^{\mathrm{T}}\|_r \leq C\|y\|_{W_D^{1,r}}$ then follows with C given by the following expression (r' denotes the dual exponent of r)

$$C = \big(\tilde{C}_1(\|D\|_{Y^*})^{r'} + \tilde{C}_2^{r'}\big)^{1/r'} . \qquad \square$$

A.2 Proofs of Lemmas 4.4, 4.5 and 4.6

Proof of Lemma 4.4. Denote by $q_1, \ldots, q_{d-1}$ a set of orthonormal vectors perpendicular to $q(x)$ such that $\{q_1, \ldots q_{d-1}, q(x)\}$ becomes a positive orthogonal basis. By local Lipschitz continuity (with Lipschitz constant C_1), there exists an open neighborhood $\Omega_0' \subset \Omega$ of x such that for each $x' \in \Omega_0'$, the vectors $\{q_1, \ldots, q_{d-1}, q(x')\}$ are still linearly independent and fulfilling (16) in Ω_0'. Observe that q is Lipschitz continuous in a neighborhood of x by virtue of Proposition 2.3, so let j be defined as the solution of the ordinary differential equation

$$\frac{\partial j}{\partial \xi_d}(\bar{\xi}, \xi_d) = q\big(j(\bar{\xi}, \xi_d)\big) \quad , \quad j(\bar{\xi}, 0) = \sum_{i=1}^{d-1} \bar{\xi}_i q_i + x ,$$

which makes sense on the product $U_0 \times \,]-\varepsilon, \varepsilon[$ where $U_0 \subset \mathbb{R}^{d-1}$ is a neighborhood of 0 and ε a sufficiently small constant which arises in the existence theorem of Picard and Lindelöf (and does not depend on $\bar{\xi}$, see [3]).

Let $\bar{\xi}_1, \bar{\xi}_2 \in U_0$, $|\xi_d| < \varepsilon$ and estimate

$$\begin{aligned} |j(\bar{\xi}_1, \xi_d) - j(\bar{\xi}_2, \xi_d)| &\leq |j(\bar{\xi}_1, 0) - j(\bar{\xi}_2, 0)| + \Big| \int_0^{\xi_{1,d}} \big|q\big(j(\bar{\xi}_1, s)\big) - q\big(j(\bar{\xi}_2, s)\big)\big| \,\mathrm{d}s \Big| \\ &\leq |\bar{\xi}_1 - \bar{\xi}_2| + C_1 \Big| \int_0^{\xi_{1,d}} |j(\bar{\xi}_1, s) - j(\bar{\xi}_2, s)| \,\mathrm{d}s \Big| \end{aligned}$$

so $|j(\bar{\xi}_1, \xi_d) - j(\bar{\xi}_2, \xi_d)| \leq \mathrm{e}^{\varepsilon C_1} |\bar{\xi}_1 - \bar{\xi}_2|$ with the help of Gronwall's inequality. Consequently, for $\xi_k = (\bar{\xi}_k, \xi_{k,d})$ where $k = 1, 2$ and $|\xi_{k,d}| < \varepsilon$, we get

$$\begin{aligned} |j(\xi_1) - j(\xi_2)| &\leq |j(\bar{\xi}_1, \xi_{1,d}) - j(\bar{\xi}_2, \xi_{1,d})| + |j(\bar{\xi}_2, \xi_{1,d}) - j(\bar{\xi}_2, \xi_{2,d})| \\ &\leq \mathrm{e}^{\varepsilon C_1} |\bar{\xi}_1 - \bar{\xi}_2| + \Big| \int_{\xi_{1,d}}^{\xi_{2,d}} \big|q\big(j(\bar{\xi}_2, s)\big)\big| \,\mathrm{d}s \Big| \\ &\leq \mathrm{e}^{\varepsilon C_1} |\bar{\xi}_1 - \bar{\xi}_2| + \|q\|_\infty |\xi_{1,d} - \xi_{2,d}| \leq C_2 |\xi_1 - \xi_2| \end{aligned}$$

with a suitable $C_2 > 0$. This yields the Lipschitz continuity of j in $U_0 \times \,]-\varepsilon, \varepsilon[$. Note that in this domain, j is also partially differentiable with respect to ξ_d with derivative $\frac{\partial j}{\partial \xi_d}(\bar{\xi}, \xi_d) = q\big(j(\bar{\xi}, \xi_d)\big)$ satisfying

$$\Big| \frac{\partial j}{\partial \xi_d}(\xi_1) - \frac{\partial j}{\partial \xi_d}(\xi_2) \Big| = \big|q\big(j(\xi_1)\big) - q\big(j(\xi_2)\big)\big| \leq C_1 |j(\xi_1) - j(\xi_2)| \leq C_1 C_2 |\xi_1 - \xi_2| ,$$

meaning that $\frac{\partial j}{\partial \xi_d}$ is Lipschitz continuous. The partial derivatives of j with respect to $\bar{\xi}$ in 0 trivially exist, leading to

$$\nabla j(0) = \begin{bmatrix} q_1 & \ldots & q_{d-1} & q(x) \end{bmatrix}$$

if j is also (totally) differentiable in 0. But this follows from

$$\begin{aligned} |j(\xi) - j(0) - \nabla j(0)\xi| &= |j(\bar{\xi}, \xi_d) - j(\bar{\xi}, 0) - \xi_d q(x)| \\ &= \Big| \int_0^1 \Big(q\big(j(\bar{\xi}, s)\big) - q\big(j(0,0)\big) \Big) \xi_d \,\mathrm{d}s \Big| \leq C_1 C_2 |\xi| |\xi_d| . \end{aligned}$$

In particular, j is Hadamard differentiable in 0 meaning that the generalized Jacobian is a singleton and moreover non-singular by the above. This allows us to use the Lipschitz inverse function theorem (see Theorem 3.12 and Chapter 2 in [8] for details) to deduce the existence of a neighborhood $U = \tilde{U}_0 \times]-\tilde{\varepsilon}, \tilde{\varepsilon}[$ of 0 such that the restriction $j : U \to j(U) = \Omega' \subset \Omega_0'$ is invertible with Lipschitz-continuous inverse. Eventually, (16) still holds on Ω'. □

Proof of Lemma 4.5. We prove that $\frac{\partial}{\partial \xi_d} \nabla j(\bar{\xi}, \xi_d) = \nabla q\big(j(\bar{\xi}, \xi_d)\big) \nabla j(\bar{\xi}, \xi_d)$ with certain initial conditions holds almost everywhere in U so the desired identities follow from Liouville's formula for Wronski determinants [3].

First recall from the proof of Lemma 4.4 that $\frac{\partial j}{\partial \xi_d} = q\big(j(\bar{\xi}, \xi_d)\big)$ is Lipschitz continuous. Thus, one can differentiate almost everywhere with respect to ξ_d and gets

$$\frac{\partial^2 j}{\partial \xi_d^2}(\bar{\xi}, \xi_d) = \nabla q\big(j(\bar{\xi}, \xi_d)\big) \frac{\partial j}{\partial \xi_d}(\bar{\xi}, \xi_d) \quad , \quad \frac{\partial j}{\partial \xi_d}(\bar{\xi}, 0) = q\Big(x + \sum_{i=1}^{d-1} \xi_i q_i\Big)$$

with the q_i already chosen in the proof of Lemma 4.4. Now consider the initial value problems

$$\frac{\partial \varphi_i}{\partial \xi_d}(\bar{\xi}, \xi_d) = \nabla q\big(j(\bar{\xi}, \xi_d)\big) \varphi_i(\bar{\xi}, \xi_d) \quad , \quad \varphi_i(\bar{\xi}, 0) = q_i \quad , \quad i = 1, \ldots, d-1 \ .$$

These are time-variant linear equations. The set of $\bar{\xi}$ where $\nabla q\big(j(\bar{\xi}, \xi_d)\big)$ does not exist on a non-null set cannot have positive measure since q would not be almost everywhere differentiable anymore. Thus, for almost every $\bar{\xi}$ the derivative ∇q along the trajectory $j(\bar{\xi}, \xi_d)$ exists almost everywhere with respect to ξ_d. Hence, the above initial value problems admit solutions in the sense of Carathéodory almost everywhere in U [26]. Moreover, due to the boundedness of ∇q, the solutions have to be Lipschitz continuous with respect to ξ_d with a Lipschitz constant bound common to almost every $\bar{\xi}$.

Pick a $\bar{\xi}$ such that the trajectory $j(\bar{\xi}, \,\cdot\,)$ meets almost everywhere points where q is differentiable. By the same arguments as above, we can also suppose that $\frac{\partial j}{\partial \xi_i}(\bar{\xi}, \,\cdot\,)$ with $i = 1, \ldots, d-1$ exists almost everywhere since this has to be the case for almost all $\bar{\xi}$. Now

$$\begin{aligned} &\frac{1}{\varepsilon}\big(j(\bar{\xi} + \varepsilon e_i, \xi_d) - j(\bar{\xi}, \xi_d)\big) - \varphi_i(\bar{\xi}, \xi_d) \\ &\qquad = \int_0^{\xi_d} \frac{1}{\varepsilon}\Big(q\big(j(\bar{\xi} + \varepsilon e_i, s)\big) - q\big(j(\bar{\xi}, s)\big)\Big) - \nabla q\big(j(\bar{\xi}, s)\big)\varphi_i(\bar{\xi}, s) \,\mathrm{d}s \end{aligned}$$

which becomes

$$\frac{\partial j}{\partial \xi_i}(\bar{\xi}, \xi_d) - \varphi_i(\bar{\xi}, \xi_d) = \int_0^{\xi_d} \nabla q\big(j(\bar{\xi}, s)\big)\Big(\frac{\partial j}{\partial \xi_i}(\bar{\xi}, s) - \varphi_i(\bar{\xi}, s)\Big) \,\mathrm{d}s$$

in the limit due to a.e. pointwise convergence and the estimate

$$\Big|\frac{1}{\varepsilon}\Big(q\big(j(\bar{\xi} + \varepsilon e_i, \xi_d)\big) - q\big(j(\bar{\xi}, \xi_d)\big)\Big)\Big| \leq C_1 C_2$$

which follows easily from the Lipschitz estimate in the proof of Lemma 4.4 and allows to use Lebesgue's dominated-convergence theorem. With Gronwall's inequality, it follows that $\varphi_i(\bar{\xi}, \xi_d) = \frac{\partial j}{\partial \xi_i}(\bar{\xi}, \xi_d)$ almost everywhere in U.

Finally, denoting

$$q_d = \frac{q(x)}{|q(x)|} \quad , \quad \bar{q}(\bar{\xi}) = q\big(j(\bar{\xi}, 0)\big) \quad , \quad \big(\tilde{q}(\bar{\xi})\big)_i = q_i \cdot \bar{q}(\bar{\xi})$$

leads to, remembering that $q_1, \ldots, q_d$ are orthonormal vectors,

$$\nabla j(\bar{\xi}, 0) = \begin{bmatrix} q_1 & \ldots & q_{d-1} & \bar{q}(\xi) \end{bmatrix} = \begin{bmatrix} q_1 & \ldots & q_d \end{bmatrix} \begin{bmatrix} e_1 & \ldots & e_{d-1} & \tilde{q}(\xi) \end{bmatrix}$$
$$\Rightarrow \qquad \det \nabla j(\bar{\xi}, 0) = q_d \cdot \bar{q}(\xi) = \frac{q(x)}{|q(x)|} \cdot q\big(j(\bar{\xi}, 0)\big)$$

by the determinant product formula, implying the desired initial-value problem (18) as well as (19) by the solution formula for scalar time-variant ordinary differential equations. □

Proof of Lemma 4.6. The proof is analog to the proof presented for Lemma 3.6. First, observe that J according to Lemma 4.5 is weakly differentiable with respect to ξ_d with $\frac{\partial J}{\partial \xi_d} \in L^\infty(U)$ by (18). Moreover, we can obtain, with the help of (19) the a.e. bounds

$$c_k \mathrm{e}^{-\sqrt{d}\|\nabla q\|_\infty |\xi_d|} \leq |J(\bar{\xi}, \xi_d)| \leq \|q\|_\infty \mathrm{e}^{\sqrt{d}\|\nabla q\|_\infty |\xi_d|}$$

for $\xi = (\bar{\xi}, \xi_d)$, hence $J^{-1} = 1/J$ belongs to $L^\infty(U)$ and is weakly differentiable with respect to ξ_d with $\frac{\partial J^{-1}}{\partial \xi_d} \in L^\infty(U)$.

Suppose that $\frac{\partial y}{\partial \xi_d} \in L^1_{\mathrm{loc}}(U)$ exists and fix a $\tilde{z} = J(z \circ j)$ with $z \in \mathcal{C}_0^\infty(j(U))$. Take the dilated versions G_ε of the standard mollifier G and $\varepsilon > 0$ small enough such that all $\tilde{z}_\varepsilon = \tilde{z} * G_\varepsilon$ have their supports inside of a $U' \subset\subset U$. Its easy to see that we have pointwise a.e. convergence and boundedness of

$$\tilde{z}_\varepsilon \to \tilde{z} \quad , \quad \frac{\partial \tilde{z}_\varepsilon}{\partial \xi_d} \to \frac{\partial \tilde{z}}{\partial \xi_d} \qquad , \qquad \|\tilde{z}_\varepsilon\|_\infty \leq \|\tilde{z}\|_\infty \quad , \quad \Big\|\frac{\partial \tilde{z}_\varepsilon}{\partial \xi_d}\Big\|_\infty \leq \Big\|\frac{\partial \tilde{z}}{\partial \xi_d}\Big\|_\infty$$

and in particular $\tilde{z}_\varepsilon \overset{*}{\rightharpoonup} \tilde{z}$ as well as $\frac{\partial \tilde{z}_\varepsilon}{\partial \xi_d} \overset{*}{\rightharpoonup} \frac{\partial \tilde{z}}{\partial \xi_d}$ in $L^\infty(U')$. Thus,

$$\int_U y \frac{\partial \tilde{z}}{\partial \xi_d} \,\mathrm{d}\xi = \lim_{\varepsilon \to 0} \int_U y \frac{\partial \tilde{z}_\varepsilon}{\partial \xi_d} \,\mathrm{d}\xi = -\lim_{\varepsilon \to 0} \int_U \frac{\partial y}{\partial \xi_d} \tilde{z}_\varepsilon \,\mathrm{d}\xi = -\int_U \frac{\partial y}{\partial \xi_d} \tilde{z} \,\mathrm{d}\xi \tag{32}$$

meaning that testing with $\tilde{z}$ is still possible.

The other way around, suppose that a function we denote by $\frac{\partial y}{\partial \xi_d} \in L^1_{\mathrm{loc}}(\Omega)$ satisfies (20) for each function of the type $J(z \circ j)$ with $z \in \mathcal{C}_0^\infty(j(U))$. For a given $\tilde{z} \in \mathcal{C}_0^\infty(U)$, we approximate, for $\varepsilon > 0$ such that $\operatorname{supp} z + B_\varepsilon(0) \subset U' \subset\subset U$, as follows:

$$\tilde{z}_\varepsilon = J(z_\varepsilon \circ j) \qquad , \qquad z_\varepsilon = \Big(\frac{\tilde{z}}{J} \circ j^{-1}\Big) * G_\varepsilon \ .$$

It is clear that $z_\varepsilon \to (J^{-1}\tilde{z}) \circ j^{-1}$ with $\|z_\varepsilon\|_\infty \leq \|(J^{-1}\tilde{z}) \circ j^{-1}\|_\infty$ pointwise almost everywhere. Since j is a Lipeomorphism, $\tilde{z}_\varepsilon \to \tilde{z}$ with $\|\tilde{z}_\varepsilon\|_\infty \leq \|\tilde{z}\|_\infty$ pointwise a.e. Showing the pointwise a.e. boundedness of $\frac{\partial \tilde{z}_\varepsilon}{\partial \xi_d}$ is a little more involved: Note that for $\xi \in U$, we can write, applying the change of variables thereom for Lipeomorphisms (see, e.g. [11]),

$$z_\varepsilon\big(j(\xi)\big) = \int_U G_\varepsilon\big(j(\xi) - j(\zeta)\big) z(\zeta) \,\mathrm{d}\zeta$$

so $\frac{\partial(z_\varepsilon\circ j)}{\partial\xi_d}(\xi)$ becomes

$$\begin{aligned}\frac{\partial(z_\varepsilon\circ j)}{\partial\xi_d}(\xi) &= \int_U \nabla G_\varepsilon\big(j(\xi)-j(\zeta)\big)\cdot\frac{\partial j}{\partial\xi_d}(\xi)\tilde z(\zeta)\,\mathrm{d}\zeta\\ &= \int_U \nabla G_\varepsilon\big(j(\xi)-j(\zeta)\big)\cdot\frac{\partial j}{\partial\xi_d}(\zeta)\tilde z(\zeta)\,\mathrm{d}\zeta\\ &\quad+\int_U \nabla G_\varepsilon\big(j(\xi)-j(\zeta)\big)\cdot\Big(\frac{\partial j}{\partial\xi_d}(\xi)-\frac{\partial j}{\partial\xi_d}(\zeta)\Big)\tilde z(\zeta)\,\mathrm{d}\zeta\\ &= \int_U G_\varepsilon\big(j(\xi)-j(\zeta)\big)\frac{\partial\tilde z}{\partial\xi_d}(\zeta)\,\mathrm{d}\zeta\\ &\quad+\int_U \nabla G_\varepsilon\big(j(\xi)-j(\zeta)\big)\cdot\Big(q\big(j(\xi)\big)-q\big(j(\zeta)\big)\Big)\tilde z(\zeta)\,\mathrm{d}\zeta\end{aligned}$$

while the first summand obeys

$$\begin{aligned}\Big|\int_U G_\varepsilon\big(j(\xi)-j(\zeta)\big)\frac{\partial\tilde z}{\partial\xi_d}(\zeta)\,\mathrm{d}\zeta\Big| &\leq \Big\|\frac{\partial\tilde z}{\partial\xi_d}\Big\|_\infty\int_{\Omega'}\big|G_\varepsilon\big(j(\xi)-x\big)\big|J^{-1}\big(j^{-1}(x)\big)\,\mathrm{d}x\\ &\leq \Big\|\frac{\partial\tilde z}{\partial\xi_d}\Big\|_\infty\|J^{-1}\|_\infty\ .\end{aligned}$$

Exploiting that q is Lipschitz continuous with constant estimate $C_{j(U')}\|\nabla q\|_\infty$, see Proposition 2.3, the second summand can be estimated by

$$\begin{aligned}&\Big|\int_U \nabla G_\varepsilon\big(j(\xi)-j(\zeta)\big)\cdot\Big(q\big(j(\xi)\big)-q\big(j(\zeta)\big)\Big)\tilde z(\zeta)\,\mathrm{d}\zeta\Big|\\ &\qquad\leq \|\tilde z\|_\infty\int_{\Omega'}\big|\nabla G_\varepsilon\big(j(\xi)-x\big)\big|\big|q\big(j(\xi)\big)-q(x)\big|J^{-1}\big(j^{-1}(x)\big)\,\mathrm{d}x\\ &\qquad\leq C_{j(U')}\|\tilde z\|_\infty\|\nabla q\|_\infty\|J^{-1}\|_\infty\int_{\Omega'}\varepsilon^{-d}\Big|\nabla G\Big(\frac{j(\xi)-x}{\varepsilon}\Big)\Big|\,\mathrm{d}x\ .\end{aligned}$$

Together, we have

$$\Big\|\frac{\partial\tilde z_\varepsilon}{\partial\xi_d}\Big\|_\infty\leq\Big\|\frac{\partial J}{\partial\xi_d}\Big\|_\infty\|\tilde z\|_\infty+\|J\|_\infty\|J^{-1}\|_\infty\Big(\Big\|\frac{\partial\tilde z}{\partial\xi_d}\Big\|_\infty+C_{j(U')}\|\tilde z\|_\infty\|\nabla q\|_\infty\|\nabla G\|_1\Big)\ .$$

It follows that $\{\frac{\partial\tilde z_\varepsilon}{\partial\xi_d}\}$ admits a subsequence for which $\frac{\partial\tilde z_\varepsilon}{\partial\xi_d}\overset{*}{\rightharpoonup}\frac{\partial\tilde z}{\partial\xi_d}$ in $L^\infty(U')$ because of the weak*-closedness of $\frac{\partial}{\partial\xi_d}$. Actually, this holds for the whole sequence, hence (32) is valid, so $\frac{\partial y}{\partial\xi_d}$ is indeed the weak derivative with respect to ξ_d. □

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics*. Academic Press, 2nd edition, 2003.

[2] H. W. Alt. *Lineare Funktionalanalysis*. Springer, 3rd edition, 1999.

[3] H. Amann. *Ordinary Differential Equations: An Introduction to Nonlinear Analysis*, volume 13 of *de Gruyter Studies in Mathematics*. de Gruyter, 1990.

[4] J. Appell and P. P. Zabrejko. *Nonlinear Superposition Operators*, volume 95 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 1990.

[5] K. Bredies. *Optimal control of degenerate parabolic equations in image processing: Analysis of evolution equations with variable degeneracy and associated minimization problems*. PhD thesis, University of Bremen, 2007.

[6] K. Bredies. Weak solutions of linear degenerate parabolic equations and an application in image processing. *Communications on Pure and Applied Analysis*, 8(4):1203–1229, 2009.

[7] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.

[8] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*, volume 178 of *Graduate Texts in Mathematics*. Springer, 1998.

[9] J. Deny and J.-L. Lions. Les espaces du type de Beppo Levi. *Annales de l'institut Fourier*, 5:305–370, 1954.
[10] N. Dunford and J. T. Schwartz. *Linear Operators. Part I: General Theory*, volume VII of *Pure and Applied Mathematics*. Wiley Interscience, 1957.
[11] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
[12] T. Frankel. *The Geometry of Physics: An Introduction*. Cambridge University Press, Cambridge, 1997.
[13] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Number 224 in Grundlehren der mathematischen Wissenschaften. Springer, 2nd edition, 1998.
[14] C. E. Gutiérrez. Harnack's inequality for degenerate Schrödinger operators. *Transactions of the American Mathematical Society*, 312(1):403–419, 1989.
[15] T. Kilpeläinen. Weighted Sobolev spaces and capacity. *Annales Academiæ Scientiarum Fennicæ*, 19:95–113, 1994.
[16] T. Kilpeläinen. Smooth approximation in weighted Sobolev spaces. *Commentationes Mathematicae Universitas Carolinae*, 38(1):29–35, 1997.
[17] A. Kufner. *Weighted Sobolev Spaces*, volume 31 of *Teubner-Texte zur Mathematik*. Teubner, 1980.
[18] S. Lang. *Fundamentals of Differential Geometry*. Number 191 in Graduate Texts in Mathematics. Springer, New York, 1999.
[19] N. G. Meyers and J. Serrin. $H = W$. *Proceedings of the National Academy of Sciences of the United States of America*, 51(6):1055–1056, 1964.
[20] N. Miller. Weighted Sobolev spaces and pseudodifferential operators with smooth symbols. *Transactions of the American Mathematical Society*, 269(1):91–109, 1982.
[21] B. Muckenhoupt. Weighted norm inequalities for the Hardy maximal function. *Transactions of the American Mathematical Society*, 165:207–226, 1972.
[22] O. A. Oleĭnik and E. V. Radkevič. *Second Order Equations With Nonnegative Characteristic Form*. American Mathematical Society, 1973.
[23] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
[24] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. Johann Ambrosius Barth, 2nd edition, 1995.
[25] B. O. Turesson. *Nonlinear Potential Theory and Weighted Sobolev Spaces*. Number 1736 in Lecture Notes in Mathematics. Springer, Berlin, 2000.
[26] W. Walter. *Gewöhnliche Differentialgleichungen*. Springer, 6th edition, 1996.
[27] J. Weickert. *Anisotropic diffusion in image processing*. Teubner, 1998.
[28] K. Yosida. *Functional Analysis*. Springer, 6th edition, 1980.
[29] P. Zamboni. Hölder continuity for solutions of linear degenerate elliptic equations under minimal assumptions. *Journal of Differential Equations*, 182(1):121–140, 2002.
[30] E. Zeidler. *Nonlinear Functional Analysis and its Applications I*. Springer-Verlag, 1986.
[31] V. V. Zhikov. Weighted Sobolev spaces. *Sbornik: Mathematics*, 189(8):1139–1170, 1998.