Dedicated to Jim Douglas, Jr., on the occasion of his 75th birthday

APPROXIMATION BY PIECEWISE CONSTANT FUNCTIONS IN A BV METRIC

PAVEL BĚLÍK AND MITCHELL LUSKIN

ABSTRACT. We study the approximation properties of piecewise constant functions with respect to triangular and rectangular finite elements in a metric defined on functions of bounded variation. We apply our results to a thin film model for martensitic crystals and to the approximation of deformations with microstructure.

1. INTRODUCTION

In mathematical models for crystal microstructure [1,2,27], the deformation gradient is nearly piecewise constant in space to enable the deformation to attain a low energy. The length scale of the microstructure is limited by a surface energy associated with the transition from one piecewise constant variant phase to another piecewise constant variant phase [1, 27]. Motivated by these models, numerical methods have been developed and utilized that approximate the deformation gradient by piecewise constant functions and that minimize an energy that includes both an elastic energy and a surface energy proportional to the total variation of the deformation gradient [8].

In this paper, we give approximation results for piecewise constant functions in a metric for functions of bounded variation. We consider these results in the context of a thin film model for martensitic crystals.

We have developed numerical methods for the computation of microstructure in martensitic and ferromagnetic crystals and validated these methods by the development of a numerical analysis of microstructure [4, 6, 7, 14, 27–30]. Related results are given in [10, 12, 14, 20–26, 31, 32].

2. DEFINITIONS OF BV SPACES

In this section, we give the relevant definitions for functions of bounded variation and establish some notation used throughout this paper. Recall that, for a bounded

Lipschitz domain $\Omega \subset \mathbb{R}^n$ and for a given $m \in \mathbb{N}$, the space $BV(\Omega)$ of $L^1(\Omega)$-functions of bounded variation mapping Ω to $\mathbb{R}^m$ can be defined [9] as the set of functions $w \in L^1(\Omega; \mathbb{R}^m)$ such that $\int_\Omega |Dw|_k < \infty$, where

$$\int_\Omega |Dw|_k = \sup \left\{ \sum_{i=1}^m \sum_{j=1}^n \int_\Omega w_i(x)\psi_{ij,j}(x)\, dx : \right.$$
$$\left. \psi \in \mathcal{C}_c^1(\Omega; \mathbb{R}^{m \times n}), \ |\psi(x)|_{k'} \le 1 \text{ for all } x \in \Omega \right\}. \tag{2.1}$$

Here, $|.|_{k'}$ denotes the $\ell_{k'}$ vector norm and $k' = k/(k-1)$ is the dual exponent to k. (Whenever the index k is omitted, it is understood that $k = 2$.) Due to equivalence of norms on finite-dimensional vector spaces, the above definition of $BV(\Omega)$ is independent of k, which may be assumed to take values in $[1, \infty]$. Equivalently, if we define the norms

$$\|w\|_{BV_{q,k}(\Omega)} = \|w\|_{L^q(\Omega)} + \int_\Omega |Dw|_k \quad \text{for } 1 \le q < \infty,$$

then $w \in BV(\Omega)$ if and only if $\|w\|_{BV_{1,k}(\Omega)} < \infty$ for some $k \in [1, \infty]$. The space $BV(\Omega)$, when equipped with the norm $\|.\|_{BV_{1,k}(\Omega)}$, is a Banach space. The range of the functions in the spaces $L^q(\Omega)$ and $BV(\Omega)$ should be clear from context and will usually not be denoted explicitly.

More generally, we can define the spaces $BV_q(\Omega)$ for $1 \le q < \infty$ by

$$BV_q(\Omega) = L^q(\Omega) \cap BV(\Omega).$$

With this notation, we have that $BV(\Omega) = BV_1(\Omega)$.

Since $\ell_{k'}$ is the dual space to ℓ_k, we have for smooth functions w that

$$\int_\Omega |Dw|_k = \int_\Omega |\nabla w|_k,$$

where

$$|\nabla w|_k = \left(\sum_{i,j} |w_{i,j}|^k \right)^{1/k}.$$

If w jumps across a smooth interface $\mathcal{S}$ which divides Ω into two parts, Ω_1 and Ω_2, on each of which w is smooth, then

$$\int_\Omega w_i(x)\psi_{ij,j}(x)\, dx = \int_\mathcal{S} [\![\, w_i(\sigma)\,]\!]\psi_{ij}(\sigma)n_j(\sigma)\, d\sigma - \sum_{l=1}^2 \int_{\Omega_l} w_{i,j}(x)\psi_{ij}(x)\, dx,$$

where $[\![\, w_i(\sigma)\,]\!]$ denotes the difference of the traces of $w_i|_{\Omega_2}$ and $w_i|_{\Omega_1}$ on $\mathcal{S}$ and n denotes the normal vector to $\mathcal{S}$ pointing out of Ω_2 such that $|n|_2 = 1$. If w is constant on Ω_1 and Ω_2, then we have that

$$\int_\Omega w_i(x)\psi_{ij,j}(x)\, dx = [\![\, w_i\,]\!]\big|_\mathcal{S} \int_\mathcal{S} \psi_{ij}(\sigma)n_j(\sigma)\, d\sigma,$$

and since $\ell_{k'}$ is the dual space to ℓ_k, we have in this case that

$$\int_\Omega |Dw|_k = \big| [\![\, w\,]\!]\big|_\mathcal{S}\big|_k \int_\mathcal{S} |n(\sigma)|_k\, d\sigma.$$

If, in addition, $\mathcal{S}$ is planar, then n is independent of $\sigma \in \mathcal{S}$ and we have that

$$\int_\Omega |Dw|_k = \left| [\![\, w \,]\!]_{\mathcal{S}} \right|_k |n|_k \, \text{area} \, \mathcal{S}.$$

We note that $|n|_k$ for $|n|_2 = 1$ depends on the orientation of n unless $k = 2$. This corresponds to the fact that $\int_\Omega |Dw|_k$ is frame-indifferent only if $k = 2$.

3. Approximation results for the thin-film model

In this section, we address some approximation issues concerning the total-variation thin-film model rigorously derived in [9] from a three-dimensional energy with the surface energy modeled by the total variation of the deformation gradient. For our total-variation surface energy model, the bulk energy for a film of thickness $h > 0$ with reference configuration $\Omega_h \equiv \Omega \times (-h/2, h/2)$, where $\Omega \subset \mathbb{R}^2$ is a domain with a Lipschitz continuous boundary $\partial\Omega$, is given by the sum of the surface energy and the elastic energy

$$\kappa \int_{\Omega_h} |D(\nabla u)| + \int_{\Omega_h} \phi(\nabla u, \theta), \tag{3.1}$$

where $u : \Omega_h \to \mathbb{R}^3$ is a deformation, $\int_{\Omega_h} |D(\nabla u)|$ is the total variation of the deformation gradient, κ is a small positive constant, θ is the fixed temperature, and $\phi(\nabla u, \theta)$ is the energy density. We assume that the film adheres to a rigid material on its edge $\Gamma_h = \partial\Omega \times (-h/2, h/2)$, and is thus constrained by

$$u(x_1, x_2, x_3) = y_0(x_1, x_2) + b_0(x_1, x_2)x_3 \quad \text{for } (x_1, x_2, x_3) \in \Gamma_h,$$

where $y_0, b_0 \in W^{1,p}(\Omega; \mathbb{R}^3)$ are such that $\nabla y_0, \nabla b_0 \in BV(\Omega)$.

We have shown in [9] that energy-minimizing deformations u of the bulk energy (3.1) are asymptotically of the form

$$u(x_1, x_2, x_3) = y(x_1, x_2) + b(x_1, x_2)x_3 + o(x_3^2) \quad \text{for } (x_1, x_2, x_3) \in \Omega_h,$$

where (y, b) minimizes the thin film energy

$$\mathcal{E}(y, b) = \kappa \left(\int_\Omega |D(\nabla y | \sqrt{2}b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0| \right) + \int_\Omega \phi(\nabla y | b, \theta) \tag{3.2}$$

over all deformations of finite energy such that $y = y_0$ on $\partial\Omega$. Since the temperature θ will be considered as a given variable in this paper, we will not explicitly denote the dependence of the energy $\mathcal{E}(y, b)$ on the temperature. The graphical interpretation of the pair (y, b) is given in Figure 1. We denote by $(\nabla y | b) \in \mathbb{R}^{3 \times 3}$ the matrix whose first two columns are given by the columns of ∇y and the last column by b. In the above equation, $\int_\Omega |D(\nabla y | \sqrt{2}b)|$ is the total variation of the vector-valued function $(\nabla y | \sqrt{2}b) : \Omega \to \mathbb{R}^{3 \times 3}$.

We assume that the energy density $\phi : \mathbb{R}^{3 \times 3} \times \mathbb{R} \to [0, \infty)$ is a continuous function satisfying the growth condition

$$c_1(|F|^p - 1) \leq \phi(F, \theta) \leq c_2(|F|^p + 1) \quad \text{for all } F \in \mathbb{R}^{3 \times 3} \text{ and } \theta \in \mathbb{R}, \tag{3.3}$$

where c_1 and c_2 are fixed positive constants. We will also assume that $p > 3$ to ensure that deformations with finite energy are uniformly continuous. We then denote the space $\mathcal{A}$ of admissible deformations of the thin film to be

$$\mathcal{A} = \{(y, b) \in W^{1,p}(\Omega; \mathbb{R}^3) \times L^p(\Omega; \mathbb{R}^3) : \nabla y, b \in BV_p(\Omega), \ y = y_0 \text{ on } \partial\Omega\}. \tag{3.4}$$
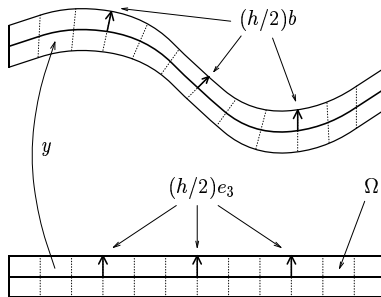
FIGURE 1. A graphical description of the deformation of a three-dimensional thin film of thickness $0 < h \ll 1$. The map y gives the deformation of the planar mid-section, Ω, of the film, while b describes the deformation of the cross-sections.

Numerical simulations using finite element methods should be designed so that they approximate both the energy of the deformed film, $\mathcal{E}(y, b)$, and the deformation of the thin film, $(y, b) \in \mathcal{A}$. While it is clear what it means to approximate the energy, it is not obvious in what sense one should approximate the deformation pair (y, b). To understand this better, we will make several observations.

Using the definition (3.4) of the space $\mathcal{A}$ of admissible deformations of the thin film, one might consider approximating the solution in a variant of the $BV_{p,2}(\Omega)$-norm. However, this would fail even in the simplest case of approximating $w(x) = x$ for $x \in (0, 1)$ by piecewise constant functions—no matter how the approximating functions are chosen, on each subinterval of $(0, 1)$ where the approximating function, $\tilde{w}$, is constant, the difference $w - \tilde{w}$ has variation equal to the length of the subinterval, hence the total variation of $w - \tilde{w}$ is at least 1. Another possibility for $w = (y, b) \in \mathcal{A}$ would be to approximate y in $W^{1,p}(\Omega)$ and b in $L^p(\Omega)$, but the total-variation part of the energy would then not be controlled.

To study the numerical approximation of (3.2), we will consider the metrics on $BV_q(\Omega)$ given by

$$\rho_{q,k}(w_1, w_2) = \|w_1 - w_2\|_{L^q(\Omega)} + \left| \int_\Omega |Dw_1|_k - \int_\Omega |Dw_2|_k \right| \tag{3.5}$$

for $k \in [1, \infty]$. This is partially justified by the conclusion of Lemma 4.1 that functions in $BV_q(\Omega)$ can be approximated arbitrarily closely in this metric by functions in $\mathcal{C}^\infty(\Omega) \cap BV_q(\Omega)$. Also, for the example $w(x) = x$ above, if w_h is defined as, say, the L^2-projection of w into the space of piecewise constant functions defined on a partition of $(0, 1)$ into disjoint intervals of length at most $h > 0$, then clearly $\lim_{h \to 0} \rho_{q,k}(w, w_h) = 0$.

We shall now provide some results that will shed light on the way in which one should approximate the thin-film problem. We start with two lemmas that will be used in the theorem below. These results have been proved in [9] (the first one in a slightly more general form).

Lemma 3.1 (Lower semi-continuity). *Let* $\Omega \subset \mathbb{R}^2$ *be a bounded domain with a Lipschitz continuous boundary. Let* $b, b_k \in BV(\Omega)$, $k = 1, 2, \ldots,$ *satisfy*

$$\lim_{k \to \infty} \|b_k - b\|_{L^1(\Omega)} = 0,$$

and assume that there exists $b_0 \in BV(\Omega)$ *such that* $b_k = b_0$ *on* $\partial\Omega$ *for all* k. *Then, for any* $v \in BV(\Omega)$, *we have*

$$\int_\Omega |D(v|\sqrt{2}\,b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0| \leq \liminf_{k \to \infty} \int_\Omega |D(v|\sqrt{2}\,b_k)|.$$

Lemma 3.2 (Approximation). *Let* $1 \leq q < \infty$ *and let* $\Omega \subset \mathbb{R}^2$ *be a bounded domain with a Lipschitz continuous boundary. Let* $b_0 \in W^{1,q}(\Omega)$ *be such that* $\nabla b_0 \in BV_q(\Omega)$, *let* $b \in BV_q(\Omega)$, *and let* $v \in BV_q(\Omega)$. *Then there exists a family* $\{b_\varepsilon\} \subset W^{1,q}(\Omega)$ *with* $\nabla b_\varepsilon \in BV_q(\Omega)$ *such that* $b_\varepsilon = b_0$ *on* $\partial\Omega$ *for every* $\varepsilon > 0$, *and*

$$\lim_{\varepsilon \to 0} \|b_\varepsilon - b\|_{L^q(\Omega)} = 0,$$

$$\lim_{\varepsilon \to 0} \int_\Omega |D(v|\sqrt{2}\,b_\varepsilon)| = \int_\Omega |D(v|\sqrt{2}\,b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0|.$$

Theorem 3.1. *Let* $\{(y_k, b_k)\} \subset \mathcal{A}$ *be a given sequence of admissible deformations of the thin film.*

(a) Assume that

$$(y_k, b_k) \to (y, b) \quad \text{in } W^{1,p}(\Omega) \times L^p(\Omega),$$

$$\int_\Omega |D(\nabla y_k | \sqrt{2}\,b_k)| \to \int_\Omega |D(\nabla y | \sqrt{2}\,b)| \tag{3.6}$$

as $k \to \infty$ *or, equivalently,*

$$\rho_{p,2}\left((\nabla y_k, b_k), (\nabla y, b)\right) \to 0 \quad \text{as } k \to \infty.$$

Then we have convergence of the energy, that is, there exists a subsequence of $\{(y_k, b_k)\}$, *again denoted by* $\{(y_k, b_k)\}$, *such that*

$$\mathcal{E}(y_k, b_k) \to \mathcal{E}(y, b) \quad \text{as } k \to \infty.$$

(b) Assume that $\{(y_k, b_k)\}$ *is an energy-minimizing sequence, that is,*

$$\mathcal{E}(y_k, b_k) \to \min_{(\tilde{y}, \tilde{b}) \in \mathcal{A}} \mathcal{E}(\tilde{y}, \tilde{b}) \quad \text{as } k \to \infty.$$

Then there exists a subsequence of $\{(y_k, b_k)\}$, *again denoted by* $\{(y_k, b_k)\}$, *and a pair* $(y, b) \in \mathcal{A}$ *such that*

$$(y_k, b_k) \rightharpoonup (y, b) \quad \text{in } W^{1,p}(\Omega) \times L^p(\Omega),$$

$$(y_k, b_k) \to (y, b) \quad \text{in } W^{1,1}(\Omega) \times L^1(\Omega),$$

$$\int_\Omega |D(\nabla y_k|\sqrt{2}\,b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| \to \int_\Omega |D(\nabla y|\sqrt{2}\,b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0| \tag{3.7}$$

as $k \to \infty$. *Moreover,* (y, b) *minimizes the energy functional, that is,*

$$\mathcal{E}(y, b) = \min_{(\tilde{y}, \tilde{b}) \in \mathcal{A}} \mathcal{E}(\tilde{y}, \tilde{b}) = \lim_{k \to \infty} \mathcal{E}(y_k, b_k).$$

Proof. (a) That (3.6) implies (3.7) follows from Theorem 2.11 of [19], which says that the convergence of a sequence in $L^1(\Omega)$ and the convergence of its total variation implies the L^1-convergence of the traces on $\partial\Omega$ of the sequence. The convergence of the elastic part of the energy

$$\int_\Omega \phi(\nabla y_k | b_k) \to \int_\Omega \phi(\nabla y | b) \quad \text{as } k \to \infty.$$

follows from the dominated convergence theorem using the upper bound (3.3) from growth of the energy density ϕ and the assumption that $(\nabla y_k, b_k) \to (\nabla y, b)$ in $L^p(\Omega) \times L^p(\Omega)$.

(b) We give here a sketch of the proof for which more details can be found in [9]. Using the lower bound from the growth estimate (3.3) on the energy density ϕ, the compactness theorem for functions of bounded variation, and the trace theorem, we have that there exists a pair $(y, b) \in \mathcal{A}$ and a subsequence of $\{(y_k, b_k)\}$, again denoted by $\{(y_k, b_k)\}$, such that

$$
\begin{aligned}
(y_k, b_k) &\rightharpoonup (y, b) &&\text{in } W^{1,p}(\Omega) \times L^p(\Omega), \\
(y_k, b_k) &\to (y, b) &&\text{in } W^{1,1}(\Omega) \times L^1(\Omega), \\
(\nabla y_k, b_k) &\to (\nabla y, b) &&\text{almost everywhere in } \Omega.
\end{aligned}
$$

Using the approximation Lemma 3.2, we can construct, for every k, a one-parameter family

$$\{b_k^\varepsilon \in W^{1,p}(\Omega) : \ \varepsilon > 0, \ \nabla b_k^\varepsilon \in BV(\Omega), \ b_k^\varepsilon = b_0 \text{ on } \partial\Omega\}$$

such that

$$
\left.
\begin{aligned}
b_k^\varepsilon &\to b_k \quad \text{in } L^p(\Omega), \\
\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k^\varepsilon)| &\to \int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0|
\end{aligned}
\right\} \quad \text{as } \varepsilon \to 0.
$$

For every k, we can find $\varepsilon_k > 0$ such that

$$\lim_{k\to\infty} \|b_k^{\varepsilon_k} - b\|_{L^1(\Omega)} = 0$$

and

$$\left| \int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| - \int_\Omega |D(\nabla y_k^{\varepsilon_k} | \sqrt{2}\, b_k^{\varepsilon_k})| \right| \to 0 \qquad (3.8)$$

as $k \to \infty$. Then, since $b_k^{\varepsilon_k} = b_0$ on $\partial\Omega$, we can use the semi-continuity Lemma 3.1 and (3.8) to obtain

$$\int_\Omega |D(\nabla y | \sqrt{2}\, b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0| \leq \liminf_{k\to\infty} \int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k^{\varepsilon_k})|$$

$$= \liminf_{k\to\infty} \left[\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| \right].$$

Similarly, using the almost-everywhere convergence of $(\nabla y_k, b_k)$, continuity of the energy density ϕ, and Fatou's lemma, we have

$$\int_\Omega \phi(\nabla y | b) \leq \liminf_{k\to\infty} \int_\Omega \phi(\nabla y_k | b_k).$$

Therefore, since $\{(y_k, b_k)\}$ is an energy-minimizing sequence, we have that

$$\mathcal{E}(y, b) \leq \liminf_{k \to \infty} \left[\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| \right] + \liminf_{k \to \infty} \int_\Omega \phi(\nabla y_k | b_k)$$

$$\leq \liminf_{k \to \infty} \left[\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| + \int_\Omega \phi(\nabla y_k | b_k) \right]$$

$$\leq \liminf_{k \to \infty} \mathcal{E}(y_k, b_k) \leq \mathcal{E}(y, b),$$

and that (y, b) is a minimizer of $\mathcal{E}$. It then also follows that, in fact, for a further subsequence of $\{(y_k, b_k)\}$ we have that

$$\lim_{k \to \infty} \left[\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| \right] = \int_\Omega |D(\nabla y | \sqrt{2}\, b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0|$$

and

$$\lim_{k \to \infty} \int_\Omega \phi(\nabla y_k | b_k) = \int_\Omega \phi(\nabla y | b).$$

$\square$

Remark 3.1. *An example that (3.7) does not imply (3.6) can be constructed as follows. Let $\Omega = (0, 1) \times (0, 1)$, $b_0 \equiv 1$, $b \equiv 0$, and $b_k(x_1, x_2) = 2^{k-1} |x_1 - 1/2|^k$ for $(x_1, x_2) \in \Omega$. Then $b_k \to b$ in $L^p(\Omega)$ for $1 \leq p < \infty$ and for $y_k(x) = y(x) = x$ we have*

$$\int_\Omega |D(\nabla y | \sqrt{2}\, b)| = 0, \qquad \int_{\partial\Omega} |b - b_0| = 4,$$

$$\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| = \sqrt{2}, \quad \int_{\partial\Omega} |b_k - b_0| \to 3,$$

so that

$$\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| + \sqrt{2} \int_{\partial\Omega} |b_k - b_0| \to \int_\Omega |D(\nabla y | \sqrt{2}\, b)| + \sqrt{2} \int_{\partial\Omega} |b - b_0|,$$

but

$$\int_\Omega |D(\nabla y_k | \sqrt{2}\, b_k)| \nrightarrow \int_\Omega |D(\nabla y | \sqrt{2}\, b)|.$$

Remark 3.2. *It follows from the first part of Theorem 3.1 that if one uses finite element subspaces of $\mathcal{A}$ that allow approximating elements of $\mathcal{A}$ in the metric $\rho_{p,2}(\cdot, \cdot)$ as defined in (3.5), where $w_1 = (\nabla y_1 | b_1)$ and $w_2 = (\nabla y_2 | b_2)$, then one can also approximate the energy, $\mathcal{E}$.*

On the other hand, if one only knows that the minimum energy can be approximated in the sense of the assumption of part (b) of Theorem 3.1, then one can only conclude that (3.7) holds, but not necessarily (3.6), as was illustrated in Remark 3.1.

4. APPROXIMATION OF BV FUNCTIONS BY PIECEWISE POLYNOMIAL FUNCTIONS

In this section, we address the issue of approximating functions in $BV(\Omega)$ by piecewise polynomial finite element functions. Many papers have been devoted to the approximation of functions of bounded variation in the metrics $\rho_{q,k}(\cdot, \cdot)$, mostly for $q = 1, 2$ and $k = 1$ (for example, [11, 17, 18]). Reasons to justify the choice $k = 1$ have been given, and we will mention them below. However, there are also good reasons to use the metrics corresponding to $k = 2$. One of them is the fact that

the total-variation part of the energy functional (3.2) is frame-indifferent only when $k = 2$, and frame indifference is an important assumption in continuum mechanics.

A related reason, equally important, is that the total variation is used to model surface energy and the surface area is measured correctly only when $k = 2$ [19]. More precisely, if $E \subset \mathbb{R}^n$ is a bounded set with a boundary ∂E that is of class $\mathcal{C}^2$ and if χ_E is the characteristic function of E, then $\chi_E \in BV(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} |D\chi_E| = \mathcal{H}_{n-1}(\partial E),$$

where $\mathcal{H}_{n-1}$ denotes the $(n-1)$-dimensional Hausdorff measure. This would not, in general, be the case for $k \neq 2$. An important consequence of this fact is the appearance of the boundary integral in the 2-dimensional thin-film model (3.2).

Much of the research on the approximation of the total variation has been devoted to image processing and to rectangular Cartesian grids [11, 15, 17, 33]. On the other hand, we shall be mainly concerned with polygonal domains $\Omega \subset \mathbb{R}^2$ and triangular finite elements with applications in continuum mechanics. This allows more flexibility of the underlying geometry and also allows adjusting the finite element mesh to features of the solution. As an example, in our earlier simulations of the behavior of martensitic thin films [3,6,8], the solution was expected to change phase across the diagonals of the square computational domain. If a feature like this is known *a priori*, there is no reason to use a mesh that does not conform to this feature. However, it is not possible to simultaneously orient rectangular elements to these variant phase boundaries and to the crystal boundary. Hence, much finer Cartesian meshes are needed to capture the variant phase boundary to the accuracy that can be obtained from appropriate triangular meshes. Moreover, as we will see later, rectangular finite elements would, in general, fail to approximate the energy (3.2).

Our ultimate goal is to understand the approximation properties of piecewise constant functions in metrics using the total variation. However, we first address the case of approximating functions of bounded variation by continuous piecewise linear functions. This case turns out to be much simpler than for piecewise constant functions. Let us first recall some general results (see, for example, [11, 19]).

Lemma 4.1 (Approximation in $BV(\Omega)$ by smooth functions). *Let $1 \leq q < \infty$, let $1 \leq k \leq \infty$, and let Ω be a bounded open set in $\mathbb{R}^n$ with a Lipschitz continuous boundary, $\partial\Omega$.*

(a) For every $v \in BV_q(\Omega)$, there exists a sequence $\{v_j\} \subset \mathcal{C}^\infty(\Omega)$ such that

$$\lim_{j \to \infty} \|v_j - v\|_{L^q(\Omega)} = 0,$$

$$\lim_{j \to \infty} \int_\Omega |Dv_j|_k = \int_\Omega |Dv|_k$$

or equivalently,

$$\lim_{j \to \infty} \rho_{q,k}(v_j, v) = 0.$$

Moreover, the trace of each v_j on $\partial\Omega$ coincides with the trace of v.

(b) For every $v \in BV_q(\Omega)$, there exists a sequence $\{\bar{v}_j\} \subset \mathcal{C}^\infty(\bar{\Omega})$ such that

$$\lim_{j\to\infty} \|\bar{v}_j - v\|_{L^q(\Omega)} = 0,$$

$$\lim_{j\to\infty} \int_\Omega |D\bar{v}_j|_k = \int_\Omega |Dv|_k,$$

or equivalently,

$$\lim_{j\to\infty} \rho_{q,k}(v_j, v) = 0.$$

Remark 4.1. *Note that in part (b) we cannot have the result on the traces of $\bar{v}_j$ as in part (a). However, it follows from Theorem 2.11 of [19] that the traces of $\bar{v}_j$ converge to the trace of v in $L^1(\partial\Omega)$.*

With this lemma in hand, we can now prove the following result (see [11, 15]). The finite-element terminology can be found in [5, 13].

Theorem 4.1. *Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain, let $k \in [1, \infty]$, let $h_0 > 0$, and assume that a regular (or quasi-uniform) family $\{\tau_h : 0 < h \leq h_0\}$ of triangulations of Ω is given such that the mesh-size of τ_h is at most h. Let $\mathcal{A}_h^1$ denote the space of continuous piecewise linear functions corresponding to the triangulation τ_h. Let $u \in BV_q(\Omega)$ for some $q \in [1, \infty)$. Then there exists a family $\{u_h \in \mathcal{A}_h^1 : 0 < h \leq h_0\}$ such that*

$$\rho_{q,k}(u, u_h) \to 0 \quad as \ h \to 0.$$

Proof. Recall first that if $v \in \mathcal{C}^\infty(\bar{\Omega})$ and if $\Pi_h v \in \mathcal{A}_h^1$ is its piecewise linear interpolant, that is, $v = \Pi_h v$ at the vertices of the triangulation τ_h, then there exists a constant $C(v)$, depending on v but independent of h, such that [13]

$$\|v - \Pi_h v\|_{W^{1,q}(\Omega)} \leq C(v)h. \tag{4.1}$$

Also, given $\varepsilon > 0$, we can use Lemma 4.1 and choose $v_\varepsilon \in \mathcal{C}^\infty(\bar{\Omega})$ such that

$$\rho_{q,k}(u, v_\varepsilon) < \varepsilon/2.$$

Now, in view of (4.1), we can find $h_0(\varepsilon) > 0$ such that

$$\rho_{q,k}(v_\varepsilon, \Pi_h v_\varepsilon) < \varepsilon/2 \quad \text{for all } h \leq h_0(\varepsilon).$$

The theorem now follows from the triangle inequality. $\square$

Remark 4.2. *Since the finite element spaces of discontinuous piecewise linear functions contain the spaces of continuous piecewise linear functions, Theorem 4.1 also holds in the discontinuous piecewise linear case. Theorem 4.1 will be used in proving a similar result for piecewise constant functions when $k = 2$ (Theorem 4.3).*

We now proceed with the analysis for approximating functions of bounded variation by piecewise constant functions. Let us first note that since

$$|x|_q \leq |x|_p \quad \text{for } x \in \mathbb{R}^n \text{ and } 1 \leq p < q \leq \infty,$$

we also have

$$\int_\Omega |Du|_q \leq \int_\Omega |Du|_p \quad \text{for } u \in BV(\Omega) \text{ and } 1 \leq p < q \leq \infty. \tag{4.2}$$

These inequalities can be strict; for instance, if $u(x_1, x_2) = x_1 + x_2$ for $(x_1, x_2) \in \Omega = (0,1) \times (0,1)$, then it can be checked that

$$\int_\Omega |Du|_k = 2^{1/k} \quad \text{for } 1 \le k \le \infty,$$

where, as usual, $1/\infty$ is defined as 0.

The inequality (4.2) is the reason one cannot approximate a general function u of bounded variation by piecewise constant functions u_h corresponding to Cartesian grids (of mesh-size h) in any of the metrics $\rho_{q,k}$ with $k > 1$ [11, 17]. Due to the fact that all of the vector norms $|\,.\,|_k$, $1 \le k \le \infty$, give identical values for the vectors of the canonical basis of $\mathbb{R}^n$, it is easily seen that on such meshes we have for all scalar-valued functions u_h that

$$\int_\Omega |Du_h|_{k_1} = \int_\Omega |Du_h|_{k_2} \quad \text{for all } 1 \le k_1, k_2 \le \infty.$$

Hence, if u is such that $\int_\Omega |Du|_{k_0} < \int_\Omega |Du|_1$ for some $k_0 > 1$ and if $u_h \to u$ in $L^1(\Omega)$, then

$$\int_\Omega |Du|_k < \int_\Omega |Du|_1 \le \liminf_{h \to 0} \int_\Omega |Du_h|_1 = \liminf_{h \to 0} \int_\Omega |Du_h|_k \qquad (4.3)$$

for all $k \in [k_0, \infty]$. The second inequality in (4.3) follows from the lower semi-continuity of the total variation [16, 19]. Moreover, we see in the following Theorem 4.2 that for a special choice of u_h, this inequality can be turned into equality [11, 17].

Theorem 4.2. *Let $h_0 > 0$. Let $\{\tau_h : 0 < h \le h_0\}$ be a family of Cartesian triangulations of Ω such that the mesh-size of τ_h is at most h. Let $\mathcal{V}_h$, $0 < h \le h_0$, denote the spaces of piecewise constant functions corresponding to the grids τ_h. Let $u \in BV_q(\Omega)$ for some $q \in [1, \infty)$. Then there exists a family $\{u_h \in \mathcal{V}_h : 0 < h \le h_0\}$ such that*

$$\rho_{q,1}(u, u_h) \to 0 \quad \text{as } h \to 0.$$

The above theorem and the considerations preceding it provide a complete description of approximation of functions of bounded variation by piecewise constant functions on Cartesian grids. The following example shows that this theorem fails for non-Cartesian rectangular grids.

Example 4.1. Consider the domain $\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| < 1\}$ and the function $u : \Omega \to \mathbb{R}$ defined by $u(x) = 0$ for $x_1 < 0$ and $u(x) = 1$ for $x_1 > 0$. Then

$$\int_\Omega |Du|_k = 2 \quad \text{for } 1 \le k \le \infty.$$

For each $n \in \mathbb{N}$, let $h = \sqrt{2}/2^n$ and define a grid on Ω by subdividing it into square elements of edge length h (Figure 2). Then the normal vectors to each edge are given by $n = (1/\sqrt{2}, \pm 1/\sqrt{2})$ so that
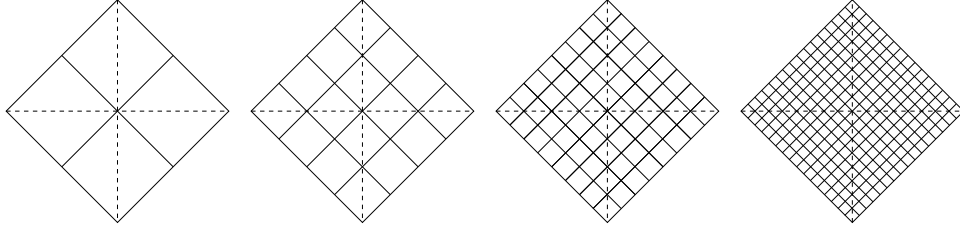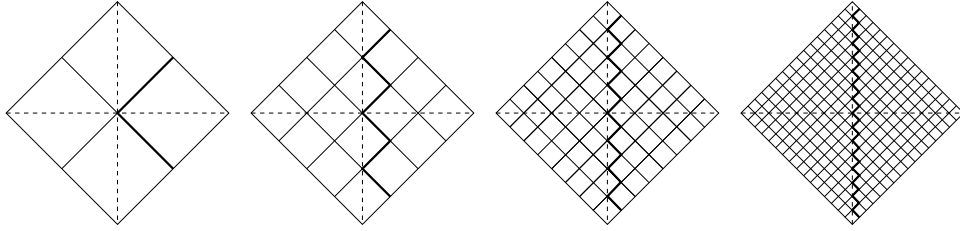
$$|n|_k = 2^{1/k - 1/2},$$

FIGURE 2. Rectangular meshes defined in Example 4.1.

FIGURE 3. Division of Ω into two disjoint parts to obtain convergence in Example 4.1.

and we have that $|n|_{k_2} < |n|_{k_1}$ for $1 \leq k_1 < k_2 \leq \infty$. We then have for any (non-constant) piecewise constant function u_h on this grid that

$$\int_\Omega |Du_h|_{k_2} < \int_\Omega |Du_h|_{k_1} \quad \text{for } 1 \leq k_1 < k_2 \leq \infty,$$

and we thus have for such a sequence $u_h \to u$ in $L^1(\Omega)$ as $h \to 0$ that

$$\int_\Omega |Du|_1 = \int_\Omega |Du|_{k_2} \leq \liminf_{h \to 0} \int_\Omega |Du_h|_{k_2} < \liminf_{h \to 0} \int_\Omega |Du_h|_{k_1}$$

for $1 \leq k_1 < k_2 \leq \infty$. Hence, we see that in this case the only candidate k to obtain convergence of u_h to u in the metric $\rho_{1,k}$ is $k = \infty$. Dividing Ω into two parts as shown in Figure 3 and defining $u_h = 0$ in the left part of Ω and $u_h = 1$ in the right part of Ω, we see that $|[\![u_h]\!]| = 1$ across the dividing curve, whose length converges to $2\sqrt{2}$. Since $|n|_\infty = 1/\sqrt{2}$, we have for $1 \leq q < \infty$ that

$$\left.\begin{array}{l} u_h \to u \text{ in } L^q(\Omega), \\[2mm] \displaystyle\int_\Omega |Du_h|_\infty \to 2 = \int_\Omega |Du|_\infty \end{array}\right\} \quad \text{as } h \to 0 \text{ for } h = \sqrt{2}/2^n.$$

We now proceed with the case of triangular meshes. This issue has been raised in [11] and to some extent answered in [17]. While it is clear that the same approximation result as in Theorem 4.2 holds for triangular meshes obtained by partitioning Cartesian grids, it is more interesting to allow more general grids and also allow metrics other than $\rho_{q,k}$ with $k = 1$.

Example 4.2. We can obtain a counterexample to the approximation in $\rho_{q,k}$ for $1 \leq k < \infty$ by piecewise constant functions on triangular meshes by considering
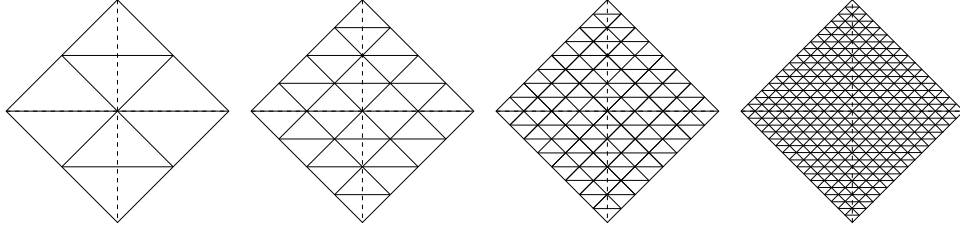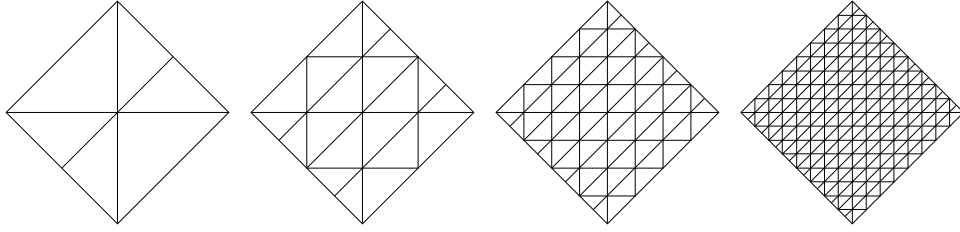
FIGURE 4. Triangular meshes used in Example 4.2.

FIGURE 5. Meshes $\tilde{\tau}_1$, $\tilde{\tau}_2$, $\tilde{\tau}_3$, and $\tilde{\tau}_4$ defined in Example 4.3.

the function $u(x)$ defined in Example 4.1 and by dividing each square element in Figure 2 into two triangular elements by a horizontal edge as shown in Figure 4.

We then see that

$$\int_\Omega |Du_h|_{k_2} < \int_\Omega |Du_h|_{k_1} \quad \text{for } 1 \leq k_1 < k_2 \leq \infty$$

unless u_h is constant on each of the strips $jh < x_2 < (j+1)h$ for $j = -2^n, \ldots, 2^n - 1$. If $u_h \to u$ in $L^1(\Omega)$ as $h \to 0$, we have again that

$$\int_\Omega |Du|_1 = \int_\Omega |Du|_{k_2} \leq \liminf_{h \to 0} \int_\Omega |Du_h|_{k_2} < \liminf_{h \to 0} \int_\Omega |Du_h|_{k_1}$$

for $1 \leq k_1 < k_2 \leq \infty$.

Finally, the sequence u_h of piecewise constant functions with respect to the rectangular mesh constructed in Example 4.1 such that $u_h \to u$ in $\rho_{1,\infty}$ can also be used for this example since each u_h is also a piecewise constant function on the triangulation.

Example 4.3. We next show that a statement similar to that of Theorem 4.2 cannot hold for triangular meshes and $1 < k \leq \infty$. We shall consider the domain $\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| < 1\}$ and the function $u : \Omega \to \mathbb{R}$ defined by $u(x) = x_1 + x_2$. We note that $\nabla u(x) \equiv (1, 1)$ so that

$$\int_\Omega |Du|_k = 2^{1/k}|\Omega| = 2^{1+1/k} \quad \text{for } 1 \leq k \leq \infty.$$

We shall assume that a sequence of triangulations $\{\tilde{\tau}_n\}_{n=0}^{\infty}$ is given such that the edges of $\tilde{\tau}_n$ are formed by the boundary of Ω and the lines

$$\left.\begin{array}{c} x_1 = i/2^n \\ x_2 = i/2^n \\ x_2 = x_1 + i/2^n \end{array}\right\} \quad \text{for } i = -2^n + 1, \ldots, 0, \ldots, 2^n - 1.$$

(See Figure 5.) Notice that each $\tilde{\tau}_{n+1}$ is a refinement of $\tilde{\tau}_n$, that the diameter of the largest triangle of $\tilde{\tau}_n$ is $2^{1/2-n}$, and that all triangles in each triangulation are similar; hence, if $h_0 > 0$, we can define a regular (or quasi-uniform) family of triangulations $\{\tau_h : \ 0 < h \leq h_0\}$ by

$$\tau_h = \begin{cases} \tilde{\tau}_0 & \text{if } 2^{1/2} \leq h \leq h_0, \\ \tilde{\tau}_n & \text{if } 2^{1/2-n} \leq h < 2^{3/2-n} \quad \text{for } n \in \mathbb{N}. \end{cases}$$

We now consider the spaces $\mathcal{A}_h^0$ of piecewise constant functions corresponding to the triangulations τ_h. We claim that

$$0 = \lim_{h \to 0} \inf_{\tilde{u}_h \in \mathcal{A}_h^0} \rho_{q,1}(u, \tilde{u}_h) < \lim_{h \to 0} \inf_{\tilde{u}_h \in \mathcal{A}_h^0} \rho_{q,k}(u, \tilde{u}_h) \quad \text{for } k > 1. \tag{4.4}$$

Clearly, the equality in (4.4) follows from Theorem 4.2. To show the strict inequality in (4.4), we shall assume the existence of a family $\{\bar{u}_h \in \mathcal{A}_h^0\}$ such that

$$\lim_{h \to 0} \rho_{q,k}(u, \bar{u}_h) = 0 \tag{4.5}$$

and obtain a contradiction. Since the sequence $\{\bar{u}_h\}$ converges to u in $L^q(\Omega)$, it does so also in $L^1(\Omega)$. If we can modify each $\bar{u}_h$ to obtain $u_h \in \mathcal{A}_h^0$ such that u_h is constant across the diagonal edges $x_2 = x_1 \pm i/2^n$ of the triangulation $\tilde{\tau}_n$ corresponding to $\mathcal{A}_h^0$, and such that (4.5) implies

$$\lim_{h \to 0} \rho_{1,k}(u, u_h) = 0, \tag{4.6}$$

then we obtain a contradiction, since, in view of (4.3), the total-variation part of $\rho_{1,k}(u, u_h)$ cannot converge to zero.

To construct u_h from $\bar{u}_h$, let $N = 2^{n+1}$ and consider the N parallel strips

$$S_i \equiv \{(x_1, x_2) \in \Omega : \ x_1 + i/2^n < x_2 < x_1 + (i+1)/2^n\} \quad \text{for } i = -2^n, \ldots, 2^n - 1.$$
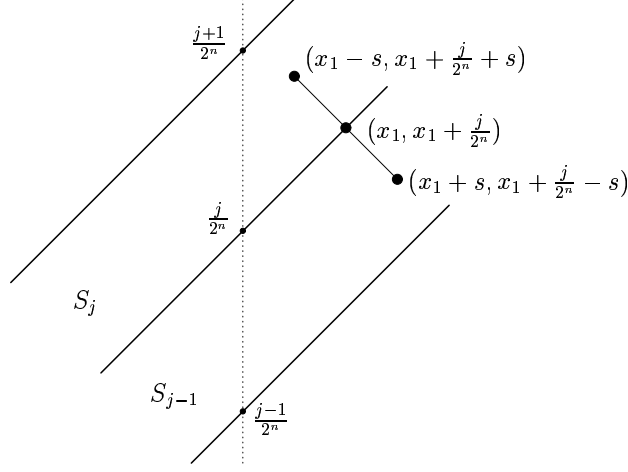
We define

$$v_i|_{S_i} = \bar{u}_h|_{S_i} \quad \text{for } i = -2^n, \ldots, 2^n - 1,$$

and then extend each v_i to Ω by repeated reflection about the lines $x_2 = x_1 + j/2^n$ for $j = -2^n + 1, \ldots, 0, \ldots, 2^n - 1$. That is, for $(x_1 - s, x_1 + \frac{j}{2^n} + s) \in S_j$ with $0 < s < 1/2^{n+1}$, we have that $(x_1 + s, x_1 + \frac{j}{2^n} - s) \in S_{j-1}$ and (see Figure 6)

$$v_i\left(x_1 - s, x_1 + \frac{j}{2^n} + s\right) = v_i\left(x_1 + s, x_1 + \frac{j}{2^n} - s\right). \tag{4.7}$$

Finally, we define u_h to be the average of the v_i,

$$u_h = \frac{1}{N} \sum_{i=1}^{N} v_i.$$

FIGURE 6. The construction of v_i by (4.7).

Then, using the symmetries of u and the v_i, we have that

$$
\int_\Omega |u - u_h| = \int_\Omega \left| u - \frac{1}{N} \sum_{i=1}^{N} v_i \right|
$$

$$
= \sum_{i=1}^{N} \frac{1}{N} \int_\Omega |u - v_i|
$$

$$
= \sum_{i=1}^{N} \int_{S_i} |u - v_i|
$$

$$
= \sum_{i=1}^{N} \int_{S_i} |u - \bar{u}_h|
$$

$$
= \int_\Omega |u - \bar{u}_h|,
$$

so $u_h \to u$ in $L^1(\Omega)$ since $\bar{u}_h \to u$ in $L^1(\Omega)$.

Also, since u_h was constructed so that it has no jumps across the diagonal edges, all of the variation of u_h is associated with the edges parallel to the coordinate axes. We can associate these edges parallel to the coordinate axes into the piecewise linear curves

$$
L_\ell = \cup_{i=-2^n}^{2^n-1} e_{\ell i} \quad \text{for } \ell = 1, \ldots, 2^{n+1}
$$

obtained by reflecting an edge $e_{\ell i} \subset S_i$ about the lines $x_2 = x_1 + i/2^n$ for $i = -2^n + 1, \ldots, 0, \ldots, 2^n - 1$ to obtain $e_{\ell,i-1} \subset S_{i-1}$. Then since $|e_{\ell i}|$ is independent

of ℓ and i, we have that

$$\int_\Omega |Du_h|_k = \sum_{\ell=1}^{2^{n+1}} \sum_{i=1}^{2^n-1} |[\![u_h]\!]_{e_{\ell i}}| |e_{\ell i}|$$

$$= \sum_{\ell=1}^{2^{n+1}} \sum_{i=1}^{2^n-1} \left| \left[\!\!\left[\frac{1}{N} \sum_{j=1}^N v_j \right]\!\!\right]_{e_{\ell i}} \right| |e_{\ell i}|$$

$$\leq \sum_{\ell=1}^{2^{n+1}} \sum_{i=1}^{2^n-1} \frac{1}{N} \sum_{j=1}^N \left| [\![v_j]\!]_{e_{\ell i}} \right| |e_{\ell i}|$$

$$\leq \sum_{\ell=1}^{2^{n+1}} \sum_{i=1}^{2^n-1} \left| [\![v_i]\!]_{e_{\ell i}} \right| |e_{\ell i}|$$

$$\leq \sum_{\ell=1}^{2^{n+1}} \sum_{i=1}^{2^n-1} \left| [\![\bar{u}_h]\!]_{e_{\ell i}} \right| |e_{\ell i}|$$

$$= \int_\Omega |D\bar{u}_h|_k.$$

In view of the lower semi-continuity result $\int_\Omega |Du|_k \leq \liminf_{h\to 0} \int_\Omega |Du_h|_k$ and the convergence $\int_\Omega |D\bar{u}_h|_k \to \int_\Omega |Du|_k$, we have that

$$\int_\Omega |Du_h|_k \to \int_\Omega |Du|_k \quad \text{as } h \to 0.$$

We have thus demonstrated that (4.6) holds.

From Example 4.3, it is clear that it is not enough to just keep refining the mesh to obtain better results. However, it is not immediately obvious whether there exist families of triangular meshes for which a variant of Theorem 4.1 with $\mathcal{A}_h^1$ replaced by $\mathcal{A}_h^0$ holds. It is shown in [17] that if Ω is discretized with a smooth curvilinear grid and the grid conforms to the level curves of u, then such an approximation result holds with $k = 2$. We present a different approach, providing a basis for a numerical algorithm. For the sake of clarity of exposition, we first present the result with $k = 2$ and then the general result.

Theorem 4.3. *Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain and let $h_0 > 0$. Then, given a scalar-valued $u \in BV_q(\Omega)$ for some $q \in [1, \infty)$, there exists a family $\{\tau_h : 0 < h \leq h_0\}$ of triangulations of Ω such that the mesh-size of τ_h is at most h, and functions $u_h \in \mathcal{A}_h^0$, where $\mathcal{A}_h^0$ denotes the space of piecewise constant functions corresponding to the triangulation τ_h, such that*

$$\rho_{q,2}(u, u_h) \to 0 \quad \text{as } h \to 0.$$

Proof. From Theorem 4.1, we know that for any regular family $\{\tilde{\tau}_h : 0 < h \leq h_0\}$ of triangulations of Ω there exists a family of functions $\tilde{u}_h \in \mathcal{A}_h^1$, where $\mathcal{A}_h^1$ denotes the space of continuous piecewise linear functions corresponding to the triangulation $\tilde{\tau}_h$, such that

$$\rho_{q,2}(u, \tilde{u}_h) \to 0 \quad \text{as } h \to 0.$$

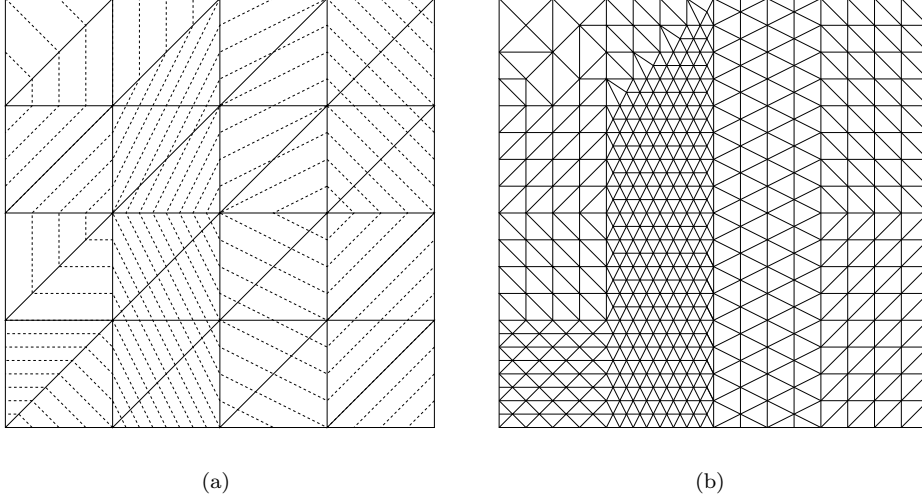(a)   (b)

FIGURE 7. (a) The piecewise linear level curves L_i^n. (b) A new triangulation obtained from the piecewise linear level curves.

For each h, we show that there exists another triangular mesh, τ_h, and a function $u_h \in \mathcal{A}_h^0$, where $\mathcal{A}_h^0$ is the space of piecewise constant functions corresponding to τ_h, such that u_h approximates $\tilde{u}_h$ arbitrarily closely in the metric $\rho_{q,2}$. This will establish the claim of the theorem.

We define $m, M \in \mathbb{R}$ by

$$m = \min_{x \in \bar{\Omega}} \tilde{u}_h(x) \qquad \text{and} \qquad M = \max_{x \in \bar{\Omega}} \tilde{u}_h(x).$$

For $n \in \mathbb{N}$, we define the (piecewise linear) level curves (see Figure 7(a))

$$L_i^n = \partial\{x \in \bar{\Omega} : \ \tilde{u}_h(x) = m + \frac{i}{2^n}(M - m)\} \quad \text{for } i = 0, \ldots, 2^n.$$

We note that we have defined L_i^n to be the *boundaries* of the level sets of $\tilde{u}_h$. The reason is that a new triangulation τ_h^n will be defined for all n such that the union $\cup_{i=0}^{2^n} L_i^n$ will be a subset of the set of all edges of the new triangulation, and hence it is necessary to avoid cases when the level sets have nonempty interior. However, this only happens when $\nabla \tilde{u}_h \equiv 0$ on some triangles of $\tilde{\tau}_h$.

Next, we define a piecewise constant function u_h^n in the following way. If the level set $\{x \in \bar{\Omega} : \ \tilde{u}_h(x) = m + \frac{i}{2^n}(M - m)\}$ has a nonempty interior, set u_h^n equal to $m + \frac{i}{2^n}(M - m)$ in the interior of the set. For all $x \in \Omega$ such that $m + \frac{i}{2^n}(M - m) < \tilde{u}_h(x) < m + \frac{i+1}{2^n}(M - m)$, we set $u_h^n(x)$ equal to $m + \frac{2i+1}{2^{n+1}}(M - m)$, that is, the average of the values on the level curves. Then, clearly, $u_h^n \to \tilde{u}_h$ uniformly in Ω as $n \to \infty$.

Let us denote by T the (open) triangles and by e the edges of the triangulation $\tilde{\tau}_h$ and write $e \subset \Omega$ for the internal (non-boundary) edges of $\tilde{\tau}_h$. We note that since $\nabla \tilde{u}_h$ is constant within each triangle $T \in \tilde{\tau}_h$, the level curves L_i^n are parallel and separated by a uniform distance h_T^n within each T (see Figure 8). Let also $|T|$
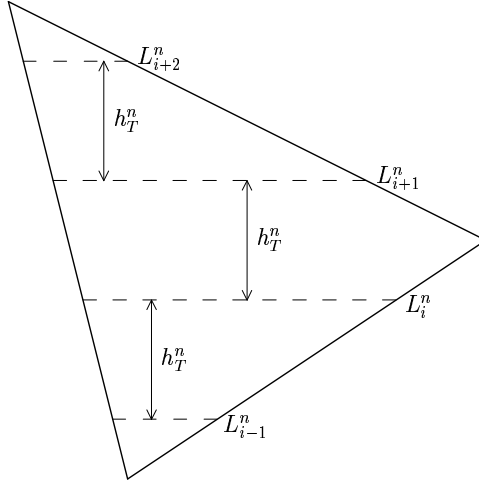
FIGURE 8. The level curves L_i^n separated by the distance h_T^n within each triangle $T \in \tilde{\tau}_h$.

denote the area of T, and let $|L_i^n|$ and $|L_i^n \cap e|$ denote the length of L_i^n and $L_i^n \cap e$, respectively. We then have

$$
\int_\Omega |Du_h^n|_2 = \sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |L_i^n|
$$

$$
= \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |L_i^n \cap T| \right] + \sum_{e \subset \Omega} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |L_i^n \cap e| \right]
$$

$$
= \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |L_i^n \cap T| \right] + O\left(\frac{M-m}{2^n} \sum_{e \subset \Omega} |e| \right)
$$

$$
= \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| \nabla \tilde{u}_h|_T \right|_2 \cdot h_T^n \cdot |L_i^n \cap T| \right] + o(1) \quad \text{as } n \to \infty
$$

$$
\to \sum_{T \in \tilde{\tau}_h} \left| \nabla \tilde{u}_h|_T \right|_2 \cdot |T| \quad \text{as } n \to \infty
$$

$$
= \int_\Omega |D\tilde{u}_h|_2
$$

in view of the continuity of $\tilde{u}_h$.

Finally, define any triangulation τ_h^n with mesh-size at most h and such that the union $\cup_{i=0}^{2^n} L_i^n$ forms a subset of the edges of τ_h^n (see Figure 7(b)). From the above considerations, it is clear that we can find $n_h \in \mathbb{N}$, a triangulation $\tau_h^{n_h}$, and $u_h \equiv u_h^{n_h}$ such that $\rho_{q,2}(u, u_h) \to 0$ as $h \to 0$. $\qquad \square$

The theorem for general k follows.

Theorem 4.4. *Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain and let $h_0 > 0$. Then, given a scalar-valued $u \in BV_q(\Omega)$ for some $q \in [1, \infty)$, there exists a family $\{\tau_h :\ 0 < h \le$*

$h_0\}$ of triangulations of Ω such that the mesh-size of τ_h is at most h, and functions $u_h \in \mathcal{A}_h^0$, where $\mathcal{A}_h^0$ denotes the space of piecewise constant functions corresponding to the triangulation τ_h, such that for any $1 \le k \le \infty$ we have

$$\rho_{q,k}(u, u_h) \to 0 \quad as\ h \to 0.$$

Proof. The proof is almost exactly the same as the proof of Theorem 4.3. Using the same notation, the only difference is to show that

$$\int_\Omega |Du_h^n|_k \to \int_\Omega |D\tilde{u}_h|_k \quad as\ n \to \infty.$$

To this end, we denote by ν the ℓ_2-unit normal vector to the corresponding line segment $L_i^n \cap T$ or $L_i^n \cap e$ and we have

$$\int_\Omega |Du_h^n|_k = \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |\nu|_k \cdot |L_i^n \cap T| \right] + \sum_{e \subset \Omega} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |\nu|_k \cdot |L_i^n \cap e| \right]$$

$$= \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| [\![u_h^n]\!]_{L_i^n} \right| \cdot |\nu|_k \cdot |L_i^n \cap T| \right] + O\left(\frac{M-m}{2^n} \cdot |\nu|_k \sum_{e \subset \Omega} |e| \right)$$

$$= \sum_{T \in \tilde{\tau}_h} \left[\sum_{i=0}^{2^n} \left| \nabla \tilde{u}_h|_T \right|_2 \cdot h_T^n \cdot \frac{\left| \nabla \tilde{u}_h|_T \right|_k}{\left| \nabla \tilde{u}_h|_T \right|_2} \cdot |L_i^n \cap T| \right] + o(1) \quad as\ n \to \infty$$

$$\to \sum_{T \in \tilde{\tau}_h} \left| \nabla \tilde{u}_h|_T \right|_k \cdot |T| \quad as\ n \to \infty$$

$$= \int_\Omega |D\tilde{u}_h|_k.$$

$\square$

Remark 4.3. *It is not clear to us how to generalize Theorems 4.3 and 4.4 to the case of vector-valued functions w with the definition of total variation $\int_\Omega |Dw|_k$ given by (2.1). However, if the definition is modified to*

$$\int_\Omega |Dw|_k = \sum_{j=1}^m \int_\Omega |Dw_j|_k, \tag{4.8}$$

then one can take the union of the level curves L_i^n for all components of w and define a new mesh including all of these level curves as a subset of the edges. It is then easy to see that Theorems 4.3 and 4.4 hold since they hold for each component of w. We note, however, that the definition (4.8) does not give a frame-indifferent energy when used to model the surface energy in (3.1) even if $k = 2$.

The above proofs suggest several computational strategies. One can try to follow the proofs and solve on a coarse mesh with continuous finite element basis functions and then modify the mesh to correspond to the level sets of the solution and resolve on this finer mesh with piecewise constants. For an application of this strategy for the thin film problem, this means using $\mathcal{C}^1$ functions for y on a coarse mesh to determine the fine mesh and then using continuous piecewise linear functions for y on the fine mesh (which gives piecewise constant functions for the deformation gradient).

An alternative strategy is a moving-mesh method allowing the nodes defining the triangulation to move so that the total variation of the deformation gradient is better approximated by piecewise constant functions.

References

[1] J. Ball and R. James. Fine phase mixtures as minimizers of energy. *Arch. Rat. Mech. Anal.*, 100:13–52, 1987.

[2] J. Ball and R. James. Proposed experimental tests of a theory of fine microstructure and the two-well problem. *Phil. Trans. R. Soc. Lond. A*, 338:389–450, 1992.

[3] P. Bělík. *Computational methods for martensitic thin films.* PhD thesis, University of Minnesota, 2000.

[4] K. Bhattacharya, B. Li, and M. Luskin. The simply laminated microstructure in martensitic crystals that undergo a cubic to orthorhombic phase transformation. *Arch. Rat. Mech. Anal.*, 149:123–154, 1999.

[5] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, New York, 1994.

[6] P. Bělík, T. Brule, and M. Luskin. On the numerical modeling of deformations of pressurized martensitic thin films. *Math. Model. Numer. Anal.*, 35:525–548, 2001.

[7] P. Bělík and M. Luskin. Stability of microstructure for tetragonal to monoclinic martensitic transformations. *Math. Model. Numer. Anal.*, 34:663–685, 2000.

[8] P. Bělík and M. Luskin. A computational model for the indentation and phase transformation of a martensitic thin film. *J. Mech. Phys. Solids*, 50:1789–1815, 2002.

[9] P. Bělík and M. Luskin. A total-variation surface energy model for thin films of martensitic crystals. *Interfaces and Free Boundaries*, 4(1):71–88, 2002.

[10] C. Carstensen and P. Plecháč. Numerical analysis of compatible phase transitions in elastic solids. *SIAM J. Numer. Anal.*, 37:2061–2081, 2000.

[11] E. Casas, K. Kunisch, and C. Pola. Regularization by functions of bounded variation and applications to image enhancement. *Appl. Math. Optim.*, 40(2):229–257, 1999.

[12] M. Chipot, C. Collins, and D. Kinderlehrer. Numerical analysis of oscillations in multiple well problems. *Numer. Math.*, 70:259–282, 1995.

[13] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.

[14] C. Collins. Convergence of a reduced integration method for computing microstructures. *SIAM. J. Numer. Anal.*, 35:1271–1298, 1998.

[15] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34(5):1779–1791, 1997.

[16] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[17] B. G. Fitzpatrick and S. L. Keeling. On approximation in total variation penalization for image reconstruction and inverse problems. *Numer. Funct. Anal. Optim.*, 18(9-10):941–958, 1997.

[18] I. Fonseca. Phase transitions of elastic solid materials. *Arch. Rat. Mech. Anal.*, 107:195–223, 1989.

[19] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*, volume 80 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel-Boston, Mass., 1984.

[20] M. Gobbert and A. Prohl. A discontinuous finite element method for solving a multiwell problem. *SIAM J. Numer. Anal.*, 37:246–268, 1999.

[21] M. Kružík. Numerical approach to double well problems. *SIAM J. Numer. Anal.*, 35(5):1833–1849, 1998.

[22] M. Kružík and A. Prohl. Young measure approximation in micromagnetics. *Numer. Math.*, 90:291–307, 2001.

[23] B. Li. Approximation of martensitic microstructure with general homogeneous boundary data. *J. Math. Anal. Appl.*, 266:451–467, 2002.

[24] B. Li. Finite element analysis of a class of stress-free martensitic microstructures. *Math. Comp.*, to appear.

[25] Z. Li. Simultaneous numerical approximation of microstructures and relaxed minimizers. *Numer. Math.*, 78:21–38, 1997.

[26] Z. Li. Rotational transformation method and some numerical techniques for computing microstructures. *Math. Models Methods Appl. Sci.*, 8:985–1002, 1998.

[27] M. Luskin. On the computation of crystalline microstructure. *Acta Numerica*, 5:191–258, 1996.

[28] M. Luskin. Approximation of a laminated microstructure for a rotationally invariant, double well energy density. *Numer. Math.*, 75:205–221, 1997.

[29] M. Luskin. Computational modeling of microstructure. In *Proceedings of the International Congress of Mathematicians, ICM 2002, Beijing*, volume III, pages 707–716, 2002.

[30] M. Luskin and L. Ma. Analysis of the finite element approximation of microstructure in micromagnetics. *SIAM J. Numer. Anal.*, 29:320–331, 1992.

[31] P. Plecháč and T. Roubíček. Visco-elasto-plastic model for martensitic phase transformation in shape-memory alloys. *Math. Methods Appl. Sci.*, 25:1281–1298, 2002.

[32] A. Prohl. An adaptive finite element method for solving a double well problem describing crystalline microstructure. *Math. Model. Numer. Anal.*, 33:781–796, 1999.

[33] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.

PAVEL BĚLÍK, SCHOOL OF MATHEMATICS, 206 CHURCH STREET SE, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455, U.S.A.

E-mail address: `belik@math.umn.edu`

MITCHELL LUSKIN, SCHOOL OF MATHEMATICS, 206 CHURCH STREET SE, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455, U.S.A.

E-mail address: `luskin@math.umn.edu`