# A Convex, Lower Semi-Continuous Approximation of Euler's Elastica Energy

K. Bredies        T. Pock        B. Wirth

A–8010 GRAZ,  HEINRICHSTRASSE 36,  AUSTRIA

FWF Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)

- **University of Graz**

- **Graz University of Technology**

- **Medical University of Graz**

- **Government of Styria**

- **City of Graz**

# A CONVEX, LOWER SEMI-CONTINUOUS APPROXIMATION OF EULER'S ELASTICA ENERGY

KRISTIAN BREDIES[*], THOMAS POCK[†], AND BENEDIKT WIRTH[‡]

**Abstract.** We propose a convex, lower semi-continuous, coercive approximation of Euler's elastica energy for images, which is thus very well-suited as a regularizer in image processing. The approximation is not quite the convex relaxation, and we discuss its close relation to the exact convex relaxation as well as the difficulties associated with computing the latter. Interestingly, the convex relaxation of the elastica energy reduces to constantly zero if the total variation part of the elastica energy is neglected.

Our convex approximation arises via functional lifting of the image gradient into a Radon measure on the four-dimensional space $\Omega \times S^1 \times \mathbb{R}$, $\Omega \subset \mathbb{R}^2$, of which the first two coordinates represent the image domain and the last two the normal and curvature of the image level lines. It can be expressed as a linear program which also admits a predual representation. Combined with a tailored discretization of measures via linear combinations of short line measures, the proposed functional becomes an efficient regularizer for image processing tasks, whose feasibility and effectiveness is verified in a series of exemplary applications.

**Key words.** curvature regularization, elastica energy, Willmore energy, convex relaxation, functional lifting, measure-based regularization

**AMS subject classifications.** 49J45, 49M29, 68U10

**1. Introduction.** In various different applications in image processing, curvature-dependent functionals have gained importance during the past decade. In particular, for tasks such as image reconstruction, inpainting, denoising, or segmentation it seems often a valid a priori assumption that level lines are mostly smooth and of rather small curvature, which suggests the use of curvature-penalizing priors. We here consider the particular case of Euler's elastica energy,

$$(1.1) \qquad R_{\mathrm{el}}(u) = \int_\Omega (a\kappa_u^2(x) + b)\,\mathrm{d}|\nabla u|(x)\,,$$

for an image $u \in \mathrm{BV}(\Omega; \mathbb{R})$ and parameters $a, b > 0$, where $\kappa_u(x)$ denotes the curvature of the image level line passing through $x$. For smooth functions $u \in \mathcal{C}^2(\Omega; \mathbb{R})$ the energy can also be expressed in the more common form

$$(1.2) \qquad R_{\mathrm{el}}(u) = \int_\Omega \left( a \left| \mathrm{div} \frac{\nabla u}{|\nabla u|} \right|^2 + b \right) \mathrm{d}|\nabla u|\,.$$

Since by Euler's approximation the elastic bending energy of a thin rod is proportional to its squared curvature, $R_{\mathrm{el}}(u)$ may be interpreted as the accumulated bending energy of all level lines of the image $u$, hence the name.

Euler's elastica energy can be employed as a regularizing prior to devise variational models for specific image processing tasks, which typically take the form

$$(1.3) \qquad \mathcal{F}(u) + R_{\mathrm{el}}(u) \to \min!$$

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

[†]Institute for Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, A-8010 Graz, Austria (`pock@icg.tugraz.at`)

[‡]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York 10012, USA (`benedikt.wirth@cims.nyu.edu`)

for a so-called fidelity term $\mathcal{F}(u)$ (in denoising, for example, $\mathcal{F}(u)$ could be the $L^2$-difference between the given noisy image and its noise-free approximation $u$). Due to the high nonlinearity of $R_{\mathrm{el}}$, such models typically feature a large number of local minima and are thus difficult to minimize. One therefore strives to replace $R_{\mathrm{el}}(u)$ by a convex energy with similar behavior, for example its convex relaxation, which is the largest convex, lower semi-continuous functional below $R_{\mathrm{el}}$ (the lower semi-continuity is required to ensure existence of minimizers). The new model might not have the exact same minimizers, but serves as a good and viable approximation.

Sometimes, the convex relaxation of an energy functional such as $R_{\mathrm{el}}(u)$ can be obtained by adding one or more space dimensions and representing each of the functions $u$ by a function on this higher-dimensional space, which is called functional lifting [8]. The energy functional then turns into a convex functional on the space of the higher-dimensional representatives, where the minimization is carried out. We propose a similar strategy, only we do not lift the functions $u$ themselves but rather their gradient $\nabla u$. In particular, we associate with every image $u \in \mathrm{BV}(\Omega; \mathbb{R})$ a non-negative Radon measure $v$ on $\Omega \times S^1 \times \mathbb{R}$ such that for any smooth test function $\phi : \Omega \times S^1 \times \mathbb{R} \to \mathbb{R}$ we have

$$(1.4) \qquad \int_\Omega \phi(x, \nu_u(x), \kappa_u(x)) \, \mathrm{d}|\nabla u|(x) = \int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu, \kappa) \, \mathrm{d}v(x, \nu, \kappa) \,,$$

where $\nu_u(x)$ and $\kappa_u(x)$ denote the local normal and curvature of the level lines of $u$. For illustration, think of the measure $v$ as having support only on those $(x, \nu, \kappa) \in \Omega \times S^1 \times \mathbb{R}$ where $\nu$ coincides with the image normal $\frac{\nabla u(x)}{|\nabla u(x)|}$ and $\kappa$ with the level line curvature $-\mathrm{div}\left(\frac{\nabla u(x)}{|\nabla u(x)|}\right)$. The corresponding elastica energy then reads

$$(1.5) \qquad R_{\mathrm{el}}(u) = \int_{\Omega \times S^1 \times \mathbb{R}} (a\kappa^2 + b) \, \mathrm{d}v(x, \nu, \kappa) \,.$$

A convex lower bound on its convex relaxation $R_{\mathrm{el}}^{**}$ is obtained by relaxing the restriction (1.4) on $v$ and taking the infimum over all measures from a convex admissible set $\mathcal{M}(u)$,

$$(1.6) \qquad R_{\mathrm{el}}^{**}(u) \geq \inf_{v \in \mathcal{M}(u)} \int_{\Omega \times S^1 \times \mathbb{R}} (a\kappa^2 + b) \, \mathrm{d}v(x, \nu, \kappa) \,.$$

Interestingly, for the degenerate case $b = 0$ one can show $R_{\mathrm{el}}^{**} \equiv 0$ so that $b > 0$ is required for a useful model.

The above set $\mathcal{M}(u)$ of admissible measures for a given image $u$ turns out to be rather complicated for which reason we replace it by a slightly larger set. In particular, we will allow all non-negative measures $v$ that satisfy the "compatibility" and the "consistency" condition

$$(1.7) \qquad \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu \, \mathrm{d}v(x, \nu, \kappa) = -\int_\Omega u(x) \mathrm{div}\varphi(x) \, \mathrm{d}x \quad \forall \varphi \in \mathcal{C}_0^\infty(\Omega; \mathbb{R}^2) \,,$$

$$(1.8) \quad 0 = \int_{\Omega \times S^1 \times \mathbb{R}} \nabla_x \psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x, \nu)}{\partial \nu} \kappa \, \mathrm{d}v(x, \nu, \kappa) \quad \forall \psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R}) \,,$$

of which the first has the interpretation that the projection $\int_{S^1 \times \mathbb{R}} \nu \, \mathrm{d}v(\cdot, \nu, \kappa)$ of $v$ (the reversal of the lifting) acts like $\nabla u$, and the second is motivated by (1.4) for particular

test functions. The resulting approximation of the elastica energy,

$$(1.9) \qquad R_{\mathrm{el}}^{\mathrm{ap}}(u) = \inf_{v \geq 0 \wedge (1.7) \wedge (1.8)} \int_{\Omega \times S^1 \times \mathbb{R}} (a\kappa^2 + b) \, \mathrm{d}v(x, \nu, \kappa) \,,$$

will be convex lower semi-continuous, but not the exact convex relaxation. The reason turns out to lie in the fact that we do not forbid measures $v$ that satisfy

$$(1.10) \qquad \int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu, \kappa) \, \mathrm{d}v(x, \nu, \kappa) = \int_C \phi(x, \nu(x), \kappa(x)) \, \mathrm{d}\mathcal{H}^1(x)$$

for an immersed curve $C \subset \Omega$ which crosses itself.

**1.1. Related work on curvature regularization and convex relaxation.** The elastica energy was first introduced as a prior to image processing by Nitzberg, Mumford, and Shiota [22] who used it for recovering occluded objects in an image by optimal edge continuation. Masnou employed this concept for image inpainting via level line continuation [17]. Meanwhile, the regularization of level line curvature has found its way into various image processing applications [30, 14, 29]. An alternative but strongly related regularization is given by the absolute mean curvature of the graph $\partial\{(x,t) \ : \ t < u(x)\}$ of an image $u$, which allows the preservation of image corners [33].

Due to its nonlinear and non-convex nature, the elastica energy is difficult to minimize. Chan et al. formally derive the Euler–Lagrange equation for (1.2), $0 = -\mathrm{div}V$ with a morphologically invariant flux field $V$, and they employ an $L^2$-flow in direction $|\nabla u|\mathrm{div}V$ for image inpainting [11]. Tai et al. replace (1.2) by a constrained minimization problem for which they propose an augmented Lagrange scheme with a simple and efficient alternating inner minimization [32].

To avoid local minima and to simplify the minimization, relaxations and convex approximations to the elastica energy become more and more important, and the present paper contributes to this area of research. To extend (1.2) to BV-functions, Chan et al. employ formulation (1.1) with the weak curvature notion $\kappa_u = \limsup_{\epsilon \to 0} |\mathrm{div}\frac{\nabla u_\epsilon}{|\nabla u_\epsilon|}|$ for a mollification $u_\epsilon$ of $u$ with mollifier width $\epsilon$ [11]. Ambrosio and Masnou instead extend (1.2) from $\mathcal{C}^2(\Omega; \mathbb{R})$ to $L^1(\Omega; \mathbb{R})$ by $\infty$ and then consider the lower semi-continuous envelope $\overline{R_{\mathrm{el}}}$ on $L^1(\Omega; \mathbb{R})$ [4] (they work in $n$ dimensions and with the square replaced by a general $p^{\mathrm{th}}$ power). Using the tool of integral varifolds and an associated curvature notion they derive a locality result for the mean curvature and succeed in showing $\overline{R_{\mathrm{el}}} \equiv R_{\mathrm{el}}$ on $\mathcal{C}^2(\Omega; \mathbb{R})$. Recently [19], these results could be strengthened by expressing $\overline{R_{\mathrm{el}}}$ as the minimum of $\int_\Omega (a\kappa_V^2 + b) \, \mathrm{d}(\mathrm{proj}_\Omega V)$ over a certain class of gradient Young varifolds $V$ (varifolds defined by the generalized Young measure associated with a weakly-* converging sequence in BV) which are compatible with the given image in a way similar to our compatibility condition (1.7). Note that our lifted image gradients in (1.4) can also be seen as generalized varifolds which not only encode tangent, but also curvature information. A different but related, coarea-type representation of $\overline{R_{\mathrm{el}}}$ is provided in [18] as the minimum over certain sets of nested $W^{2,p}$ curves of their accumulated elastica energy. Curvature-based models allowing global optimization have been introduced to image processing in [27, 30]. The authors use graph-based algorithms to globally maximize a discrete version of $\frac{\int_0^{L(C)} |\nabla u(C(s)) \cdot (C'(s))^\perp| \, \mathrm{d}s}{\int_0^{L(C)} a|\kappa_{C(s)}|^p + b \, \mathrm{d}s}$ for a closed segmentation curve $C : [0, L(C)] \to \mathbb{R}^2$. Goldlücke and Cremers use a convex approximation of the $p$-Menger–Melnikov curvature which can be defined for image gradient measures [14]. Schoenemann et al.

perform denoising and inpainting with elastica-type regularization on discretized images by rephrasing the problem as an integer linear program [28, 29] which can be globally solved after relaxation. Finally, in [7] a convex approximation is proposed for a functional penalizing absolute mean curvature of image level lines.

Functional lifting as a means of (convex) relaxation has been applied especially for segmentation problems. Alberti et al. derived a convex formulation of the highly non-convex Mumford–Shah functional $E(u, S) = \int_{\Omega \setminus S} |\nabla u|^2 \, \mathrm{d}x + \mathcal{H}^1(S)$ as $\sup_{\phi \in K} \int_{\Omega \times \mathbb{R}} \phi(x) D1_{u(x)>t} \, \mathrm{d}x \mathrm{d}t$ for a particular convex set $K \subset \mathcal{C}_0(\Omega \times \mathbb{R}; \mathbb{R}^2)$, where the lifted image $u$ is given by the characteristic function $1_{u(x)>t}$ of $\{(x, t) \in \Omega \times \mathbb{R} \; : \; u(x) > t\}$ [1]. Pock et al. relax the constraint that $1_{u(x)>t}$ be a characteristic function and thus approximate the image segmentation by a convex optimization problem which they solve via a primal-dual ascent-descent iteration [24]. Chambolle generalizes the above lifting approach to general BV-coercive functionals $E : L^1(\Omega; \mathbb{R}) \to [0, \infty]$ and shows that their lower semi-continuous envelope $\overline{E}(u)$ can be represented convexly as $\tilde{E}(1_{u(x)>t}) = \sup_{\phi \in K} \int_{\Omega \times \mathbb{R}} \phi(x) D1_{u(x)>t} \, \mathrm{d}x \mathrm{d}t$ for a convex $K$ [8]. However, the set $K$ typically is too complicated, and Chambolle examines for what functionals $K$ can be characterized locally. For functionals of the form $E(u) = \int_\Omega f(x, u(x), \nabla u(x)) \, \mathrm{d}x$ with $f$ convex in $\nabla u$ and satisfying some regularity conditions, $K$ takes the form $K = \{\phi \in \mathcal{C}_0(\Omega \times \mathbb{R}; \mathbb{R}^{n+1}) \; : \; \phi^t(x, t) \geq f^*(x, t, \phi^x(x, t))\}$ for the Legendre–Fenchel dual $f^*$ of $f$ [25]. Allowing also non-characteristic functions as arguments of $\tilde{E}$ one obtains a convex optimization problem, and thresholding its minimizer can be shown to yield a minimizer of the original problem.

Note that for binary image segmentation, the situation is simpler: Nikolova et al. show that thresholding the minimizer of the convex energy $\int_\Omega |\nabla u(x)| + \lambda g(x) u(x) \, \mathrm{d}x$ on $\mathrm{BV}(\Omega; [0, 1])$ yields the global minimum of the binary Mumford–Shah segmentation problem with data $g$ [21]. For multilabel segmentation, the same convex optimization and thresholding approach gives experimentally good results [16].

Chambolle et al. investigated the convex relaxation of the same multilabel segmentation [9, 23]. They derive bounds for the convex relaxation in terms of the original functional and also find the largest convex lower semi-continuous functional below the original functional which is local in the sense that it can be expressed as an integral with localized integrand. Our article has a similar aim in that we also seek a convex functional close below $R_{\mathrm{el}}$ which can still be evaluated locally. An approach closely related to ours (though inherently discrete) computes a convex approximation of a discrete elastica functional: Schoenemann et al. perform denoising and inpainting with elastica-type regularization on discretized images by rephrasing the problem as an integer linear program [28, 29] which can be globally solved after relaxation. After dividing the image into subpixel segments they introduce a binary variable for each segment (encoding fore- or background) and for each pair of adjacent image edges. A linear surface continuation constraint now ensures that foreground regions and boundary edges are compatible, and a linear boundary continuation constraint ensures that the foreground boundary is a continuous curve. Both constraints in a sense are paralleled by our continuous compatibility and consistency conditions (1.7) and (1.8), and the main source of convexification in their as well as our method lies in the representation of compatibility and consistency as a single linear constraint.

**1.2. Notation and conventions.** For convenience and not to have to go into the differential geometric details we identify, throughout the paper, normal vectors $\nu \in S^1$ with angles in the quotient space $\mathbb{R}/(2\pi\mathbb{Z})$ via the inverse mapping $e : \mathbb{R}/(2\pi\mathbb{Z}) \to S^1$, $e(t) = (\cos t, \sin t)$. In particular, $S^1 \sim \mathbb{R}/(2\pi\mathbb{Z})$ inherits its topological structure

from $S^1$ and is endowed with the group structure of $\mathbb{R}/(2\pi\mathbb{Z})$. The latter means that $(1,0)$ corresponds to $0$ and that angle sums and differences can be identified with elements in $S^1$. Moreover, the derivatives of functions $\varphi : S^1 \to \mathbb{R}$ and $\psi : \mathbb{R} \to S^1$ may be expressed as

$$\varphi'(\nu) = \lim_{h \to 0} \frac{\varphi(\nu + h) - \varphi(\nu)}{h} \in \mathbb{R}, \quad \psi'(t) = \lim_{h \to 0} \frac{\left(\psi(t + h) - \psi(t)\right) \cdot \psi(t)^\perp}{h} \in \mathbb{R},$$

respectively, where $\nu^\perp$ denotes counterclockwise rotation by $\frac{\pi}{2}$, that is, $\nu^\perp = (-\nu_2, \nu_1)$. The latter corresponds to identifying, locally around $\psi(t) \in S^1$, the range $S^1 \sim \mathbb{R}/(2\pi\mathbb{Z})$ with an interval in $\mathbb{R}$. Differentiation of $\varphi : S^1 \to S^1$ is then defined analogously. In this sense, angle addition $+ : S^1 \times S^1 \to S^1$ as well as angle negation $- : S^1 \to S^1$ are arbitrarily differentiable with first derivatives $(1,1)$ and $-1$, respectively.

We are considering spaces of vector-valued regular countably additive measures over a locally compact set $\Omega$, denoted by $\mathrm{rca}(\Omega; \mathbb{R}^n)$. Given $v \in \mathrm{rca}(\Omega; \mathbb{R}^n)$, the total-variation measure $|v|$ is then a measure on $\Omega$ which gives the norm $\|v\|_{\mathrm{TV}(\Omega;\mathbb{R}^n)} = |v|(\Omega)$ such that $\mathrm{rca}(\Omega; \mathbb{R}^n)$ becomes a Banach space. The space of functions of bounded variation $\mathrm{BV}(\Omega; \mathbb{R})$ over a domain $\Omega \subset \mathbb{R}^n$ is then the collection of all $u \in L^1(\Omega; \mathbb{R})$ for which the weak derivative $\nabla u$ can be represented, via dual pairing, by an element in $\mathrm{rca}(\Omega; \mathbb{R}^n)$. The total variation of $u$ is then defined by $|u|_{\mathrm{TV}(\Omega;\mathbb{R})} = |\nabla u|(\Omega)$.

We will also frequently make use of measure disintegration. Typically we consider $v \in \mathrm{rca}(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ on the four-dimensional space $\Omega \times S^1 \times \mathbb{R}$, where $\Omega \subset \mathbb{R}^2$ is an open Lipschitz domain. We define the projections $\pi^\Omega : \Omega \times S^1 \times \mathbb{R} \to \Omega$, $(x, \nu, \kappa) \mapsto x$, and $\pi^{\Omega \times S^1} : \Omega \times S^1 \times \mathbb{R} \to \Omega \times S^1$, $(x, \nu, \kappa) \mapsto (x, \nu)$. Given $v \in \mathrm{rca}(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$, the pushforward measure $(\pi^{\Omega \times S^1})_\# v \in \mathrm{rca}(\Omega \times S^1; \mathbb{R})$ is defined by $(\pi^{\Omega \times S^1})_\# v(E) = v \circ (\pi^{\Omega \times S^1})^{-1}(E) = v(E \times \mathbb{R})$ for all Borel sets $E \subset \Omega \times S^1$. Then by the disintegration theorem [3], for $(\pi^{\Omega \times S^1})_\# v$-almost all $(x, \nu) \in \Omega \times S^1$ there exists a probability measure $v_{(x,\nu)} \in \mathrm{rca}(\mathbb{R}; \mathbb{R})$ such that $v$ is the product measure $v = v_{(x,\nu)} \times (\pi^{\Omega \times S^1})_\# v$. Correspondingly we define the pushforward measure $(\pi^\Omega)_\# v \in \mathrm{rca}(\Omega; \mathbb{R})$ and the probability measure $v_x \in \mathrm{rca}(S^1 \times \mathbb{R}; \mathbb{R})$, $x \in \Omega$.

The remainder of the article proceeds as follows. Section 2 rigorously introduces the elastica energy $R_{\mathrm{el}}$ and its lifted version. It also includes the proof that the convex relaxation $R_{\mathrm{el}}^{**}$ is trivial if $R_{\mathrm{el}}$ only measures the curvature of image level lines. Section 3 then introduces our approximation of the elastica functional. In particular, we prove its convex lower semi-continuity as well as the existence of minimizing measures, we discuss why the approximation fails to be the exact convex relaxation, we derive the predual formulation of $R_{\mathrm{el}}^{\mathrm{ap}}$, and we characterize the minimizing measures. Finally, we propose a numerical discretization in Section 4 and show some image processing applications.

**2. Euler's elastica energy on $L^1$.** This section introduces Euler's elastica energy $R_{\mathrm{el}}$ for $L^1$-images. After defining $R_{\mathrm{el}}$ based on a generalized notion of level line curvature we briefly show that its convex relaxation is zero for certain degenerate parameters. Finally we introduce a functionally lifted version of $R_{\mathrm{el}}$ as a preparation for the convex approximation to be derived in the subsequent section.

**2.1. Generalized curvature and a coarea representation of the elastica energy.** We first define Euler's elastica functional for characteristic functions of smooth sets, after which we will introduce a weak definition of curvature in order to

define Euler's elastica energy for general images. Let $\Omega \subset \mathbb{R}^2$ be an open bounded Lipschitz domain. Let $\mathcal{O} \subset \Omega$ be simply connected and have a smooth boundary $\partial\mathcal{O} \cap \Omega$ which can be parameterized counterclockwise by arclength via some parameterization $x : [0, L] \to \partial\mathcal{O}, s \mapsto x(s) \in \partial\mathcal{O}$. The (inward) normal and curvature of $\partial\mathcal{O}$ are then given by $\nu_{\mathcal{O}}(x(s)) = \dot{x}(s)^{\perp}$ and $\kappa_{\mathcal{O}}(x(s)) = \ddot{x}(s) \cdot \nu_{\mathcal{O}}(x(s)) = \frac{\mathrm{d}}{\mathrm{d}s}\nu_{\mathcal{O}}(x(s))$, where the dot denotes differentiation with respect to arclength and $\frac{\mathrm{d}}{\mathrm{d}s}\nu_{\mathcal{O}}(x(s))$ is to be interpreted as explained in Section 1.2. Now consider the image $u = \chi_{\mathcal{O}}$, where $\chi_{\mathcal{O}} : \Omega \to \{0, 1\}$ shall be the characteristic function of $\mathcal{O}$. Euler's elastica energy of $u$ is defined as

$$(2.1) \quad R_{\mathrm{el}}(u) = \int_0^L f(x(s), \nu_{\mathcal{O}}(x(s)), \kappa_{\mathcal{O}}(x(s))) \, \mathrm{d}s$$
$$= \int_{\partial\mathcal{O}} f(x, \nu_{\mathcal{O}}(x), \kappa_{\mathcal{O}}(x)) \, \mathrm{d}\mathcal{H}^1(x) = \int_{\Omega} f(x, \nu_{\mathcal{O}}(x), \kappa_{\mathcal{O}}(x)) \, \mathrm{d}|\nabla u|(x) \,,$$

where we mostly have in mind $f : \Omega \times S^1 \times \mathbb{R} \to \mathbb{R}, (x, \nu, \kappa) \mapsto a\kappa^2 + b$ for $a, b > 0$, but other choices (in particular anisotropic ones) are also possible. Here, $\mathcal{H}^1$ denotes the one-dimensional Hausdorff measure.

For the above set $\mathcal{O}$ and parameterization of $\partial\mathcal{O}$, any smooth $\psi \in \mathcal{C}_0^{\infty}(\Omega \times S^1; \mathbb{R})$ satisfies
$$(2.2)$$
$$0 = \int_0^L \frac{\mathrm{d}}{\mathrm{d}s}\psi(x(s), \nu_{\mathcal{O}}(x(s))) \, \mathrm{d}s = \int_{\partial\mathcal{O}} \nabla_x \psi(x, \nu_{\mathcal{O}}(x)) \cdot \nu_{\mathcal{O}}(x)^{-\perp} + \frac{\partial \psi(x, \nu_{\mathcal{O}}(x))}{\partial \nu} \kappa_{\mathcal{O}}(x) \, \mathrm{d}\mathcal{H}^1(x) \,,$$

where $\nu^{-\perp} = -\nu^{\perp}$ denotes the clockwise rotation by $\frac{\pi}{2}$ and $\frac{\partial}{\partial \nu}$ the derivative with respect to the angle as introduced in Section 1.2. This relation can be employed to generalize the notion of boundary curvature to sets of finite perimeter.

DEFINITION 2.1 (Generalized curvature). *Let $\mathcal{O} \subset \Omega$ have finite perimeter, and let $\partial^*\mathcal{O} \subset \Omega$ denote its essential boundary and $\nu_{\mathcal{O}}$ its generalized inner normal [2]. $\kappa_{\mathcal{O}} \in L_{\mathrm{loc}}^1(\mathcal{H}^1 \llcorner \partial^*\mathcal{O}; \mathbb{R})$ (the space of scalar, locally integrable functions with respect to the measure $\mathcal{H}^1 \llcorner \partial^*\mathcal{O}$) is a generalized curvature of $\partial^*\mathcal{O}$, if*
$$(2.3)$$
$$0 = \int_{\partial^*\mathcal{O}} \nabla_x \psi(x, \nu_{\mathcal{O}}(x)) \cdot \nu_{\mathcal{O}}(x)^{-\perp} + \frac{\partial \psi(x, \nu_{\mathcal{O}}(x))}{\partial \nu} \kappa_{\mathcal{O}}(x) \, \mathrm{d}\mathcal{H}^1(x) \quad \forall \psi \in \mathcal{C}_0^{\infty}(\Omega \times S^1; \mathbb{R}) \,.$$

THEOREM 2.2 (Uniqueness of generalized curvature). *Given $\mathcal{O} \subset \Omega$ with finite perimeter, its generalized curvature, if it exists, is unique.*

*Proof.* Let $\kappa, \tilde{\kappa} \in L_{\mathrm{loc}}^1(\mathcal{H}^1 \llcorner \partial^*\mathcal{O}; \mathbb{R})$ satisfy (2.3); we have to show $\kappa = \tilde{\kappa}$. Let $\zeta \in \mathcal{C}_0^{\infty}(S^1; \mathbb{R})$ with derivative $\zeta'(0) = 1$ and define $\psi(x, \nu) = \varphi(x)\zeta(\nu - \nu_{\mathcal{O}}(x))$ for given $\varphi \in \mathcal{C}_0^{\infty}(\Omega; \mathbb{R})$ (where $\nu - \nu_{\mathcal{O}}(x)$ is the angle difference). Furthermore, let $m_{\varepsilon} \in \mathcal{C}_0^{\infty}(\mathbb{R}^2; \mathbb{R})$ be a Dirac sequence with support of diameter smaller than $\varepsilon > 0$ and define $\psi_{\varepsilon}(x, \nu) = (m_{\varepsilon} * \psi)(x, \nu) = \int_{\Omega} \psi(y, \nu) m_{\varepsilon}(x - y) \, \mathrm{d}y$. Obviously, $\psi_{\varepsilon} \in \mathcal{C}_0^{\infty}(\Omega \times S^1; \mathbb{R})$ with $\frac{\partial \psi_{\varepsilon}(x, \nu_{\mathcal{O}}(x))}{\partial \nu} = \left(m_{\varepsilon} * \frac{\partial \psi}{\partial \nu}\right)(x, \nu_{\mathcal{O}}(x)) = (m_{\varepsilon} * \varphi)(x)$. We thus have

$$\int_{\partial^*\mathcal{O}} \varphi(\kappa - \tilde{\kappa}) \, \mathrm{d}\mathcal{H}^1 = \lim_{\varepsilon \to 0} \int_{\partial^*\mathcal{O}} \frac{\partial \psi_{\varepsilon}(x, \nu_{\mathcal{O}}(x))}{\partial \nu} (\kappa(x) - \tilde{\kappa}(x)) \, \mathrm{d}\mathcal{H}^1(x) = 0 \,,$$

where the convergence follows from the uniform convergence $m_{\varepsilon} * \varphi \to \varphi$ and Lebesgue's dominated convergence theorem. By the arbitrariness of $\varphi$, the fundamental lemma of the calculus of variations implies $\kappa = \tilde{\kappa}$. □

REMARK 1. *By construction, the generalized curvature coincides with the classical notion of curvature for sets $\mathcal{O} \subset \Omega$ with smooth boundary as well as with the curvature of level lines in smooth images. Indeed, if $\mathcal{O} = \{u > 0\}$ for some smooth $u : \Omega \to \mathbb{R}$ and $\partial\mathcal{O}$ corresponds to the null level-set of $u$, then for any arclength parametrization $x : I \to \Omega$ on some open interval $I \subset \mathbb{R}$, which is oriented counterclockwise with respect to $\mathcal{O}$, we have for $s \in I$ that $\nabla u\big(x(s)\big) \cdot \dot{x}(s) = 0$ implying $\nu_{\mathcal{O}}(x(s)) = \dot{x}(s)^{\perp} = \frac{\nabla u(x(s))}{|\nabla u(x(s))|}$. Consequently, writing $\nu(s) = \nu_{\mathcal{O}}(x(s))$ and skipping the arguments in $\nu$, $\dot{x}$, $\nabla u$, $\nabla^2 u$, we get*

$$\frac{\mathrm{d}}{\mathrm{d}s}\nu(s) = \lim_{h \to 0} \frac{\big(\nu(s+h) - \nu(s)\big) \cdot \nu(s)^{\perp}}{h} = \frac{1}{|\nabla u|}\Big(\nabla^2 u\Big(I - \frac{\nabla u}{|\nabla u|} \otimes \frac{\nabla u}{|\nabla u|}\Big)\dot{x}\Big) \cdot \nu^{\perp}$$

$$= \frac{1}{|\nabla u|}\big(\nabla^2 u(I - \nu \otimes \nu)\nu^{-\perp}\big) \cdot \nu^{\perp} = -\frac{(\nabla^2 u)\nu^{\perp} \cdot \nu^{\perp}}{|\nabla u|}\,.$$

*On the other hand, computing $-\mathrm{div}\big(\frac{\nabla u}{|\nabla u|}\big)$ yields, since $\mathrm{tr}(A) = \mathrm{tr}(V^T A V)$ for any orthonormal $V$,*

$$-\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big) = -\frac{1}{|\nabla u|}\mathrm{tr}\Big(\nabla^2 u\Big(I - \frac{\nabla u}{|\nabla u|} \otimes \frac{\nabla u}{|\nabla u|}\Big)\Big)$$

$$= -\frac{1}{|\nabla u|}\Big(\big(\nabla^2 u(I - \nu \otimes \nu)\nu\big) \cdot \nu + \big(\nabla^2 u(I - \nu \otimes \nu)\nu^{\perp}\big) \cdot \nu^{\perp}\Big)$$

$$= -\frac{(\nabla^2 u)\nu^{\perp} \cdot \nu^{\perp}}{|\nabla u|}\,,$$

*hence $\frac{\mathrm{d}}{\mathrm{d}s}\nu\big(x(s)\big)$ and $-\mathrm{div}\big(\frac{\nabla u}{|\nabla u|}\big)\big(x(s)\big)$ coincide for all $s \in I$. Therefore, analogously to (2.2), for all $\psi \in \mathcal{C}_0^{\infty}(\Omega \times S^1; \mathbb{R})$, the identity*

$$\int_{\partial\mathcal{O}} \nabla_x \psi(x, \nu_{\mathcal{O}}(x)) \cdot \nu_{\mathcal{O}}(x)^{-\perp} - \frac{\partial\psi(x, \nu_{\mathcal{O}}(x))}{\partial\nu}\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big)(x)\,\mathrm{d}\mathcal{H}^1 = 0$$

*holds, meaning that $\kappa_{\mathcal{O}} = -\mathrm{div}\big(\frac{\nabla u}{|\nabla u|}\big)$.*

The notion of generalized curvature now allows to introduce the elastica energy with generic integrands $f$ and for general functions in $\mathrm{BV}(\Omega; \mathbb{R})$ (the space of scalar functions of bounded variation on $\Omega$).

DEFINITION 2.3 (Elastica energy). *Let $f : \Omega \times S^1 \times \mathbb{R} \to [0, \infty]$ be lower semi-continuous. For a characteristic function $u = \chi_{\mathcal{O}}$ of $\mathcal{O} \subset \Omega$ we then define*

$$(2.4) \quad R_{\mathrm{el}}(u) = \int_{\partial^*\mathcal{O}} f(x, \nu_{\mathcal{O}}(x), \kappa_{\mathcal{O}}(x))\,\mathrm{d}\mathcal{H}^1(x) = \int_{\Omega} f(x, \nu_{\mathcal{O}}(x), \kappa_{\mathcal{O}}(x))\,\mathrm{d}|\nabla u|(x)\,,$$

*if $\mathcal{O}$ has finite perimeter and $\kappa_{\mathcal{O}}$ exists, and $R_{\mathrm{el}}(u) = \infty$ else. For a general image $u \in L^1(\Omega; \mathbb{R})$ we abbreviate $\mathcal{O}_u^t := \{x \in \Omega \,:\, u(x) \geq t\}$ and define the elastica energy as*

$$(2.5) \quad\quad\quad\quad R_{\mathrm{el}}(u) = \int_{\mathbb{R}} R_{\mathrm{el}}(\chi_{\mathcal{O}_u^t})\,\mathrm{d}t\,,$$

*where the right-hand side shall be infinite if $u \notin \mathrm{BV}(\Omega; \mathbb{R})$ or $t \mapsto R_{\mathrm{el}}(\chi_{\mathcal{O}_u^t})$ is not in $L^1(\mathbb{R}; \mathbb{R})$.*

REMARK 2. *Note that the integral in (2.4) is indeed well-defined. Since $f$ is a non-negative, lower semi-continuous integrand and $\nu_{\mathcal{O}}$ as well as $\kappa_{\mathcal{O}}$ are measurable and*

uniquely determined almost everywhere with respect to $\mathcal{H}^1 \llcorner \partial^* \mathcal{O}$ (see, e. g., [2] as well as Theorem 2.2), the integrand is still measurable. Changing $\nu_{\mathcal{O}}$ and $\kappa_{\mathcal{O}}$ on an $\mathcal{H}^1$-negligible set does not affect the integral, hence $R_{\mathrm{el}}(\chi_{\mathcal{O}})$ is well-defined. Likewise, if $u \in \mathrm{BV}(\Omega; \mathbb{R})$, $\mathcal{O}_u^t$ has finite perimeter and $\partial^* \mathcal{O}_u^t$ can be uniquely defined for Lebesgue-almost every $t \in \mathbb{R}$. Therefore, $R_{\mathrm{el}}$ is also well-defined on $\mathrm{BV}(\Omega; \mathbb{R})$.

REMARK 3. *If $u : \Omega \to \mathbb{R}$ is smooth, then analogously to Remark 1 and by virtue of the coarea formula, we are able to deduce that*

$$R_{\mathrm{el}}(u) = \int_{\mathbb{R}} R_{\mathrm{el}}(\chi_{\mathcal{O}_u^t}) \, \mathrm{d}t = \int_{\mathbb{R}} \int_{\partial \mathcal{O}_u^t} f\Big(x, \frac{\nabla u}{|\nabla u|}(x), -\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big)(x)\Big) \, \mathrm{d}\mathcal{H}^1 \, \mathrm{d}t$$
$$= \int_{\Omega} f\Big(x, \frac{\nabla u}{|\nabla u|}(x), -\mathrm{div}\Big(\frac{\nabla u}{|\nabla u|}\Big)(x)\Big) |\nabla u|(x) \, \mathrm{d}x,$$

*i. e., for smooth functions, $R_{\mathrm{el}}$ indeed yields the well-known level-set free version of the elastica energy.*

**2.2. Interlude: The importance of the total variation in the elastica energy.** As mentioned earlier, we particularly have in mind the elastica integrand $f(x, \nu, \kappa) = a\kappa^2 + b$. This implies that we do not only measure the actual bending energy of the image level lines, but also their length. Put differently, the elastica energy of an image $u$ is composed of its total variation semi-norm plus the pure curvature term. This may seem undesirable at first, and one is tempted to consider the mere Willmore energy density $f(x, \nu, \kappa) = a\kappa^2$. However, surprisingly, in that case the convex relaxation $R_{\mathrm{el}}^{**}$ of $R_{\mathrm{el}}$ (the largest, convex lower semi-continuous functional below $R_{\mathrm{el}}$) vanishes completely. The reason lies in the fact that smooth images can be reproduced via convex combinations of striped images.

THEOREM 2.4. *Any image $u \in \mathcal{C}_0^\infty(\Omega; \mathbb{R})$ can be approximated in $\mathcal{C}^k(\overline{\Omega}; \mathbb{R})$ for any $k > 0$ via convex combinations of smooth images whose level lines have zero curvature, that is, for $n \in \mathbb{N}$ there exist $N_n \in \mathbb{N}$, $\alpha_i^n \in [0, 1]$, $\nu_i^n \in S^1$, $u_i^n \in \mathcal{C}^\infty(\overline{\Omega}; \mathbb{R})$ uniformly bounded, $i = 1, \ldots, N_n$, with*

$$\nabla u_i^n \cdot (\nu_i^n)^\perp = 0, \quad \sum_{i=1}^{N_n} \alpha_i^n = 1, \quad and \quad \sum_{i=1}^{N_n} \alpha_i^n u_i^n \underset{n \to \infty}{\longrightarrow} u \text{ in } \mathcal{C}^k(\overline{\Omega}; \mathbb{R}).$$

*Proof.* Without loss of generality, let $\Omega$ be the open unit ball centered around zero. Let $u \in \mathcal{C}_0^\infty(\Omega; \mathbb{R})$, then the Radon transform $Ru$ of $u$ is defined as

$$Ru : S^1 \times (-1, 1) \to \mathbb{R}, \ (\nu, s) \mapsto \int_{\{x \in \Omega \, : \, \nu \cdot x = s\}} u(x) \, \mathrm{d}\mathcal{H}^1(x).$$

Due to the density of $\mathcal{C}_0^\infty(\Omega; \mathbb{R})$ in $L^2(\Omega; \mathbb{R})$ and the estimate $\|Ru\|_{L^2(S^1 \times (-1,1); \mathbb{R})}^2 \leq 4\pi \|u\|_{L^2(\Omega; \mathbb{R})}^2$, the Radon transform can be uniquely extended to a continuous operator $R : L^2(\Omega; \mathbb{R}) \to L^2(S^1 \times (-1, 1); \mathbb{R})$ [20]. Its adjoint can be computed via

$$R^* g : \Omega \to \mathbb{R}, \ x \mapsto \int_{S^1} g(\nu, \nu \cdot x) \, \mathrm{d}\mathcal{H}^1(\nu)$$

for $g \in L^2(S^1 \times (-1, 1); \mathbb{R})$. The same holds true if the reals $\mathbb{R}$ are replaced by $\mathbb{C}$.

For a given image $u \in L^2(\Omega; \mathbb{C})$ we intend to find a preimage under $R^*$. By the Fourier slice theorem,

$$\mathcal{F}_2 Ru(\nu, \omega) = \sqrt{2\pi} \hat{u}(\omega \nu),$$

where the hat denotes the Fourier transform, and $\mathcal{F}_2 : L^2(S^1 \times (-1,1); \mathbb{C}) \to L^2(S^1 \times \mathbb{R}; \mathbb{C})$ stands for the Fourier transform only with respect to the second variable, that is,

$$\mathcal{F}_2 Ru(\nu, \omega) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} \exp(-is\omega) Ru(\nu, s) \, \mathrm{d}s \,.$$

(For the Fourier transform, $u$ is extended to $\mathbb{R}^2$ by zero.) Thus, for $\tilde{g} \in L^2(S^1 \times \mathbb{R}; \mathbb{C})$ we have

$$\int_{S^1} \int_{\mathbb{R}} \mathcal{F}_2 Ru(\nu, \omega) \overline{\tilde{g}(\nu, \omega)} \, \mathrm{d}\omega \, \mathrm{d}\nu = \sqrt{2\pi} \int_{S^1} \int_{\mathbb{R}} \hat{u}(\omega\nu) \overline{\tilde{g}(\nu, \omega)} \, \mathrm{d}\omega \, \mathrm{d}\nu$$

$$= \sqrt{2\pi} \int_{\mathbb{R}^2} \hat{u}(x) \frac{\overline{\tilde{g}(\frac{x}{|x|}, |x|) + \tilde{g}(\frac{-x}{|x|}, -|x|)}}{|x|} \, \mathrm{d}x = \sqrt{2\pi} \int_{\mathbb{R}^2} u(x) \overline{\left[ y \mapsto \frac{\tilde{g}(\frac{y}{|y|}, |y|) + \tilde{g}(\frac{-y}{|y|}, -|y|)}{|y|} \right]^{\smallsmile}(x)} \, \mathrm{d}x \,,$$

where a bar denotes the complex conjugate, $\smallsmile$ denotes the inverse Fourier transform, and in the last step we have used Parseval's identity. We thus obtain

$$R^* \mathcal{F}_2^* \tilde{g} = \sqrt{2\pi} \left[ x \mapsto \frac{\tilde{g}(\frac{x}{|x|}, |x|) + \tilde{g}(\frac{-x}{|x|}, -|x|)}{|x|} \right]^{\smallsmile}$$

which for given $u \in L^2(\Omega; \mathbb{C})$ implies

$$R^* \mathcal{F}_2^* \tilde{g} = u \qquad \text{where } \mathcal{F}_2^* \tilde{g} : (\nu, s) \mapsto \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(is\omega) \tilde{g}(\nu, \omega) \, \mathrm{d}\omega \text{ and } \tilde{g} : (\nu, \omega) \mapsto \frac{1}{\sqrt{8\pi}} \hat{u}(\omega\nu) |\omega| \,,$$

if this is well-defined. In the case of real-valued $u$ we may replace $g$ by its real part.

Now consider $u \in \mathcal{C}_0^\infty(\Omega; \mathbb{R})$ so that $\hat{u} \in \mathcal{C}^\infty(\mathbb{R}^2; \mathbb{C})$ due to the compact support of $u$. Even more, since $u$ (extended to $\mathbb{R}^2$ by zero) is a Schwartz function, $\hat{u}$ and thus also $\xi \mapsto \hat{u}(\xi)|\xi|$ decay faster than $|\xi|^{-k}$ for any $k > 0$. As a result, the mapping $\nu \mapsto \hat{u}(\cdot\nu)|\cdot|$ is continuous from $S^1$ into $L^2(\mathbb{R}, 1 + \omega^k) = \{ h \in L^2(\mathbb{R}; \mathbb{C}) : \int_{\mathbb{R}} |h(\omega)|^2 (1 + \omega^k) \, \mathrm{d}\omega < \infty \}$: For any sequence $\nu_j \to_{j\to\infty} \nu$ we have

$$\int_{\mathbb{R}} \left| \hat{u}(\omega\nu_j)|\omega| - \hat{u}(\omega\nu)|\omega| \right|^2 (1 + \omega^k) \, \mathrm{d}\omega \xrightarrow[j\to\infty]{} 0$$

by Lebesgue's theorem since the integrand converges pointwise to zero and is bounded by the integrable function $2[(\hat{u}(\cdot\nu)| \cdot |)^2 + \max_\nu (\hat{u}(\cdot\nu)| \cdot |)^2](1 + \cdot^k)$. Furthermore, the (inverse) Fourier transform is continuous from $L^2(\mathbb{R}, 1 + \omega^{2k+2})$ into $H^k(\mathbb{R}; \mathbb{C})$ so that the mapping $\nu \mapsto \mathcal{F}_2^* (\hat{u}(\cdot\nu)| \cdot |)$ is continuous from $S^1$ into $H^k(\mathbb{R}; \mathbb{C})$ (and thus into $\mathcal{C}^{k-2}([-1,1]; \mathbb{C})$) for any $k > 0$. Summarizing, there is a $g \in L^2(S^1 \times (-1,1); \mathbb{R})$ with $R^*g = u$, which is continuous in its first and smooth up to the boundary in its second argument.

Now we approximate $R^*g$ by a Riemann sum, i.e. for $n \in \mathbb{N}$ we choose $N_n = 2^n$, $\alpha_i^n = \frac{1}{N_n}$, $\nu_i^n = (\cos(2\pi \frac{i}{N_n}), \sin(2\pi \frac{i}{N_n}))^T$ and $u_i^n : x \mapsto 2\pi g(\nu_i^n, \nu_i^n \cdot x)$. If we show

$$\sum_{i=1}^{N_n} \alpha_i^n u_i^n = \frac{2\pi}{2^n} \sum_{i=1}^{2^n} g(\nu_i^n, \nu_i^n \cdot (\cdot)) \xrightarrow[n\to\infty]{} \int_{S^1} g(\nu, \nu \cdot (\cdot)) \, \mathrm{d}\mathcal{H}^1(\nu) = u \quad \text{in } \mathcal{C}^k(\overline{\Omega}; \mathbb{R}) \text{ for } k > 0$$

this concludes the proof. Indeed, $\sum_{i=1}^{N_n} \alpha_i^n u_i^n$ is a Cauchy sequence in $\mathcal{C}^k(\overline{\Omega}; \mathbb{R})$: For given $\varepsilon > 0$ and any $(x, \nu) \in \overline{\Omega} \times S^1$ there is $\delta_{(x,\nu)}$ such that $|D_x^l g(\nu, x \cdot \nu) - D_x^l g(\tilde{\nu}, \tilde{x} \cdot$

$\tilde{\nu})| < \varepsilon$ for all $0 \leq l \leq k$ and all $|(\tilde{x}, \tilde{\nu}) - (x, \nu)| < \delta_{(x,\nu)}$. The balls $B_{\delta_{(x,\nu)}}(x, \nu)$ cover $\overline{\Omega} \times S^1$ so that we can choose a finite subcover. Of this subcover, choose $\delta$ as the minimum ball radius and $n$ such that $2^{1-n} < \delta$. Then

$$\sum_{l=0}^{k} \left| D_x^l \left( \sum_{i=1}^{N_n} \alpha_i^n u_i^n(x) - \sum_{i=1}^{N_m} \alpha_i^m u_i^m(x) \right) \right| \leq 2\pi(k+1)\varepsilon \quad \forall m \geq n \, .$$

Hence, $\sum_{i=1}^{N_n} \alpha_i^n u_i^n \to_{n\to\infty} \tilde{u}$ in $\mathcal{C}^k(\overline{\Omega}; \mathbb{R})$ for some $\tilde{u} \in \mathcal{C}^k(\overline{\Omega}; \mathbb{R})$. However,

$$\tilde{u}(x) = \lim_{n\to\infty} \sum_{i=1}^{N_n} \alpha_i^n u_i^n(x) = \lim_{n\to\infty} \frac{2\pi}{2^n} \sum_{i=1}^{2^n} g(\nu_i^n, \nu_i^n \cdot x) = \int_{S^1} g(\nu, \nu \cdot x) \, d\mathcal{H}^1(\nu) = u(x) \quad \forall x \in \overline{\Omega}$$

due to the convergence of the Riemann sum. □ Since images with straight level lines have zero bending energy, Theorem 2.4 implies $R_{\mathrm{el}}^{**}(u) = 0$ for any $u \in \mathcal{C}_0^\infty(\Omega; \mathbb{R})$ and $f(x, \nu, \kappa) = a\kappa^2$. Via density, this even holds on all of $L^p(\Omega; \mathbb{R})$, $1 \leq p < \infty$.

COROLLARY 2.5 (Convex relaxation of curvature functional). *Let $1 \leq p < \infty$. If $f(x, \nu, \kappa) = a\kappa^2$, then the convex relaxation of $R_{\mathrm{el}}$ on $L^p(\Omega; \mathbb{R})$ is identically zero.*

**2.3. Functional lifting of Euler's elastica energy.** To derive a convex approximation of Euler's elastica energy, we employ the technique of functional lifting. For an image $u \in \mathrm{BV}(\Omega; \mathbb{R})$ whose super-level sets admit a generalized curvature $\kappa_{\mathcal{O}_u^t}$, consider the polar decomposition of its gradient $\nabla u$ into the normalized vector field $\nu_u = \frac{\nabla u}{|\nabla u|}$ and the scalar measure $|\nabla u|$. (Note that for almost all $t \in \mathbb{R}$ we have $\mathcal{H}^1 \llcorner \partial^* \mathcal{O}_u^t$-almost everywhere $\nu_u = \nu_{\mathcal{O}_u^t}$.) The vector component of the polar decomposition shall serve as a new dimension, and additionally we introduce the level line curvature as a dimension. Specifically, we aim to represent the measure $\nabla u$ by a non-negative measure $v[u]$ on $\Omega \times S^1 \times \mathbb{R}$ such that $v[u]$ has support on those $(x, \nu, \kappa) \in \Omega \times S^1 \times \mathbb{R}$ for which $\nu$ and $\kappa$ coincide with the normal $\nu_{\mathcal{O}_u^t}(x)$ and the curvature $\kappa_{\mathcal{O}_u^t}(x)$ of a level line $\partial^* \mathcal{O}_u^t \ni x$. Furthermore, the image gradient shall be recovered by the reversal of the lifting,

$$(2.6) \qquad \qquad \nabla u = \int_{S^1 \times \mathbb{R}} \nu \, dv[u](\cdot, \nu, \kappa)$$

(the right-hand side just represents an intuitive notation for the pushforward measure $(\pi^\Omega)_\#(\nu v[u])$ under the projection $\pi^\Omega$ defined in Section 1.2, where $\nu v[u]$ denotes the measure obtained by multiplying $v[u]$ with the continuous function $(x, \nu, \kappa) \mapsto \nu$). More specifically, we introduce the following (where $\mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ denotes the continuous functions on $\Omega \times S^1 \times \mathbb{R}$ with compact support).

DEFINITION 2.6 (Lifted image gradient). *Let $u \in \mathrm{BV}(\Omega; \mathbb{R})$ such that the generalized curvature $\kappa_{\mathcal{O}_u^t}$ is defined for almost all $t \in \mathbb{R}$. Then, by Riesz' theorem, the relation*

$$(2.7)$$
$$\int_{\mathbb{R}} \int_{\partial^* \mathcal{O}_u^t} \phi(x, \nu_{\mathcal{O}_u^t}(x), \kappa_{\mathcal{O}_u^t}(x)) \, d\mathcal{H}^1(x) \, dt = \int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu, \kappa) \, dv[u](x, \nu, \kappa) \quad \forall \phi \in \mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$$

*uniquely defines a nonnegative, regular countably additive measure $v[u] \in \mathrm{rca}(\Omega \times S^1 \times \mathbb{R})$ [26], which we shall call the lifting of $\nabla u$.* We can readily validate that (2.6) is indeed satisfied.

THEOREM 2.7 (Reversal of lifting). *Any image $u \in \mathrm{BV}(\Omega; \mathbb{R})$ which admits a lifting satisfies (2.6).*

*Proof.* By the disintegration theorem (see e. g. [3]) there exists a family $(v_x)_{x \in \Omega}$ of probability measures on $S^1 \times \mathbb{R}$ such that for all $\varphi \in \mathcal{C}_c(\Omega; \mathbb{R}^2)$ we have

$$\int_\Omega \varphi \cdot \mathrm{d}(\pi^\Omega)_\# (\nu v[u]) = \int_\Omega \int_{S^1 \times \mathbb{R}} \varphi(x) \, \mathrm{d}v_x \cdot \mathrm{d}(\pi^\Omega)_\#(\nu v[u])(x) = \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu \, \mathrm{d}v[u](x, \nu, \kappa) \, .$$

Also, by the coarea formula for functions in $\mathrm{BV}(\Omega; \mathbb{R})$,

$$\int_\Omega \varphi \cdot \mathrm{d}\nabla u = \int_\mathbb{R} \int_{\partial^* \mathcal{O}_u^t} \varphi \cdot \nu_{\mathcal{O}_u^t} \, \mathrm{d}\mathcal{H}^1 \, \mathrm{d}t \, .$$

We will show the right-hand sides of both previous equations to be equal: For $k \in \mathbb{N}$ let $\phi_k : \Omega \times S^1 \times \mathbb{R} \to \mathbb{R}$, $\phi_k(x, \nu, \kappa) = (\varphi(x) \cdot \nu) \max(0, \min(1, k - |\kappa|))$. Furthermore, let us abbreviate $[\cdot]^+ = \max(0, \cdot)$ and $[\cdot]^- = \min(0, \cdot)$. Obviously, $[\phi_k]^\pm \to [\varphi \cdot \nu]^\pm$ pointwise and monotonously. Thus,

$$\int_{\Omega \times S^1 \times \mathbb{R}} [\varphi(x) \cdot \nu]^\pm \, \mathrm{d}v[u](x, \nu, \kappa) = \lim_{k \to \infty} \int_{\Omega \times S^1 \times \mathbb{R}} [\phi_k]^\pm \, \mathrm{d}v[u]$$

$$= \lim_{k \to \infty} \int_\mathbb{R} \int_{\partial^* \mathcal{O}_u^t} [\phi_k(x, \nu_{\mathcal{O}_u^t}(x), \kappa_{\mathcal{O}_u^t}(x))]^\pm \, \mathrm{d}\mathcal{H}^1(x) \, \mathrm{d}t = \int_\mathbb{R} \int_{\partial^* \mathcal{O}_u^t} [\varphi \cdot \nu_{\mathcal{O}_u^t}]^\pm \, \mathrm{d}\mathcal{H}^1 \, \mathrm{d}t \, .$$

where the first equality is due to Beppo Levi's theorem and the last due to Lebesgue's dominated convergence theorem (note that the positive and negative part of a continuous function is again continuous). Adding the equation for the positive and the negative part yields the desired equality. $\square$

The elastica energy can now be expressed as a function of the lifted image gradient.

THEOREM 2.8 (Elastica energy in lifted space). *Let $f$ and $R_{\mathrm{el}}$ be given according to Definition 2.3. If $u \in \mathrm{BV}(\Omega; \mathbb{R})$ admits a lifting $v[u]$ (else $R_{\mathrm{el}}(u) = \infty$ anyway), then*

$$(2.8) \qquad R_{\mathrm{el}}(u) = \int_{\Omega \times S^1 \times \mathbb{R}} f(x, \nu, \kappa) \, \mathrm{d}v[u](x, \nu, \kappa) \, .$$

*Proof.* First observe that since $f$ is lower semi-continuous and non-negative, we can write it as the pointwise limit $f(x, \nu, \kappa) = \lim_{k \to \infty} Y_k f(x, \nu, \kappa)$ (see e. g. [6]) of the Lipschitz-continuous Yosida transforms

$$Y_k f(x, \nu, \kappa) = \inf \left\{ f(\tilde{x}, \tilde{\nu}, \tilde{\kappa}) + k \sqrt{|x - \tilde{x}|^2 + \mathrm{dist}(\nu, \tilde{\nu})^2 + (\kappa - \tilde{\kappa})^2} : (\tilde{x}, \tilde{\nu}, \tilde{\kappa}) \in \Omega \times S^1 \times \mathbb{R} \right\} ,$$

where $\mathrm{dist}(\nu, \tilde{\nu})$ shall be the metric distance between $\nu$ and $\tilde{\nu}$ induced by the embedding of $S^1$ in $\mathbb{R}^2$. Furthermore, we can define a monotonously increasing sequence of cutoff functions $T_k \in \mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ with $T_k \equiv 1$ on $\{x \in \Omega : \mathrm{dist}(x, \partial\Omega) \geq 1/k\} \times S^1 \times [-k, k]$. The sequence $f_k = (Y_k f) T_k$, $k \in \mathbb{N}$, then lies in $\mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ and converges pointwise and monotonously against $f$. By Beppo Levi's theorem,

$$R_{\mathrm{el}}(u) = \lim_{k \to \infty} \int_\mathbb{R} \int_{\partial^* \mathcal{O}_u^t} f_k(x, \nu_{\mathcal{O}_u^t}(x), \kappa_{\mathcal{O}_u^t}(x)) \, \mathrm{d}\mathcal{H}^1(x) \, \mathrm{d}t = \lim_{k \to \infty} \int_{\Omega \times S^1 \times \mathbb{R}} f_k \, \mathrm{d}v[u] = \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v[u] \, .$$

$\square$

In the following, let us assume $f$ to be strictly positive so that there exists a positive, continuous function $w \in \mathcal{C}(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ with $0 < w \leq f$. Let us characterize the space of the measures $v[u]$ with finite $R_{\mathrm{el}}(u)$. Equation (2.8) implies

$\|wv[u]\|_{\mathrm{TV}(\Omega \times S^1 \times \mathbb{R};\mathbb{R})} < \infty$. Hence, the measures $v[u]$ can be interpreted as elements of the Banach space

$$(2.9) \qquad \mathcal{M} = \left\{ v \text{ Borel measure } : \ wv \in \mathrm{rca}(\Omega \times S^1 \times \mathbb{R};\mathbb{R}) \right\}$$

with norm $\|v\|_{\mathcal{M}} = \|wv\|_{\mathrm{TV}(\Omega \times S^1 \times \mathbb{R};\mathbb{R})}$. By Riesz, $\mathcal{M}$ is isometrically isomorphic to the dual space of

$$(2.10) \quad \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R};\mathbb{R}) = \left\{ \phi : \Omega \times S^1 \times \mathbb{R} \to \mathbb{R} \ : \ \phi/w \in \mathcal{C}_0(\Omega \times S^1 \times \mathbb{R};\mathbb{R}) \right\}$$

with the weighted norm $\|\phi\|_w = \|\phi/w\|_\infty$. Here, $\mathcal{C}_0(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ denotes the closure of $\mathcal{C}_c(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ with respect to the supremum norm $\|\cdot\|_\infty$. Equivalently, $\mathcal{C}_w(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ can be interpreted as the closure of $\mathcal{C}_c(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ with respect to the norm $\|\cdot\|_w$. By the same standard argument as for $\mathcal{C}_0(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$, $\mathcal{C}_w(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ is a separable Banach space. In the following we will always assume $w(x,\nu,\kappa) \geq c_0(|\kappa| + 1)$ for some constant $c_0 > 0$.

For later reference, let us also state the following property of the lifted elastica energy.

THEOREM 2.9 (Lower semi-continuity with respect to lifted image). $R_{\mathrm{el}}^l(v) := \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v$ *is weak-\* (and thus also sequentially weak-\*) lower semi-continuous on* $\{v \in \mathcal{M} \ : \ v \geq 0\}$.

*Proof.* Let $f_k$ be a monotonously increasing sequence in $\mathcal{C}_c(\Omega \times S^1 \times \mathbb{R};\mathbb{R})$ converging pointwise to $f$ (as in the proof of Theorem 2.8). By Beppo-Levi's theorem, for any $v \in \mathcal{M}$ with $v \geq 0$ we have

$$R_{\mathrm{el}}^l(v) = \lim_{k \to \infty} \int_{\Omega \times S^1 \times \mathbb{R}} f_k \, \mathrm{d}v = \sup_{k \in \mathbb{N}} \int_{\Omega \times S^1 \times \mathbb{R}} f_k \, \mathrm{d}v \,,$$

which is weak-\* lower semi-continuous on $\{v \in \mathcal{M} : v \geq 0\}$ as the pointwise supremum of weak-\* continuous functionals. $\square$

**3. A convex, lower semi-continuous approximation.** Here we propose a convex approximation $R_{\mathrm{el}}^{\mathrm{ap}}$ of Euler's elastica energy, which turns out to be not quite the convex relaxation $R_{\mathrm{el}}^{**}$. We examine some of the properties of $R_{\mathrm{el}}^{\mathrm{ap}}$ and of the involved minimizing measures, and we also find a predual representation of $R_{\mathrm{el}}^{\mathrm{ap}}$.

**3.1. Relaxing the elastica energy.** The convex relaxation of $R_{\mathrm{el}}$ on $L^1(\Omega;\mathbb{R})$, i. e. the largest convex, lower semi-continuous functional on $L^1(\Omega;\mathbb{R})$ below $R_{\mathrm{el}}$, can be expressed as
$$(3.1)$$
$$R_{\mathrm{el}}^{**}(u) = \inf \left\{ \liminf_{j \to \infty} \sum_{i=1}^{n^j} \alpha_i^j R_{\mathrm{el}}(u_i^j) \ : \ n^j \in \mathbb{N}, u_i^j \in L^1(\Omega;\mathbb{R}), \alpha_i^j \in [0,1], \sum_{i=1}^{n^j} \alpha_i^j = 1, \sum_{i=1}^{n^j} \alpha_i^j u_i^j \xrightarrow{L^1(\Omega;\mathbb{R})} u \right\} .$$

Hence, for any $u \in L^1(\Omega;\mathbb{R})$ with $R_{\mathrm{el}}^{**}(u) < \infty$ there exist convex combinations $u_j = \sum_{i=1}^{n^j} \alpha_i^j u_i^j$, $j \in \mathbb{N}$, of $L^1(\Omega;\mathbb{R})$-functions with $u_j \to u$ in $L^1(\Omega;\mathbb{R})$ and $\sum_{i=1}^{n^j} \alpha_i^j \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v[u_i^j] \to R_{\mathrm{el}}^{**}(u)$. This in turn implies the boundedness of $\sum_{i=1}^{n^j} \alpha_i^j v[u_i^j]$ in $\mathcal{M}$ so that (upon choosing a subsequence) we may additionally assume $\sum_{i=1}^{n^j} \alpha_i^j v[u_i^j] \xrightarrow{*} v$. Theorem 2.9 now implies

$$(3.2) \qquad\qquad R_{\mathrm{el}}^{**}(u) \geq \inf_{v \in \mathcal{M}(u)} \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v \,,$$

where $\mathcal{M}(u)$ is the set of all measures $v$ in $\mathcal{M}$ such that there exist convex combinations $u_j = \sum_{i=1}^{n^j} \alpha_i^j u_i^j$, $j \in \mathbb{N}$, of $L^1(\Omega; \mathbb{R})$-functions with $u_j \to u$ in $L^1(\Omega; \mathbb{R})$ and $\sum_{i=1}^{n^j} \alpha_i^j v[u_i^j] \overset{*}{\rightharpoonup} v$. We will choose the right-hand side of (3.2) as starting point to develop a convex, lower semi-continuous approximation of $R_{\mathrm{el}}$. Note that it may well be that equality holds in (3.2), however, a corresponding analysis promises to be difficult. A related problem considers integral functionals on $L^p$, where Young measures can be interpreted as liftings of the $L^p$-arguments [13]. Here indeed, the convex relaxation (with respect to the weak topology) of a sufficiently coercive integral functional, evaluated at $u \in L^p$, $p > 1$, is shown to equal the infimum over a set of Young measures compatible with $u$ and satisfying a $p$-integrability condition.

THEOREM 3.1 (Characterization of $\mathcal{M}(u)$). *$\mathcal{M}(u)$ is the intersection of the closed convex hull (with respect to sequential weak-\* convergence in $\mathcal{M}$) of $\{v[\tilde{u}] : \tilde{u} \in L^1(\Omega; \mathbb{R}), R_{\mathrm{el}}(\tilde{u}) < \infty\}$ with the affine subspace of measures $v$ in $\mathcal{M}$ satisfying*

$$(3.3) \qquad \nabla u = \int_{S^1 \times \mathbb{R}} \nu \, \mathrm{d}v(\cdot, \nu, \kappa).$$

*Proof.* Let $v \in \mathcal{M}(u)$, then by definition there exist convex combinations $u_j = \sum_{i=1}^{n^j} \alpha_i^j u_i^j$, $j \in \mathbb{N}$, of $L^1(\Omega; \mathbb{R})$-functions with $u_j \to u$ in $L^1(\Omega; \mathbb{R})$ and $v_j = \sum_{i=1}^{n^j} \alpha_i^j v[u_i^j] \overset{*}{\rightharpoonup} v$ in $\mathcal{M}$ so that $v$ indeed lies in the above closed convex hull. Furthermore, by Theorem 2.7,

$$\nabla u_j = \sum_{i=1}^{n^j} \alpha_i^j (\pi^\Omega)_\# (\nu v[u_i^j]) = (\pi^\Omega)_\# (\nu v_j).$$

Since $u_j \to u$ in $L^1(\Omega; \mathbb{R})$ and boundedness of $\nabla u_j$ for BV-functions $u_j$ implies the weak-\* convergence $u_j \to u$ in $\mathrm{BV}(\Omega; \mathbb{R})$, the left-hand side of the equation converges weakly-\* to $\nabla u$. The right-hand side converges to $(\pi^\Omega)_\# (\nu v)$ in the same sense, as

$$\int_\Omega \varphi \cdot \mathrm{d}(\pi^\Omega)_\# (\nu v_j) = \int_{\Omega \times S^1 \times \mathbb{R}} \varphi \cdot \nu \, \mathrm{d}v_j \underset{j \to \infty}{\longrightarrow} \int_{\Omega \times S^1 \times \mathbb{R}} \varphi \cdot \nu \, \mathrm{d}v = \int_\Omega \varphi \cdot \mathrm{d}(\pi^\Omega)_\# (\nu v)$$

for any $\varphi \in \mathcal{C}_c(\Omega; \mathbb{R}^2)$, which concludes the proof of (3.3).

Conversely, let $v$ in the above closed convex hull satisfy (3.3). Then again, there exist convex combinations $u_j = \sum_{i=1}^{n^j} \alpha_i^j u_i^j$ with $v_j = \sum_{i=1}^{n^j} \alpha_i^j v[u_i^j] \overset{*}{\rightharpoonup} v$ (where without loss of generality we may assume the $u_i^j$ to have the same mean as $u$). By Theorem 2.7, this implies $\nabla u_j = (\pi^\Omega)_\# (\nu v_j) \overset{*}{\rightharpoonup} (\pi^\Omega)_\# (\nu v)$. Thus, by Poincaré's inequality for BV-functions, the $u_j$ converge weakly-\* in $\mathrm{BV}(\Omega; \mathbb{R})$ against some $\tilde{u} \in \mathrm{BV}(\Omega; \mathbb{R})$ with same mean as $u$ and with $\nabla \tilde{u} = (\pi^\Omega)_\# (\nu v)$. Again by Poincaré, this implies $\tilde{u} = u$. $\square$ By (2.5) and the linearity of (2.7), the closed convex hull of $\{v[\tilde{u}] : \tilde{u} \in L^1(\Omega; \mathbb{R}), R_{\mathrm{el}}(\tilde{u}) < \infty\}$ in the above characterization of $\mathcal{M}(u)$ may also be expressed as follows: Denoting by $O$ the set of all simply connected subsets $\mathcal{O} \subset \Omega$ with $R_{\mathrm{el}}(\chi_\mathcal{O}) < \infty$ and their complements, the closed convex hull of $\{v[\tilde{u}] : \tilde{u} \in L^1(\Omega; \mathbb{R}), R_{\mathrm{el}}(\tilde{u}) < \infty\}$ corresponds to the closed convex hull of

$$\mathcal{M}_O = \{v[\tilde{u}] : \tilde{u} = s\chi_\mathcal{O}, s \geq 0, \mathcal{O} \in O\}$$

$$= \Big\{v \in \mathcal{M} : v \geq 0, \exists \mathcal{O} \in O, s \geq 0 \,\forall \phi \in \mathcal{C}_w^\infty(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) :$$

$$(3.4)$$

$$\int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu, \kappa) \, \mathrm{d}v(x, \nu, \kappa) = s \int_{\partial^* \mathcal{O}} \phi(x, \nu_\mathcal{O}(x), \kappa_\mathcal{O}(x)) \, \mathrm{d}\mathcal{H}^1(x)\Big\}.$$

The set $\mathcal{M}(u)$ appears to have a complicated structure which we will elaborate on in the next section. We replace it by a simpler, slightly larger set (consequently, our approximation will definitely lie below the convex relaxation), namely the set of all non-negative measures in $\mathcal{M}$ satisfying the two conditions specified below. By Theorem 3.1, all measures $v \in \mathcal{M}(u)$ satisfy (3.3) or equivalently

$$(3.5) \qquad \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu \, \mathrm{d}v(x, \nu, \kappa) = -\int_\Omega u(x) \mathrm{div}\varphi(x) \, \mathrm{d}x \quad \forall \varphi \in \mathcal{C}_0^\infty(\Omega; \mathbb{R}^2) \,.$$

We call this the compatibility condition since it expresses that the measure $v$ is compatible with the image gradient $\nabla u$. Furthermore, by continuity (2.7) also holds for any $\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$, in particular for $\phi(x, \nu, \kappa) = \nabla_x\psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial\psi(x,\nu)}{\partial\nu}\kappa$, $\psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R})$. Thus, by the definition (2.3) of the generalized curvature we have

$$(3.6) \quad \int_{\Omega \times S^1 \times \mathbb{R}} \nabla_x\psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial\psi(x,\nu)}{\partial\nu}\kappa \, \mathrm{d}v[u](x, \nu, \kappa)$$

$$= \int_\mathbb{R} \int_{\partial^*\mathcal{O}_u^t} \nabla_x\psi(x, \nu_{\mathcal{O}_u^t}) \cdot \nu_{\mathcal{O}_u^t}^{-\perp} + \frac{\partial\psi(x, \nu_{\mathcal{O}_u^t})}{\partial\nu}\kappa_{\mathcal{O}_u^t} \, \mathrm{d}\mathcal{H}^1(x) \, \mathrm{d}t = 0 \,.$$

Since this condition is linear and weak-* closed in $v[u] \in \mathcal{M}$, all measures $v \in \mathcal{M}(u)$ also satisfy

$$(3.7) \quad 0 = \int_{\Omega \times S^1 \times \mathbb{R}} \nabla_x\psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial\psi(x,\nu)}{\partial\nu}\kappa \, \mathrm{d}v(x, \nu, \kappa) \quad \forall \psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R}) \,,$$

which we denote the consistency condition, since it expresses that the level line normals encoded in $v$ fit to the encoded curvatures.

We call a non-negative measure $v \in \mathcal{M}$ adequate for a given $u \in L^1(\Omega; \mathbb{R})$, if it satisfies the compatibility and the consistency conditions (3.5) and (3.7). Now we approximate Euler's elastica energy for $u \in L^1(\Omega; \mathbb{R})$ by

$$(3.8) \qquad\qquad R_{\mathrm{el}}^{\mathrm{ap}}(u) = \inf_{v \in \mathcal{M}, v \geq 0 \wedge (3.5) \wedge (3.7)} \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v \,,$$

if an adequate measure exists for $u$, and $R_{\mathrm{el}}^{\mathrm{ap}}(u) = \infty$ else. $R_{\mathrm{el}}^{\mathrm{ap}}$ has the form of a linear program. By (3.2) and the above, $R_{\mathrm{el}}^{\mathrm{ap}}$ is a minorant of $R_{\mathrm{el}}$.

**3.2. Examples of adequate measures.** Before we begin with the analysis of the approximation $R_{\mathrm{el}}^{\mathrm{ap}}$, let us briefly give some examples of adequate measures to provide some intuition of the image gradient lifting.

EXAMPLE 1 (Adequate measures). $v[u] \in \mathcal{M}$ is adequate for any $u \in L^1(\Omega; \mathbb{R})$ with finite $R_{\mathrm{el}}(u)$. A special, instructive case is the following.

1. Let $\mathcal{O} \subset \Omega$ be simply connected with smooth boundary, and let $\partial\mathcal{O}$ be parameterized counterclockwise by arclength, $x : [0, L] \to \partial\mathcal{O}$, $s \mapsto x(s)$. For the unit inward normal $\nu_\mathcal{O}(s) = \dot{x}(s)^\perp$ and the curvature $\kappa_\mathcal{O}(s) = \ddot{x}(s) \cdot \nu_\mathcal{O}(s) = \frac{\mathrm{d}}{\mathrm{d}s}\nu_\mathcal{O}(s)$, the measure $v[\chi_\mathcal{O}]$ is defined via

$$(3.9) \qquad\qquad \int_{\Omega \times S^1 \times \mathbb{R}} \phi \, \mathrm{d}v[\chi_\mathcal{O}] = \int_0^L \phi(x(s), \nu_\mathcal{O}(s), \kappa_\mathcal{O}(s)) \, \mathrm{d}s$$

for all test functions $\phi \in \mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$.

*More general examples and properties are given below. Let $v, v_1, v_2$ be adequate for $u, u_1, u_2$, respectively.*

2. *$v$ is adequate for $u + a$ with any constant $a \in \mathbb{R}$.*
3. *Due to the linearity of (3.5) and (3.7), $av_1 + bv_2$ is adequate for $au_1 + bu_2$ with any $a, b \geq 0$.*
4. *$v^-$ with $v^-(x, \nu, \kappa) = v(x, -\nu, -\kappa)$ in terms of a pushforward measure is adequate for $-u$.*
5. *$v + w + w^-$ is adequate for $u$, where $w$ is any non-negative measure in $\mathcal{M}$ satisfying (3.7) (take for example $w$ adequate for a different image $\hat{u}$).*
6. *Any non-negative measure $w \in \mathcal{M}$ with $(\pi^{\Omega \times S^1})_\#(\kappa w) = (\pi^{\Omega \times S^1})_\#(\kappa v)$ is adequate for $u$.*
7. *If $v_t$ is adequate for $\chi_{\mathcal{O}_u^t}$, $t \in \mathbb{R}$, define $v$ by $v(E) = \int_{\mathbb{R}} v_t(E)\, dt$ for all Borel sets $E \subset \Omega \times S^1 \times \mathbb{R}$. Then $v$ is adequate for $u$ if $v \in \mathcal{M}$.*

THEOREM 3.2 (Adequate measures with simple support). *If $v \in \mathcal{M}$ is adequate for $u \in L^1(\Omega; \mathbb{R})$ with $\operatorname{card}(\operatorname{supp}(v_x)) \leq 1$ for $(\pi^\Omega)_\# v$-almost all $x \in \Omega$, then $v$ is unique and satisfies $(\pi^\Omega)_\# v = |\nabla u|$. If in addition the lifting of $\nabla u$ exists, then $v = v[u]$.*

*Proof.* Let $(\nu(x), \kappa(x))$ denote the support of $v_x$ for $x \in \operatorname{supp}((\pi^\Omega)_\# v)$. We first show $\nu(x)(\pi^\Omega)_\# v(x) = \nabla u(x)$ and thus by polar decomposition $(\pi^\Omega)_\# v = |\nabla u|$. Indeed, by (3.5), for any $\varphi \in \mathcal{C}_0^\infty(\Omega; \mathbb{R}^2)$ we have
(3.10)
$$\int_\Omega \varphi\, d\nabla u = \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu\, dv(x, \nu, \kappa) = \int_\Omega \varphi(x) \int_{S^1 \times \mathbb{R}} \nu\, dv_x(\nu, \kappa)\, d(\pi^\Omega)_\# v(x) = \int_\Omega \varphi(x) \nu(x)\, d(\pi^\Omega)_\# v(x).$$

Now let $v^1, v^2 \in \mathcal{M}$ be adequate for $u$ with $\operatorname{card}(\operatorname{supp}(v_x^i)) \leq 1$ for $|\nabla u|$-almost all $x \in \Omega$, $i = 1, 2$, and denote the support of $v_x^i$ by $(\nu^i(x), \kappa^i(x))$. We will show $v^1 = v^2$. Equation (3.10) implies $\nu^1 = \nu^2$ $|\nabla u|$-almost everywhere. Hence, (3.7) implies $0 = \int_\Omega \int_{S^1 \times \mathbb{R}} \frac{\partial \psi(x, \nu)}{\partial \nu} \kappa\, d(v_x^1 - v_x^2)(\nu, \kappa)\, d|\nabla u|(x) = \int_\Omega \frac{\partial \psi(x, \nu^1(x))}{\partial \nu}(\kappa^1(x) - \kappa^2(x))\, d|\nabla u|(x)$ for all $\psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R})$. Now $\kappa^1 = \kappa^2$ follows as in the proof of Theorem 2.2 (note $\kappa^1, \kappa^2 \in L^1(|\nabla u|; \mathbb{R})$ due to $v^1, v^2 \in \mathcal{M}$).

For $v = v[u]$ just note that $v[u]_x$ has simple support in $(\nu_{\mathcal{O}_u^t}(x), \kappa_{\mathcal{O}_u^t}(x))$, where $\partial^* \mathcal{O}_u^t$ is the level line through $x$. $\square$

### 3.3. Properties of $R_{\mathrm{el}}^{\mathrm{ap}}$.

We proceed by showing some properties of the approximation $R_{\mathrm{el}}^{\mathrm{ap}}$. We will first show that minimizers of the optimization problem in (3.8) exist and then prove the coercivity and convex lower semi-continuity of $R_{\mathrm{el}}^{\mathrm{ap}}$. It follows that $R_{\mathrm{el}}^{\mathrm{ap}}$ is a convex lower semi-continuous functional below $R_{\mathrm{el}}$, but—as discussed earlier—not the largest one.

THEOREM 3.3 (Existence of minimizing measure). *The minimum in (3.8) is achieved.*

*Proof.* Given $u \in \mathrm{BV}(\Omega; \mathbb{R})$ with $R_{\mathrm{el}}^{\mathrm{ap}}(u) < \infty$, consider a minimizing sequence $v_n$, $n \in \mathbb{N}$, where the $v_n$ are adequate for $u$. Due to $\int_{\Omega \times S^1 \times \mathbb{R}} f\, dv_n \geq \|v_n\|_{\mathcal{M}}$ (recall that $w \leq f$ and $\|v\|_{\mathcal{M}} = \|wv\|_{\mathrm{TV}(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})}$), the $v_n$ are bounded in $\mathcal{M}$ so that without loss of generality $v_n \overset{*}{\rightharpoonup} v$ for some $v \in \mathcal{M}$. $v$ is adequate, since $v \geq 0$ and (3.5) and (3.7) are weak-* closed. Finally, $\int_{\Omega \times S^1 \times \mathbb{R}} f\, dv \leq \liminf_{n \to \infty} \int_{\Omega \times S^1 \times \mathbb{R}} f\, dv_n$ by Theorem 2.9. $\square$

THEOREM 3.4 (Bound on total variation). *For $u \in L^1(\Omega; \mathbb{R})$ and a compatible $v \in \mathcal{M}$ (in the sense of (3.5)), $|\nabla u|$ is absolutely continuous with respect to $(\pi^\Omega)_\# v$,*

*and the Radon–Nikodym derivative has norm*

$$\left\|\partial|\nabla u|/\partial(\pi^{\Omega})_{\#}v\right\|_{L^{\infty}((\pi^{\Omega})_{\#}v;\mathbb{R})} \leq 1\,.$$

*Proof.* Let $E \subset \Omega$ be an open Borel set, then

$$\int_{E} \mathrm{d}|\nabla u| = \sup_{\substack{\varphi \in \mathcal{C}_0^{\infty}(\Omega;\mathbb{R}^2) \\ |\varphi| \leq 1 \\ \mathrm{supp}(\varphi) \subset E}} \int_{\Omega} \varphi \cdot \mathrm{d}\nabla u = \sup_{\substack{\varphi \in \mathcal{C}_0^{\infty}(\Omega;\mathbb{R}^2) \\ |\varphi| \leq 1 \\ \mathrm{supp}(\varphi) \subset E}} \int_{\Omega} \int_{S^1 \times \mathbb{R}} \varphi(x) \cdot \nu \, \mathrm{d}v_x(\nu,\kappa) \, \mathrm{d}(\pi^{\Omega})_{\#}v(x) \leq \int_{E} \mathrm{d}(\pi^{\Omega})_{\#}v\,.$$

□

Example 1.5 shows that the reverse is not true in general, i.e. $(\pi^{\Omega})_{\#}v$ is not necessarily absolutely continuous with respect to $|\nabla u|$, and that the inequality cannot be turned into an equality. Recalling $f \geq w \geq c_0(1 + |\kappa|)$ for some $c_0 > 0$, we obtain the coercivity with respect to the TV-seminorm.

COROLLARY 3.5 (Coercivity of $R_{\mathrm{el}}^{\mathrm{ap}}$). *We have $R_{\mathrm{el}}^{\mathrm{ap}}(u) \geq c_0|u|_{\mathrm{TV}(\Omega;\mathbb{R})}$.*

For the ease of notation, let us introduce the constraining sets

(3.11)
$$\mathcal{M}_1 = \left\{ v \in \mathcal{M} \; : \; \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu \, \mathrm{d}v(x,\nu,\kappa) = 0 \; \forall \varphi \in \mathcal{C}_0^{\infty}(\Omega;\mathbb{R}^2) \right\},$$

(3.12)
$$\mathcal{M}_2 = \left\{ v \in \mathcal{M} \; : \; \int_{\Omega \times S^1 \times \mathbb{R}} \nabla_x \psi(x,\nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x,\nu)}{\partial \nu} \kappa \, \mathrm{d}v(x,\nu,\kappa) = 0 \; \forall \psi \in \mathcal{C}_0^{\infty}(\Omega \times S^1;\mathbb{R}) \right\},$$

and denote by $I_M$ the indicator function of a convex set $M$.

THEOREM 3.6 (Convex lower semi-continuity of $R_{\mathrm{el}}^{\mathrm{ap}}$). *$R_{\mathrm{el}}^{\mathrm{ap}}$ is convex lower semi-continuous on $L^1(\Omega;\mathbb{R})$.*

*Proof.* From Theorem 3.3 we know

$$R_{\mathrm{el}}^{\mathrm{ap}}(u) = \min_{v \in \mathcal{M}} \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v + I_{\{v \in \mathcal{M} \,:\, v \geq 0\}}(v) + I_{\mathcal{M}_1}(v - \hat{v}[u]) + I_{\mathcal{M}_2}(v)\,,$$

where $\hat{v}[u]$ is any measure satisfying (3.5) (if no such measure exists, interpret $I_{\mathcal{M}_1}(v - \hat{v}[u])$ as constantly infinity). It is readily seen that the objective function $R(v,u)$ in this minimization is convex in $(v,u) \in \mathcal{M} \times L^1(\Omega;\mathbb{R})$. However, this implies the convexity of $R_{\mathrm{el}}^{\mathrm{ap}}$ on $L^1(\Omega;\mathbb{R})$: Given $u_1, u_2 \in L^1(\Omega;\mathbb{R})$ with corresponding minimizers $v_1, v_2 \in \mathrm{rca}(\Omega \times S^1 \times \mathbb{R})$ and $\alpha \in [0,1]$,

$$R_{\mathrm{el}}^{\mathrm{ap}}(\alpha u_1 + (1-\alpha)u_2) \leq R(\alpha u_1 + (1-\alpha)u_2, \alpha v_1 + (1-\alpha)v_2)$$
$$\leq \alpha R(u_1, v_1) + (1-\alpha)R(u_2, v_2) = \alpha R_{\mathrm{el}}^{\mathrm{ap}}(u_1) + (1-\alpha)R_{\mathrm{el}}^{\mathrm{ap}}(u_2)\,.$$

Concerning the lower semi-continuity, consider a sequence $u_n \to_{n\to\infty} u$ in $L^1(\Omega;\mathbb{R})$ with $\liminf_{n\to\infty} R_{\mathrm{el}}^{\mathrm{ap}}(u_n) = C < \infty$. Upon extracting a subsequence, we may assume $\lim_{n\to\infty} R_{\mathrm{el}}^{\mathrm{ap}}(u_n) = C$. Let $v_n \in \mathcal{M}$ be such that $R(u_n, v_n) = R_{\mathrm{el}}^{\mathrm{ap}}(u_n)$. Due to $R_{\mathrm{el}}^{\mathrm{ap}}(u_n) \geq \|v_n\|_{\mathcal{M}}$, we have $v_n \overset{*}{\rightharpoonup} v$ for a subsequence and some $v \in \mathcal{M}$. Likewise, due to the coercivity of $R_{\mathrm{el}}^{\mathrm{ap}}$ (Corollary 3.5) we have $u_n \overset{*}{\rightharpoonup} u$ in $\mathrm{BV}(\Omega;\mathbb{R})$ after extracting a further subsequence (which we consider from now on). Thus $R_{\mathrm{el}}^{\mathrm{ap}}(u) \leq R(u,v) = \lim_{n\to\infty} R(u_n, v_n) = C$. □
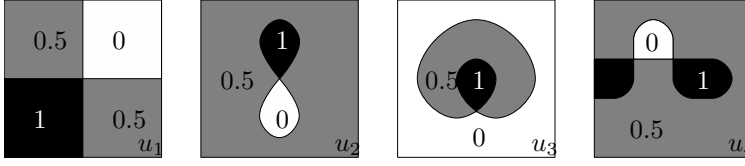
FIGURE 3.1. *Four images with $R_{\mathrm{el}}(u_i) = \infty$ and $R_{\mathrm{el}}^{\mathrm{ap}}(u_i) < \infty$, $i = 1,2,3,4$. While the first one is the convex combination of two images with finite elastica energy (and thus $R_{\mathrm{el}}^{**}(u_1) < \infty$), the last three images cannot be represented as the limit of convex combinations of images with bounded elastica energy so that $R_{\mathrm{el}}^{**}(u_2) = R_{\mathrm{el}}^{**}(u_3) = R_{\mathrm{el}}^{**}(u_4) = \infty$.*

**3.4. The missing gap to the convex relaxation.** To obtain a simple approximation to Euler's elastica energy, we have in the previous sections replaced the set $\mathcal{M}(u)$ by the set $(\mathcal{M}_1 + \hat{v}[u]) \cap \mathcal{M}_2 \cap \{v \in \mathcal{M} : v \geq 0\}$ of non-negative measures satisfying (3.5) and (3.7) (where for $u \in \mathrm{BV}(\Omega; \mathbb{R})$ the measure $\hat{v}[u]$ is any measure satisfying (3.5)). This new set is larger than $\mathcal{M}(u)$ which results in the consequence that $R_{\mathrm{el}}^{\mathrm{ap}}$ is not the convex relaxation of $R_{\mathrm{el}}$. Indeed, consider the images $u_1$ to $u_4$ in Figure 3.1. They all feature a corner in their level lines, leading to an infinite elastica energy. The first image can be decomposed into the convex combination of two images which are half black and half white, thus it satisfies $R_{\mathrm{el}}^{**}(u_1) < \infty$. The other three images, however, cannot be interpreted as the limit of convex combinations of images with bounded elastica energy so that $R_{\mathrm{el}}^{**}(u_i) = \infty$, $i = 2,3,4$. Nevertheless, they all admit a smooth (though self-intersecting) arclength parameterization $x_i : [0, L_i] \to \Omega$ of their discontinuity set with normal $\nu_i = \dot{x}_i^{\perp}$ and curvature $\kappa_i = \ddot{x}_i \cdot \nu_i = \dot{\nu}_i$. Consequently, the non-negative line measure $v_i$, defined by

$$(3.13) \quad \int_{\Omega \times S^1 \times \mathbb{R}} \phi \, dv_i = \int_0^{L_i} \phi(x_i(s), \nu_i(s), \kappa_i(s)) \, ds \quad \forall \phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}),$$

satisfies (3.5) (since the image jump across the curve stays constant) and (3.7) so that $R_{\mathrm{el}}^{\mathrm{ap}}(u_i) \leq \int_{\Omega \times S^1 \times \mathbb{R}} f \, dv_i < \infty$, $i = 2,3,4$.

The reason for this phenomenon apparently lies in the fact that we also admit line measures (3.13) which belong to smooth curves $x : [0, L] \to \Omega$ (or more precisely to smooth immersions of $S^1$ or $[0,1]$ in $\Omega$) that exhibit self-intersection. Such a line measure cannot be decomposed into separate, consistent (in the sense of (3.7)) line measures without self-intersection, each of which would be compatible to a different image so that the original image can be reproduced as a convex combination of these images. In fact, one could think of even wilder line measures with an arbitrary number of self-intersections.

Of course, such self-intersecting measures are excluded if in (3.8) we only admit measures from the (sequentially weakly-*) closed convex hull $\overline{\mathrm{co}}\mathcal{M}_O$ of $\mathcal{M}_O$. In order to distill the underlying mechanism, one might interpret $\overline{\mathrm{co}}\mathcal{M}_O$ as a subset of the intersection of $\mathcal{M}_2$ with $\overline{\mathrm{co}}\tilde{\mathcal{M}}_O$ for

$$(3.14) \quad \tilde{\mathcal{M}}_O = \left\{ v \in \mathcal{M} : v \geq 0, \exists \mathcal{O} \in O, s \geq 0 \; \forall \phi \in \mathcal{C}_0^{\infty}(\Omega \times S^1; \mathbb{R}) : \right.$$

$$\left. \int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu) \, dv(x, \nu, \kappa) = s \int_{\partial^* \mathcal{O}} \phi(x, \nu(x)) \, d\mathcal{H}^1(x) \right\},$$

using the same notation as in the previous section. $\overline{\mathrm{co}}\tilde{\mathcal{M}}_O$ indeed excludes the line measures corresponding to $u_2$ to $u_4$ in Figure 3.1. $\tilde{\mathcal{M}}_O$ is a cone so that $\overline{\mathrm{co}}\tilde{\mathcal{M}}_O$ is a

sequentially weakly-* closed convex cone. One can try to obtain a little more intuition of $\overline{\mathrm{co}}\tilde{\mathcal{M}}_O$ via duality techniques: We can express the indicator function $I_{\overline{\mathrm{co}}\tilde{\mathcal{M}}_O}$ as the biconjugate $I^{**}_{\tilde{\mathcal{M}}_O}$ with respect to $\mathcal{C}_w \times \mathcal{M}$ dual pairing, where the asterisk denotes the Legendre–Fenchel dual. For any $\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ which is independent of its third argument,

$$(3.15) \qquad I^*_{\tilde{\mathcal{M}}_O}(\phi) = \begin{cases} 0, & \text{if } \int_{\partial^* \mathcal{O}} \phi(x, \nu(x), 0)\, \mathrm{d}\mathcal{H}^1(x) \leq 0 \; \forall \mathcal{O} \in O\,, \\ \infty & \text{else,} \end{cases}$$

so a condition on the measure $v$ would have the form $\int_{\Omega \times S^1 \times \mathbb{R}} \phi\, \mathrm{d}v \leq 0$ for all functions $\phi$ satisfying $\int_{\partial \mathcal{O}} \phi(x, \nu(x), 0)\, \mathrm{d}\mathcal{H}^1(x) \leq 0 \; \forall \mathcal{O} \in O$. Such a condition obviously is global in nature and cannot be verified locally. Intuitively, for any two branches of a measure $v$ which have support at the same $x \in \Omega$ but at different $\nu \in S^1$ we have to check that they stay independent from each other throughout the image. This fact makes the true convex relaxation of $R_{\mathrm{el}}$ inadequate for computational purposes.

**3.5. A predual representation.** Sometimes it is useful to have also a predual representation of the optimization problem (3.8), for example to obtain a more adequate expression for numerical implementation or for the analysis.

THEOREM 3.7 (Predual representation of $R^{\mathrm{ap}}_{\mathrm{el}}$). *For $u \in \mathrm{BV}(\Omega; \mathbb{R})$ the minimization problem in (3.8) is the dual problem to*

(3.16)

$$\sup_{\substack{\varphi \in \mathcal{C}^\infty_0(\Omega; \mathbb{R}^2) \\ \psi \in \mathcal{C}^\infty_0(\Omega \times S^1; \mathbb{R})}} \int_\Omega \varphi \cdot \mathrm{d}\nabla u \quad s.\,t. \quad \nabla_x \psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x, \nu)}{\partial \nu}\kappa + \varphi(x) \cdot \nu \leq f(x, \nu, \kappa) \; \forall (x, \nu, \kappa) \in \Omega \times S^1 \times \mathbb{R}\,,$$

*and strong duality holds (that is, both optimization problems yield the same value).*

*Proof.* Recall the definition of the constraining sets (3.11) and (3.12) and additionally define

$$\mathcal{M}^\perp_1 = \mathrm{cl}\big\{\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \; : \; \exists \varphi \in \mathcal{C}^\infty_0(\Omega; \mathbb{R}^2) : \phi(x, \nu, \kappa) = \varphi(x) \cdot \nu\big\}\,,$$

$$\mathcal{M}^\perp_2 = \mathrm{cl}\Big\{\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \; : \; \exists \psi \in \mathcal{C}^\infty_0(\Omega \times S^1; \mathbb{R}) : \phi(x, \nu, \kappa) = \nabla_x \psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x, \nu)}{\partial \nu}\kappa\Big\}\,,$$

where the closure is taken in $\mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$. Furthermore, let $\hat{v} \in \mathcal{M}$ satisfy (3.5), that is,

$$\int_\Omega \varphi(x) \cdot \mathrm{d}\nabla u(x) = \int_{\Omega \times S^1 \times \mathbb{R}} \varphi(x) \cdot \nu\, \mathrm{d}\hat{v}(x, \nu, \kappa) \quad \forall \varphi \in \mathcal{C}^\infty_0(\Omega; \mathbb{R}^2)\,.$$

Denote the indicator function of a convex set $M$ by $I_M$ and the Legendre–Fenchel dual by an asterisk. We have by definition

$$(\hat{v} + I_{\mathcal{M}^\perp_1})^* = I_{\mathcal{M}_1}(\cdot - \hat{v})\,, \qquad (I_{\mathcal{M}^\perp_2})^* = I_{\mathcal{M}_2}\,.$$

Define

$$F : \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \to \mathbb{R}, F(\phi) = I_{\{\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \, : \, \phi \leq 0\}}(\phi - f)\,,$$

$$G : \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \to \mathbb{R}, G(\phi) = ((\hat{v} + I_{\mathcal{M}^\perp_1}) \star I_{\mathcal{M}^\perp_2})(\phi) = \inf_{\substack{\tilde{\phi} \in \mathcal{M}^\perp_1 \\ \phi - \tilde{\phi} \in \mathcal{M}^\perp_2}} \int_{\Omega \times S^1 \times \mathbb{R}} \tilde{\phi}\, \mathrm{d}\hat{v}\,.$$

Assuming the validity of the Fenchel–Rockafellar formula, we obtain

$$(3.17) \qquad \min_{v \in \mathcal{M}} F^*(v) + G^*(v) = \sup_{\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})} -F(-\phi) - G(\phi) \,,$$

where

$$F^*(v) = \int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v + I_{\{v \in \mathcal{M} \,:\, v \geq 0\}}(v) \,,$$

$$G^*(v) = I_{\mathcal{M}_1}(v - \hat{v}) + I_{\mathcal{M}_2}(v) \,.$$

According to [5], the strong duality (3.17) indeed holds, if $\bigcup_{\lambda \leq 0} \lambda(\mathrm{Dom}(F) + \mathrm{Dom}(G))$ is a closed vector space. However, $0 \in \mathrm{Dom}(G)$, and the ball $\{\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) : \|\phi\|_w \leq 1\}$ is a subset of $\mathrm{Dom}(F)$ so that $\bigcup_{\lambda \leq 0} \lambda(\mathrm{Dom}(F) + \mathrm{Dom}(G)) = \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$ fulfills the requirement.

The left-hand side of (3.17) obviously is the minimization problem in (3.8). Its right-hand side is the claimed predual problem. $\square$

As an example for its application, consider the following alternative proof of Corollary 3.5.

THEOREM 3.8 (Coercivity of $R_{\mathrm{el}}^{\mathrm{ap}}$). $R_{\mathrm{el}}^{\mathrm{ap}}(u) \geq c_0 |u|_{\mathrm{TV}(\Omega; \mathbb{R})}$ *for any* $c_0$ *with* $f \geq c_0 > 0$.

*Proof.* Let, for each $n$, $\varphi_n \in \mathcal{C}_0^\infty(\Omega; \mathbb{R}^2)$ with $\|\varphi_n\|_\infty \leq 1$ such that

$$\int_\Omega \varphi_n \cdot \mathrm{d}\nabla u \xrightarrow[n \to \infty]{} |u|_{\mathrm{TV}(\Omega; \mathbb{R})} \,.$$

Now $\psi = 0 \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R})$ and $\varphi = c_0 \varphi_n$ satisfy the constraint $\nabla_x \psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x, \nu)}{\partial \nu} \kappa + \varphi(x) \cdot \nu \leq f(x, \nu, \kappa)$ so that by (3.16), $R_{\mathrm{el}}^{\mathrm{ap}}(u) \geq c_0 \int_\Omega \varphi_n \cdot \mathrm{d}\nabla u$ for all $n \in \mathbb{N}$. The limit $n \to \infty$ yields the desired result. $\square$

**3.6. Some characterization of minimizing measures.** In this section we intend to analyze the properties of the minimizing measures $v$ in (3.8). While the measures can be shown to have support only at one particular curvature $\kappa$ for each $(x, \nu) \in \Omega \times S^1$, the support in the $\nu$-dimension is generally non-unique. If however, for a simply connected set $\mathcal{O} \subset \Omega$ with smooth boundary, $(\pi^\Omega)_\# v$ is absolutely continuous with respect to $\nabla \chi_\mathcal{O}$ (hence $\mathrm{supp}(v) \subset \mathrm{supp}(\nabla \chi_\mathcal{O}) \times S^1 \times \mathbb{R}$), then also the support in the $\nu$-direction is unique, and $v$ is the line measure $v[\chi_\mathcal{O}]$. Finally, under certain conditions, we can show $v$ to decompose into a set of simple line measures, and we expect this behavior also to hold in the general case.

THEOREM 3.9 (Uniqueness of curvature support). *If $f$ is strictly convex in $\kappa$, then* $(x, \nu, \kappa_1) \in \mathrm{supp}(v_1)$, $(x, \nu, \kappa_2) \in \mathrm{supp}(v_2)$ *for any two (possibly non-distinct) minimizers* $v_1, v_2$ *of* (3.8) *implies* $\kappa_1 = \kappa_2$ *for* $((\pi^{\Omega \times S^1})_\# v_1 + (\pi^{\Omega \times S^1})_\# v_2)$*-almost all* $(x, \nu) \in \Omega \times S^1$.

*Proof.* First consider the case $v_1 = v_2 = v$. Jensen's inequality then implies for $(\pi^{\Omega \times S^1})_\# v$-almost all $(x, \nu) \in \Omega \times S^1$

$$\int_\mathbb{R} f(x, \nu, \kappa) \, \mathrm{d}\delta_{\kappa_{(x, \nu)}} = f\left(x, \nu, \int_\mathbb{R} \kappa \, \mathrm{d}v_{(x, \nu)}(\kappa)\right) \leq \int_\mathbb{R} f(x, \nu, \kappa) \, \mathrm{d}v_{(x, \nu)}(\kappa) \,,$$

where $\delta_{\kappa_{(x, \nu)}}$ denotes the Dirac measure centered around $\kappa_{(x, \nu)} = \int_\mathbb{R} \kappa \, \mathrm{d}v_{(x, \nu)}(\kappa)$. The inequality is strict if $v_{(x, \nu)}$ has support in more than one point. Since $v$ is a minimizer of

$$\int_{\Omega \times S^1 \times \mathbb{R}} f \, \mathrm{d}v = \int_{\Omega \times S^1} \int_\mathbb{R} f(x, \nu, \kappa) \, \mathrm{d}v_{(x, \nu)}(\kappa) \, \mathrm{d}(\pi^{\Omega \times S^1})_\# v(x, \nu)$$
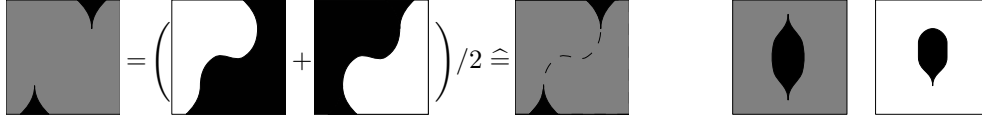
FIGURE 3.2. *Left: The left-most image u has infinite elastica energy, but can be expressed as the convex combination of two images with finite elastica energy. Hence, $R_{\mathrm{el}}^{\mathrm{ap}}(u) < \infty$, where the support of $(\pi^\Omega)_\# v$ for an adequate measure $v$ is indicated by the dotted line. Along that line, $v$ has support at the normal $\nu$ and at $-\nu$. Right: Two more examples of the same type, where the second example additionally exhibits a self-intersection.*

we thus must have $v_{(x,\nu)} = \delta_{\kappa_{(x,\nu)}}$ for $(\pi^{\Omega \times S^1})_\# v$-almost all $(x,\nu) \in \Omega \times S^1$ (note that this $v$ is adequate).

Now let $v_1, v_2$ be distinct minimizers, then $v = \frac{1}{2} v_1 + \frac{1}{2} v_2$ is also a minimizer. The above implies that $v_{(x,\nu)}$ has support only at a single $\kappa$-value for $(\pi^{\Omega \times S^1})_\# v$-almost all $(x,\nu) \in \Omega \times S^1$. Likewise, $(v_1)_{(x,\nu)}$ and $(v_2)_{(x,\nu)}$ only have support at a single $\kappa$-value, which implies that all these $\kappa$-values must coincide. $\square$

While the previous result implies that there is a unique curvature $\kappa(x,\nu)$ associated with every point $(x,\nu) \in \Omega \times S^1$, there is in general no unique normal $\nu$ associated with $x \in \Omega$. This is closely linked to the fact that the support of $(\pi^\Omega)_\# v$ for the minimizing $v$ may be larger than $\mathrm{supp}(\nabla u)$. As an example, consider the images in Figure 3.2. Here, we cannot have $\mathrm{supp}((\pi^\Omega)_\# v) \subset \mathrm{supp}(\nabla u)$ without violating the consistency condition (3.7). Hence, ghost edges will occur in the minimizing measure $v$ which cancel each other during integration in (3.5), i.e., $\int_{S^1 \times \mathbb{R}} \nu \, dv_x(\nu, \kappa) = 0$ on these edges.

An even more subtle example is related to Theorem 2.4. Consider $u \in \mathcal{C}_0^\infty(\Omega; \mathbb{R})$. By the proof of Theorem 2.4, $u$ can be written as an integral $u = \int_{S^1} u_\nu \, d\mathcal{H}^1(\nu)$ of images $u_\nu$ with straight, parallel level lines and uniformly bounded BV-norm. The corresponding adequate measure $v = \int_{S^1} v[u_\nu] \, d\mathcal{H}^1(\nu)$ typically has support $\mathrm{supp}(v) = \Omega \times S^1 \times \{0\}$, that is, neither is the support of $(\pi^\Omega)_\# v$ restricted to $\mathrm{supp}(\nabla u)$ nor is there a unique normal $\nu$ associated with any $x \in \Omega$. If the curvature term in the elastica energy density $f$ is weighted strongly enough, one may expect this $v$ to be the minimizing measure, since it satisfies $\kappa = 0$ all over its support.

Despite the above, if characteristic functions of sets $\mathcal{O} \subset \Omega$ with smooth boundary are considered, then $\mathrm{supp}((\pi^\Omega)_\# v) \subset \mathrm{supp}(\nabla \chi_\mathcal{O})$ does imply that for $|\nabla \chi_\mathcal{O}|$-almost each $x \in \Omega$, $v_x$ has support at a unique $(\nu(x), \kappa(x))$.

THEOREM 3.10 (Adequate measures for characteristic functions). *Let $u = \chi_\mathcal{O}$ for $\mathcal{O} \subset \Omega$ simply connected with smooth boundary and $f$ be strictly convex in $\kappa$. By Theorem 3.9 it suffices to consider only adequate measures with unique support in the $\kappa$-dimension. Then any such measure $v$ with $\mathrm{supp}((\pi^\Omega)_\# v) \subset \mathrm{supp}(\nabla u) = \partial \mathcal{O}$ satisfies $v = (1 + c)v[u] + cv[u]^-$ for some constant $c \geq 0$ (where $v[u]^-(x, \nu, \kappa) = v[u](x, -\nu, -\kappa)$ interpreted as a pushforward measure).*

*Proof.* Denote by $\mathrm{sd}_{\partial \mathcal{O}}$ the signed distance function to $\partial \mathcal{O}$ and define

$$\psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R}), \ \psi(x, \nu) = T(x) \mathrm{sd}_{\partial \mathcal{O}}(x) \left( \nabla_x \mathrm{sd}_{\partial \mathcal{O}}(x) \cdot \nu^{-\perp} \right),$$

where $T : \Omega \to [0, 1]$ is a smooth cutoff function which is zero outside a neighborhood of $\partial \mathcal{O}$, one inside a neighborhood of $\partial \mathcal{O}$, and which ensures $\psi \in \mathcal{C}_0^\infty(\Omega \times S^1)$. The

consistency condition (3.7) then implies

$$0 = \int_{\partial\mathcal{O}\times S^1\times\mathbb{R}} \nabla_x\psi(x,\nu)\cdot\nu^{-\perp} + \frac{\partial\psi(x,\nu)}{\partial\nu}\kappa\,\mathrm{d}v(x,\nu,\kappa)$$

$$= \int_{\partial\mathcal{O}\times S^1\times\mathbb{R}} \left(\nabla_x\mathrm{sd}_{\partial\mathcal{O}}(x)\cdot\nu^{-\perp}\right)^2\,\mathrm{d}v(x,\nu,\kappa) = \int_{\partial\mathcal{O}\times S^1} \left(\nabla_x\mathrm{sd}_{\partial\mathcal{O}}(x)\cdot\nu^{-\perp}\right)^2\,\mathrm{d}(\pi^{\Omega\times S^1})_\#v(x,\nu)\,.$$

We deduce $\mathrm{supp}((\pi^{\Omega\times S^1})_\#v) \subset \{(x,\nu) \in \Omega \times S^1 \ : \ x \in \partial\mathcal{O}, \nu = \pm\nu(x)\}$, where $\nu(x)$ is the inward normal to $\partial\mathcal{O}$ at $x$. Since the support in the $\kappa$-dimension is unique, we obtain that $\mathrm{supp}(v)$ is a subset of two non-intersecting curves $\gamma_\pm : [0,L] \to \Omega\times S^1\times\mathbb{R}, s \mapsto (x(s),\pm\nu(x(s)),\kappa_\pm(s))$, where $x(s) : [0,L] \to \Omega$ denotes the arc length parameterization of $\partial\mathcal{O}$. The consistency condition (3.7) now implies that actually $\kappa_\pm(s) = \pm\kappa(x(s))$ is the curvature of $\partial\mathcal{O}$, and furthermore that $v = c_1 v[u] + c_2 v[u]^-$ for two constants $c_1, c_2 \geq 0$. $c_1 \geq 1$ and $c_1 - c_2 = 1$ then follows from the compatibility condition (3.5). $\square$

It is obvious that taking $v = v[\chi_\mathcal{O}]$ in the above theorem reduces the integral $\int_{\Omega\times S^1\times\mathbb{R}} f\,\mathrm{d}v$ so that $\mathrm{supp}((\pi^\Omega)_\#v) \subset \partial\mathcal{O}$ for the minimizing $v$ implies $v = v[\chi_\mathcal{O}]$. Another case where the minimizing measure $v$ can be identified with the lifting $v[u]$ is when $v$ has simple support in the $\nu$-coordinate (cf. Theorem 3.2). This can be used to assess the tightness of the relaxation $R_\mathrm{el}^\mathrm{ap}$.

THEOREM 3.11 (A posteriori characterization of tight relaxation). *If $u$ admits a lifting $v[u]$ and if the minimizer $v$ of (3.8) has simple support in the $\nu$-coordinate (in the sense $\mathrm{card}(\pi^{S^1}(\mathrm{supp}(v_x))) \leq 1$) for $(\pi^\Omega)_\#v$-almost every $x \in \Omega$, then $v = v[u]$ and thus $R_\mathrm{el}(u) = R_\mathrm{el}^{**}(u) = R_\mathrm{el}^\mathrm{ap}(u)$.*

*Proof.* By Theorem 3.9 we have $\mathrm{card}(\mathrm{supp}(v_x)) \leq 1$ which by Theorem 3.2 implies $v = v[u]$ and thus $R_\mathrm{el}^\mathrm{ap}(u) = R_\mathrm{el}(u)$. $\square$

Finally we would like to show that under certain conditions the minimizing measure in (3.8) decomposes into a set of line measures. Let $\mathrm{Imm}(S^1;\Omega)$ denote the set of immersions from $S^1$ into $\Omega$ and define $\mathrm{Imm}((0,1);\Omega)$ to be the set of immersions $C : (0,1) \to \Omega$ with $C(0),C(1) \in \partial\Omega$. Furthermore, let $|C|$ denote the image of an immersion $C$ and introduce the set

$$(3.18) \qquad \mathfrak{C} = \{|C| \subset \Omega \ : \ C \in \mathrm{Imm}(S^1;\Omega) \cup \mathrm{Imm}((0,1);\Omega), \mathcal{H}^1(|C|) < \infty\}\,.$$

THEOREM 3.12 (Decomposition of minimizing measure). *By Theorem 3.9 the minimizing measure $v$ in (3.8) can be represented by a curvature function $\kappa : \Omega\times S^1 \to \mathbb{R}$ and the projected measure $(\pi^{\Omega\times S^1})_\#v$. If $\kappa$ is Lipschitz-continuous on $\overline{\Omega}\times S^1$ and $\Omega$ has smooth boundary, then there exists a set $\mathfrak{C}_\kappa \subset \mathfrak{C}$ and a measure $\sigma$ on $\mathfrak{C}_\kappa$ such that for all $\phi \in \mathcal{C}_w(\Omega\times S^1\times\mathbb{R};\mathbb{R})$*

$$(3.19) \qquad \int_{\Omega\times S^1\times\mathbb{R}} \phi\,\mathrm{d}v = \int_{\mathfrak{C}_\kappa}\int_{|C|} \phi(x,\nu_C(x),\kappa_C(x))\,\mathrm{d}\mathcal{H}^1(x)\,\mathrm{d}\sigma(|C|)\,,$$

*where $\nu_C$ and $\kappa_C$ are the normal and curvature to the curve $|C|$, and $\kappa_C(x) = \kappa(x,\nu_C(x))$.*

*Proof.* For given $(x_0,\nu_0) \in \Omega\times S^1$, define the characteristic $(x_{(x_0,\nu_0)},\nu_{(x_0,\nu_0)})$ as the solution to the autonomous ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_{(x_0,\nu_0)} \\ \nu_{(x_0,\nu_0)} \end{pmatrix} = \begin{pmatrix} \nu_{(x_0,\nu_0)}^{-\perp} \\ \kappa(x_{(x_0,\nu_0)},\nu_{(x_0,\nu_0)}) \end{pmatrix}, \qquad \begin{pmatrix} x_{(x_0,\nu_0)} \\ \nu_{(x_0,\nu_0)} \end{pmatrix}\bigg|_{t=0} = \begin{pmatrix} x_0 \\ \nu_0 \end{pmatrix},$$

which is well-defined in an interval around 0 due to the Lipschitz-continuity of the right-hand side. Even more, the solution can be continued for positive and negative times until $(x_{(x_0,\nu_0)}(t), \nu_{(x_0,\nu_0)}(t))$ reaches $\partial\Omega \times S^1$. In particular, it is also well-defined for $(x_0, \nu_0) \in \partial\Omega \times S^1$.

Next consider a copy $\Omega^- \times S^{1^-}$ of $\Omega \times S^1$ and the isometry

$$i : \overline{\Omega} \times S^1 \to \overline{\Omega^-} \times S^{1^-}, \quad (x,\nu) \mapsto (x^-, \nu^-) = (x, -\nu).$$

For each characteristic $(x_{(x_0,\nu_0)}, \nu_{(x_0,\nu_0)}) : [-t_l, t_r] \to \Omega \times S^1$ with $(x_0, \nu_0) \in \Omega \times S^1$ define an associated characteristic $(x^-_{i(x_0,\nu_0)}, \nu^-_{i(x_0,\nu_0)}) : [-t_r, t_l] \to \Omega^- \times S^{1^-}$ via $(x^-_{i(x_0,\nu_0)}, \nu^-_{i(x_0,\nu_0)})|_t = (x_{(x_0,\nu_0)}, -\nu_{(x_0,\nu_0)})|_{-t}$. Furthermore, we will identify points $(x,\nu) \in \partial\Omega \times S^1$ with points $(x^-, \nu^-) \in \partial\Omega^- \times S^{1^-}$ via the restriction $j \equiv i|_{\partial\Omega \times S^1} : \partial\Omega \times S^1 \to \partial\Omega^- \times S^{1^-}$. Now, whenever a characteristic on $\Omega \times S^1$ or $\Omega^- \times S^{1^-}$ reaches the boundary, we continue it on $\Omega^- \times S^{1^-}$ or $\Omega \times S^1$ with its associated characteristic. In this way, any characteristic can be continued to a continuous curve that maps the whole real line into the disjoint union $(\overline{\Omega} \times S^1) \sqcup_j (\overline{\Omega^-} \times S^1) =: D$. This curve is differentiable everywhere except when it crosses $\partial\Omega \times S^1 \equiv \partial\Omega^- \times S^{1^-}$.

For $(x_0, \nu_0) \in D$, denote the above curve running through $(x_0, \nu_0)$ at $t = 0$ by

$$(x_{(x_0,\nu_0)}, \nu_{(x_0,\nu_0)}) : \mathbb{R} \to D.$$

For $t \in \mathbb{R}$ we now define the transformations

$$T_t : D \to D, \ (x,\nu) \mapsto (x_{(x,\nu)}, \nu_{(x,\nu)})|_t,$$

which form an Abelian automorphism group of measurable flows [31]. Furthermore, we define a measure $\mu \in \mathrm{rca}(D)$ by

$$\mu(E) = (\pi^{\Omega \times S^1})_\# v\big(E \cap (\Omega \times S^1)\big) + (\pi^{\Omega^- \times S^{1^-}})_\# v^-\big(E \cap (\Omega^- \times S^{1^-})\big)$$

for any Borel set $E \subset D$, where $v^- \in \mathrm{rca}(\Omega^- \times S^{1^-})$ is related to $v$ by $v^-(x, \nu, \kappa) = v(x, -\nu, -\kappa)$.

We intend to show in the following that $\mu$ is $T_t$-invariant. For this purpose let $\psi \in C_0^\infty(\Omega \times S^1; \mathbb{R})$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_D \psi \circ T_t \, \mathrm{d}\mu \Big|_{t=0} = \lim_{t \to 0} \int_{\Omega \times S^1} \frac{\psi \circ T_t - \psi}{t} \, \mathrm{d}(\pi^{\Omega \times S^1})_\# v \Big|_{t=0}$$

$$= \int_{\Omega \times S^1} \frac{\mathrm{d}}{\mathrm{d}t} \psi \circ T_t \Big|_{t=0} \mathrm{d}(\pi^{\Omega \times S^1})_\# v = \int_{\Omega \times S^1} \nabla_x \psi(x,\nu) \cdot \nu^{-\perp} + \frac{\partial\psi(x,\nu)}{\partial\nu} \kappa(x,\nu) \, \mathrm{d}(\pi^{\Omega \times S^1})_\# v(x,\nu) = 0,$$

where the integral converges due to Lebesgue's dominated convergence theorem, since the integrand converges pointwise and is bounded by the Lipschitz constant $L_\psi$ of $\psi$ times the Lipschitz constant $L_{T_t}$ of $T_t$ in $t$. The same holds true on the other half of $D$.

Next consider a general $\psi \in C(D; \mathbb{R})$ which is smooth up to the boundary on $(\Omega \times S^1) \sqcup (\Omega^- \times S^{1^-})$. For $\varepsilon > 0$, let $T \in C(D; \mathbb{R})$ be a smooth cutoff function with $T|_{\Omega \times S^1}(x,\nu) = T|_{\Omega^- \times S^{1-}}(x, -\nu)$ which is one in an $\frac{\varepsilon}{2}$-neighborhood of $\partial\Omega \times S^1$ and

zero outside an $\varepsilon$-neighborhood. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \psi \circ T_t \,\mathrm{d}\mu\bigg|_{t=0}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\int_D (T\psi)\circ T_t\,\mathrm{d}\mu\bigg|_{t=0} + \frac{\mathrm{d}}{\mathrm{d}t}\int_D ((1-T)\psi)\circ T_t\,\mathrm{d}\mu\bigg|_{t=0} = \lim_{t\to 0}\int_D \frac{(\psi T)\circ T_t - \psi T}{t}\,\mathrm{d}\mu\bigg|_{t=0}$$

$$= \lim_{t\to 0}\left(\int_D \frac{\psi \circ T_t - \psi}{t}T\,\mathrm{d}\mu + \int_D \psi\frac{T\circ T_t - T}{t}\,\mathrm{d}\mu + t\int_D \frac{(\psi\circ T_t - \psi)(T\circ T_t - T)}{t^2}\,\mathrm{d}\mu\right)\bigg|_{t=0} ,$$

where in the first step we used $\psi = T\psi + (1-T)\psi$, in the second step we exploited the $T_t$-invariance of $\mu$ in the interior of $\Omega \times S^1$ and $\Omega^- \times S^{1^-}$, and the third step used the trivial identity $ab - cd = (a-c)d + c(b-d) + (a-c)(b-d)$. We can now estimate the different integrals as follows. The first term is bounded by

$$\int_D \left|\frac{\psi\circ T_t - \psi}{t}T\right|\mathrm{d}\mu \le 2\mu(B_\varepsilon(\partial\Omega)\times S^1)\sup_{(x,\nu)\in D}\left|\frac{\psi\circ T_t(x,\nu) - \psi(x,\nu)}{t}\right|$$

$$\le 2\mu(B_\varepsilon(\partial\Omega)\times S^1)L_\psi\sup_{(x,\nu)\in D}\left|\frac{T_t(x,\nu) - T_0(x,\nu)}{t}\right| \le 2\mu(B_\varepsilon(\partial\Omega)\times S^1)L_\psi L_{T_t} ,$$

where $B_\varepsilon(\partial\Omega) = \{x \in \Omega : \mathrm{dist}(x,\partial\Omega) < \varepsilon\}$ denotes the $\varepsilon$-ball around $\partial\Omega$. By Lebesgue's dominated convergence theorem, the absolute value of the second term converges against

$$\left|\int_D \psi\frac{\mathrm{d}}{\mathrm{d}t}(T\circ T_t)\,\mathrm{d}\mu\right| = \left|\int_{B_\varepsilon(\partial\Omega)\times S^1}(\psi(x,\nu) - \psi(i(x,\nu)))\frac{\mathrm{d}}{\mathrm{d}t}(T\circ T_t)(x,\nu)\,\mathrm{d}(\pi^{\Omega\times S^1})_\# v\right|$$

$$\le \mu(B_\varepsilon(\partial\Omega)\times S^1)\sup_{(x,\nu)\in D}|\psi(x,\nu) - \psi(i(x,\nu))|\left|\frac{\mathrm{d}}{\mathrm{d}t}(T\circ T_t)(x,\nu)\right| \le \mu(B_\varepsilon(\partial\Omega)\times S^1)2\varepsilon L_\psi L_T L_{T_t}$$

where we can choose $T$ such that $L_T$, the Lipschitz constant of $T$, is $O(\frac{1}{\varepsilon})$ as $\varepsilon \to 0$. The expression is thus bounded by some constant times $\mu(B_\varepsilon(\partial\Omega)\times S^1)L_\psi L_{T_t}$. Finally, the third term is bounded by $t\mu(D)(L_\psi L_{T_t})(L_T L_{T_t})$ and thus converges to 0 as $t \to 0$. Now noting that $\mu$ is regular and has no support on $\partial\Omega \times S^1$ so that $\mu(B_\varepsilon(\partial\Omega)\times S^1) \to 0$ as $\varepsilon \to 0$, by letting $\varepsilon \to 0$ we see

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_D \psi\circ T_t\,\mathrm{d}\mu\bigg|_{t=0} = 0 .$$

Hence, for any $\hat{t} \in \mathbb{R}$, $\frac{\mathrm{d}}{\mathrm{d}t}\int_D \psi\circ T_t\,\mathrm{d}\mu\big|_{t=\hat{t}} = \frac{\mathrm{d}}{\mathrm{d}t}\int_D (\psi\circ T_{\hat{t}})\circ T_t\,\mathrm{d}\mu\big|_{t=0} = 0$ and thus

$$\int_D \psi\,\mathrm{d}(T_t)_\#\mu = \int_D \psi\circ T_t\,\mathrm{d}\mu = \int_D \psi\,\mathrm{d}\mu$$

so that $\mu$ is indeed $T_t$-invariant.

Now define $|C|_{(x_0,\nu_0)} = \{(x_{(x_0,\nu_0)},\nu_{(x_0,\nu_0)})|_t : t \in \mathbb{R}\}$ for any $(x_0,\nu_0) \in D$ as well as the set of measures

$$\hat{\mathfrak{C}}_\kappa = \{(\mathcal{H}^1 \otimes \mathcal{H}^0)\llcorner|C|_{(x_0,\nu_0)} : (x_0,\nu_0) \in D\} .$$

The measures in $\hat{\mathfrak{C}}_\kappa$ are also $T_t$-invariant, and they are even ergodic. Thus they form the extremal points of all $T_t$-invariant measures. Furthermore, $\hat{\mathfrak{C}}_\kappa$ comprises the only $T_t$-ergodic measures, for let $q \notin \hat{\mathfrak{C}}_\kappa$ be a $T_t$-ergodic measure with $q(A) > 0$ for some $A \subset D$, then $q(\hat{A}) > 0$ for $\hat{A} := \bigcup_{(x_0,\nu_0) \in A} |C|_{(x_0,\nu_0)}$ which contradicts the fact that ergodic measures are mutually singular (in particular, $q$ must be mutually singular with all measures in $\hat{\mathfrak{C}}_\kappa$). Thus, by the ergodic decomposition theorem (see e. g. [31, Lect. 2, Thm. 2'], [12, Thms. 4.8&8.2], or [15] for different versions), there is a measure $\hat{\sigma}$ on $\hat{\mathfrak{C}}_\kappa$ or equivalently a measure $\sigma$ on the set $\mathfrak{C}_\kappa$ of the corresponding curves $|C| = \{x_{(x_0,\nu_0)}|_t : t \in \mathbb{R}\} \cap \Omega$, $(x_0,\nu_0) \in D$, such that

$$\mu = \int_{\hat{\mathfrak{C}}_\kappa} \hat{C} \, \mathrm{d}\hat{\sigma}(\hat{C}) \quad \text{and thus} \quad \int_{\Omega \times S^1 \times \mathbb{R}} \phi \, \mathrm{d}v = \int_{\mathfrak{C}_\kappa} \int_{|C|} \phi(x, \nu_C(x), \kappa_C(x)) \, \mathrm{d}\mathcal{H}^1(x) \, \mathrm{d}\sigma(|C|)$$

for all $\phi \in \mathcal{C}_w(\Omega \times S^1 \times \mathbb{R}; \mathbb{R})$.

The curves in $\mathfrak{C}_\kappa$ are $\sigma$-almost everywhere of finite length since otherwise, due to Theorem 3.8, the resulting elastica energy would be infinite. Hence, since up to $\partial\Omega$ they can be arclength parameterized by the characteristics $x_{(x_0,\nu_0)}|_t$, they must be either closed or have both ends in $\partial\Omega$ $\sigma$-almost everywhere. $\square$

The proof can obviously be adapted for $\Omega$ having only piecewise smooth boundary. Also, if $(\pi^{\Omega \times S^1})_\# v$ has compact support in $\Omega \times S^1$, we see that all curves in $\mathfrak{C}_\kappa$ must be closed.

**4. Discretization and applications.** We first discuss how to numerically discretize problem (3.8) after which we provide the results from applying our functional to various image processing tasks. The exposition and experiments are for $f = a\kappa^2 + b$, however, other choices work analogously.

**4.1. Numerical discretization.** In order to implement and apply the proposed functional to real images we have to find an adequate discretization of measures $v \in \mathcal{M}$ and of the constraints (3.5) and (3.7). This is by no means trivial since we expect the optimal measure $v$ to be composed of line measures with support only on one-dimensional sets in $\Omega \times S^1 \times \mathbb{R}$ (compare previous section), which are difficult to represent. Furthermore, constraint (3.7) has the interpretation that at any point $(x, \nu, \kappa)$ the distributional derivative of the measure $v$ in the direction $(\nu^{-\perp}, \kappa, 0)$ integrates up to zero in the $\kappa$-direction. We thus need an accurate approximation of directional derivatives, preferably involving no interpolation between different nodal values of the discretized measure $v$ (otherwise the discretized consistency condition (3.7) would prevent the measure from concentrating on one-dimensional lines). At the same time, the discretized compatibility condition (3.5) must remain satisfiable.

Our approach is based on discretizing the space of directions and curvatures, $S^1 \times \mathbb{R}$, such that each direction $\nu \in S^1$ corresponds to a lattice vector of the underlying grid on $\Omega$ and such that each discrete curvature value $\kappa$ is represented by a pair $(\nu, \tilde{\nu}) \in S^1 \times S^1$. Constraints (3.5) and (3.7) are then discretized taking care that discrete images representing characteristic functions $u : \Omega \to \mathbb{R}$ and the corresponding line measures $v$ are admissible pairs. Below we provide the details for a discretization based on 16 discrete angles (which seems sufficient in applications); a higher number of angles can be dealt with analogously.

For the ease of exposition we consider the square domain $\Omega = (0, 1)^2$. A discrete pixel image $U$ is regarded as the $L^2$-projection of an original image $u : \Omega \to \mathbb{R}$ onto the space $\mathcal{U}_h$ of piecewise constant pixel images with pixel width $h = \frac{1}{\ell}$,

$$\mathcal{U}_h = \{U : \Omega \to \mathbb{R} \mid U = \text{const. on } (ih, ih + h) \times (jh, jh + h) \text{ for } i, j = 0, \dots, \ell - 1\}.$$
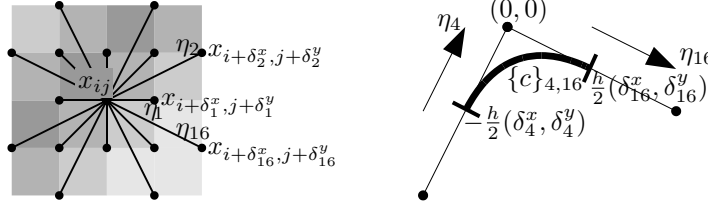
FIGURE 4.1. *Discrete angles and local variational problems for the discretization of a measure $v \in \mathcal{M}$. The gray squares represent the image pixels, the lines are in the directions of the discrete angles $\eta_1, \ldots, \eta_{16}$. $\{c\}_{4,16}$ is the optimal curve connecting two lines in directions $\eta_4$ and $\eta_{16}$.*

The discretized measure $V$ will be regarded as a weak-* approximation of a non-discrete measure $v \in \mathcal{M}$. We first consider the collection of discrete grid points $x_{ij} = (ih, jh)$, $i, j = 0, \ldots, \ell$ covering $\overline{\Omega}$. Next we consider the 16 discrete directions $\eta_1, \ldots, \eta_{16} \in S^1$ as shown in Figure 4.1, left, where $\eta_1 = (1,0)^T$ and the angles are numbered counterclockwise (the discretized normal directions will be $\eta_1^\perp, \ldots, \eta_{16}^\perp$). Each $\eta_k$ is associated with a discrete grid offset $(\delta_k^x, \delta_k^y)^T = \alpha_k \eta_k$ where $\alpha_k$ is the smallest positive number such that $\alpha_k \eta_k \in \mathbb{Z} \times \mathbb{Z}$. Finally we associate with each pair $(\eta_k, \eta_l)$ the variational problem of finding an optimal smooth curve segment $c$ connecting two lines with directions $\eta_k$ and $\eta_l$, respectively (see Figure 4.1, right),

$$\min \left\{ \int_{\{c\}} a\kappa^2 + b \, d\mathcal{H}^1 \;\middle|\; c : [0,1] \to \mathbb{R}^2, \; c(0) = -\tfrac{h}{2}(\delta_k^x, \delta_k^y), \; c(1) = \tfrac{h}{2}(\delta_l^x, \delta_l^y), \right.$$
$$\left. c'(0) \,\|\, \eta_k, \; c'(1) \,\|\, \eta_l, \; \mathrm{sgn}(\eta_k \cdot \eta_l^\perp) \, \kappa \geq 0 \text{ on } \{c\} \right\}.$$

Here $\{c\}$ denotes the image of $c$, that is, the actual curve segment, and $\kappa$ is its curvature. Let us denote an optimal line segment by $\{c\}_{kl}$ (in case of multiple minimizers we pick one) and set $f_{kl} = \int_{\{c\}_{kl}} a\kappa^2 + b \, d\mathcal{H}^1$. Measures $v \in \mathcal{M}$ are now approximated by discrete versions of the form

$$V = \sum_{i,j=0}^{\ell} \sum_{k,l=1}^{16} V_{ijkl} v_{ijkl} \,,$$

where $V_{ijkl} \geq 0$ and $v_{ijkl}$ is the lifting of $\mathcal{H}^1 \llcorner (x_{ij} + c_{kl})$, that is, the measure defined via

$$\int_{\Omega \times S^1 \times \mathbb{R}} \phi(x, \nu, \kappa) \, dv_{ijkl}(x, \nu, \kappa) = \int_{x_{ij} + \{c\}_{kl}} \phi(x, \nu(x), \kappa(x)) \, d\mathcal{H}^1(x) \quad \forall \phi \in \mathcal{C}_c(\Omega \times S^1 \times \mathbb{R}; \mathbb{R}) \,,$$

where $\nu(x)$ and $\kappa(x)$ are the normal and curvature to $x_{ij} + \{c\}_{kl}$ at $x$.

Inserting the discrete form of $v$ into the consistency condition (3.7) and using

$$\int_{\Omega \times S^1 \times \mathbb{R}} \nabla_x \psi(x, \nu) \cdot \nu^{-\perp} + \frac{\partial \psi(x, \nu)}{\partial \nu} \kappa \, dv_{ijkl}(x, \nu, \kappa) = \psi(x_{ij} + \tfrac{h}{2}(\delta_l^x, \delta_l^y), \eta_l^\perp) - \psi(x_{ij} - \tfrac{h}{2}(\delta_k^x, \delta_k^y), \eta_k^\perp)$$

for all $\psi \in \mathcal{C}_0^\infty(\Omega \times S^1; \mathbb{R})$, we obtain the discretized consistency condition as the set of equations

$$(4.1) \qquad 0 = \sum_{l=1}^{16} V_{ijkl} - \sum_{l=1}^{16} V_{i - \delta_k^x, j - \delta_k^y, l, k} \,, \qquad i, j = 0, \ldots, \ell, \quad k = 1, \ldots, 16 \,.$$

For the discretization of the compatibility condition, note that the best available approximation of the gradient of the image $u$ is based on difference quotients between two pixels of the pixel approximation $U$. Hence, let us define

$$g(s) = \begin{cases} h + s, & -h \le s \le 0, \\ h - s, & 0 \le s \le h, \\ 0 & \text{else}, \end{cases}$$

$$\varphi_{ij}^x(x) = \tfrac{1}{h^3} \chi_{(jh,jh+h)}(x_2) \left(g(x_1 - ih), 0\right)^T,$$

$$\varphi_{ij}^y(x) = \tfrac{1}{h^3} \chi_{(ih,ih+h)}(x_1) \left(0, g(x_2 - jh)\right)^T$$

for $i, j \in \mathbb{N}$. Upon integrating by parts we obtain $\int_\Omega \nabla u(x) \cdot \varphi_{ij}^x(x) \, dx = \frac{U_{ij} - U_{i-1,j}}{h}$ and $\int_\Omega \nabla u(x) \cdot \varphi_{ij}^y(x) \, dx = \frac{U_{ij} - U_{i,j-1}}{h}$, where $U_{ij} = \frac{1}{h^2} \int_{[ih,ih+h] \times [jh,jh+h]} u(x) \, dx$ is the $(i,j)$ image pixel value. Now, inserting the discrete form of $v$ into the compatibility condition (3.5) and using the test functions $\varphi_{ij}^x$ and $\varphi_{ij}^y$ (which is allowed by a mollifying argument), we obtain the discrete compatibility condition

$$(4.2) \qquad \sum_{m,n=0}^{\ell} \sum_{k,l=1}^{16} V_{mnkl} \begin{pmatrix} \Delta_{i-m,j-n,k,l}^x \\ \Delta_{i-m,j-n,k,l}^y \end{pmatrix} = \begin{pmatrix} \frac{U_{ij} - U_{i-1,j}}{h} \\ \frac{U_{ij} - U_{i,j-1}}{h} \end{pmatrix}, \qquad i, j = 1, \dots, \ell,$$

with $\Delta_{ijkl}^{x,y} = \int_{\{c\}_{kl}} \varphi_{ij}^{x,y}(x) \cdot \nu(x) \, d\mathcal{H}^1(x)$. Obviously, $\Delta_{ijkl}^x = 0$ for $i \notin \{-1, 0, 1\}$ or $j \notin \{-1, 0\}$, and likewise $\Delta_{ijkl}^y = 0$ for $i \notin \{-1, 0\}$ or $j \notin \{-1, 0, 1\}$, which makes the above condition sparse.

Finally, the discretized relaxed elastica functional is obtained by inserting the discrete form of $V$ into (3.8),

$$(4.3) \qquad R_{\text{el},h}^{\text{ap}}(U) = \min_{V_{ijkl} \ge 0 \wedge (4.2) \wedge (4.1)} \sum_{i,j=0}^{\ell} \sum_{k,l=1}^{16} V_{ijkl} f_{kl}.$$

In order to use this discretization, one just needs to precompute the values $f_{kl}$ and $\Delta_{ijkl}^{x,y}$.

To solve the optimization problem associated with $\{c\}_{kl}$ we represent the curve segment $c$ as a polynomial $c(t) = c_0 + t c_1 + \dots + t^N c_N$ of degree $N$ and then numerically optimize for the coefficients $c_0, \dots, c_N$. Note that for $N = 2$ the constraints on $c$ already fully specify all coefficients, which can be found by solving a linear system of equations. An example of optimized curve segments for $N = 4$, $h = 1$, and $a = b = 1$ is shown in Figure 4.2. $f_{kl}$ and $\Delta_{ijkl}^{x,y}$ are then obtained via numerical quadrature.

**4.2. Applications to image processing problems.** In this section we present numerical results on several problems in image processing. First we define a class of convex optimization problems appropriate for many image processing problems,

$$(4.4) \qquad \min_U G(U) + R_{\text{el},h}^{\text{ap}}(U),$$

where $R_{\text{el},h}^{\text{ap}}(U)$ is the discretized convex approximation of the elastica energy defined in (4.3) and $G$ is a proper, convex, lower semi-continuous function with simple to compute proximal map

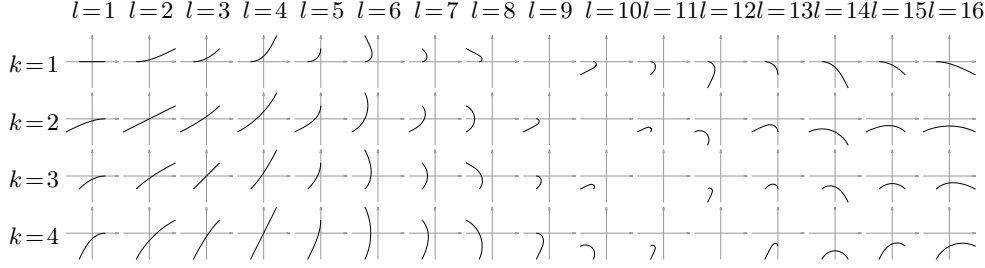$$\text{prox}_G^\tau(\hat{U}) = \arg\min_U \frac{\|U - \hat{U}\|_2^2}{2\tau} + G(U),$$

$l=1$ $l=2$ $l=3$ $l=4$ $l=5$ $l=6$ $l=7$ $l=8$ $l=9$ $l=10$ $l=11$ $l=12$ $l=13$ $l=14$ $l=15$ $l=16$



FIGURE 4.2. *Optimal curve segments $\{c\}_{kl}$ for $h = 1$, and $a = b = 1$, approximated by fourth order polynomials. The curve segments for $k = 5,\ldots,16$ are obtained as rotated versions of the displayed curve segments.*

for some $\hat{U}$. We proceed by organizing both, the lifted measure $V$ and the image $U$ into row vectors $V \in \mathbb{R}^{16^2 wh}$ and $U \in \mathbb{R}^{wh}$, where $w$ and $h$ refer to the width and the height of the image $U$. Note that in the previous section, we have presented the discretization for the elastica energy only for $w = h = \ell$, but a generalization to the setting $w \neq h$ is straightforward. Next, we rewrite (4.4) as the following linearly constrained optimization problem jointly in $U$ and $V$.

$$(4.5) \qquad \min_{V,U} G(U) + F^T V,$$
$$\text{s.t.} V \geq 0, \ \mathbf{A}V = 0, \ \mathbf{B}V - \mathbf{C}U = 0.$$

The cost vector $F \in \mathbb{R}^{16^2 wh}$ is computed such that $F^T V$ yields (4.3). The sparse matrices $\mathbf{A} \in \mathbb{R}^{16wh \times 16^2 wh}$, $\mathbf{B} \in \mathbb{R}^{2wh \times 16^2 wh}$ and $\mathbf{C} \in \mathbb{R}^{2wh \times wh}$ are computed such that they implement the consistency condition (4.2) and the compatibility condition (4.1), respectively. Additionally, we introduce multipliers (or dual vectors) $P \in \mathbb{R}^{16wh}$ to account for the constraint $\mathbf{A}V = 0$ and $Q \in \mathbb{R}^{2wh}$ to account for the constraint $\mathbf{B}V - \mathbf{C}U = 0$. Furthermore we define the matrix

$$\mathbf{K} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & -\mathbf{C} \end{pmatrix},$$

and denote by $L = \|\mathbf{K}\|$, the operator norm of $\mathbf{K}$. In practice, an approximation to $L$ can be easily computed using for example the power iteration algorithm.

We can now rewrite (4.5) as a primal-dual saddle point problem of the form

$$(4.6) \qquad \min_{U,V \geq 0} \max_{P,Q} G(U) + F^T V + \begin{pmatrix} P \\ Q \end{pmatrix}^T \mathbf{K} \begin{pmatrix} V \\ U \end{pmatrix},$$

which we solve by the first-order primal-dual algorithm [10]. The iterates of the primal-dual algorithm are then as follows: Choose $\tau, \sigma > 0$ such that $\tau\sigma L^2 < 1$, $V^0 = 0$, $U^0 = 0$, $P^0 = 0$, $Q^0 = 0$. Then for all $n \geq 0$ set

$$(4.7)$$
$$\begin{cases} \begin{pmatrix} P^{n+1} \\ Q^{n+1} \end{pmatrix} = \begin{pmatrix} P^n \\ Q^n \end{pmatrix} + \sigma\, \mathbf{K} \begin{pmatrix} V^n \\ U^n \end{pmatrix}, \\[2mm] \begin{pmatrix} V^{n+1} \\ U^{n+1} \end{pmatrix} = \begin{pmatrix} \max(0, \cdot) \\ \mathrm{prox}_G^\tau(\cdot) \end{pmatrix} \circ \left( \begin{pmatrix} V^n \\ U^n \end{pmatrix} - \tau \begin{pmatrix} F \\ 0 \end{pmatrix} - \tau\, \mathbf{K}^T \begin{pmatrix} 2P^{n+1} - P^n \\ 2Q^{n+1} - Q^n \end{pmatrix} \right). \end{cases}$$

We stop the iterations if the root-mean-squared primal feasibility error

$$e^n = \sqrt{\frac{\left\| \mathbf{K} \begin{pmatrix} V^n \\ U^n \end{pmatrix} \right\|_2^2}{(16+2)wh}}$$

is less then a tolerance of $\mathrm{tol} = 10^{-3}$.

In order to adapt the above formulation to different image processing applications, it remains to specify different functions $G$ and to implement the proximal mappings for the respective functions. This will be covered in the next few paragraphs.

**4.2.1. Binary image segmentation.** In the first example, we apply the elastica energy to the task of binary image segmentation. We choose the well-known cameraman image, rescale the intensity values $I_{i,j}$ to the interval $[0,1]$ and specify the mean values of the foreground and the background as $\mu_f = 0.5$ and $\mu_b = 0.0$. Based on that we defined the data for the segmentation energy as

$$G_s(U) = \lambda \sum_{i,j} U_{i,j}(I_{i,j} - \mu_b)^2 + (1 - U_{i,j})(I_{i,j} - \mu_f)^2 \,, \text{ s.t. } 0 \leq U_{i,j} \leq 1 \,,$$

where $\lambda$ is a tuning parameter. The proximal map is easily computed as

$$\mathrm{prox}_{G_s}^\tau(\hat{U}) = \max(0, \min(1, \hat{U} - \tau\lambda((I - \mu_b)^2 - (I - \mu_f)^2))) \,,$$

where all operations in the proximal map are carried out pointwise. Figure 4.3 compares the results of using pure length-based regularization ($a = 0$) and elastica-based regularization ($a = 10$) for different values of $\lambda$. Observe that pure length-based regularization causes a strong shrinkage of the segmented object, while elastica-based regularization leads to a better preservation of elongated structures.

**4.2.2. Image denoising.** In the next example, we apply the elastica-based energy to image denoising problems. We consider two different kinds of noise: (i) Gaussian noise, where a $\ell_2$ data-term is most adequate and (ii) impulse-noise for which it is known that a $\ell_1$ data term performs better. Let us again denote by $I$ the input image having the intensity values scaled to the interval $[0,1]$. The $\ell_2$ data term and its proximal map are given by

$$G_{\ell_2}(U) = \frac{\lambda}{2}\|U - I\|_2^2 \,, \quad \mathrm{prox}_{G_{\ell_2}}^\tau(\hat{U}) = \frac{\hat{U} + \tau\lambda I}{1 + \tau\lambda} \,,$$

where again the operations in the proximal map are understood pointwise and $\lambda$ is a tuning parameter. Likewise, the $\ell_1$ data term and its proximal map are given by

$$G_{\ell_1}(U) = \lambda\|U - I\|_1 \,, \quad \mathrm{prox}_{G_{\ell_1}}^\tau(\hat{U}) = \max(0, |\hat{U} - I| - \tau\lambda)\,\mathrm{sgn}(\hat{U} - I) \,.$$

Figure 4.4 shows results of using pure length-based regularization (TV) and elastica-based regularization for denoising an image that has been degraded by Gaussian noise of standard deviation $\sigma = 0.1$. One can observe that the elastica-based regularizer leads to a better restoration of line-like structures such as the wings of the wind mill as well as textured regions such as the stone wall in the foreground. The elastica model also produces some artifacts in homogeneous regions which are explained by the property of the elastica model to hallucinate low-contrast structures in regions, whose contrast variations are dominated by the noise component.
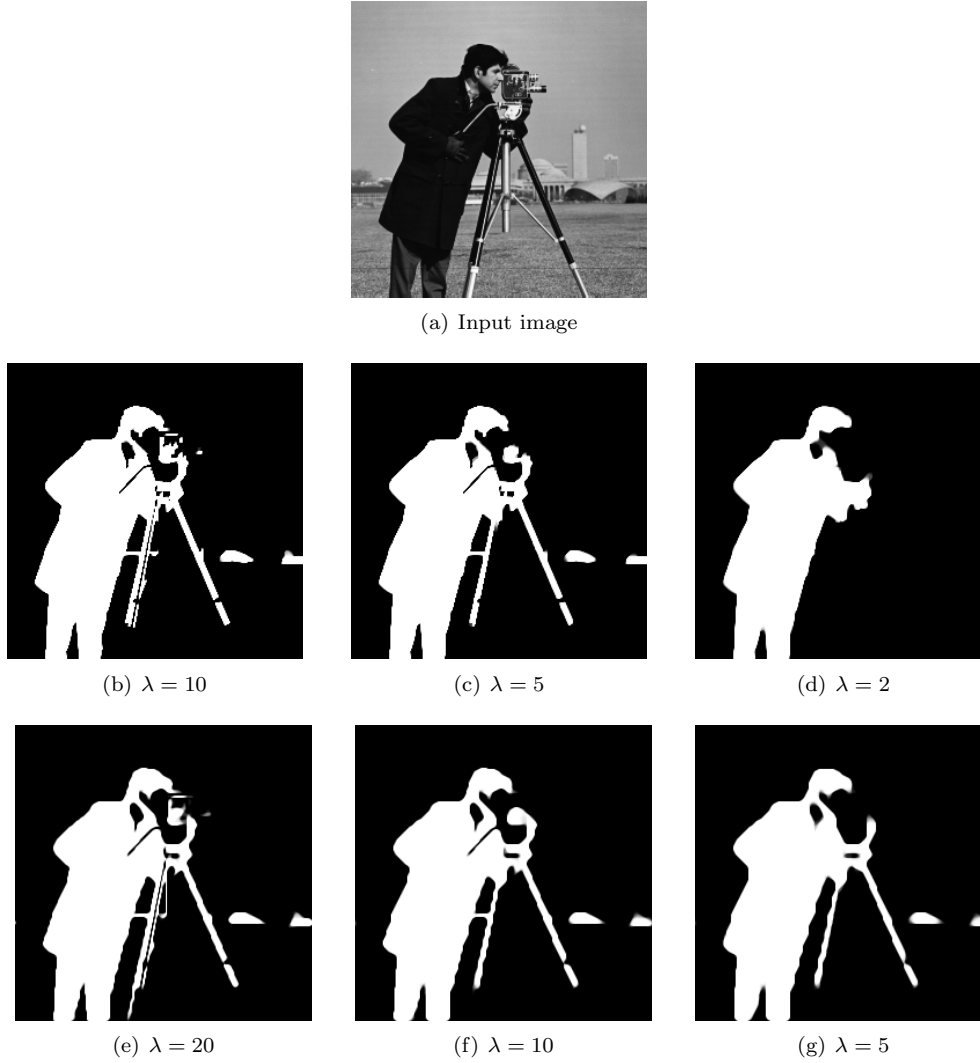
(a) Input image



(b) $\lambda = 10$          (c) $\lambda = 5$          (d) $\lambda = 2$



(e) $\lambda = 20$          (f) $\lambda = 10$          (g) $\lambda = 5$

FIGURE 4.3. *Application to binary image segmentation for different data weights $\lambda$: The first row shows the input image, the second row shows the results for pure length-based regularization ($a = 0$), and the third row shows the results for elastica-based regularization ($a = 10$). Note that for smaller values of the data weight, elastica-based regularization leads to a better preservation of elongated structures.*

Figure 4.5 shows the results of applying the elastica model to the restoration of an image that has been degraded with 25% salt-and-pepper noise. Here, the elastica-based model leads to significantly better results compared to the pure-length-based model, in particular, the line-like structures of the glass pyramid are much better preserved. The reason why the elastica model performs much better on impulse noise is that the strong outliers come along with strong local contributions to the elastica energy and hence are efficiently regularized by the elastica model. The pure length-based model can not distinguish well between outliers caused by the noise and line-like structures and hence leads to inferior restoration results.
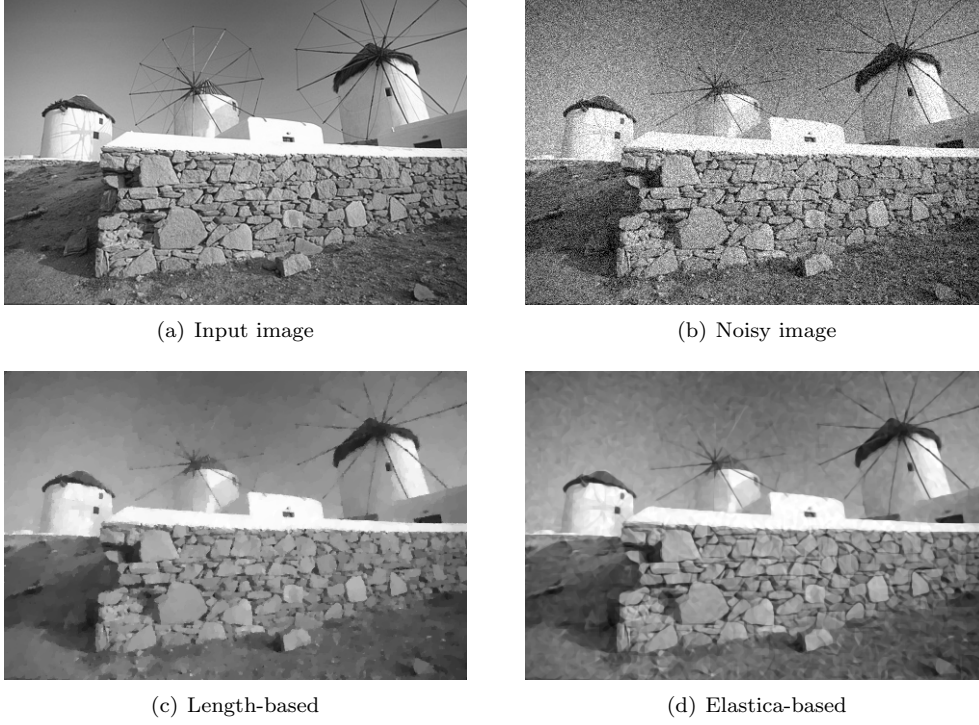
(a) Input image



(b) Noisy image



(c) Length-based



(d) Elastica-based

FIGURE 4.4. *Application to image denoising in case of Gaussian noise. One can see that the elastica regularization ($a = 10$) leads to a better preservation of line-like structures compared to pure TV regularization ($a = 0$). Note also, that in homogeneous regions, the elastica regularization starts to "hallucinate" new structures.*

**4.2.3. Image deconvolution.** Now, we apply the elastica-based regularizer to image deconvolution. We assume that the point spread function (or blur kernel), written as a linear operator $\mathbf{L} \in \mathbb{R}^{wh \times wh}$, is known in advance and assume that the image $I$ has been additionally degraded by Gaussian noise. Hence, we consider a $\ell_2$ data term of the form

$$G_D(U) = \frac{\lambda}{2}\|\mathbf{L}U - I\|_2^2\,,$$

where $\lambda$ is again a tuning parameter. The proximal map is computed as

$$\mathrm{prox}_{G_D}^\tau(\hat{U}) = \left(\mathbf{id} + \tau\lambda\mathbf{L}^T\mathbf{L}\right)^{-1}\left(\hat{U} + \tau\lambda\mathbf{L}^T I\right)\,,$$

which amounts to solving a linear system of equations. We solve this system approximately by using the conjugate gradient (CG) algorithm. We use of a warm-start strategy by always initializing the CG algorithm with the last iterate $U^n$. Using this strategy, only a few iterations (2-10) of the CG algorithm are needed per iteration of the primal-dual algorithm in order to compute the proximal map with sufficient accuracy, say $10^{-6}$.

Figure 4.6 shows the result for restoring van Gogh's sunflowers painting, that has been blurred by a motion kernel (angle $\theta = \pi/4$, length $l = 21$) and additionally degraded by Gaussian noise of standard deviation $\sigma = 0.01$. The comparison between
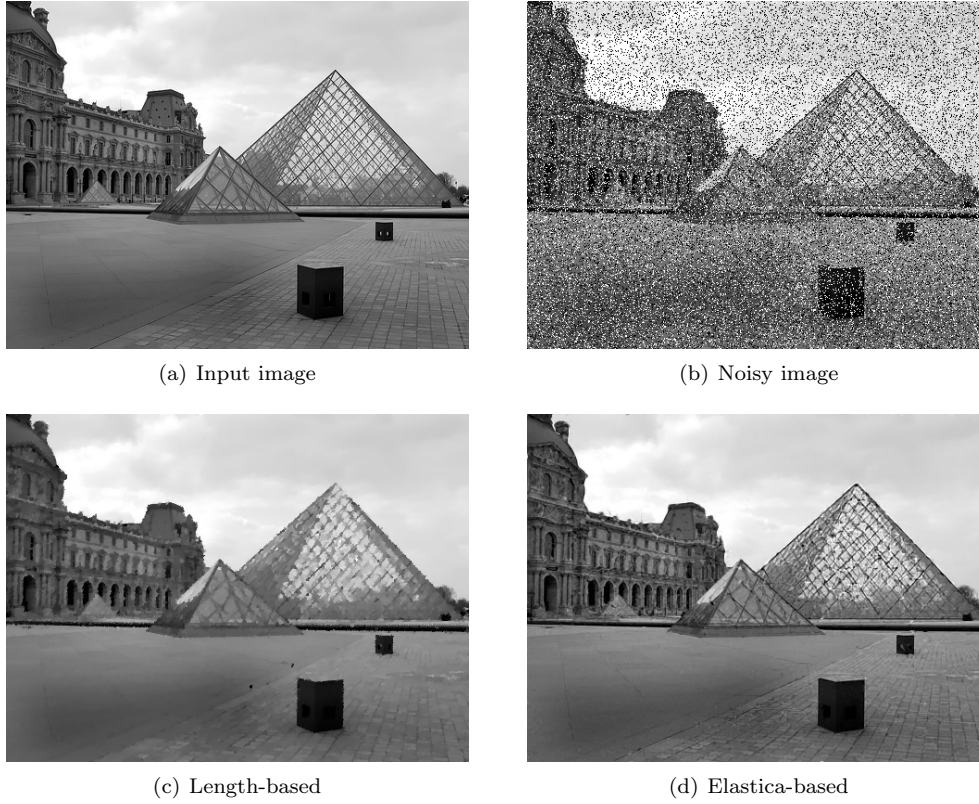
(a) Input image



(b) Noisy image



(c) Length-based



(d) Elastica-based

FIGURE 4.5. *Application to image denoising in case of salt and pepper noise. Compared to pure length-based regularization (a = 0), elastica-based regularization (a = 10) leads to significantly better restoration results. In particular the line-like structures of the glass pyramid are very well preserved.*

pure length-based regularization and elastica regularization shows that the elastica regularization can better restore fine details and in particular line like structures. Overall, elastica-based regularization clearly leads to visually more appealing results.

**4.2.4. Image inpainting.** In the last example, we apply the elastica energy to image inpainting, i.e. the restoration of lost image information. We point out that the elastica energy was initially proposed for image inpainting and indeed, we will see that the advantages of the elastica model over the TV model become most evident in this application. For image inpainting we consider an input image $I$ together with a set $\mathcal{I}$ of pixel locations $(i,j)$ for which the intensity values $I_{i,j}$ are known. An appropriate data term and its proximal map in this setting are given by

$$G_I(U) = \sum_{(i,j)\in\mathcal{I}} \begin{cases} 0 & \text{if } U_{i,j} = I_{i,j} \\ \infty & \text{else} \end{cases}, \quad \text{prox}_{G_I}^\tau(\hat{U}_{i,j}) = \begin{cases} I_{i,j} & \text{if } (i,j)\in\mathcal{I} \\ \hat{U}_{i,j} & \text{else} \end{cases}.$$

Figure 4.7 depicts the results of length- and elastica-based image inpainting. One can clearly see that the elastica-based model is much more successful in restoring the original input image than the pure length-based (TV) model. The additional assumption of the elastica-model that the level lines should have low curvature helps a lot in
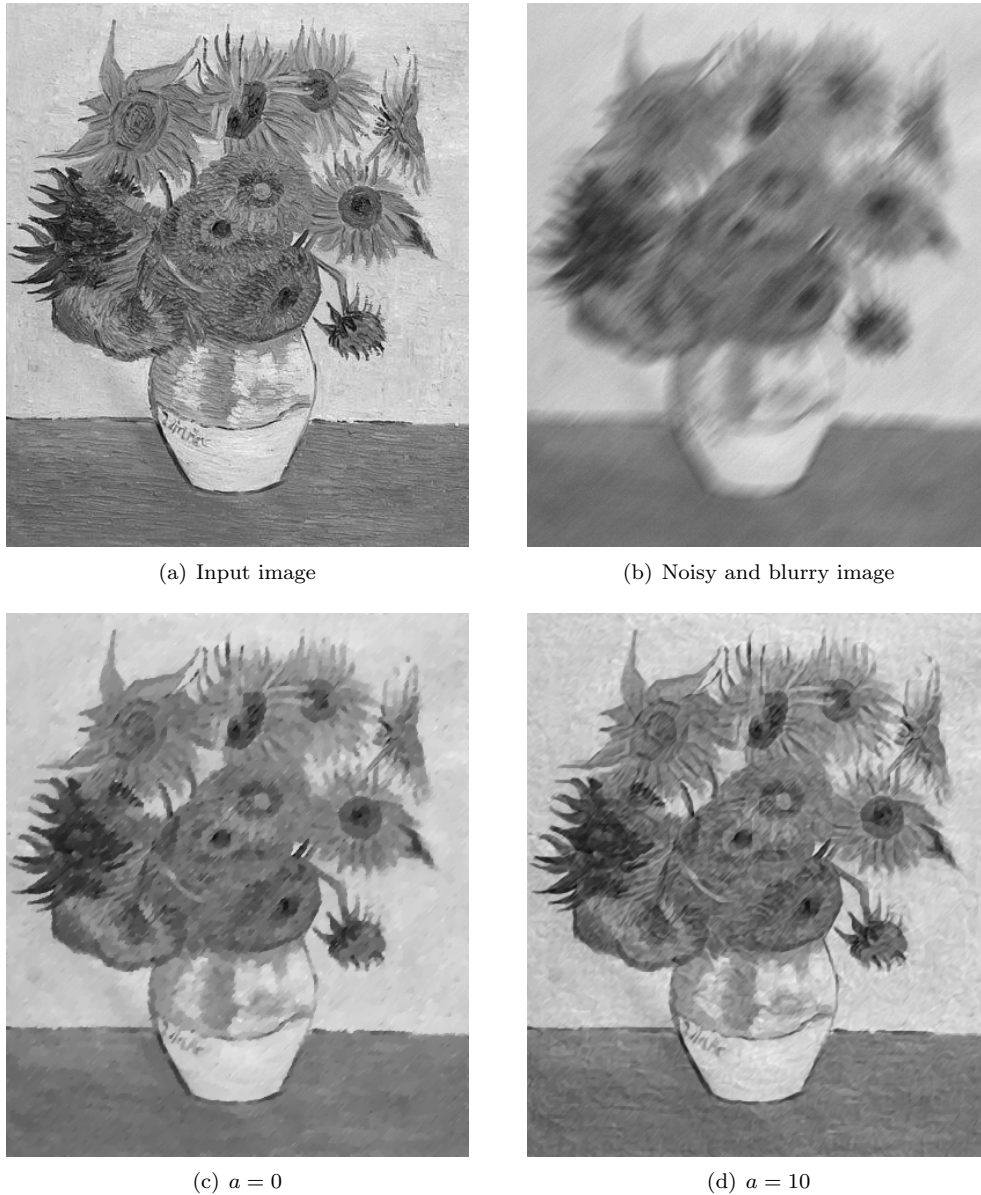
<center>(a) Input image</center>



<center>(b) Noisy and blurry image</center>



<center>(c) $a = 0$</center>



<center>(d) $a = 10$</center>

FIGURE 4.6. *Application to image deconvolution. In contrast to pure length-based regularization ($a = 0$), elastica regularization ($a = 10$) can better restore van Gogh's gestural brush strokes.*

reconstructing the small structures of the image and at the same time avoids small isolated pixels, which are contained in the result of pure length-based reconstruction.

**5. Conclusion.** In this work we proposed a convex approximation of Euler's elastica functional $R_{\mathrm{el}}$. Even though our approximation turns out to lie below the convex relaxation of $R_{\mathrm{el}}$, it still appears to have very good regularizing properties for image processing tasks as numerical tests confirm. The difficulty with the true convex relaxation seems to be that it has components to it which are rather of a global
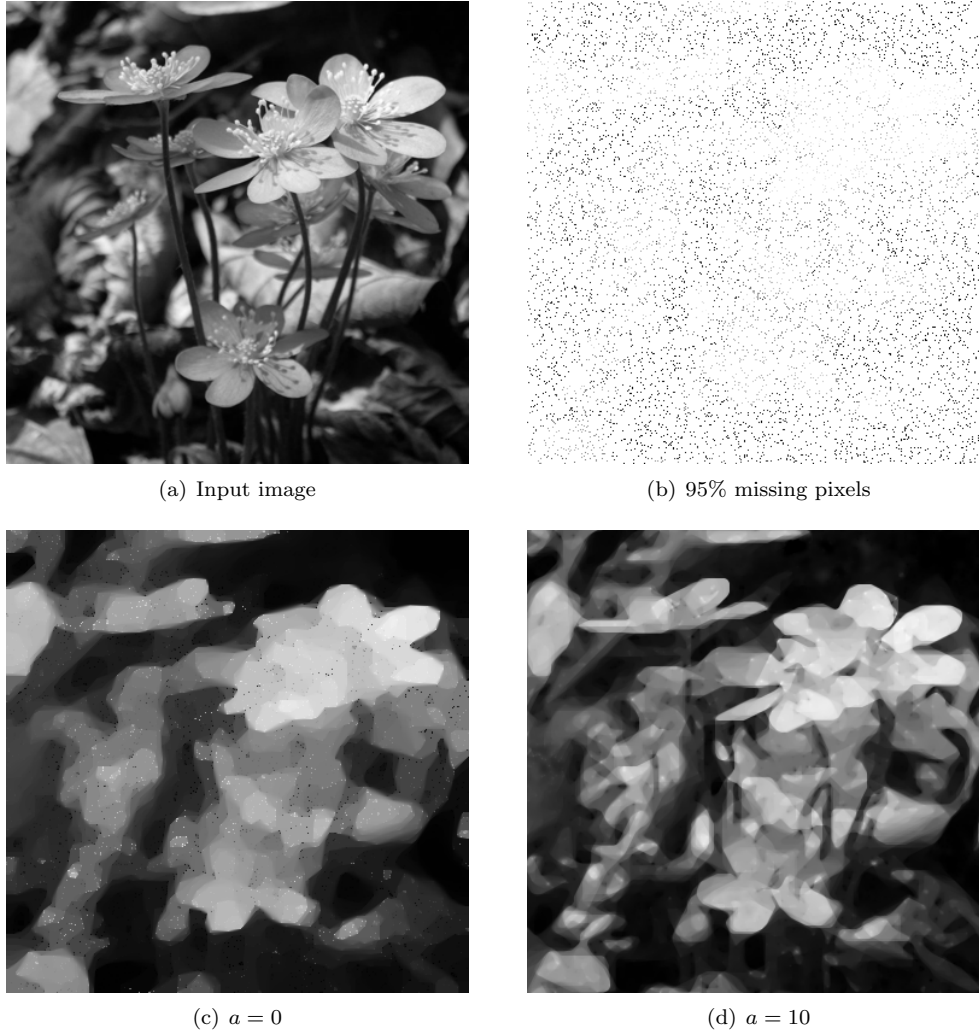
(a) Input image

(b) 95% missing pixels

(c) $a = 0$

(d) $a = 10$

FIGURE 4.7. *Application to image inpainting. This experiment compares the inpainting capabilities between pure length-based regularization ($a = 0$) and elastica-based regularization ($a = 10$). One can clearly see that elastica-based regularization leads to significantly better inpainting results.*

nature: While our convex approximation essentially reproduces the level lines of a given image by convex combinations of smooth 2D curves, the convex relaxation additionally distinguishes between 2D curves with and without self-intersection (where the detection of self-intersections would require to trace the complete curve). In contrast, our functional is fully based on local information. Furthermore, its evaluation amounts to a simple linear optimization, which makes computations very efficient.

The convexity of our functional comes at the expense of introducing additional dimensions to the image processing problem at hand: The two image dimensions are complemented with two more dimensions, one encoding normals of image level lines, the other encoding curvatures, and each of them has to be discretized with at least 16 points to obtain good results. This downside is shared by all functional lifting

approaches for the convexification of energy functionals, however, typically only one dimension needs to be added if the energy is not curvature-based. Nevertheless, the proposed model is still computationally feasible.

There are still several open questions, in particular it is not fully understood how close a computationally feasible model could get to the convex relaxation of the elastica energy. Also it is not clear how a generalization of our approach to 3D images and mean curvature regularization could be accomplished.

## REFERENCES

[1] Giovanni Alberti, Guy Bouchitté, and Gianni Dal Maso, *The calibration method for the Mumford-Shah functional and free-discontinuity problems*, Calculus of Variations and Partial Differential Equations, 16 (2003), pp. 299–333.

[2] Luigi Ambrosio, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[4] Luigi Ambrosio and Simon Masnou, *A direct variational approach to a problem arising in image reconstruction*, Interfaces and Free Boundaries, 5 (2003), pp. 63–81.

[5] Hédy Attouch and Haïm Brezis, *Duality for the sum of convex functions in general Banach spaces*, in Aspects of mathematics and its applications, vol. 34 of North-Holland Math. Library, North-Holland, Amsterdam, 1986, pp. 125–133.

[6] Andrea Braides and Anneliese Defranceschi, *Homogenization of multiple integrals*, vol. 12 of Oxford Lecture Series in Mathematics and its Applications, The Clarendon Press Oxford University Press, New York, 1998.

[7] Kristian Bredies, Thomas Pock, and Benedikt Wirth, *Convex relaxation of a class of vertex penalizing functionals*, Journal of Mathematical Imaging and Vision, 47 (2013), pp. 278–302.

[8] Antonin Chambolle, *Convex representation for lower semicontinuous envelopes of functionals in $L^1$*, Journal of Convex Analysis, 8 (2001), pp. 149–170.

[9] Antonin Chambolle, Daniel Cremers, and Thomas Pock, *A convex approach to minimal partitions*, SIAM Journal on Imaging Sciences, 5 (2012), pp. 1113–1158.

[10] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), pp. 120–145.

[11] Tony F. Chan, Sung Ha Kang, and Jianhong Shen, *Euler's elastica and curvature-based inpainting*, SIAM Journal on Applied Mathematics, 63 (2002), pp. 564–592.

[12] Manfred Einsiedler and Thomas Ward, *Ergodic theory with a view towards number theory*, vol. 259 of Graduate Texts in Mathematics, Springer-Verlag London Ltd., London, 2011.

[13] Irene Fonseca and Giovanni Leoni, *Modern methods in the calculus of variations: $L^p$ spaces*, Springer Monographs in Mathematics, Springer, New York, 2007.

[14] B. Goldlücke and D. Cremers, *Introducing total curvature for image processing*, in IEEE International Conference on Computer Vision (ICCV), 2011.

[15] Gernot Greschonig and Klaus Schmidt, *Ergodic decomposition of quasi-invariant probability measures*, Colloq. Math., 84/85 (2000), pp. 495–514. Dedicated to the memory of Anzelm Iwanik.

[16] Jan Lellmann, Jörg Kappes, Jing Yuan, Florian Becker, and Christoph Schnörr, *Convex multi-class image labeling by simplex-constrained total variation*, in Scale Space and Variational Methods in Computer Vision, vol. 5567 of Lecture Notes in Computer Science, Springer Berlin/Heidelberg, 2009, pp. 150–162.

[17] SIMON MASNOU, *Disocclusion: a variational approach using level lines*, IEEE Transactions on Image Processing, 11 (2002), pp. 68–76.

[18] SIMON MASNOU AND GIACOMO NARDI, *A coarea-type formula for the relaxation of a generalized elastica functional*, Journal of Convex Analysis, (2012).

[19] ———, *Gradient Young measures, varifolds, and a generalized Willmore functional*, Advances in Calculus of Variations, (2012).

[20] FRANK NATTERER, *The mathematics of computerized tomography*, vol. 32 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Reprint of the 1986 original.

[21] MILA NIKOLOVA, SELIM ESEDOGLU, AND TONY F. CHAN, *Algorithms for finding global minimizers of image segmentation and denoising models*, SIAM Journal on Applied Mathematics, 66 (2006), pp. 1632–1648.

[22] M. NITZBERG, D. MUMFORD, AND T. SHIOTA, *Filtering, segmentation and depth*, vol. 662 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1993.

[23] THOMAS POCK, ANTONIN CHAMBOLLE, DANIEL CREMERS, AND HORST BISCHOF, *A convex relaxation approach for computing minimal partitions*, in IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 810–817.

[24] THOMAS POCK, DANIEL CREMERS, HORST BISCHOF, AND ANTONIN CHAMBOLLE, *An algorithm for minimizing the Mumford-Shah functional*, in IEEE 12th International Conference on Computer Vision, 2009, pp. 1133–1140.

[25] ———, *Global solutions of variational models with convex regularization*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 1122–1145.

[26] WALTER RUDIN, *Real and complex analysis*, McGraw-Hill Book Co., New York, third ed., 1987.

[27] THOMAS SCHOENEMANN AND DANIEL CREMERS, *Introducing curvature into globally optimal image segmentation: Minimum ratio cycles on product graphs*, in International Conference on Computer Vision (ICCV2007), Rio de Janeiro, Brazil, October 2007.

[28] THOMAS SCHOENEMANN, FREDERIK KAHL, AND DANIEL CREMERS, *Curvature regularity for region-based image segmentation and inpainting: A linear programming relaxation*, in IEEE 12th International Conference on Computer Vision, 2009, pp. 17–23.

[29] THOMAS SCHOENEMANN, FREDERIK KAHL, SIMON MASNOU, AND DANIEL CREMERS, *A linear framework for region-based image segmentation and inpainting involving curvature penalization*, International Journal of Computer Vision, 99 (2012), pp. 53–68.

[30] THOMAS SCHOENEMANN, SIMON MASNOU, AND DANIEL CREMERS, *The elastic ratio: Introducing curvature into ratio-based image segmentation*, IEEE Transactions on Image Processing, 20 (2011), pp. 2565–2581.

[31] YA. G. SINAĬ, *Topics in ergodic theory*, vol. 44 of Princeton Mathematical Series, Princeton University Press, Princeton, NJ, 1994.

[32] XUE-CHENG TAI, JOOYOUNG HAHN, AND GINMO JASON CHUNG, *A fast algorithm for Euler's elastica model using augmented Lagrangian method*, SIAM Journal on Imaging Sciences, 4 (2012), pp. 313–344.

[33] W. ZHU AND T. CHAN, *Image denoising using mean curvature of image surface*, SIAM Journal on Imaging Sciences, 5 (2012), pp. 1–32.