VRIJE
UNIVERSITEIT
BRUSSEL

THE MUMFORD-SHAH FUNCTIONAL

Othman El Hammouchi

2021-2022

Promotors: Prof. Dr. Kenny De Commer, Prof. Dr. Andreas Debrouwere

Sciences and Bioengineering Sciences

1 Introduction

Then *segmentation problem* in computer vision and digital image processing refers to the search for efficient and accurate methods to decompose an image into its principal constituent regions in order to obtain a representation which is easier to analyse and manipulate downstream. Here, a (grayscale) image is represented as a measurable function $g : \Omega \to [0, 1]$ giving the light intensity at every point of a flat display $\Omega \subseteq \mathbb{R}^2$, which we shall take to be an open rectangle. Segmenting the image then involves finding disjoint open subregions $R_k \subseteq \Omega$, $k = 1, \dots, n$, corresponding to contiguous surfaces of objects, and separated by a one-dimensional boundary $K = \overline{\Omega} \setminus \bigcup_{k=1}^{n} R_k$ giving the edges of the image. In their 1989 paper, mathematicians David Mumford and Jayant Shah attempted to tackle this challenge by variational methods, proposing to obtain the segmentation as the solution to an optimisation problem involving their newly-introduced eponymous functional. In this framework, g is approximated by a piecewise smooth model function $u \in C^1(\Omega \setminus K)$ minimising

$$E(u, K) := \int_{\Omega} (u - g)^2 \, dx + \int_{\Omega \setminus K} \|\nabla u\|^2 \, dx + \mathcal{H}^1(K)\,, \tag{1}$$

where $\mathcal{H}^1$ denotes the 1-dimensional Hausdorff measure and (u, K) is required to belong to the set

$$\mathcal{A} := \left\{(u, K) \mid K \subseteq \overline{\Omega}\,, u \in C^1(\Omega \setminus K)\right\}$$

of admissible pairs. The first term ensures proximity of the model to the original image, the second controls the variation of u on each contiguous subregion, and the final one constrains the length of the discontinuity set. All three are necessary to make the problem interesting, and yield a solution which behaves regularly within each subregion and exhibits discontinuities when transitioning between them. These heuristics are only meaningful, however, if we know that a minimiser exists in the first place. For while it is always possible to use a numerical approach, discretising the functional and minimising it on a grid, in the absence of existence results, we cannot know whether solutions thus obtained do in fact constitute approximations to the original problem, and thus cannot benefit from theoretical properties of the latter. Our chief endeavour in this paper will therefore be the exposition of an existence theory for the Mumford-Shah problem, whose challenging nature has proven to be a powerful impetus for the development of a rich subfield of mathematics: the study of so-called *free discontinuity problems*, variational problems where the set on which the solution is allowed to exhibit pathological behaviour is itself an unknown quantity.

Following methods in the calculus of variations, our exposition shall proceed in two parts. First, we will relax (1) on a larger space with suitable compactness properties to obtain an existence result for this weaker formulation. We then proceed to show that the minimiser thus obtained in fact belongs to the original space of admissible solutions $\mathcal{A}$, which is a regularity result on this weak solution. Of paramount importance for this approach is the following result, known as the *direct method in the calculus of variations*.

Theorem 1.1. *Let $J : V \to \mathbb{R} \cup \{+\infty\}$ be a functional defined on a topological vector space V and bounded from below, i.e.* $\inf_{u \in V} J(u) > -\infty$, *and suppose the following conditions are satisfied:*

(i) For every $M \in \mathbb{R}$, the sublevel set $\{J \leq M\}$ is relatively sequentially compact for the topology on V.

(ii) J is lower semicontinuous with respect to the topology on V.

Then J realises its infimum, i.e. there exists $u_0 \in V$ such that $J(u_0) = \inf_{u \in V} J(u)$.

Proof. By definition, there exists a sequence $(u_n)_n \subseteq V$ such that $J(u_n) \xrightarrow{n \to \infty} \inf_{u \in V} J(u)$. If we choose any $M > \inf_{u \in V} J(u)$, then we will eventually have $J(u_n) \leq M$ for n sufficiently large, so that we may suppose $(J(u_n))_n$ to lie in $\{J \leq M\}$ from the outset. The relative sequential compactness of this sublevel set then allows us to extract a subsequence $(u_{k_n})_n$ converging to some $u \in V$. Lower semicontinuity of J now implies

$$J(u) \leq \liminf_{n \to \infty} J(u_{k_n}) ,$$

from which we conclude that u realises the desired infimum. QED

2 Existence

2.1 The function space $\mathbf{BV}(\Omega)$

In the search for a suitable function space on which to relax the Mumford-Shah problem, we can draw inspiration from the heuristic description of its internal logic given in the introduction. As stated earlier, this is determined by the competition between the strength of the oscillations and the length of the discontinuity set of a candidate solution. A natural solution space, then, would be one whose members are essentially *characterised* by their ability to exhibit both kinds of behaviour: smooth variation on most of their domain, and jump discontinuities along a transition set. This leads us to the following definition.

Definition 2.1. For Ω an open subset of $\mathbb{R}^n$, we say that a function $u \colon \Omega \to \mathbb{R}$ in $L^1(\Omega)$ is *of bounded variation* if its distributional derivative Du belongs to the space $\mathcal{M}(\Omega; \mathbb{R}^n)$ of vector-valued finite Radon measures, i.e.

$$\int_\Omega u(x) \mathrm{div} \phi(x) \, dx = -\int_\Omega \langle \phi(x), Du \rangle = -\sum_{k=1}^{n} \int_\Omega \phi_k(x) \, dD_k u(x)$$

for every ϕ in $C_c^\infty(\Omega; \mathbb{R}^n)$. We denote the space of such functions by $BV(\Omega)$.

Intuitively, $BV(\Omega)$ contains functions whose oscillation is allowed to be even more irregular than functions in the Sobolev space $W^{1,1}(\Omega)$[1]; in fact, the latter forms a subset of $BV(\Omega)$. This is expressed in the weaker requirement on the distributional derivative, as measures can have point masses and other forms of concentration, allowing for more rapid variation. Indeed, for any function $u \in L^1(\Omega)$, we define the *total variation* of u as

$$V(u,\Omega) \coloneqq \sup\left\{ \int_\Omega u(x)\mathrm{div}\phi(x)\,dx \mid \phi \in C_c^\infty(\Omega;\mathbb{R}^n)\,, \|\phi\|_\infty \le 1 \right\}.$$

We can then equivalently characterise $BV(\Omega)$ as the set of functions in $L^1(\Omega)$ for which $V(u,\Omega) < +\infty$, as we have

$$\left| \int_\Omega u(x)\mathrm{div}\phi(x)\,dx \right| \le V(u,\Omega)\|\phi\|_\infty\,,$$

from which it follows that $\phi \mapsto \int_\Omega u(x)\mathrm{div}\phi(x)\,dx$ defines a continuous linear functional on $C_c^\infty(\Omega;\mathbb{R}^n)$. Because $C_c^\infty(\Omega;\mathbb{R}^n) \subset C_0(\Omega;\mathbb{R}^n)$ is a linear subspace, the Hahn-Banach theorem allows us to extend this functional to all of $C_0(\Omega;\mathbb{R}^n)$, and the Riesz representation theorem then implies that it defines a finite Radon measure. Conversely, if $u \in BV(\Omega)$, then

$$\begin{aligned} |\int_\Omega u(x)\mathrm{div}\phi(x)\,dx| &\le \int_\Omega |\langle \phi, Du\rangle|\,dx \\ &\le \|\phi\|_\infty |Du|(\Omega) < \infty\,, \end{aligned}$$

so that $V(u,\Omega)$ is finite.

It can be shown that $BV(\Omega)$ is a Banach space when equipped with the norm

$$\|u\|_{BV(\Omega)} \coloneqq \|u\|_{L^1(\Omega)} + |Du|(\Omega)\,.$$

The norm topology is too strong for our applications, however, and so we will not make use of it in the sequel. For our purposes, the following topology will be prove to be much more interesting.

Definition 2.2. A sequence $(u_n)_n$ *converges weakly-** to $u \in BV(\Omega)$ (notation: $u_n \overset{*}{\rightharpoonup} u$) if $(u_n)_n$ converges to u in $L^1(\Omega)$ and $Du_n \overset{*}{\rightharpoonup} Du$.

We remark in passing that $BV(\Omega)$ can be shown to correspond to the dual of a certain space such that this definition coincides with the ordinary notion of weak-* convergence induced by its predual. A more useful characterisation of weak-* convergence is given by the following result.

[1] The Sobolev spaces $W^{k,p}$ consist of L^p-functions whose distributional derivatives of order $\le k$ are in L^p as well.

Theorem 2.3. *A sequence $(u_k)_k \subset BV(\Omega)$ converges weakly-* to $u \in BV(\Omega)$ if and only if $u_k \xrightarrow{L^1} u$ and $(u_k)_k$ is uniformly bounded in BV-norm.*

The usefulness of the weak-* topology on $BV(\Omega)$ for minimisation problems stems from the following compactness property, which simultaneously strengthens our confidence in the suitability of $BV(\Omega)$ as a solution space.

Theorem 2.4. *If Ω is a bounded open subset of $\mathbb{R}^n$, then any sequence $(u_n)_n$ in $BV(\Omega)$ satisfying*

$$\sup_{n\geq 1}\left\{\int_A |u_n|\,dx + |Du_n|(A)\right\} < +\infty \quad \forall A \Subset \Omega \text{ open}$$

admits a subsequence $(u_{n_k})_k$ converging in $L^1(\Omega)$ to some u in $BV(\Omega)$. Moreover, if the boundary $\partial\Omega$ of the domain is smooth and $(u_n)_n$ is bounded in $BV(\Omega)$, then $u \in BV(\Omega)$ and $u_{n_k} \overset{*}{\rightharpoonup} u$ in $BV(\Omega)$.

In order to exploit these properties for our existence theory, we will first need to establish a special representation for $BV(\Omega)$. This is the object of the next section.

2.2 Structure theorem for BV(Ω)

The defining property of $BV(\Omega)$ turns out to induce a rich structure on its members, which is not only useful for applications, but also helps us gain greater insight into the space. This can be most easily illustrated in the univariate case, i.e. with Ω an open subset of $\mathbb{R}$, which we will assume to be an interval (a, b) for simplicity. Here we can find for any element of $BV(\Omega)$ a member of its equivalence class expressing this special structure.

Definition 2.5. For any $u \in BV(\Omega)$, we call a member $\tilde{u}$ of this equivalence class a *good representative* if its *pointwise variation*

$$\sup\left\{\sum_{i=1}^{n-1}|\tilde{u}(t_{i+1}) - \tilde{u}(t_i)| \;\middle|\; n \geq 2\,, \quad a < t_1 < \cdots < t_n < b\right\}$$

is equal to the total variation $V(u, \Omega)$.

Intuitively, a good representative translates the oscillation of its equivalence class, which is allowed to vary arbitrarily on null sets, into a pointwise oscillation. Such a function could not have removable singularities, and we expect it to be "mostly continuous" with occasional jumps where Du has a point mass. In general, a good representative can be decomposed into a càdlàg function whose oscillation "lives" on discrete points, and a continuous function which is a.e. differentiable.

Theorem 2.6. *Let $\tilde{u}$ be a good representative of any $u \in BV(\Omega)$, and define the set of* atoms *of u as*

$$A := \{y \in \Omega \mid Du(\{y\}) \neq 0\} .$$

Then we have the following properties:

(i) For every $x \in A$, the function $\tilde{u}$ has a jump discontinuity at x, i.e. the left and right limits exist and

$$\lim_{y \downarrow x} \tilde{u}(y) \neq \lim_{y \uparrow x} \tilde{u}(y) .$$

(ii) $\tilde{u}$ is continuous in $\Omega \setminus A$.

(iii) $\tilde{u}$ is differentiable at $\lambda-$a.e. point of Ω, where λ denotes the one-dimensional Lebesgue measure.

We can use the structure of a good representative, as expressed by its decomposition, as a guide to analyse the distributional derivative of its equivalence. The pointwise decomposition of $\tilde{u}$ turns out to have a counterpart in Du. To show this, we will need some results from measure theory.

Theorem 2.7 (Lebesgue-Radon-Nikdoym)**.** *Let $(X, \mathcal{F})$ be a measurable space, and consider a finite measure ν and a σ-finite positive measure μ on this space. Then there exist μ-a.e. unique measures ν^{ac} and ν^s, with $\nu^{ac} \ll \mu$ and $\nu^s \perp \mu$, such that $\nu = \nu^{ac} + \nu^s$. Moreover, there exists a μ-a.e. unique function $f \in L^1(X)$ such that $\nu^{ac} = f\mu$. We denote this f by $\frac{d\nu^{ac}}{d\mu}$ and call it the* Radon-Nikodym derivative *of ν^{ac} with respect to μ.*

In the special case where $(X, \mathcal{F}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and μ and ν are Radon measures playing nicely with the underlying topology, we can even give an explicit characterisation of $\frac{d\nu^{ac}}{d\mu}$ as well as the singular part.

Theorem 2.8 (Besicovitch)**.** *Let μ and ν be two Radon measures on $\mathbb{R}^n$, then the limit*

$$f(x) := \lim_{r \downarrow 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}$$

exists for μ-a.e. $x \in \mathbb{R}^n$ and $f \in L^1_{loc}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$. Moreover, if we set

$$S := \left\{ x \in \operatorname{supp}(\mu) \mid \lim_{r \downarrow 0} \frac{|\nu|(B_r(x))}{\mu(B_r(x))} = +\infty \right\} \cup \mathbb{R}^n \setminus \operatorname{supp}(\mu) ,$$

then $\nu = f\mu + \nu^s$, where $\nu^s = \nu \llcorner S$ denotes the restriction of ν to S.

Combining these theorems, we can now give a constructive decomposition of any σ-finite measure μ on Ω with respect to the Lebesgue measure λ. Writing $\mu = \mu^{ac} + \mu^s$ and denoting the set of atoms of ν by A once again, Besicovitch's theorem then implies

$$A \subseteq \left\{ x \in \Omega \mid \lim_{r \downarrow 0} \frac{|\mu|(B_r(x))}{r} = +\infty \right\} =: S\,,$$

so that μ^s can be further decomposed as the sum of the mutually orthogonal measures

$$\mu^j := \mu \llcorner A \qquad \mu^c := \mu \llcorner (S \setminus A)\,.$$

The measures μ^j, μ^c and μ^{ac} are called the *jump*, *Cantor* and *absolutely continuous* parts of μ respectively.

Applying the above to the distributional derivative of any $u \in BV(\Omega)$, the following decomposition theorem can be shown to hold.

Theorem 2.9. *Every $u \in BV(\Omega)$ has a representation as a sum $u^{ac} + u^j + u^c$ with $u^{ac} \in W^{1,1}(\Omega)$, u^j a jump function and u^c a Cantor function. The variation of u has the corresponding decomposition*

$$\begin{aligned} |Du|(\Omega) &= |Du^{ac}|(\Omega) + |Du^j|(\Omega) + |Du^c|(\Omega) \\ &= \int_{\Omega \setminus S} |\tilde{u}'|\ dx + \sum_{x \in A} |\tilde{u}(x^+) - \tilde{u}(x^-)| + |D\tilde{u}^c|(\Omega)\,, \end{aligned}$$

where $\tilde{u}$ is any good representative of u.

The natural next step would now be to extend this representation to the two-dimensional case of interest to us. Unfortunately, it turns out that this cannot be done in general. For any open set $\Omega \subseteq \mathbb{R}^2$, we can find functions in $BV(\Omega)$ for which every representative is discontinuous on a set of strictly positive Lebesgue measure, and thus it follows from 2.6(iii) that they cannot have a "good representative" in the sense we've seen so far. This kind of strange behaviour turns out to be a staple of $BV(\Omega)$ in higher dimensions: the added flexibility also creates room for some very pathological constructions. To illustrate this, we must first introduce a few concepts from geometric measure theory.

Definition 2.10. For $\Omega \underset{\text{open}}{\subseteq} \mathbb{R}^n$, the *perimeter* of a measurable set $E \subseteq \mathbb{R}^n$ in Ω is defined as

$$P(E, \Omega) := \sup_{\substack{\phi \in C_c^1(\Omega;\mathbb{R}^n) \\ \|\phi\| \leq 1}} \int_E \operatorname{div}\phi\ dx\,.$$

E is said to be of *finite perimeter* in Ω if $P(E, \Omega) < +\infty$.

Sets of finite perimeter are intimately linked to BV-functions. This can be seen from the equality $P(E, \Omega) = V(I_E, \Omega) = |DI_E|(\Omega)$, which, for sets with $\lambda(E \cap \Omega)$, implies that E is of finite perimeter in Ω if and only if $I_E \in BV(\Omega)$. More generally, the indicator function of a set of finite perimeter E will belong to the space of functions of *locally bounded variation*

$$BV_{\text{loc}}(\Omega) := \left\{ u \in L^1_{\text{loc}}(\Omega) \mid V(u, U) < +\infty \ , \forall U \underset{open}{\Subset} \Omega \right\} .$$

Conversely, if $I_E \in BV_{\text{loc}}(\Omega)$, then E has finite perimeter in every $U \underset{open}{\Subset} \Omega$; such sets are said to have *locally finite perimeter*. It can be shown that perimeter is subadditive and lower semicontinuous with respect to so-called *local convergence in measure* of sets:

$$\lambda(A \cap (E_n \Delta E)) \xrightarrow{n\to\infty} 0 \quad \forall A \Subset \Omega \Longrightarrow P(E, \Omega) \le \liminf_{n\to\infty} P(E_n, \Omega) \, ,$$

for any sequence $(E_n)_n$ of measurable sets.

Sets of locally finite perimeter are also called *Caccioppoli sets*. They include all sets with a C^1 boundary of finite length (in which case $P(E, U) = \mathcal{H}^{n-1}(\partial E \cap U) \quad \forall U \Subset \Omega$), but can also provide examples of the pathological behaviour we discussed earlier.

Example. Let $(q_k)_k$ be an enumeration of $\mathbb{Q}^n$ and consider, for any $\epsilon > 0$, a sequence of radii $(r_k)_k \subset (0, \epsilon)$ such that $\sum_{k=1}^{\infty} n\omega_n r_k^{n-1} \le 1$, where ω_l denotes the volume (Lebesgue measure) of the l-dimensional unit ball. If we define $A_n := \bigcup_{k=1}^{n} B_{r_k}(q_k)$, then this sequence converges to the union $\bigcup_{k=1}^{\infty} A_n =: A$ in measure. From the formulas for the areas and volumes of hyperspheres, we now obtain

$$P(A, \mathbb{R}^n) \le \liminf_{n\to\infty} P(A_n, \mathbb{R}^n) \le \sum_{k=1}^{\infty} P(B_{r_k}(q_k), \mathbb{R}^n) \le 1$$

and

$$\lambda(A) \le \sum_{k=1}^{\infty} \omega_n r_k^n \le \epsilon \, .$$

As $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$, we have $\text{cl}(A) = \mathbb{R}^n$, so that $\mathbb{R}^n \setminus A \subset \partial A$ and $\lambda(\partial A) = +\infty$. Thus, the indicator function I_A, being integrable and of finite variation, belongs to $BV(\Omega)$. However, its set of discontinuity points, the topological boundary ∂A, has non-vanishing Lebesgue measure. Moreover, if I'_A is a representative of I_A which is the indicator function of a measurable set, then A' will be equivalent to A, i.e. $\lambda(A \triangle A') = 0$. It thus follows that A' is dense in $\mathbb{R}^n$, and its interior will be contained in that of A, up to Lebesgue null sets. This suggests that I_A has no continuous representative. ♣

Despite this negative result, it is still possible to recover a structure theorem under certain conditions. The key is to relax our notions of continuity and differentiability by disregarding not only sets of zero measure, but also those of zero *density.*

Definition 2.11. A measurable subset $E \subseteq \Omega$ is said to have density $t \in [0,1]$ at a point $x \in \Omega$ if

$$\lim_{r\downarrow 0} \frac{\lambda(E \cap B_r(x))}{\lambda(B_r(x))} = t\,.$$

We denote the set of points at which E has density t by E^t. The *essential boundary* $\partial^* E$ of E consists of the points in Ω whose at which the density of E is neither 0 nor 1, i.e. $\partial^* E \coloneqq \Omega \setminus (E^0 \cap E^1)$.

The sets E^0 and E^1 represent the measure-theoretical analogue of the interior and exterior of E. Intuitively, $\partial^* E$ further excludes the "lower-dimensional pieces" of E from the topological boundary. We can refine this notion further by restricting attention to the boundary points of at which a measure-theoretical counterpart of a normal vector field can be defined.

Definition 2.12. For a Caccioppoli set $E \subseteq \Omega$, we define *reduced boundary* $\mathcal{F}E$ as the set of points $x \in \Omega$ such that the limit

$$\nu(x) \coloneqq \lim_{r\downarrow 0} \frac{DI_E(B_r(x))}{|DI_E|(B_r(x))}$$

exists and has unit length. We call ν the *generalised inner normal* to E.

We can derive a decomposition for higher-dimensional BV-functions similar to Theorem 2.9, but this will only hold at the level of the distributional derivative and will use weakened notions of regularity. These new "approximate" conditions are obtained through local averaging of the classical ones, and as such are unaffected by the behaviour of functions on zero-density sets. Excluding these is sufficient to eliminate pathological examples such as the one we saw earlier.

Definition 2.13. We say that a function $u \in L^1_{\text{loc}}(\Omega)$ has an *approximate limit* at $x \in \Omega$ if there exists a $z \in \mathbb{R}$ such that

$$\lim_{r\downarrow 0} \fint_{B_r(x)} |u(y) - z| \; dy = 0\,,$$

where the notation $\fint_{B_r(x)}$ is shorthand for the average $\frac{1}{\lambda(B_r(x))} \int_{B_r(x)}$. In that case we write $z = \tilde{u}(x)$. The set of points where no approximate limit exists for u is called its *approximate discontinuity set*, denoted by S_u. If the approximate limit of u at a given point coincides with its value, i.e. x is a Lebesgue point of u, we say that u is *approximately continuous* at this point. Finally, we call u *approximately differentiable* at $x \in \Omega \setminus S_u$ if there exists a vector $z \in \mathbb{R}^n$ such that

$$\lim_{r \downarrow 0} \fint_{B_r(x)} \frac{|u(y) - \tilde{u}(x) - z \cdot (y - x)|}{r} \, dy = 0 \,,$$

in which case z is called the *approximate differential* of u at x, which we write as $\nabla u(x)$.

To obtain the higher-dimensional structure theorem for $BV(\Omega)$, we now follow the same logic as in the univariate case, distinguishing different components of a function based on their oscillation. We begin by defining the counterpart for the atoms, which describe the discrete jumps of the function.

Definition 2.14. A point $x \in \Omega$ is called an *approximate jump point* of u if there exist $a \neq b$ in $\mathbb{R}$ and a unit vector $\nu \in S^{n-1}$ such that

$$\lim_{r \downarrow 0} \fint_{B_r^+(x,\nu)} |u(y) - a| \, dy = 0 \qquad \lim_{r \downarrow 0} \fint_{B_r^-(x,\nu)} |u(y) - b| \, dy = 0 \,,$$

where we integrate over the sets

$$B_r^+(x,\nu) \coloneqq \{y \in B_r(x) \mid \langle y - x, \nu \rangle > 0\}$$
$$B_r^-(x,\nu) \coloneqq \{y \in B_r(x) \mid \langle y - x, \nu \rangle < 0\} \,.$$

The set of approximate jump points of u is denoted by J_u.

It can be shown that (a, b, ν) as in the above definition can be chosen at every approximate jump point of a function $u \in BV(\Omega)$ in such a way as to define a Borel-measurable mapping $J_u \to \mathbb{R} \times \mathbb{R} \times S^1 : x \mapsto (u^+(x), u^-(x), \nu_u(x)) \coloneqq (a, b, \nu)$, giving approximate right and left limits of u at x in the direction determined by $\nu(x)$.

For any function $u \in BV(\Omega)$, it is clearly the case that $J_u \subseteq S_u$. The converse turns out to almost be true as well: $\mathcal{H}^{n-1}$-a.e. point of S_u is an approximate jump point. These points therefore represent, in a way, the quintessential singularities for BV-functions. This important fact is expressed by the following theorem, which also allows us to obtain a concrete representation of the distributional derivative Du on the jump set.

Definition 2.15. A set $E \subset \mathbb{R}^n$ is called countably $\mathcal{H}^m$-rectifiable if there exists a countable family of Lipschitz-mappings $f_k : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\mathcal{H}^m(E \setminus \bigcup_{k \geq 1} f_k(\mathbb{R}^m)) = 0 \,.$$

Theorem 2.16 (Federer-Vol'pert). *For $u \in BV(\Omega)$, the approximate discontinuity set S_u is countably $\mathcal{H}^{n-1}$-rectifiable and $\mathcal{H}^{n-1}(S_u \setminus J_u) = 0$. Moreover, the distributional derivative can be represented on the jump set as*

$$Du \llcorner J_u = (u^+ - u^-)\, \nu_u \, \mathcal{H}^{n-1} \llcorner J_u \, .$$

We will not endeavour to demonstrate this fact here. The proof necessitates the introduction of formal machinery from geometric measure theory which cannot be feasibly covered within the scope of the present exposition. The interested reader my find both proof and background in [1].

With this characterisation of the jump behaviour of BV-functions in hand, we next consider to the absolutely continuous component of Du. Here, the approximate differential will turn out to play the same role as the a.e.-derivative of a good representative in the 1-dimensional case.

Lemma 2.17. *For $u \in BV(B_r(x))$ having an approximate limit at x, we have the inequality*

$$\int_{B_r(x)} \frac{|u(y) - \tilde{u}(x)|}{|y - x|} \, dy \leq \int_0^1 \frac{|Du|(B_{tr}(x))}{t^n} \, dt \, .$$

Proof. Without loss of generality, we can take $x = 0$. If u is a smooth function, then for any $y \in \Omega$ and $\rho \in (0, 1)$ we have

$$u(y) - u(\rho y) = \int_\rho^1 \nabla u(ty) \cdot y \, dt \, ,$$

whence the inequality

$$\frac{|u(y) - u(y\rho)|}{\|y\|} \leq \int_\rho^1 \|\nabla u(ty)\| \, dt \, .$$

Integrating both sides and using Fubini's theorem, we obtain

$$\int_{B_r} \frac{|u(y) - u(y\rho)|}{\|y\|} \, dy \leq \int_\rho^1 \int_{B_r} \|\nabla u(ty)\| \, dy \, dt = \int_\rho^1 \frac{|Du|(B_{tr})}{t^n} \, dt \, .$$

Choosing a mollifier ϕ and setting $u_\epsilon := \phi_\epsilon * u$, it can be shown that the previous inequality generalizes to any $u \in BV(\Omega)$. By assumption, $0 \notin S_u$, and therefore we have

$$\begin{aligned} \lim_{\rho \downarrow 0} \|u(\rho y) - \tilde{u}(0)\|_{L^1(\Omega)} &= \lim_{\rho \downarrow 0} \int_{B_r} |u(\rho y) - \tilde{u}(0)| \, dy \\ &= \lim_{\rho \downarrow 0} \frac{1}{\rho^n} \int_{B_{\rho r}} |u(z) - \tilde{u}(0)| \, dz \\ &= 0 \, . \end{aligned}$$

It follows from standard measure theory that we can find a sequence $(\rho_k)_k \subset (0,1)$ such that $u(\rho_k y) \xrightarrow{k\to\infty} \tilde{u}(0)$ almost everywhere. Applying Fatou's lemma, we finally obtain

$$\begin{aligned}\int_{B_r} \frac{|u(y)-\tilde{u}(0)|}{\|y\|}\,dy &\leq \liminf_{k\to\infty} \int_{B_r} \frac{|u(y)-u(y\rho_k)|}{\|y\|}\,dy \\ &\leq \liminf_{k\to\infty} \int_{\rho_k}^1 \frac{|Du|(B_{tr})}{t^n}\,dt \\ &\leq \int_0^1 \frac{|Du|(B_{tr})}{t^n}\,dt\,,\end{aligned}$$

which was the desired inequality. QED

Theorem 2.18 (Calderon-Zygmund)**.** *Any $u \in BV(\Omega)$ is approximately differentiable $\lambda - a.e.$ and $Du^{ac} = \nabla u \lambda$.*

Proof. By Radon-Nikodym, we can decompose Du with respect to the Lebesgue measure as the sum $D^{ac}u + D^s u$. Writing v for the density of $D^{ac}u$, we can consider for any $x_0 \in \Omega \setminus (S_u \cup S_v)$ the function

$$w(x) \coloneqq u(x) - \tilde{u}(x_0) - \tilde{v}(x_0)\cdot(x-x_0).$$

Using the linearity of the distributional derivative, we immediately obtain

$$\begin{aligned}Dw &= Du - \tilde{v}(x_0)\cdot D(x-x_0) \\ &= D^s u(x) + v(x)\lambda - \tilde{v}(x_0)\lambda \\ &= D^s u(x) + (v(x) - \tilde{v}(x_0))\lambda\,.\end{aligned}$$

Applying the previous lemma then yields the inequality

$$\begin{aligned}\frac{1}{r^n}\int_{B_r(x_0)} \frac{|u(x)-\tilde{u}(x_0)-\tilde{v}(x_0)\cdot(x-x_0)|}{|x-x_0|}\,dx &\leq \sup_{t\in(0,1)} \frac{|Dw|(B_{tr}(x_0))}{(tr)^n} \\ &= \sup_{t\in(0,1)} \frac{1}{(tr)^n}\left(\int_{B_{tr}(x_0)} |v(x)-\tilde{v}(x_0)|\,dx + |D^s u|(B_{tr}(x_0))\right).\end{aligned}$$

By Theorem 2.8, $|D^s u|(B_r(x_0)) = o(r^n)$ for λ-a.e point $x_0 \in \Omega$. As $S_u \cup S_v$ is a λ-null set, it follows from passing to the limit $r \downarrow 0$ in the previous inequality that u is approximately differentiable λ-a.e. and $\nabla u(x_0) = v(x)$. QED

Using these results, we are finally ready to derive the structure theorem. Using the Radon-Nikodym theorem once again, we obtain Du as the sum of measures $D^{ac}u$ and $D^s u$. We can now define

$$Du^j \coloneqq Du^s \llcorner J_u \qquad Du^c \coloneqq Du^s \llcorner \Omega \setminus S_u\,,$$

respectively the *jump* and *Cantor* parts of the distributional derivative. It can be shown that Du vanishes on $S_u \setminus J_u$, and these measures together therefore completely exhaust Du. Using Theorem 2.18 to characterise $D^{ac}u$, we finally obtain the following decomposition.

Theorem 2.19. *For any function $u \in BV(\Omega)$, the distributional derivative Du can be decomposed as the sum*

$$\begin{aligned} Du &= D^{ac}u + D^j u + D^c u \\ &= \nabla u \lambda + (u^+ - u^-)\nu \mathcal{H}^{n-1} + Du^c\,, \end{aligned}$$

with $D^{ac}u \perp D^j u \perp D^c u$.

Finally, we mention a result known, in analogy to its classical counterpart, as the *chain rule* in BV, which will prove to be useful later.

Theorem 2.20. *If $u \in BV(\Omega)$ and $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz function additionally satisfying $f(0) = 0$ if $\lambda(\Omega) = +\infty$. Then $v := f \circ u$ again belongs to $BV(\Omega)$ and*

$$Dv = f'(u)\nabla u \lambda + (f(u^+) - f(u^-))\, \nu_u\, \mathcal{H}^{n-1} \llcorner J_u + f'(\tilde{u}) D^c u\,.$$

2.3 The function space SBV(Ω)

In order to show our desired existence result, we will ultimately need to further restrict ourselves to a subclass of $BV(\Omega)$, the so-called *special functions of bounded variation*. These are BV-functions u for which the Cantor part $D^c u$ of the distributional derivative vanishes in the decomposition from Theorem 2.19:

$$Du = \nabla u \lambda + (u^+ - u^-)\nu_u \mathcal{H}^{n-1} \llcorner J_u\,.$$

The space of such functions is denoted by $SBV(\Omega)$, and this turns out to be the right ambient space in which to apply the direct method for a class of optimisation problems where both surface and volume energies are involved, of which the Mumford-Shah functional is a prime example. In order to leverage this, we must first show that this space possesses the requisite properties to guarantee the existence of a minimising sequence. This is the object of the present section. We begin by deriving a characterisation of $SBV(\Omega)$ which will prove to be useful in the sequel.

Lemma 2.21. *Let $\theta : (0, +\infty) \to (0, +\infty)$ be an increasing function satisfying $\theta(t)/t \xrightarrow[t \to 0]{} +\infty$ and define*

$$\|\psi\|_\theta := \sup \left\{ \frac{|\psi(s) - \psi(t)|}{\theta(|t-s|)} \mid s, t \in \mathbb{R}, s \neq t \right\},$$

for any function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$. *Additionally, consider a function* $\gamma \in W^{1,\infty}(\mathbb{R})$ *and, rescaling and shifting its argument by* $r > 0$ *and* $a \in \mathbb{R}$ *respectively, write* $\psi_r(t) := \gamma(\frac{t-a}{r})$. *Then we have*

$$\lim_{r \downarrow 0} r\|\psi_r\|_\theta = 0 .$$

Proof. Writing out the last expression, we obtain

$$r\|\psi_r\|_\theta = r \sup_{s \neq t} \frac{|\psi_r(s) - \psi_r(t)|}{\theta(|t-s|)} = \sup_{s' \neq t'} \frac{r|\gamma(s') - \gamma(t')|}{\theta(r|t'-s'|)} ,$$

where we used the substitutions $t' := (t-a)/r$ and $s' := (s-a)/r$. If $|s'-t'| \geq 1$, we can use the boundedness of γ and the monotonicity of θ to obtain an upper estimate for this expression. The same can be done for $|s'-t'| < 1$ by using the asymptotic property of θ and the boundedness of γ'. Thus we obtain

$$\begin{aligned} r\|\psi_r\|_\theta &\leq \sup_{|s'-t'| \geq 1} \frac{2r}{\theta(r|s'-t'|)} \|\gamma\|_\infty + \sup_{|s'-t'|<1} \frac{r}{\theta(r|s'-t'|)} \int_{t'}^{s'} |\gamma'(x)|\, dx \\ &\leq \frac{2r}{\theta(r)} \|\gamma\|_\infty + \sup_{\tau \in (0,1)} \frac{r\tau}{\theta(r\tau)} \|\gamma'\|_\infty , \end{aligned}$$

which goes to 0 as $r \downarrow 0$. QED

Now, if we know that $u \in SBV(\Omega)$, then, for any Lipschitz function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$, we can use the BV-chain rule (Theorem 2.20) to write

$$\begin{aligned} D(\psi \circ u) &= D^{ac}(\psi \circ u) + D^c(\psi \circ u) + D^j(\psi \circ u) \\ &= \psi'(u)\nabla u \lambda + (\psi(u^+) - \psi(u^-))\nu_u \mathcal{H}^{n-1} \llcorner J_u . \end{aligned}$$

We thus obtain the following bound on the variation of $D\psi - \psi'(u)\nabla u$:

$$|D\psi(u) - \psi'(u)\nabla u \lambda| \leq \|\psi\|_\theta \, \theta(|u^+ - u^-|)\mathcal{H}^{n-1} \llcorner J_u .$$

It turns out that this condition is sufficient as well.

Theorem 2.22 (Characterisation of SBV)**.** *Let* $\Omega \subseteq \mathbb{R}^n$ *be open and bounded and consider a function* u *in* $BV(\Omega)$. *If* $\theta : (0,+\infty) \to (0,+\infty)$ *is an increasing function such that* $\theta(t)/t \underset{t \to 0}{\longrightarrow} +\infty$ *and there exist* $a \in L^1(\Omega, \mathbb{R}^n)$ *and* μ *a finite positive measure with* $\operatorname{supp} \mu \subseteq \Omega$ *such that*

$$|D\psi(u) - \psi'(u)a\lambda| \leq \|\psi\|_\theta \mu \tag{2}$$

for every $\psi \in W^{1,\infty}(\mathbb{R}) \cap C^1(\mathbb{R})$, *then* $u \in SBV(\Omega)$, *and moreover* $a = \nabla u$ λ*-a.e.*

Proof. First we will show that ∇u and a coincide λ-a.e. Let $x_0 \in \Omega$ be a Lebesgue point of both a and u for which, additionally, u is approximately differentiable and $\lim_{r\downarrow 0} \mu(B_r(x_0))/r^n < +\infty$. As a and u are in $L^1(\Omega)$, λ-a.e. point of Ω will be a Lebesgue point of both. Combining this with Theorem 2.18 and Theorem 2.8, we can conclude that λ-a.e. point of Ω possesses the aforementioned properties, so that it suffices to show $\nabla u(x_0) = a(x_0)$.

To do this, we set $u_0(y) := \langle \nabla u(x_0), y\rangle$, and consider functions $\gamma \in C_c^1(\mathbb{R})$ and $\phi \in C_c^1(B_1)$, where B_1 denotes the open unit ball in $\mathbb{R}^n$, γ coincides with the identity on $u_0(B_1)$ and ϕ is positive and non-zero. Rescaling the arguments of both, we additionally define

$$\psi_r(t) := \gamma\left(\frac{t-u(x_0)}{r}\right) \qquad \xi_r(x) := \phi\left(\frac{x-x_0}{r}\right).$$

Now consider the pairing between ψ_r and $\nabla\xi_r$. Using a change of variables, we can write this as

$$\begin{aligned}\int_\Omega \psi_r(u)\nabla\xi_r dx &= \frac{1}{r}\int_\Omega \gamma\left(\frac{u(x)-u(x_0)}{r}\right)\nabla\phi\left(\frac{x-x_0}{r}\right)dx \\ &= r^{n-1}\int_{B_1}\gamma(u_r(y))\nabla\phi(y)dy\,,\end{aligned}$$

where we introduced the notation $u_r(y) := (u(x_0+ry)-u(x_0))/r$. On the other hand, we also have

$$\begin{aligned}\int_\Omega \psi_r(u)\nabla\xi_r dx &= -\int_\Omega \xi_r dD\psi_r(u) \\ &= -\int_\Omega \xi_r d\left[D\psi_r(u) - a\psi_r'(u)\lambda\right] - \int_\Omega \xi_r a\psi_r'(u)dx\end{aligned}$$

We can now use the previous lemma to obtain

$$\lim_{r\downarrow 0}\frac{\|\psi_r\|_\theta\,\mu(B_r(x_0))}{r^{n-1}} = \lim_{r\downarrow 0} r\|\psi_r\|_\theta\frac{\mu(B_r(x_0))}{r^n} = 0\,,$$

from which it follows, using (2), that

$$\left|\int_\Omega \xi_r d\left[D\psi_r(u) - a\psi_r'(u)\lambda\right]\right| \le \|\phi\|_\infty\|\psi_r\|_\theta\mu(B_r(x_0)) = o(r^{n-1})\,.$$

Thus, we arrive at the final estimate

$$\int_\Omega \psi_r(u)\nabla\xi_r dx = o(r^{n-1}) - r^{n-1}\int_{B_1}\phi(y)a(x_0+ry)\gamma'(u_r(y))dy\,,$$

and combining both expressions, we get

$$\begin{aligned}\int_{B_1} \gamma(u_r(y))\nabla\phi(y)dy &= -\int_{B_1} \phi(y)a(x_0+ry)\gamma'(u_r(y))dy + o(1)\\ &= -\int_{B_1} \phi(y)a(x_0)\gamma'(u_r(y))dy + o(1)\,.\end{aligned}$$

The last equality can be obtained by writing

$$\begin{aligned}&\int_{B_1} \phi(y)a(x_0+ry)\gamma'(u_r(y))dy\\ &= \int_{B_1} \phi(y)a(x_0)\gamma'(u_r(y))\,dy + \int_{B_1} \phi(y)[a(x_0+ry)-a(x_0)]\gamma'(u_r(y))\,dy\end{aligned}$$

which, by virtue of x_0 being a Lebesgue point of a, combined with the boundedness of ϕ and γ', yields

$$\begin{aligned}&\left|\int_{B_1} \phi(y)[a(x_0+ry)-a(x_0)]\gamma'(u_r(y))\,dy\right|\\ &\leq \|\phi\|_\infty\|\gamma'\|_\infty \int_{B_1} |a(x_0+ry)-a(x_0)|\,dy\\ &= \|\phi\|_\infty\|\gamma'\|_\infty\omega_n \fint_{B_r(x_0)} |a(z)-a(x_0)|\,dz \xrightarrow{r\downarrow 0} 0\,.\end{aligned}$$

Noticing that we have convergence of u_r to u_0 in $L^1(B_1)$ as $r \downarrow 0$:

$$\begin{aligned}\int_{B_1} |u_r(y)-u_0(y)|\,dy &= \int_{B_1} \frac{|u(x_0+ry)-u(x_0)-ru_0(y)|}{r}\,dy\\ &= \omega_n \fint_{B_r(x_0)} \frac{|u(z)-u(x_0)-\nabla u(x_0)\cdot(z-x_0)|}{r}\,dz \xrightarrow{r\downarrow 0} 0\,,\end{aligned}$$

this gives us

$$\int_{B_1} \gamma(u_0(y))\nabla\phi(y)\,dy = -\int_{B_1} \gamma'(u_0(y))\phi(y)a(x_0)\,dy\,.$$

Integrating by parts, we finally obtain

$$[\nabla u(x_0) - a(x_0)]\int_{B_1} \gamma'(u_0(y))\phi(y)dy = 0\,,$$

which implies $a(x_0) = \nabla u(x_0)$, as $\gamma' = 1$ on the range of u_0.

It remains for us to show that $u \in SBV(\Omega)$. By the BV-chain rule, we can write $D^{ac}\psi(u) = \psi'(u)\nabla u\lambda$ for the absolutely continuous part of $D\psi(u)$. From the first part, we then arrive at the bound

$$|D^s\psi(u)| \leq \|\psi\|_\theta \mu$$

for $\psi \in W^{1,\infty}(\mathbb{R}) \cap C^1(\mathbb{R})$. Restricting to $E := \Omega \setminus S_u$, this yields

$$|\psi'(\tilde{u})||D^c u| \leq \|\psi\|_\theta \mu \llcorner E\,.$$

If we now set $\psi^1_\epsilon(t) := \sin(t/\epsilon)$ and $\psi^2_\epsilon(t) := \cos(t/\epsilon)$, then both functions belong to $W^{1,\infty}(\mathbb{R}) \cap C^1(\mathbb{R})$, and consequently

$$\frac{1}{\epsilon}|\sin(\tilde{u}/\epsilon)||D^c u| \leq \|\psi^2_\epsilon\|_\theta \mu \llcorner E \qquad \frac{1}{\epsilon}|\cos(\tilde{u}/\epsilon)||D^c u| \leq \|\psi^1_\epsilon\|_\theta \mu \llcorner E\,.$$

Observing that $|\sin(t)| + |\cos(t)| \geq 1$ for any $t \in \mathbb{R}$, we can write

$$|D^c u| \leq \epsilon \left(\|\psi^1_\epsilon\|_\theta + \|\psi^2_\epsilon\|_\theta\right) \mu \llcorner E\,,$$

where the right term vanishes in the limit $\epsilon \downarrow 0$ by the previous lemma. It follows that $|D^c u| = 0$, and thus $u \in SBV(\Omega)$. QED

We now want to use this criterion to prove a closure property for $SBV(\Omega)$ under certain assumptions. For this we still need a few additional results, which we mention without proof.

Theorem 2.23. *(Weak* compactness of Radon measures) Every bounded sequence of finite radon measures $(\mu_k)_k$ on a locally compact separable metric space X has a weak-* convergent subsequence: if we have*

$$\sup_{k\geq 1} |\mu_k|(X) < +\infty$$

then there exists a finite radon measure μ and a subsequence $(\mu_{k_n})_n$ such that $\mu_{k_n} \overset{*}{\rightharpoonup} \mu$.

Theorem 2.24. *(Dunford-Pettis) A bounded class of integrable functions $\mathcal{F} \subset L^1$ is uniformly integrable if and only if it is relatively weakly sequentially compact.*

Theorem 2.25. *(de la Vallée Poussin) A bounded family $\mathcal{F} \subset L^1(X,\mu)$ of integrable functions on a finite measure space is uniformly integrable if and only if there exists an increasing function $\phi : (0,+\infty) \to (0,+\infty)$ satisfying $\phi(t)/t \xrightarrow{t\to+\infty} +\infty$ such that*

$$\sup_{f\in\mathcal{F}} \int_X \phi(|f|)d\mu < +\infty\,.$$

Lemma 2.26. *Let $(\mu_k)_k \subseteq \mathcal{M}(X;\mathbb{R}^n)$ be a sequence of Radon measures on a locally compact separable metric space X locally weakly-* converging to μ. If the total variation measures $|\mu_k|$ also converge locally weakly* to some positive measure ν, then $\nu \geq |\mu|$.*

We are now ready for the first main result of this section.

Theorem 2.27 (Closure of SBV)**.** *Let Ω be an open and bounded subset of $\mathbb{R}^n$, and suppose $\phi : [0,+\infty) \to [0,+\infty]$ and $\theta : (0,+\infty) \to (0,+\infty]$ are lower semi-continuous increasing functions satisfying*

$$\frac{\phi(t)}{t} \xrightarrow{t\to+\infty} +\infty \qquad \frac{\theta(t)}{t} \xrightarrow{t\to 0} +\infty\,.$$

If $(u_k)_k$ is a sequence in $SBV(\Omega)$ weakly- converging in $BV(\Omega)$ to u, and*

$$\sup_{k\geq 1}\left\{\int_\Omega \phi(\|\nabla u_k\|)\,dx + \int_{J_{u_k}} \theta(|u_k^+ - u_k^-|)\,d\mathcal{H}^{n-1}\right\} < +\infty\,, \tag{3}$$

then $u \in SBV(\Omega)$. Moreover, up to a subsequence, we have weak convergence of the approximate gradients $\nabla u_k \rightharpoonup \nabla u$, weak- convergence of the jump parts of the distributional derivative $D^j u_k \overset{*}{\rightharpoonup} D^j u$ and*

$$\int_\Omega \phi(\|\nabla u\|)\,dx \leq \liminf_{k\to\infty}\int_\Omega \phi(\|\nabla u_k\|)\,dx \quad \textit{if } \phi \textit{ is convex} \tag{4}$$

$$\int_{J_{u_k}} \theta(|u^+ - u^-|)\,d\mathcal{H}^{n-1} \leq \liminf_{k\to\infty}\int_\Omega \theta(|u_k^+ - u_k^-|)\,d\mathcal{H}^{n-1} \quad \textit{if } \theta \textit{ is concave}\,. \tag{5}$$

Proof. We will only prove the first claim; the interested reader may find the others in [1]. Without loss of generality, we may assume that $\theta < +\infty$ everywhere; otherwise we can simply consider the truncation $\theta \wedge 1$. We have to verify that u satisfies (2) in Theorem 2.22. To this end, we begin by showing that there exists $a \in L^1(\Omega)$ such that $\psi'(u_k)\nabla u_k \rightharpoonup \psi'(u)a$ in L^1 for any Lipschitz function $\psi \in C^1(\mathbb{R})$. Observing that $(\nabla u_k)_k$ is bounded in L^1,

$$\sup_{k\geq 1}\int_\Omega \|\nabla u_k\|\,dx \leq \sup_{k\geq 1}|Du_k|(\Omega) \leq \sup_{k\geq 1}\|u_k\|_{BV(\Omega)} < +\infty\,,$$

and that ϕ is asymptotically supralinear for $t \to +\infty$, we can deduce from (3) and Theorem 2.25 that $(\nabla u_k)_k$ is uniformly integrable. It follows from Theorem 2.24 that $(\nabla u_k)_k$ has subsequence which converges weakly to some $a \in L^1(\Omega)$, which, for simplicity, we shall not relabel. Now we write

$$\psi'(u_k)\nabla u_k = [(\psi'(u_k) - \psi'(u))\nabla u_k] + \psi'(u)\nabla u_k\,,$$

and consider the term between brackets. We claim that this goes to 0 in L^1. Fix $\epsilon > 0$, and denote the Lipschitz constant of ψ by K. By uniform integrability of $(\nabla u_k)_k$, there exists $\delta > 0$ such that

$$\sup_{k\geq 1} \int_E \|\nabla u_k\| \, dx \leq \epsilon/4K$$

whenever $\lambda(E) \leq \delta$. From Theorem 2.3, we know that $u_n \overset{*}{\rightharpoonup} u$ implies $u_n \to u$ in $L^1(\Omega)$, from which we can obtain $u_n \to u$ a.e. by pass to an (unrelabeled) subsequence. The continuity of ψ' then gives us $\psi'(u_n) \to \psi'(u)$ a.e., and we can use Egorov's theorem to obtain a measurable subset $G \subseteq \Omega$ such that $\lambda(\Omega \setminus G) \leq \delta$ and $\psi'(u_k) \xrightarrow{L^\infty} \psi'(u)$ on G. Now we can split the integral

$$\int_\Omega |\psi'(u_k) - \psi'(u)| \|\nabla u_k\| \, dx$$

over G and its complement. For the first component, we use the uniform convergence and the boundedness of the L^1-norms of $(\nabla u_k)_k$ to obtain

$$\int_G |\psi'(u_k) - \psi'(u)| \|\nabla u_k\| \, dx \leq \|\psi'(u_k) - \psi'(u)\|_\infty \sup_{k\geq 1} \int_G \|\nabla u_k\| \, dx \,,$$

and thus we have $\int_G |\psi'(u_k) - \psi'(u)| \|\nabla u_k\| \, dx \leq \epsilon/2$ for k sufficiently large. The second component can be estimated by using the boundedness of ψ' and the fact that $\lambda(\Omega \setminus G) \leq \delta$:

$$\int_G |\psi'(u_k) - \psi'(u)| \|\nabla u_k\| \, dx \leq 2K \sup_{k\geq 1} \int_G \|\nabla u_k\| \, dx \leq \epsilon/2$$

Thus we may conclude

$$\lim_{k\to\infty} \int_\Omega \xi \psi'(u_k) \nabla u_k \, dx = \lim_{k\to\infty} \int_\Omega \xi \psi'(u) \nabla u_k \, dx = \int_\Omega \xi \psi'(u) a \, dx \,,$$

for any $\xi \in L^\infty(\Omega)$.

Turning our attention to the distributional gradients $D\psi(u_k)$, we can use the BV-chain rule to obtain the equiboundedness of this sequence:

$$\begin{aligned} \sup_{k\geq 1} |D\psi(u_k)|(\Omega) &\leq \sup_{k\geq 1} \left(\int_\Omega |\psi'(u_k) \nabla u_k| \, dx + \int_{J_{u_k}} |\psi(u_k^+) - \psi(u_k^-)| \, d\mathcal{H}^{n-1} \right) \\ &\leq K \sup_{k\geq 1} \int_\Omega |\nabla u_k| \, dx + \|\psi\|_\theta \sup_{k\geq 1} \int_{J_{u_k}} \theta(|u_k^+ - u_k^-|) \, d\mathcal{H}^{n-1} \\ &< +\infty \,. \end{aligned}$$

Using the L^1 convergence of $\psi(u_k)$ to $\psi(u)$, which follows from

$$\int |\psi(u_k) - \psi(u)|\, dx \leq K \int |u - u_k|\, dx \xrightarrow{k\to+\infty} 0\,.$$

we can conclude from Theorem 2.3 that we have weak-* convergence of $D\psi(u_k)$ to $D\psi(u)$.

From the inclusion $C_0(\Omega) \subseteq C_b(\Omega) \subseteq L^\infty(\Omega)$, it follows that the weak convergence of $\psi'(u_k)\nabla u_k$ to $\psi'(u)a$ in L^1 implies weak-* convergence of the corresponding measures, hence:

$$D\psi(u_k) - \psi'(u_k)\nabla u_k \lambda \overset{*}{\rightharpoonup} D\psi(u) - \psi'(u)\nabla u\lambda$$

Considering the total variations $|D\psi(u_k) - \psi'(u_k)\nabla u_k\lambda|$, we can use Theorem 2.23 to find a weakly* converging subsequence (again keeping the same notation) with limit a positive finite Radon measure σ. Now consider the expression

$$|D\psi(u_k) - \psi'(u_k)\nabla u_k\lambda| \leq \|\psi\|_\theta \mu_k\,,$$

with $\mu_k = \theta(|u_k^+ - u_k^-|)$ as in Theorem 2.22. Using Theorem 2.4 (possibly passing to a subsequence) to obtain a limit μ, and taking account of Lemma 2.26, we can pass to the limit in the previous expression, yielding

$$|D\psi(u) - \psi'(u)\nabla u\lambda| \leq \sigma \leq \|\psi\|_\theta \mu\,.$$

This finally shows, by Theorem 2.22, that $u \in SBV(\Omega)$. QED

Theorem 2.28 (Compactness of SBV)**.** *Under the same assumptions as Theorem 2.27, let $(u_k)_k$ be a sequence in $SBV(\Omega)$ additionally satisfying uniform boundedness in k, i.e. there exists a constant $L > 0$ such that $\sup_{k\geq 1}\|u_k\|_\infty < L$. Then there exists a subsequence $(u_{k_n})_n$ weakly-* converging in $BV(\Omega)$ to some $u \in SBV(\Omega)$.*

Proof. Writing M for $\sup_{k\geq 1}\|u_k\|_\infty$, we can find constants $\alpha \in \mathbb{R}$ and $\beta \in (0, +\infty)$ such that

$$\phi(t) \geq t + \alpha \quad \forall t \in [0, +\infty) \qquad \theta(t) \geq \beta t \quad \forall t \in (0, 2M]\,,$$

which we can use to compute an uniform upper bound on the variations of u_k:

$$\begin{aligned} |Du_k|(\Omega) &= \int_\Omega \|\nabla u_k\|\, dx + \int_{J_{u_k}} |u_k^+ - u_k^-|\, d\mathcal{H}^{n-1} \\ &\leq \int_\Omega \phi(\|\nabla u_k\|)\, dx - \alpha\lambda(\Omega) + \frac{1}{\beta}\int_{J_{u_k}} \theta(|u_k^+ - u_k^-|)\, d\mathcal{H}^{n-1}\,. \end{aligned}$$

Thus we can use Theorem 2.4 to extract a subsequence which converges in $L^1(\Omega)$ to some $u \in BV(\Omega)$. Furthermore, as Ω is bounded and $\|u_n\|_\infty + |Du_n|(\Omega)$ are uniformly bounded, it follows that $u \in L^\infty(\Omega)$, $u_{n_k} \xrightarrow{k\to\infty} u$ in L^1 and $|Du|(\Omega) < +\infty$, and consequently $u \in BV(\Omega)$ and $u_{n_k} \overset{*}{\rightharpoonup} u$. By Theorem 2.27, this implies $u \in SBV(\Omega)$. QED

With these results in hand, we are finally ready to return to (1). First we construct a weak formulation of the original problem by defining the auxiliary functional

$$\tilde{E}(u) := \int_\Omega (u-g)^2 dx + \int_\Omega \|\nabla u\|^2 \, dx + \mathcal{H}^1(S_u) \tag{6}$$

on the larger space $SBV(\Omega)$, and show that this relaxed problem does indeed have a solution.

Theorem 2.29. *The functional (6) has a minimiser.*

Proof. Without loss of generality, we may assume that all admissible candidates belong to $L^\infty(\Omega)$. Indeed, suppose u were a minimiser of $\tilde{E}$ and not bounded, and consider the truncation

$$u^M := (M \vee u) \wedge -M$$

with $M = \|g\|_\infty$. This function again belongs to $SBV(\Omega)$, and we have

$$\begin{cases} \int_\Omega \|\nabla u^M\|^2 \, dx \le \int_\Omega \|\nabla u\|^2 \, dx \\ \int_\Omega (u^M - g)^2 \, dx \le \int_\Omega (u-g)^2 \, dx \\ \mathcal{H}^{n-1}(S_{u_M}) \le \mathcal{H}^{n-1}(S_u) \, , \end{cases}$$

from which it follows that u^M defeats u, contradicting our initial assumption. Now consider a minimising sequence $(u_n)_n$ for $\tilde{E}$. We will apply the direct method of the calculus of variations (Theorem 1.1) to obtain a minimiser from it. From the previous inequalities, it follows that

$$\inf_{v\in V} \tilde{E}(v) \le \tilde{E}(u_n^M) \le \tilde{E}(u_n) \, ,$$

and we may conclude that $(u_n^M)_n$ is again a minimising sequence, uniformly bounded by $\|g\|_\infty$. We may then apply Theorem 2.28 with the $\theta(t) = 1$ (a concave function) and $\phi(t) = t^2$ (a convex function) to obtain a subsequence $(u_{k_n})_n$ converging weakly-* to some $u^* \in SBV(\Omega)$. From functional analysis, it is known that the

L^2-norm is weakly lower semicontinuous, and we can combine this with latter part of Theorem 2.27 to write

$$\begin{aligned}\tilde{E}(u^*) &= \int_\Omega (u^* - g)^2\,dx + \int_\Omega \|\nabla u^*\|^2\,dx + \mathcal{H}^{n-1}(S_{u^*}) \\ &\leq \liminf_{n\to\infty} \int_\Omega (u_{k_n} - g)^2\,dx + \int_\Omega \|\nabla u_{k_n}\|^2\,dx + \mathcal{H}^{n-1}(S_{u_{k_n}}) \\ &= \liminf_{n\to\infty} \tilde{E}(u_{k_n})\,,\end{aligned}$$

showing that u^* is a minimiser, as desired. QED

Of course, this result is not yet satisfactory, as we want u to satisfy the original constraints. Furthermore, it is not immediately clear what set corresponds to K in this weak formulation. The next section will be concerned with these regularity questions.

3 Regularity

We will develop the regularity theory for SBV-minimisers of the relaxed Mumford-Shah functional in two parts. First, we show that for all admissible pairs (u, K) of the original problem (1), the functions u belong to $SBV(\Omega)$. This guarantees that $SBV(\Omega)$ extends the original space of admissible functions, so that $\min_{u\in SBV(\Omega)} \tilde{E}(u) \leq \inf_{(u,K)\in\mathcal{A}} E(u, K)$. Then we will show that if u^* is a minimiser of the relaxed problem (6), the pair $(u, \overline{S}_{u^*}$ belongs to the original set of candidate solutions. From these results, it will then follow that minimisers of the relaxed functional realise the infimum in the original problem, i.e

$$\min_{u\in SBV(\Omega)} \tilde{E}(u) = \inf_{(u,K)\in\mathcal{A}} E(u, K)\,.$$

We will need a few preliminaries first.

Theorem 3.1. *For $u, v \in BV(\Omega)$ and E is a set of finite perimeter in Ω with reduced boundary $\mathcal{F}E$ oriented by ν_E, define $w := uI_E + vI_{\Omega\setminus E}$. Then w again belongs to $BV(\Omega)$ if and only if*

$$\int_{\mathcal{F}E} |u^+_{\mathcal{F}E} - v^-_{\mathcal{F}E}|\,d\mathcal{H}^{n-1} < \infty\,.$$

In that case, we can write

$$Dw = Du \llcorner E^1 + (u^+_{\mathcal{F}E} - v^-_{\mathcal{F}E})\nu_E \mathcal{H}^{n-1} \llcorner \mathcal{F}E + Dv \llcorner E^0\,.$$

Lemma 3.2. *A function $u \in BV(\Omega)$ belong to $SBV(\Omega)$ if and only if $D^s u$ is concentrated on a $\mathcal{H}^{n-1}$-σ-finite Borel set.*

Lemma 3.3. *Suppose $\Omega \in \mathbb{R}^n$ is open and bounded and let K be a closed subset of Ω such that $\mathcal{H}^{n-1}(K) < \infty$. Then any function $u : \Omega \longrightarrow \mathbb{R}^n$ which belongs to $W^{1,1}(\Omega \setminus K) \cap L^\infty(\Omega \setminus K)$ belongs to $SBV(\Omega)$ as well and $\mathcal{H}^{n-1}(S_u \setminus K) < 0$.*

Proof. Recall that the d-dimensional Hausdorff measure of a set $K \subseteq \mathbb{R}^n$ can be defined as follows:

$$\mathcal{H}^d_\delta(K) = \inf \left\{ \frac{\omega_d}{2^d} \sum_{k=1}^{\infty} \operatorname{diam}(U_k) \;\middle|\; \bigcup_{k=1}^{\infty} U_k \supseteq K,\, \operatorname{diam}(U_k) < \delta \right\}$$
$$\mathcal{H}^d(K) := \lim_{\delta \to 0} \mathcal{H}^d_\delta(K)\,.$$

If $\mathcal{H}^{n-1}(K) < \infty$, we can therefore find, for any $h \geq 1$, an open cover $(B^h_{r_i}(x_i))_{i \in I_h}$ of K such that $r_i \leq \frac{1}{h}$ and

$$\sum_{i \in I_n} \omega_{n-1} r_i^{n-1} = \sum_{i \in I_n} \frac{\omega_{n-1}}{2^{n-1}} (2r_i)^{n-1} \leq \mathcal{H}^{n-1}(K) + 1\,,$$

from which, using the subadditivity of perimeter, we obtain

$$\begin{aligned} P(K, \Omega) &\leq \sum_{i \in I_h} P(B^h_{r_i}(x_i)) \\ &= \sum_{i \in I_h} \mathcal{H}^{n-1}(\partial B^h_{r_i}(x_i)) \\ &\leq \frac{n\omega_n}{\omega_{n-1}} (\mathcal{H}^{n-1}(K) + 1)\,. \end{aligned}$$

Now define the following sequence of functions:

$$u_h(x) := \begin{cases} u(x) \text{ if } x \in \Omega \setminus \cup_{i \in I_h} B^h_{r_i}(x_i) \\ 0 \text{ otherwise}\,. \end{cases}$$

Then we clearly have $u_h \xrightarrow{L^1} u$, and it follows from Theorem 3.1 that $u_n \in BV(\Omega)$ and

$$|Du_n|(\Omega) \leq \int_{\Omega \setminus K} \|\nabla u\| \, dx + \|u\|_\infty \frac{n\omega_n}{\omega_{n-1}} (\mathcal{H}^{n-1}(K) + 1)\,.$$

From lower semicontinuity of the total variation, it then follows that $u \in BV(\Omega)$. Finally, we can use the fact that $u \in W^{1,1}(\Omega \setminus K)$ combined with Lemma 3.2 to conclude that $u \in SBV(\Omega)$. QED

Theorem 3.4. *If (u, K) is an admissible the Mumford-Shah problem (1), then $u \in SBV(\Omega)$ and consequently*

$$\min_{u \in SBV(\Omega)} \tilde{E}(u) \leq \inf_{(u,K) \in \mathcal{A}} E(u, K)\,.$$

Proof. Without loss of generality, we may assume that $\mathcal{H}^1(K) < \infty$ and $u \in W^{1,2}(\Omega \setminus K)$, as any function not satisfying these requirements can be defeated by a constant function. Because $\Omega \setminus K$ is a finite measure space, it then immediately follows that $u \in W^{1,1}(\Omega \setminus K)$. We may further assume that $u \in L^\infty(\Omega \setminus K)$. Indeed, if u is unbounded, we can use a smoothed version of the argument from Theorem 2.29 to obtain a competitor which defeats it. Consider a function $\phi \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ which satisfies $0 \leq \phi' \leq 1$ and $\phi(t) = t$ for $|t| \leq \|g\|_\infty$. Then $\phi \circ u$ will again belong to $C^1(\Omega \setminus K) \cap L^\infty(\Omega \setminus K)$ and we have

$$\int_\Omega (\phi \circ u - g)^2\, dx \leq \int_\Omega (u - g)^2\, dx$$
$$\int_{\Omega \setminus K} \|\nabla(\phi \circ u)\|\, dx = \int_{\Omega \setminus K} \|\phi'(u)\| \|\nabla u\|\, dx \leq \int_{\Omega \setminus K} \|\nabla u\|\, dx\,,$$

so that $E(\phi \circ u, K) \leq E(u, K)$. Hence we it follows that

$$\left\{\, E(u,K) \;\middle|\; F \subset \Omega \,\text{closed}\,, u \in C^1(\Omega \setminus K) \,\right\}$$
$$= \left\{\, E(u,K) \;\middle|\; F \subset \Omega \,\text{closed}\,, u \in C^1(\Omega \setminus K) \cap L^\infty(\Omega) \,\right\}\,,$$

and we obtain from Lemma 3.3 that $u \in SBV(\Omega)$. QED

We are now ready to move on to the second part of the regularity theory, which we will approach as follows. Given a minimiser u^* of the relaxed problem, we cannot not know a priori whether the corresponding singularity set S_{u^*} is closed. Therefore, we will consider its closure instead, obtaining the pair $(u^*, \overline{S}_{u^*})$. If we can show that u^* has a representative which is continuous on $\Omega \setminus \overline{S}_{u^*}$, then $(u^*, \overline{S}_{u^*})$ belongs to the set $\mathcal{A}$ of admissible pairs. It must then still be shown that this modification does not increase the value of the functional. This is a deep result which will not prove here; the interested reader is directed to the original article [4] in which it was first discussed.

Theorem 3.5. *If $u^* \in SBV(\Omega)$ is a minimiser for the relaxed Mumford-Shah problem (6), then $u^* \in C^1(\Omega \setminus \overline{S}_{u^*})$ and $\mathcal{H}^{n-1}(\overline{S}_{u^*} \setminus S_{u^*}) = 0$, i.e. S_{u^*} is essentially closed. Consequently,*

$$\inf_{(u,K) \in \mathcal{A}} E(u, K) \leq \min_{u \in SBV(\Omega)} \tilde{E}(u)\,.$$

Proof. Let x be any point in $\Omega \setminus \overline{S}_{u^*}$ and consider an open ball $B_r(x) \subset \Omega \setminus \overline{S}_{u^*}$. Then we have $u^* \in W^{1,2}(B_r(x))$, and moreover u^* minimises the problem

$$u^* = \operatorname{argmin} \left\{ \int_{B_r(x)} |\nabla v|^2 \, dx + \int_{B_r(x)} (v-g)^2 \, dx \;\middle|\; v \in u^* + W_0^{1,2}(B_r(x)) \right\} ,$$

from which it follows, by the standard techniques of variational calculus, that u^* is a weak solution for the Dirichlet problem

$$\begin{cases} -\Delta v + v = g & \text{in } B_r(x) \\ v = u^* & \text{on } \partial B_r(x) \,, \end{cases}$$

where the appropriate restrictions are assumed but suppressed for notational clarity. Arguing as in the previous theorem, we may assume $v \in L^\infty(\Omega)$ and can thus rewrite the first equation as

$$\delta v = v - g =: f \,,$$

with $f \in L^\infty(\Omega)$. As the fundamental solution of the Laplace operator is known to be given by

$$\Delta \left(\frac{1}{2\pi} \log(\|x\|) \right) = \delta(x)$$

we can obtain a solution to this equation through the convolution

$$F(y) = \frac{1}{2\pi} \int_{\overline{B_r}(x)} f(z) \log(\|y-z\|) \, dz \,.$$

For every fixed $y \in \overline{B_r}(x)$, the mapping $z \mapsto f(z)\log(|y-z|)$ is in $L^1(\overline{B_r}(x))$ by virtue of the boundedness of f and the compactness of the domain of integration. Differentiating under the integral sign, we thus obtain

$$\frac{\partial F}{\partial y_i}(y) = \int_{\overline{B_r}(x)} g(z) \frac{y_i}{\|y-z\|} \, dz \,,$$

from which is follows that $F \in C^1(\overline{B_r}(x))$. If v is any other solution, it holds that $\Delta(F - v) = 0$ so that $F - v$ lies in the kernel $C^\infty(\overline{B_r}(x))$ of the Laplace operator. Consequently, $v \in C^1(\overline{B_r}(x))$ and the combination of these local results yields $u^* \in C^1(\Omega \setminus \overline{S}_{u^*})$.

Finally, it can be shown that $\mathcal{H}^1(\overline{S}_{u^*} \setminus S_{u^*}) = 0$, from which it follows that

$$\min_{u \in SBV(\Omega)} \tilde{E}(u^*) = \inf_{(u,K) \in \mathcal{A}} E(u^*, \overline{S}_{u^*}) \,,$$

demonstrating that $(u^*, \overline{S}_{u^*}) \in \mathcal{A}$ minimisation the original Mumford-Shah problem. QED

References

[1] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of Bounded Variationand Free Discontinuity Problems.* Oxford Mathematical Monographs. Clarendon Press, 2000. ISBN: 0-19-850245-1.

[2] Hedy Attouch, Giuseppe Buttazzo, and Gérard Michaille. *Variational Analysis in Sobolev and BV Spaces.* MPSSIAM Series on Optimization. Mathematical Programming Society, Society for Industrial, and Applied Mathematics, 2006. ISBN: 0-89871-600-4.

[3] Haïm Brezis. *Analyse Fonctionnelle. Théorie et applications.* Collection Mathématiques appliquées pour la maîtrise. Masson, 1983. ISBN: 2-225-77198-7.

[4] Ennio De Giorgi, Michele Carriero, and Antonio Leaci. "Existence Theorem for a Minimum Problem with Free Discontinuity Set". In: *Archive for Rational Mechanics and Analysis* 108 (4 1989), 195–218.

[5] Antoine Lemenant. "A selective review on Mumford-Shah minimizers." In: *Bollettino dell'Unione Matematica Italiana* 9 (1 2016), pp. 69–113.

[6] Francesco Maggi. *Sets of Finite Perimeter and Geometric Variational Problems.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2012. ISBN: 978-1-107-02103-7.

[7] David Mumford and Jayant Shah. "Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems". In: *Communications on Pure and Applied Mathematics* 42 (5 1989), pp. 577–685.