Regularity of gradient vector fields giving rise to finite Caccioppoli partitions

Roger Moser

January 20, 2022

Abstract

For a finite set $A \subseteq \mathbb{R}^n$, consider a function $u \in BV_{loc}^2(\mathbb{R}^n)$ such that $\nabla u \in A$ almost everywhere. If $A$ is convex independent, then it follows that $u$ is piecewise affine away from a closed, countably $H^{n-1}$-rectifiable set. If $A$ is affinely independent, then $u$ is piecewise affine away from a closed $H^{n-1}$-null set.

1 Introduction

For $n \in \mathbb{N}$, consider a finite set $A \subseteq \mathbb{R}^n$. We study continuous functions $u : \mathbb{R}^n \rightarrow \mathbb{R}$ such that the weak gradient $\nabla u$ satisfies $\nabla u \in BV_{loc}(\mathbb{R}^n; \mathbb{R}^n)$ and $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. This means that whenever $\Omega \subseteq \mathbb{R}^n$ is open and bounded, the sets $\{x \in \Omega : \nabla u(x) = a\}$, for $a \in A$, form a Caccioppoli partition of $\Omega$ as discussed, e.g., by Ambrosio, Fusco, and Pallara [1, Section 4.4]. The theory of Caccioppoli partitions therefore applies and gives some information on the structure of $\nabla u$ and of $u$. The fact that we are dealing with a gradient, however, gives rise to a better theory, especially under additional assumptions on the geometry of $A$. We work with the following notions in this paper.

Definition 1. A set $A \subseteq \mathbb{R}^n$ is called convex independent if any $a \in A$ does not belong to the convex hull of $A \setminus \{a\}$. It is called affinely independent if any $a \in A$ does not belong to the affine span of $A \setminus \{a\}$.

If either of these conditions is satisfied, then we can prove statements on the regularity of $u$ that finite Caccioppoli partitions do not share in general. In fact, we will see that $u$ is locally piecewise affine away from a closed, countably $H^{n-1}$-rectifiable set (if $A$ is convex independent) or away from a closed $H^{n-1}$-null set (if $A$ is affinely independent).

In order to make this more precise, we introduce some notation. Given $r > 0$ and $x \in \mathbb{R}^n$, we write $B_r(x)$ for the open ball of radius $r$ centred at $x$. Given $a \in \mathbb{R}^n$, the function $\lambda_a : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined by $\lambda_a(x) = a \cdot x$ for $x \in \mathbb{R}^n$. Given two functions $v, w : \mathbb{R}^n \rightarrow \mathbb{R}$, we write $v \wedge w$ and $v \vee w$, respectively, for the functions with $(v \wedge w)(x) = \min\{v(x), w(x)\}$ and $(v \vee w)(x) = \max\{v(x), w(x)\}$ for $x \in \mathbb{R}^n$. 

*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK. E-mail: r.moser@bath.ac.uk
Defi on 2. Given a function \( u : \mathbb{R}^n \to \mathbb{R} \), the regular set of \( u \), denoted by \( \mathcal{R}(u) \), consists of all \( x \in \mathbb{R}^n \) such that there exist \( a, b \in \mathbb{R}^n \), \( c \in \mathbb{R} \), and \( r > 0 \) with \( u = \lambda_a \wedge \lambda_b + c \) in \( B_r(x) \) or \( u = \lambda_a \vee \lambda_b + c \) in \( B_r(x) \). The singular set of \( u \) is its complement \( \mathcal{S}(u) = \mathbb{R}^n \setminus \mathcal{R}(u) \).

The condition for \( \mathcal{R}(u) \) allows the possibility that \( a = b \), in which case \( u \) is affine near \( x \). If \( a \neq b \), then it is still piecewise affine near \( x \). Obviously \( \mathcal{R}(u) \) is an open set and \( \mathcal{S}(u) \) is closed.

It would be reasonable to include functions consisting of more than two affine pieces in the definition of \( \mathcal{R}(u) \), for example \( (\lambda_{a_1} \wedge \lambda_{a_2}) \vee \lambda_{a_3} + c \) for \( a_1, a_2, a_3 \in \mathbb{R}^n \) and \( c \in \mathbb{R} \). For the results of this paper, however, this would make no difference, therefore we choose the simpler definition.

For \( s \geq 0 \), we denote the \( s \)-dimensional Hausdorff measure in \( \mathbb{R}^n \) by \( \mathcal{H}^s \).

The notation \( \text{BV}^2_{\text{loc}}(\mathbb{R}^n) \) is used for the space of functions with weak gradient in \( \text{BV}^2_{\text{loc}}(\mathbb{R}^n; \mathbb{R}^n) \). Thus the hypotheses of the following theorems are identical to the assumptions at the beginning of the introduction.

**Theorem 3.** Suppose that \( A \) is a finite, convex independent set. Let \( u \in \text{BV}^2_{\text{loc}}(\mathbb{R}^n) \) with \( \nabla u(x) \in A \) for almost every \( x \in \mathbb{R}^n \). Then \( \mathcal{S}(u) \) is countably \( \mathcal{H}^{n-1} \)-rectifiable.

**Theorem 4.** Suppose that \( A \) is a finite, affinely independent set. Let \( u \in \text{BV}^2_{\text{loc}}(\mathbb{R}^n) \) with \( \nabla u(x) \in A \) for almost every \( x \in \mathbb{R}^n \). Then \( \mathcal{H}^{n-1}(\mathcal{S}(u)) = 0 \).

For \( n = 2 \), Theorem 4 was proved in a previous paper [10]. For higher dimensions, the result is new. Theorem 4 is new even for \( n = 1 \). For \( n = 1 \), both statements are easy to prove.

The results are optimal in terms of the Hausdorff measures involved. Furthermore, the assumption of convex/affine independence is necessary. Indeed, there are examples of finite sets \( A \subseteq \mathbb{R}^2 \) and functions \( u \in \text{BV}^2_{\text{loc}}(\mathbb{R}^2) \) with \( \nabla u(x) \in A \) almost everywhere such that

- \( \mathcal{H}^2(\mathcal{S}(u)) > 0 \); or
- \( \mathcal{H}^1(\mathcal{S}(u)) > 0 \) and \( A \) is convex independent; or
- \( \mathcal{H}^s(\mathcal{S}(u)) = \infty \) for any \( s < 1 \) and \( A \) is affinely independent.

All of these can be found in the author’s previous paper [10].

Apart from being of obvious geometric interest, functions as described above appear in problems from materials science. They naturally arise as limits in \( \Gamma \)-convergence theories in the spirit of Modica and Mortola [8, 9] for quantities such as

\[
\int_\Omega \left( \epsilon |\nabla^2 u|^2 + \frac{W(\nabla u)}{\epsilon} \right) \, dx,
\]

where \( \Omega \subseteq \mathbb{R}^n \) is an open set and \( W : \mathbb{R}^n \to [0, \infty) \) is a function with \( A = W^{-1}(\{0\}) \). Functional of this sort appear in certain models for the surface energy of nanocrystals [13, 14, 15]. For \( \Omega \subseteq \mathbb{R}^2 \), functions \( u \in \text{BV}^2(\Omega) \) with \( \nabla u \in \{ \pm 1, 0 \} \) have also been used by Cicalese, Forster, and Orlando [3] for a different sort of \( \Gamma \)-limit arising from a model for frustrated spin systems.

Functional similar to (1), but for maps \( u : \Omega \to \mathbb{R}^n \), also appear in certain models for phase transitions in elastic materials (see, e.g., the seminal paper...
of Ball and James [2] or the introduction into the theory by Müller [11]. In this context, due to the frame indifference of the underlying models, the set $W^{-1}(\{0\})$ is typically not finite. Sometimes, however, the frame indifference is disregarded (as in the paper by Conti, Fonseca, and Leoni [4]), or the theory gives a limit with $\nabla u \in BV(\Omega; A)$ for a finite set $A \subseteq \mathbb{R}^{n \times n}$ anyway (such as in recent results of Davoli and Friedrich [6, 5]). In such a case, Theorem 3 and Theorem 4 are potentially useful, as they apply to the components (or other one-dimensional projections) of $u$.

In the proof of Theorem 4, we use some of the tools from the author’s previous paper [10]. In particular, we will analyse the intersections of the graph of $u$ with certain hyperplanes in $\mathbb{R}^{n+1}$. We will see that these intersections correspond to the graphs of functions with $(n-1)$-dimensional domains and with properties similar to $u$. The key ideas from the previous paper, however, are specific to $\mathbb{R}^2$, so we eventually use different arguments. In this paper, we use the theory of $BV_{\text{loc}}(\mathbb{R}^n; \mathbb{R}^n)$ to a much greater extent. The central argument will consider approximate jump points of $\nabla u$. Near such a point, we know that $u$ is close to a piecewise affine function in a measure theoretic sense by definition. We then use an induction argument (with induction over $n$) to show that $u$ is in fact piecewise affine near $\mathcal{H}^{n-1}$-almost every approximate jump point.

We also need to analyse points where $u$ has an approximate limit, and they are of interest for the proofs of both Theorem 4 and Theorem 3. This part of the analysis is significantly simpler and relies on the fact that for any $a \in A$, the function $v(x) = u(x) - a \cdot x$ has some monotonicity properties.

In the rest of the paper, we study a fixed function $u \in BV_{\text{loc}}(\mathbb{R}^n)$ with $\nabla u(x) \in A$ for almost every $x \in \mathbb{R}^n$. Since we are interested only in the local properties of $u$, we may assume that it is also bounded. (Otherwise we can modify it outside of a bounded set with the construction described in [10, Section 6].) We define the function $U: \mathbb{R}^n \to \mathbb{R}^{n+1}$ by

$$U(x) = \left( \begin{array}{c} x \\ u(x) \end{array} \right), \quad x \in \mathbb{R}^n.$$ 

We use the notation $\text{graph}(u) = U(\mathbb{R}^n)$ for the graph of $u$.

As we sometimes work with points in $\mathbb{R}^{n+1}$ (especially points on graph($u$)) and their projections onto $\mathbb{R}^n$ simultaneously, we use the following notation. A generic point in $\mathbb{R}^{n+1}$ is denoted by $x = (x_1, \ldots, x_{n+1})^T$, and then we write $x = (x_1, \ldots, x_n)^T$. Thus $x = (x_{n+1})$. We think of elements of $\mathbb{R}^n$ and of $\mathbb{R}^{n+1}$ as column vectors, and this is sometimes important, as we use them as columns in certain matrices.

As our function satisfies in particular the condition $\nabla u \in BV_{\text{loc}}(\mathbb{R}^n; \mathbb{R}^n)$, the theory of this space will of course be helpful. In this context, we mostly follow the notation and terminology of Ambrosio, Fusco, and Pallara [11]. We also use several of the results found in this book.

## 2 Approximate faces and edges of the graph

In this section, we decompose $\mathbb{R}^n$ into three sets $F$, $E$, and $N$. These are defined such that we expect regularity in $F$ under the assumptions of either of the main theorems, and also in $E$ under the assumptions of Theorem 4. The third set, $N$, will be an $\mathcal{H}^{n-1}$-null set. The sets $F$ and $E$ characterised, up to $\mathcal{H}^{n-1}$-null sets,
by the condition that $\nabla u$ has an approximate limit or an approximate jump, respectively. Since much of our analysis examines $\text{graph}(u)$, it is also convenient to think of $\mathcal{F}$ as the set of points where the graph behaves approximately like the ($n$-dimensional) faces of a polyhedral surface, whereas $\mathcal{E}$ corresponds to approximate ($n - 1$)-dimensional edges.

First we define the set $\mathcal{F} \subseteq \mathbb{R}^n$, comprising all points $x \in \mathbb{R}^n$ such that there exists $a \in \mathbb{R}^n$ satisfying

$$\lim_{r \searrow 0} \int_{B_r(x)} |\nabla u - a| \, dH^n = 0.$$  

In other words, this is the set of all points where $\nabla u$ has an approximate limit $a$. It is then clear that $a \in A$. The complement $\mathbb{R}^n \setminus \mathcal{F}$ is called the approximate discontinuity set of $\nabla u$.

Furthermore, let $\mathcal{E}$ be the set of all $x \in \mathbb{R}^n$ such that there exist $a_-, a_+ \in \mathbb{R}^n$ with $a_- \neq a_+$ and there exists $\eta \in S^{n-1}$ such that

$$\lim_{r \searrow 0} \frac{1}{r} \int_{B_r(x)} (\tilde{x} - x) \cdot \eta > 0$$

and

$$\lim_{r \searrow 0} \frac{1}{r} \int_{B_r(x)} (\tilde{x} - x) \cdot \eta < 0.$$  

This is the approximate jump set of $\nabla u$. Again, the points $a_-, a_+$ will always belong to $A$.

According to a result by Federer and Vol’pert (which can be found in the book by Ambrosio, Fusco, and Pallara [1, Theorem 3.78]), there exists an $H^{n-1}$-null set $\mathcal{N} \subseteq \mathbb{R}^n$ such that

$$\mathbb{R}^n = \mathcal{F} \cup \mathcal{E} \cup \mathcal{N}.$$  

Furthermore, the set $\mathcal{E}$ is countably $H^{n-1}$-rectifiable.

Given $x \in \mathbb{R}^n$ and $\rho > 0$, we define the function $u_{x,\rho} : \mathbb{R}^n \to \mathbb{R}$ with

$$u_{x,\rho}(\tilde{x}) = \frac{1}{\rho} (u(x + \rho \tilde{x}) - u(x))$$

for $\tilde{x} \in \mathbb{R}^n$. For $x$ fixed, the family of functions $(u_{x,\rho})_{\rho > 0}$ is clearly bounded in $C^0(\overline{K})$ for any compact set $K \subseteq \mathbb{R}^n$. Therefore, the theorem of Arzelà–Ascoli implies that there exists a sequence $\rho_k \searrow 0$ such that $u_{x,\rho_k}$ converges locally uniformly. If we have in fact a limit for $\rho \searrow 0$, then we write

$$T_x u = \lim_{\rho \searrow 0} u_{x,\rho}$$

and call this limit the tangent function of $u$ at $x$.

If $x \in \mathcal{F}$ and $a \in A$ is the approximate limit of $\nabla u$ at $x$, then for any sequence $\rho_k \searrow 0$, the limit of $u_{x,\rho_k}$ can only be $\lambda a$. Hence in this case, there exists a tangent function $T_x u$, which is exactly this function. Similarly, if $x \in \mathcal{E}$, then $T_x u$ exists and

$$T_x u(\tilde{x}) = \begin{cases} 
\lambda a_-(\tilde{x}) & \text{if } \tilde{x} \cdot \eta < 0, \\
\lambda a_+(\tilde{x}) & \text{if } \tilde{x} \cdot \eta \geq 0.
\end{cases}$$  

4
Because $T_x u$ is a continuous function, this means that

$$\eta = \pm \frac{a_+ - a_-}{|a_+ - a_-|}.$$  

Then we conclude that $T_x u = \lambda_{a_-} \wedge \lambda_{a_+}$ or $T_x u = \lambda_{a_-} \vee \lambda_{a_+}$, depending on the sign.

If we consider the functions $a_-, a_+: E' \to A$ and $\eta: E' \to S^{n-1}$ such that (2) and (3) are satisfied on $E'$, then the previously used result [1, Theorem 3.78] also implies that

$$D\nabla u \mathbf{L} E' = (a_+ - a_-) \otimes \eta \mathcal{H}^{n-1} \mathbf{L} E'.$$

Let $\gamma = \min \{|a - b|: a, b \in A\}$. Then for any Borel set $\Omega \subseteq \mathbb{R}^n$, we conclude that

$$|D\nabla u|(\Omega) \geq \gamma \mathcal{H}^{n-1}(E' \cap \Omega).$$

Now define

$$\mathcal{F} = \left\{ x \in E': \lim_{\rho \to 0} \rho^{1-n} |D\nabla u|(B_\rho(x)) = 0 \right\}.$$  

Then standard results [1, Theorem 2.56 and Lemma 3.76] imply that $\mathcal{H}^{n-1}(\mathcal{F} \setminus \mathcal{F}) = 0$.

Recall the map $U: \mathbb{R}^n \to \mathbb{R}^{n+1}$ defined in the introduction. Set $\mathcal{F}^* = U(\mathcal{F})$ and $\mathcal{E}^* = U(E^*)$. Then $\mathcal{E}^*$ is a countably $\mathcal{H}^{n-1}$-rectifiable subset of $\mathbb{R}^{n+1}$. Hence at $\mathcal{H}^{n-1}$-almost every $x \in \mathcal{E}^*$, the measure $\mathcal{H}^{n-1} \mathbf{L} \mathcal{E}^*$ has a tangent measure [1, Theorem 2.83] of the form $\mathcal{H}^{n-1} \mathbf{L} T_x \mathcal{E}^*$, where $T_x \mathcal{E}^*$ is an $(n-1)$-dimensional linear subspace of $\mathbb{R}^{n+1}$ (the approximate tangent space of $\mathcal{E}^*$ at $x$). Let $\mathcal{E}^*$ be the set of all $x \in \mathcal{E}^*$ where this is the case. Furthermore, let $\mathcal{E} = U^{-1}(\mathcal{E}^*)$. Then $\mathcal{E}^* \setminus \mathcal{E}$ is an $\mathcal{H}^{n-1}$-null set.

Thus if we define $\mathcal{N} = \mathbb{R}^n \setminus (\mathcal{F} \cup \mathcal{E})$, then $\mathcal{N}$ is an $\mathcal{H}^{n-1}$-null set and we have the disjoint decomposition

$$\mathbb{R}^n = \mathcal{F} \cup \mathcal{E} \cup \mathcal{N}.$$  

3 Proof of Theorem 3

In this section we prove our first main result, Theorem 3. The proof is based on the following proposition, which will also be useful for the proof of Theorem 4 later on.

**Proposition 5.** Suppose that $A \subseteq \mathbb{R}^n$ is finite and convex independent. Let $u \in BV_{loc}^2(\mathbb{R}^n)$ be a function with $\nabla u(x) \in A$ for almost all $x \in \mathbb{R}^n$. Then there exist $r > 0$ and $\epsilon > 0$ with the following property. Suppose that there exists $a \in A$ such that

$$\mathcal{H}^n(\{x \in B_1(0): \nabla u(x) \neq a\}) \leq \epsilon$$  

and

$$|D\nabla u|(B_1(0)) \leq \epsilon.$$  

Then $\nabla u(x) = a$ for almost every $x \in B_r(0)$.  

5
Proof. Because $A$ is convex independent, there exists $\omega \in S^{n-1}$ such that
\[ a \cdot \omega < \min_{b \in A \setminus \{a\}} b \cdot \omega. \]
As $A$ is finite, there also exists $\delta \in (0,1)$ such that the inequality $a \cdot \xi \leq \min_{b \in A \setminus \{a\}} (b \cdot \xi)$ holds even for $\xi$ in the cone
\[ C = \{ \xi \in \mathbb{R}^n : \xi \cdot \omega \geq \delta |\xi| \}. \]

Consider the function $v : \mathbb{R}^n \to \mathbb{R}$ with $v(x) = u(x) - a \cdot x$ for $x \in \mathbb{R}^n$. Then for any $\xi \in C$,
\[ \xi \cdot \nabla v(x) = \xi \cdot \nabla u(x) - a \cdot \xi \geq 0 \]
almost everywhere. Thus $v$ is monotone along lines parallel to $\xi$. (This is true for every such line by the continuity of $v$.) Furthermore, for almost every $x \in \mathbb{R}^n$, we find that either $\nabla u(x) = a$ or $\omega \cdot \nabla v(x) > 0$.

Suppose that $\nabla u = a$ does not hold almost everywhere in $B_r(0)$. Then there exist $x_-, x_+ \in B_r(0)$ with $v(x_-) < v(x_+)$. Define
\[ C_+ = (x_- - C) \cap B_1(0) \quad \text{and} \quad C_- = (x_+ + C) \cap B_1(0). \]

Then for any $x' \in C_-$ and $x'' \in C_+$, we conclude that
\[ v(x') \leq v(x_-) < v(x_+) \leq v(x''). \]

We now foliate a part of $B_1(0)$ by line segments parallel to $\omega$. For $R \in (0,1]$, let $Z_R = \{ x \in B_R(0) : \omega \cdot x = 0 \}$. For every $z \in Z_R$, consider the line segment
\[ L_z = \left\{ z + t\omega : -\frac{1}{2} \leq t \leq \frac{1}{2} \right\}. \]
Provided that $r$ is chosen sufficiently small, we can find $R \in (0,1]$ such that
\[ \left\{ z - \frac{\omega}{2} : z \in Z_R \right\} \subseteq C_- \quad \text{and} \quad \left\{ z + \frac{\omega}{2} : z \in Z_R \right\} \subseteq C_+. \]

Hence for any $z \in Z_R$,
\[ v \left( z + \frac{\omega}{2} \right) - v \left( z - \frac{\omega}{2} \right) \geq v(x_+) - v(x_-) > 0. \]

In particular, the restriction of $v$ to the line segment $L_z$ is not constant. For $z \in Z_R$, define $L_z = \{ x \in L_z : \nabla u(x) = a \}$. Then it follows that $H^1(L_z) < 1$ for $H^{n-1}$-almost all $z \in Z_R$.

On the other hand, because of [1], we also know that
\[ H^{n-1} \left( \{ z \in Z_R : H^1(L_z) = 0 \} \right) \leq \epsilon. \]
Thus if we define $Z' = \{ z \in Z_R : 0 < H^1(L_z) < 1 \}$, then
\[ H^{n-1}(Z') \geq H^{n-1}(Z_R) - \epsilon. \]

Set $c = \min_{b \in A} |a - b|$. For $H^{n-1}$-almost any $z \in Z'$, the function $t \mapsto \nabla u(z + t\omega)$ belongs to BV$((-\frac{1}{2}, \frac{1}{2}); \mathbb{R}^n)$ and its total variation is at least $c$. Hence [3, Theorem 3.103]
\[ |D\nabla u|(B_1(0)) \geq c H^{n-1}(Z') \geq c(H^{n-1}(Z_R) - \epsilon). \]
If $\epsilon$ is sufficiently small, then this means in particular that $|D\nabla u|(B_1(0)) > \epsilon$. Thus we have proved the contrapositive of Proposition [5].
Proof of Theorem 3. We show that $\mathcal{F} \subseteq \mathcal{R}(u)$. To this end, fix $x \in \mathcal{F}$ and consider the rescaled functions $u_{x, \rho}$ for $\rho > 0$. Since $x \in \mathcal{F}$, we know that $\nabla u_{x, \rho} \rightarrow a$ in $L^1(B_1(0))$ as $\rho \searrow 0$ for some $a \in A$. Furthermore, since

$$|D\nabla u_{x, \rho}|(B_1(0)) = \rho^{1-n}|D\nabla u|(B_\rho(x)) \rightarrow 0$$

as $\rho \searrow 0$, the function $u_{x, \rho}$ satisfies the inequalities of Proposition 5 for $\rho$ sufficiently small. Hence $\nabla u_{x, \rho}(\tilde{x}) = a$ for almost every $\tilde{x} \in B_{\rho}(0)$, which implies that $u(\tilde{x}) = u(x) + a \cdot (\tilde{x} - x)$ for all $\tilde{x} \in B_{\rho}(x)$. Hence $x \in \mathcal{R}(u)$. Theorem 3 now follows from the observations in Section 2.

4 Specialising to a regular $n$-simplex

The rest of the paper is devoted to the proof of Theorem 4. Instead of considering any affinely independent set $A$, we now assume that $a_0, \ldots, a_n \in \mathbb{R}^n$ are the corners of a regular $n$-simplex of side length $\sqrt{2n+2}$ centred at 0, and that $A = \{a_0, \ldots, a_n\}$. We further assume that the matrix with columns $a_0 - a_1, \ldots, a_0 - a_n$ has a positive determinant. Theorem 4 can then be reduced to this situation by composing $u$ with an affine transformation. The details are given on page 24 below.

As it is sometimes convenient to permute $a_0, \ldots, a_n$ cyclically, we regard 0, $\ldots$, $n$ as members of $\mathbb{Z}_{n+1} = \mathbb{Z}/(n+1)\mathbb{Z}$ in this context. Thus $a_{i+n+1} = a_i$.

The condition that our simplex has side length $\sqrt{2n+2}$ means that $|a_i| = \sqrt{n}$ for every $i \in \mathbb{Z}_{n+1}$. Indeed, by the calculations of Parks and Wills [12], the dihedral angle of the regular $n$-simplex is $\arccos \frac{1}{n}$. As each $a_i$ is orthogonal to one of the faces, this means that $a_i \cdot a_j = -\frac{1}{n}|a_i||a_j|$ for $i \neq j$, and therefore $2n+2 = |a_i - a_j|^2 = \frac{2n+2}{n}|a_i||a_j|$. From this we conclude that $|a_i| = \sqrt{n}$ for $i \in \mathbb{Z}_{n+1}$ and $a_i \cdot a_j = -1$ for $i \neq j$.

For $i \in \mathbb{Z}_{n+1}$, we now define the vector $\nu_i \in \mathbb{R}^{n+1}$ by

$$\nu_i = \frac{1}{\sqrt{n+1}} \begin{pmatrix} -a_i \\ 1 \end{pmatrix}.$$

Then

$$|\nu_i|^2 = \frac{|a_i|^2 + 1}{n+1} = 1,$$

whereas for $i \neq j$,

$$\nu_i \cdot \nu_j = \frac{a_i \cdot a_j + 1}{n+1} = 0.$$

Hence $(\nu_1, \ldots, \nu_{n+1})$ is an orthonormal basis of $\mathbb{R}^{n+1}$. (This is the reason why we choose $A$ as above.) Furthermore,

$$\det \begin{pmatrix} -a_1 & \cdots & -a_{n+1} \\ 1 & \cdots & 1 \end{pmatrix} = \det \begin{pmatrix} a_0 - a_1 & \cdots & a_0 - a_n & -a_0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} a_0 - a_1 & \cdots & a_0 - a_n \\ 0 & \cdots & 0 \end{pmatrix}.$$
(In the first step, we have used the fact that \( a_{n+1} = a_0 \) and subtracted the last column from each of the other columns of the matrix.) Hence the above assumption guarantees that the basis \( (\nu_1, \ldots, \nu_{n+1}) \) gives the standard orientation of \( \mathbb{R}^{n+1} \).

We now use the notation \( \lambda_i = \lambda_{a_i}, \) recalling that this is the linear function with \( \lambda_i(x) = a_i \cdot x \) for \( x \in \mathbb{R}^n \). For \( i \in \mathbb{Z}_{n+1} \), we set

\[ F_i = \{ x \in F : T_x u = \lambda_i \}. \]

Thus we have the disjoint decomposition

\[ F = \bigcup_{i \in \mathbb{Z}_{n+1}} F_i. \]

Furthermore, we define \( F^*_i = U(F_i) \).

Of course \( U : \mathbb{R}^n \to \text{graph}(u) \) is a bi-Lipschitz map. Thus in order to understand \( F, E, \) or \( F_i, \) it suffices to study \( F^*, E^*, \) or \( F^*_i \) and how \( U^{-1} \) transforms them. In particular, the following is true.

**Lemma 6.** For any Borel set \( \Omega \subseteq \mathbb{R}^n \),

\[ H^{n-1}(E^* \cap (\Omega \times \mathbb{R})) = \frac{\sqrt{n+1}}{2} H^{n-1}(E \cap \Omega) = \frac{1}{2} |D\nabla u|(\Omega). \]

**Proof.** We use the area formula [1, Theorem 2.91]. Hence we need to calculate the Jacobian of \( U \) restricted to the approximate tangent spaces of \( E \).

More precisely, since \( E \) is countably \( H^{n-1} \)-rectifiable, there exists an approximate tangent space \( T_x E \) at \( H^{n-1} \)-almost every \( x \in E \). Because \( U \) is Lipschitz continuous, the tangential derivative \( dU(x) \) exists at \( H^{n-1} \)-almost every \( x \in E \) [1, Theorem 2.90]. We write \( L^* \) for the adjoint of a linear operator \( L \). Then

\[ J_{E^*} U(x) = \sqrt{\det((dU(x))^* \circ dU(x))} \]

is the Jacobian of \( U \) at \( x \) with respect to \( T_x E \). The area formula implies that

\[ H^{n-1}(U(E \cap \Omega)) = \int_{E \cap \Omega} J_{E^*} U(x) dH^{n-1}(x). \]

Thus in order to prove the first identity, it suffices to show that

\[ J_{E^*} U(x) = \frac{\sqrt{n+1}}{2} \]

for \( H^{n-1} \)-almost every \( x \in E \).

To this end, consider \( x \in E \). Note that \( T_x E = (a_i - a_j)^\perp \) for some \( i, j \in \mathbb{Z}_{n+1} \) with \( i \neq j \) at \( H^{n-1} \)-almost every such point. For \( \xi \in (a_i - a_j)^\perp \), we know that

\[ \frac{1}{\rho}(u(x + \rho \xi) - u(x)) = u_{x,\rho}(\xi) \to T_x u(\xi) \]

as \( \rho \searrow 0 \). The convergence is in fact uniform on compact subsets of \( (a_i - a_j)^\perp \). Moreover, since \( T_x u = \lambda_i \land \lambda_j \) or \( T_x u = \lambda_i \lor \lambda_j \), its restriction to \( (a_i - a_j)^\perp \)
is linear with \( T_2u(\xi) = a_i \cdot \xi \). Hence \( d^F u(x) \) exists, and so does \( d^F U(x) \). We calculate
\[
d^F U(x)\xi = \begin{pmatrix} \xi \\ a_i \cdot \xi \end{pmatrix}.\]

For simplicity, we assume that \( i = n - 1 \) and \( j = n \). The space \((a_i - a_j)^\perp\) is spanned by the vectors \(a_0, \ldots, a_{n-2}\). Suppose that we choose an orthonormal basis \((\epsilon_0, \ldots, \epsilon_{n-2})\) of \(T_xE\). Let \( L : T_xE \to T_xE \) denote the linear operator that maps \( \epsilon_i \) to \( a_i \) for \( i = 0, \ldots, n-2 \). Then \( d^F U(x) \circ L \) is represented by the matrix
\[
M_1 = \begin{pmatrix} a_0 & \cdots & a_{n-2} \\ a_0 \cdot a_{n-1} & \cdots & a_{n-2} \cdot a_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & \cdots & a_{n-2} \\ -1 & \cdots & -1 \end{pmatrix}
\]
with respect to the above basis. Hence
\[
J_xU(x) = \sqrt{\frac{\det(M_1^T M_1)}{\det(L^* \circ L)}}.
\]

We write \( I_k \) for the identity \( k \times k \)-matrix. Then
\[
M_1^T M_1 = \begin{pmatrix} a_0 \cdot a_0 + 1 & \cdots & a_0 \cdot a_{n-2} + 1 \\ \vdots & \ddots & \vdots \\ a_{n-2} \cdot a_0 + 1 & \cdots & a_{n-2} \cdot a_{n-2} + 1 \end{pmatrix} = (n+1)I_{n-1}
\]
and \( \det(M_1^T M_1) = (n+1)^{n-1} \).

As \( L \) maps an \((n-1)\)-cube of side length 1 to the parallelepiped spanned by \( a_0, \ldots, a_{n-2} \), we know that \( \det(L^* \circ L) \) is the \((n-1)\)-volume of the latter. Thus if \( M_2 \) is the \( n \times (n-1) \)-matrix with columns \( a_0, \ldots, a_{n-2} \), then
\[
\det(L^* \circ L) = \det(M_2^T M_2).
\]

We further compute
\[
M_2^T M_2 = \begin{pmatrix} a_0 \cdot a_0 & \cdots & a_0 \cdot a_{n-2} \\ \vdots & \ddots & \vdots \\ a_{n-2} \cdot a_0 & \cdots & a_{n-2} \cdot a_{n-2} \end{pmatrix} = \begin{pmatrix} n & -1 & \cdots & -1 \\ -1 & n & \vdots \\ \vdots & \ddots & 0 \\ -1 & \cdots & -1 & n \end{pmatrix}.
\]

In order to calculate the determinant, we first subtract the first row of this matrix from each of the other rows. We obtain
\[
\det(M_2^T M_2) = \det \begin{pmatrix} n & -1 & \cdots & -1 \\ -(n+1) & n+1 & 0 & \cdots & 0 \\ -1 & n & \vdots \\ \vdots & \ddots & 0 \\ -1 & \cdots & 0 & n+1 \end{pmatrix} = (n+1)^{n-2} \det \begin{pmatrix} n & -1 & \cdots & -1 \\ -1 & n+1 & 0 & \cdots & 0 \\ -1 & 0 & \vdots \\ \vdots & \ddots & 0 \\ -1 & 0 & \cdots & 0 \end{pmatrix}.
\]
In the last matrix, we now add to the first row the sum of all the other rows. Thus

\[
\det(M^T M) = (n + 1)^{n-2} \det \begin{pmatrix}
2 & 0 & \cdots & 0 \\
-1 & 1 & 0 & \cdots & 0 \\
-1 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & 0 \\
-1 & 0 & \cdots & 0 & 1
\end{pmatrix} = 2(n + 1)^{n-2}.
\]

Hence

\[
J \mathcal{E}(x) = \sqrt{\det(M^T M_1)} = \sqrt{\frac{n + 1}{2}}.
\]

In order to prove the second identity, we recall that \( |a_i - a_j| = \sqrt{2(n + 2)} \) for \( i \neq j \). Hence \( |D \nabla u|(\Omega) = \sqrt{2n + 2} H^{n-1}(E \cap \Omega) = 2 H^{n-1}(E^* \cap (\Omega \times \mathbb{R})). \)

5 Slicing the graph

We still assume that \( A \) consists of the corners of the regular \( n \)-simplex from Section 4 and we assume that \( u \in BV^2_{\text{loc}}(\mathbb{R}^n) \) is bounded and satisfies \( \nabla u(x) \in A \) for almost every \( x \in \mathbb{R}^n \). In this section, we analyse the graph of \( u \). In particular, we examine intersections of graph(\( u \)) with hyperplanes perpendicular to one of the vectors \( \nu_i \). We will see that almost all such intersections can be represented as the graphs of functions in \( BV^2_{\text{loc}}(P) \), where

\[
P = \{ y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0 \},
\]

and with gradient taking one of \( n \) different values almost everywhere. That is, we have a function with properties similar to \( u \), but with an \( (n - 1) \)-dimensional domain. This observation will eventually make it possible to prove Theorem 4 with the help of an induction argument.

We use some tools from the author’s previous paper [10] in this section. Given \( i \in \mathbb{Z}^n_{n+1} \), let \( \Phi_i : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1} \) be the linear map with

\[
\Phi_i(x) = \begin{pmatrix}
\nu_i+1 \cdot x \\
\vdots \\
\nu_i+n+1 \cdot x
\end{pmatrix},
\]

so that \( \Phi_i(\nu_{i+k}) \) is the \( k \)-th standard basis vector in \( \mathbb{R}^{n+1} \). For \( t \in \mathbb{R} \), let

\[
\Gamma_i(t) = \left \{ y \in \mathbb{R}^n : \begin{pmatrix} y \\ t \end{pmatrix} \in \Phi_i(\text{graph}(u)) \right \}.
\]

This corresponds to the intersection of \( \text{graph}(u) \) with a hyperplane orthogonal to \( \nu_i \) after rotation by \( \Phi_i \), or in other words, a slice of \( \text{graph}(u) \).

We further define the functions

\[
\mathcal{G}_i(y) = \sup \left \{ t \in \mathbb{R} : u(t \nu_i + y_1 \nu_{i+1} + \cdots + y_n \nu_{i+n}) > \frac{t + y_1 + \cdots + y_n}{\sqrt{n + 1}} \right \}
\]
and
\[ \overline{g}_i(y) = \inf \left\{ t \in \mathbb{R} : u(t \nu_i + y_1 \nu_{i+1} + \cdots + y_n \nu_{i+n}) < \frac{t + y_1 + \cdots + y_n}{\sqrt{n+1}} \right\}. \]

Note that for a fixed \( y \in \mathbb{R}^n \), the set
\[ \left\{ t \in \mathbb{R} : u(t \nu_i + y_1 \nu_{i+1} + \cdots + y_n \nu_{i+n}) = \frac{t + y_1 + \cdots + y_n}{\sqrt{n+1}} \right\} \]
corresponds to the intersection of \( \text{graph}(u) \) with a line parallel to \( \nu_i \), so the functions \( g_i \) and \( \overline{g}_i \) tell us something about the geometry of \( \text{graph}(u) \) as well.

The following properties of \( g_i \) and \( \overline{g}_i \) have been proved elsewhere for \( n = 2 \) \[10, Lemma 16\]. The proof carries over to higher dimensions as well. We therefore do not repeat it here.

**Lemma 7.** For any \( i \in \mathbb{Z}_{n+1} \), the following statements hold true.

(i) The function \( g_i \) is lower semicontinuous and \( \overline{g}_i \) is upper semicontinuous.

(ii) The identity \( g_i = \overline{g}_i \) holds almost everywhere in \( \mathbb{R}^n \).

(iii) For any \( y \in \mathbb{R}^n \), the inequality \( g_i(y) \leq \overline{g}_i(y) \) holds true and
\[ \{y\} \times [g_i(y), \overline{g}_i(y)] \subseteq \Phi_i(\text{graph}(u)). \]

(iv) Let \( t \in \mathbb{R} \) and \( y \in \mathbb{R}^n \). Then \( y \in \Gamma_i(t) \) if, and only if, \( g_i(y) \leq t \leq \overline{g}_i(y) \).

(v) For all \( y \in \mathbb{R}^n \) and all \( \zeta \in (0, \infty)^n \), the inequality \( \overline{g}_i(y + \zeta) \leq g_i(y) \) is satisfied; and if equality holds, then
\[ g_i(y) = g_i(y + s\zeta) = \overline{g}_i(y + s\zeta) = \overline{g}_i(y + \zeta) \]
for all \( s \in (0, 1) \).

(vi) For all \( y \in \mathbb{R}^n \) and all \( \zeta \in [0, \infty)^n \), the inequalities \( g_i(y) \geq g_i(y + \zeta) \) and \( \overline{g}_i(y) \geq \overline{g}_i(y + \zeta) \) are satisfied.

Now consider the hyperplane \( P \subseteq \mathbb{R}^n \) given by
\[ P = \{y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0\} \]
and its unit normal vector
\[ \sigma = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n. \]

Let \( e_1, \ldots, e_n \) be the standard basis vectors of \( \mathbb{R}^n \) and define
\[ b_i = \sigma - \sqrt{n}e_i \]
for \( i = 1, \ldots, n \). Then
\[ |b_i|^2 = n - 1 \]
and
\[ b_i \cdot b_j = -1 \]
for \( i \neq j \). Hence \( b_1, \ldots, b_n \) are the corners of a regular \((n - 1)\)-simplex in \( P \) centred at 0 with side length \( \sqrt{2a} \). (Indeed the construction is similar to the standard \((n - 1)\)-simplex.) Thus they are the \((n - 1)\)-dimensional counterparts to \( a_0, \ldots, a_n \).

Given a function \( f: P \times \mathbb{R} \to \mathbb{R} \), we write \( \hat{\nabla} f \) for its gradient with respect to the variable \( p \in P \). We want to show the following.

**Proposition 8.** Let \( i \in \mathbb{Z}_{n+1} \). Then there exists a function \( f_i: P \times \mathbb{R} \to \mathbb{R} \) such that for almost every \( t \in \mathbb{R} \),

- the function \( p \mapsto f_i(p, t) \) belongs to \( \text{BV}^2_{\text{loc}}(P) \) and \( \hat{\nabla} f_i(p, t) \in \{ b_1, \ldots, b_n \} \) for \( \mathcal{H}^{n-1} \)-almost every \( p \in P \); and
- its graph is \( \Gamma_i(t) \), that is, \( \Gamma_i(t) = \{ p + f_i(p, t)\sigma: p \in P \} \).

Before we can prove this result, we need a few lemmas.

**Lemma 9.** Let \( i \in \mathbb{Z}_{n+1} \). Suppose that \( t \in \mathbb{R} \) and \( y_-, y_+ \in \Gamma_i(t) \). Then
\[
(y_- + [0, \infty)^n) \cap (y_+ - [0, \infty)^n) \subseteq \Gamma_i(t).
\]

**Proof.** We first prove that
\[
(y_- + (0, \infty)^n) \cap (y_+ - (0, \infty)^n) \subseteq \Gamma_i(t).
\]

Let
\[
y = (y_- + (0, \infty)^n) \cap (y_+ - (0, \infty)^n).
\]
Define \( \zeta_- = y - y_- \) and \( \zeta_+ = y_+ - y \). Then \( \zeta_, \zeta_+ \in (0, \infty)^n \). According to Lemma 4, this means that
\[
t \geq g(y_-) \geq \underline{g}(y_- + \zeta_-) = \underline{g}(y) \geq g(y) = g(y_+ - \zeta_+) \geq \underline{g}(y_+) \geq t.
\]
Hence \( y \in \Gamma_i(t) \). By the semicontinuity of \( g \) and \( \underline{g} \), we also conclude that
\[
g(y) \leq t \leq \underline{g}(t)
\]
for all \( y \in (y_- + [0, \infty)^n) \cap (y_+ - [0, \infty)^n) \).

**Lemma 10.** Let \( i \in \mathbb{Z}_{n+1} \). Let \( t \in \mathbb{R} \) and \( p \in P \). Suppose that
\[
\{ s \in \mathbb{R}: p + s\sigma \in \Gamma_i(t) \} = [s_, s_+].
\]
Then
\[
\Gamma_i(t) \cap (p + s_-\sigma - (0, \infty)^n) = \emptyset
\]
and
\[
\Gamma_i(t) \cap (p + s_+\sigma + (0, \infty)^n) = \emptyset.
\]

**Proof.** Let \( y \in p + s_-\sigma - (0, \infty)^n \). Choose \( s < s_- \) such that \( y \in p + s\sigma - (0, \infty)^n \) as well. Then Lemma 4 implies that
\[
g(y) \geq \underline{g}(p + s\sigma) \geq g(p + s\sigma) > t.
\]
Hence \( y \notin \Gamma_i(t) \). The proof of the second statement is similar.
Lemma 11. There exists a constant $C$ such that the following holds true. Suppose that $v : \mathbb{R}^n \to \mathbb{R}$ is smooth and bounded with $a_j \cdot \nabla v > -1$ for all $j \in \mathbb{Z}_{n+1}$ and $\sup_{\mathbb{R}^n} |v| \leq M$. Let $i \in \mathbb{Z}_{n+1}$. Let $\phi : P \times \mathbb{R} \to \mathbb{R}$ be the unique function such that

$$\left( p + \phi(p, t) \sigma \right) t \in \Phi_i(\text{graph}(v))$$

for $p \in P$ and $t \in \mathbb{R}$. Then

$$|\nabla \phi(p, t)| \leq \sqrt{n}$$

for all $p \in P$ and $t \in \mathbb{R}$. Moreover, for any $R > 0$,

$$\int_{-R}^R \int_{P \cap B_R(0)} |\nabla^2 \phi| d\mathcal{H}^{n-1} dt \leq C \int_{B_{C(M+2)}(0)} |\nabla^2 v| dx.$$

Since the proof of this statement is lengthy, we postpone it to the next section. We now prove Proposition 3.7.

Proof of Proposition 3.7. Let $t \in \mathbb{R}$ and $p \in P$. Since $u$ is bounded, the line

$$\left\{ t \nu_i + \sum_{k=1}^n (p_k + s \sigma_k) \nu_{i+k} : s \in \mathbb{R} \right\}$$

must intersect graph$(u)$. Hence there exists $s \in \mathbb{R}$ with $p + s \sigma \in \Gamma_i(t)$.

If there are $s_-, s_+ \in \mathbb{R}$ with $s_- < s_+$ such that $p + s_- \sigma \in \Gamma_i(t)$ and $p + s_+ \sigma \in \Gamma_i(t)$, then Lemma 3 implies that $\Gamma_i(t)$ has non-empty interior, denoted by $\tilde{\Gamma}_i(t)$. Because of Lemma 2(v) we know that $q(y) = \tilde{q}(y) = t$ for every $y \in \tilde{\Gamma}_i(t)$. Hence for $t_1 \neq t_2$, it follows that $\tilde{\Gamma}_i(t_1) \cap \tilde{\Gamma}_i(t_2) = \emptyset$. Therefore, there can only be countably many $t \in \mathbb{R}$ such that $\tilde{\Gamma}_i(t) \neq \emptyset$. For all other values, we see that $\Gamma_i(t)$ is a graph of a function over $P$. We denote this function by $f_i(\cdot, t)$.

We extend $f_i$ arbitrarily to the remaining values of $t$.

If $t$ is such that $\Gamma_i(t) = \emptyset$, then Lemma 10 shows that for every $y \in \Gamma_i(t)$, the set $\Gamma_i(t)$ is between the cones $y + (0, \infty)^n$ and $y - (0, \infty)^n$. It follows that $f_i(\cdot, t)$ is Lipschitz continuous.

Next we employ an approximation argument in conjunction with Lemma 11. Using a standard mollifier, we can find a sequence of smooth, uniformly bounded functions $v_k : \mathbb{R}^n \to \mathbb{R}$ such that $v_k \to u$ locally uniformly as $k \to \infty$ and $\lim \frac{\partial v_k}{\partial x_j}$ when $\Omega \subset \mathbb{R}^n$ is an open, bounded set with $|\nabla u| = 0$. It is then easy to modify $v_k$ such that in addition, it satisfies $a_j \cdot \nabla v_k > -1$ in $\mathbb{R}^n$ for every $j \in \mathbb{Z}_{n+1}$. Hence Lemma 11 applies to $v_k$.

From the above convergence, it follows that for any sequence of points $x_k \in \text{graph}(v_k)$, if $x_k \to x$ as $k \to \infty$, then $x \in \text{graph}(u)$. If we define $\phi_k$ as in Lemma 11 then for any fixed $t \in \mathbb{R}$, the functions $\phi_k(\cdot, t)$ are uniformly bounded in $C^0(\overline{P \cap B_R(0)})$ for any $R > 0$. Hence there is a subsequence that converges locally uniformly. If $t$ is such that $\Gamma_i(t)$ is the graph of $f_i(\cdot, t)$, then it is clear that the limit of any such subsequence must coincide with $f_i(\cdot, t)$. 

13
Hence in this case, we have the locally uniform convergence $\phi_k(\cdot, t) \to f_t(\cdot, t)$ as $k \to \infty$. The second inequality in Lemma 11 implies that

$$\limsup_{k \to \infty} \int_{-R}^{R} \int_{P \cap B_R(0)} |\nabla^2 \phi_k| d\mathcal{H}^{n-1} dt < \infty$$

for any $R > 0$. By Fatou’s lemma,

$$\int_{-R}^{R} \liminf_{k \to \infty} \int_{P \cap B_R(0)} |\nabla^2 \phi_k| d\mathcal{H}^{n-1} dt < \infty.$$  

Therefore, for almost every $t \in (-R, R)$, there exists a subsequence $(\phi_{k_n}(\cdot, t))_{n \in \mathbb{N}}$ converging to $f_t(\cdot, t)$ locally uniformly and such that

$$\limsup_{t \to \infty} \int_{P \cap B_R(0)} |\nabla^2 \phi_{k_n}| d\mathcal{H}^{n-1} < \infty.$$  

We conclude that $f_t(\cdot, t) \in \text{BV}_2^p(P)$ for almost all $t \in \mathbb{R}$.

We finally need to show that $\nabla f_t(p, t) \in \{b_1, \ldots, b_n\}$ for almost every $t \in \mathbb{R}$ and $\mathcal{H}^{n-1}$-almost every $p \in P$.

Consider the function $w_i : \mathbb{R}^n \to \mathbb{R}$ with

$$w_i(x) = \frac{u(x) - a_i \cdot x}{\sqrt{n + 1}}, \quad x \in \mathbb{R}^n.$$  

Then for every $t \in \mathbb{R}$,

$$\Gamma_i(t) \times \{t\} = \Phi_i(\{x \in \text{graph}(u) : x \cdot \nu_i = t\}) = \Phi_i \left( \left\{ \left( \frac{x}{u(x)} \right) : x \in \mathbb{R}^n \text{ with } w_i(x) = t \right\} \right).$$

Note further that $\mathcal{F}_i$ coincides up to an $\mathcal{H}^n$-null set with $\{x \in \mathbb{R}^n : \nabla w_i(x) = 0\}$. Let $Z \subset \mathbb{R}^n$ denote the set of all points where $u$ is not differentiable. By Rademacher’s theorem, this is an $\mathcal{H}^n$-null set. Hence the coarea formula gives

$$0 = \int_{\mathcal{F}_i \cup Z} |\nabla w_i| \, dx = \int_{-\infty}^{\infty} \mathcal{H}^{n-1}(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup Z)) \, dt.$$  

In particular, for almost all $t \in \mathbb{R}$,

$$\mathcal{H}^{n-1}(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup Z)) = 0.$$  

As the map $U$ (defined in the introduction) is Lipschitz continuous, we conclude that $U(w_i^{-1}(\{t\}) \cap (\mathcal{F}_i \cup Z))$ is an $\mathcal{H}^{n-1}$-null set, too. Therefore, for $\mathcal{H}^{n-1}$-almost all $y \in \Gamma_i(t)$, the unique point $x \in \mathbb{R}$ with

$$\Phi_i(U(x)) = \left( \frac{y}{t} \right)$$

belongs to $\mathbb{R}^n \setminus Z$ and satisfies $\nabla u(x) \in A \setminus \{a_i\}$.

To put it differently, for almost every $t \in \mathbb{R}$, the following holds true: for $\mathcal{H}^{n-1}$-almost every $p \in P$ the derivative of $u$ exists at the point

$$\Theta(p, t) = tv_a + \sum_{k=1}^{n} (p_k + f_i(p, t)\sigma_k)\nu_{i+k}.$$  

14
and belongs to $A \setminus \{a_i\}$. Furthermore, we know that $f_i(\cdot, t)$ is differentiable at $H^{n-1}$-almost every $p$ by Rademacher’s theorem. At a point $p \in P$ where both statements hold true, we can differentiate the equation

$$u(\Theta(p, t)) = \frac{t + \sqrt{n} f_i(p, t)}{\sqrt{n+1}}.$$  

(The right-hand side is the $(n+1)$-st component of

$$t \nu_i + \sum_{k=1}^{n} (p_k + f_i(p, t) \sigma_k) \nu_{i+k} = \Phi_i^{-1} \left( \frac{p + f_i(p, t) \sigma}{t} \right)$$

because $p \in P$ and by the definition of $\sigma$.) For any $\varpi \in P$, we thus obtain

$$-(n+1) \varpi_{j-i} - \frac{1}{\sqrt{n}} \varpi \cdot \nabla f_i(p, t) = \sqrt{n} \varpi \cdot \nabla f_i(p, t).$$

Hence

$$\varpi \cdot \nabla f_i(p, t) = -\sqrt{n} \varpi_{j-i} = b_{j-i} \cdot \varpi.$$  

We therefore conclude that $\nabla f_i(p, t) = b_{j-i}$ at such a point. \hfill \Box

## 6 Proof of Lemma 11

In this section we give the postponed proof of Lemma 11. To this end, we first need another lemma.

**Lemma 12.** Let $\Lambda$ denote the $(n \times n)$-matrix with columns

$$\sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik} a_i, \quad k = 1, \ldots, n.$$

Then

$$\det(\Lambda) = (-1)^n (n+1)^{\frac{n+1}{2}} \det \begin{pmatrix} \gamma_{01} & \cdots & \gamma_{0n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \end{pmatrix}.$$  

**Proof.** Let $M$ denote the $((n+1) \times (n+1))$-matrix with columns

$$\sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik} \nu_i, \quad k = 1, \ldots, n, \quad \text{and} \quad \sum_{i \in \mathbb{Z}_{n+1}} \nu_i.$$

Then, since $(\nu_1, \ldots, \nu_{n+1})$ is a positively oriented basis of $\mathbb{R}^{n+1}$, we conclude that

$$\det(M) = \det \begin{pmatrix} \gamma_{01} & \cdots & \gamma_{0n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \end{pmatrix}.$$  

15
On the other hand,
\[
M = \frac{1}{\sqrt{n+1}} \begin{pmatrix}
    0 & & \\
    -\Lambda & \ddots & 0 \\
    m_1 & \cdots & m_n + 1
\end{pmatrix},
\]
where \( m_k = \sum_{i \in \mathbb{Z}_{n+1}} \gamma_{ik} \). Hence
\[
\det(M) = (-1)^n(n + 1)^{n-1} \det(\Lambda).
\]
The claim follows immediately.

**Proof of Lemma 11.** First we note that by the assumptions on \( v \), the intersection of \( \text{graph}(v) \) with the hyperplane \( \{ \mathbf{x} \in \mathbb{R}^{n+1} : \mathbf{x} \cdot \nu_i = t \} \) is a smooth \((n-1)\)-dimensional manifold for every \( t \in \mathbb{R} \). Furthermore, the function \( \phi \) is smooth. If we define \( \Xi : P \times \mathbb{R}^2 \to \mathbb{R}^{n+1} \) such that
\[
\Xi(p, s, t) = t \nu_i + \sum_{k=1}^n (p_k + s \sigma_k) \nu_{i+k}
\]
for \( p \in P \) and \( s, t \in \mathbb{R} \), then \( \phi \) is characterised by the condition that
\[
\Xi(p, \phi(p, t), t) \in \text{graph}(v)
\]
for all \( t \in \mathbb{R} \) and \( p \in P \). Hence
\[
\nu(\Xi(p, \phi(p, t), t)) = \Xi_{n+1}(p, \phi(p, t), t). \tag{6}
\]
We now differentiate this equation.

We compute
\[
\frac{\partial \Xi}{\partial t} = \nu_i = - \frac{a_i}{\sqrt{n+1}}, \quad \frac{\partial \Xi_{n+1}}{\partial t} = \frac{1}{\sqrt{n+1}}.
\]
For \( \omega \in P \),
\[
\omega \cdot \tilde{\nabla} \Xi = - \frac{1}{\sqrt{n+1}} \sum_{k=1}^n \omega_k a_{i+k}, \quad \omega \cdot \tilde{\nabla} \Xi_{n+1} = \frac{1}{\sqrt{n+1}} \sum_{k=1}^n \omega_k = 0.
\]
Finally,
\[
\frac{\partial \Xi}{\partial s} = \sum_{k=1}^n \sigma_k \nu_{i+k} = - \frac{1}{\sqrt{n^2 + n}} \sum_{k=1}^n a_{i+k} = \frac{a_i}{\sqrt{n^2 + n}}, \quad \frac{\partial \Xi_{n+1}}{\partial s} = \frac{1}{\sqrt{n+1}}.
\]
We define \( \Theta(p, t) = \Xi(p, \phi(p, t), t) \). Differentiating (6), we now conclude that
\[
\left( \frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial t}(p, t) - 1 \right) a_i \cdot \nabla v(\Theta(p, t)) = \sqrt{n} \frac{\partial \phi}{\partial t}(p, t) + 1
\]
and
\[
\left( \frac{1}{\sqrt{n}} \omega \cdot \tilde{\nabla} \phi(p, t) a_i - \sum_{k=1}^n \omega_k a_{i+k} \right) \cdot \nabla v(\Theta(p, t)) = \sqrt{n} \omega \cdot \tilde{\nabla} \phi(p, t). \tag{7}
\]
Hence
\[ \frac{\partial \phi}{\partial t}(p, t) = \sqrt{n} \frac{a_i \cdot \nabla v(\Theta(p, t)) + 1}{a_i \cdot \nabla v(\Theta(p, t))} - n \]  \hspace{1cm} (8) \]

and
\[ \varpi \cdot \tilde{\nabla} \phi(p, t) = \sqrt{n} \sum_{k=1}^{n} \omega_k a_{i+k} \cdot \nabla v(\Theta(p, t)) \]

Fix \( t \in \mathbb{R} \) and \( p \in P \). Since \( \nabla v(\Theta(p, t)) \) is in the interior of the convex hull of the set \( \{ a_j : j \in \mathbb{Z}_{n+1} \} \), there exist \( \tau_j \in (0, 1) \) for \( j \in \mathbb{Z}_{n+1} \) such that
\[ \sum_{j \in \mathbb{Z}_{n+1}} \tau_j = 1 \]

and
\[ \nabla v(\Theta(p, t)) = \sum_{j \in \mathbb{Z}_{n+1}} \tau_j a_j. \]

Then
\[ a_i \cdot \nabla v(\Theta(p, t)) - n = n \tau_i - \sum_{j \neq i} \tau_j - n = (n + 1)(\tau_i - 1), \]

while
\[ \sum_{k=1}^{n} \omega_k a_{i+k} \cdot \nabla v(\Theta(p, t)) = \sum_{k=1}^{n} \omega_k \left( n \tau_{i+k} - \sum_{j \neq i+k} \tau_j \right) = (n + 1) \sum_{k=1}^{n} \omega_k \tau_{i+k}. \]

We further note that
\[ \tau_{i+1}^2 + \cdots + \tau_{i+n}^2 \leq (\tau_{i+1} + \cdots + \tau_{i+n})^2 = (1 - \tau_i)^2. \]

The Cauchy-Schwarz inequality therefore implies that
\[ \left| \sum_{k=1}^{n} \omega_k a_{i+k} \cdot \nabla v(\Theta(p, t)) \right| \leq (n + 1)(1 - \tau_i)|\varpi|. \]

It follows that
\[ |\varpi \cdot \tilde{\nabla} \phi(p, t)| \leq \sqrt{n}|\varpi|, \]

and inequality (9) is proved.

In order to prove the second statement of Lemma (1), we need to differentiate (7) again with respect to \( p \). We write \( \Lambda : M \) for the Frobenius inner product between two matrices \( \Lambda \) and \( M \). We also drop the arguments \((p, t)\) in the derivatives of \( \phi \) and in \( \Theta \). Then for all \( \varpi, \xi \in P^*, \)
\[ \sqrt{\frac{n+1}{n}} (\xi \otimes \varpi) : \tilde{\nabla}^2 \phi \]
\[ = \left( \frac{\xi \otimes \varpi}{\sqrt{n}} a_i - \sum_{k=1}^{n} \xi_k a_{i+k} \right) \otimes \left( \frac{\varpi \otimes \varpi}{\sqrt{n}} a_i - \sum_{k=1}^{n} \varpi_k a_{i+k} \right) : \nabla^2 v(\Theta). \]

As we have already seen that \( |\tilde{\nabla} \phi| \leq \sqrt{n} \), it follows that there is a constant \( C_1 = C_1(n) \) such that
\[ |\tilde{\nabla} \phi| \leq C_1 |\nabla^2 v(\Theta)|, \]

and inequality (10) is proved.
Choose an orthonormal basis \((\eta_1, \ldots, \eta_{n-1})\) of \(P\). Next we examine the derivative \(d\Theta\), and more specifically, its determinant.

Let \(\eta_1, \ldots, \eta_n\) denote the components of \(\eta_k\). For \(t \in \mathbb{R}\) and \(p \in P\), we also define
\[
\eta_{n+1,k}(p, t) = -\frac{1}{\sqrt{n}} \eta_k \cdot \hat{\nabla} \phi(p, t), \quad k = 1, \ldots, n-1,
\]
and
\[
\eta_{n+1,n}(p, t) = 1 - \frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial t}(p, t).
\]
Finally, we set \(\eta_{\ell n} = 0\) for \(\ell = 1, \ldots, n\). We compute
\[
\eta_k \cdot \hat{\nabla} \Theta(p, t) = \frac{1}{\sqrt{n}+1} \left( \frac{1}{\sqrt{n}} \eta_k \cdot \hat{\nabla} \phi(p, t) a_i - \sum_{\ell=1}^{n} \eta_{k \ell} a_{i+\ell} \right)
\]
and
\[
\frac{\partial \Theta}{\partial t}(p, t) = \left( \frac{1}{\sqrt{n}} \frac{\partial \phi}{\partial t}(p, t) - 1 \right) \frac{a_i}{\sqrt{n}+1}.
\]
Hence we can represent \(d\Theta\) by the matrix with columns
\[
-\sum_{\ell=1}^{n+1} \eta_{k \ell} a_{i+\ell}, \quad k = 1, \ldots, n,
\]
with respect to the basis of \(P \times \mathbb{R}\) generated by \(\eta_1, \ldots, \eta_{n-1}\). Lemma \([12]\) now tells us that
\[
\det(d\Theta) = \pm \frac{1}{\sqrt{n}+1} \det \begin{pmatrix}
\eta_{11} & \cdots & \eta_{1n} & 1 \\
\vdots & \ddots & \vdots & \vdots \\
\eta_{n+1,1} & \cdots & \eta_{n+1,n} & 1
\end{pmatrix}
\]
\[
= \pm \frac{1}{\sqrt{n}+1} \det \begin{pmatrix}
\eta_{11} & \cdots & \eta_{1,n-1} & 0 & 1 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
\eta_{n1} & \cdots & \eta_{n,n-1} & 0 & 1 \\
\eta_{n+1,1} & \cdots & \eta_{n+1,n-1} & \eta_{n+1,n} & 1
\end{pmatrix}
\]
\[
= \pm \sqrt{n+1} \eta_{n+1,n} \det \begin{pmatrix}
\eta_{11} & \cdots & \eta_{1,n-1} & \sigma_1 \\
\vdots & \ddots & \vdots & \vdots \\
\eta_{n1} & \cdots & \eta_{n,n-1} & \sigma_n
\end{pmatrix}.
\]
As \((\eta_1, \ldots, \eta_{n-1}, \sigma)\) form an orthonormal basis of \(\mathbb{R}^n\), we find that
\[
|\det(d\Theta)| = \sqrt{n+1} |\eta_{n+1,n}| = \frac{1}{\sqrt{n}+1} \sqrt{n} - \frac{\partial \phi}{\partial t}.
\]
Recalling \([8]\), we now obtain
\[
|\det(d\Theta)| = \frac{\sqrt{n^2+n}}{n - a_i \cdot \nabla v(\Theta)}.
\]
We also note that the map Θ is injective. Given $R > 0$, we therefore compute

$$\int_{-R}^{R} \int_{P \cap B_R(0)} |\nabla^2 \phi| \, dH^{n-1} \, dt \leq C_1 \int_{-R}^{R} \int_{P \cap B_R(0)} \frac{|\nabla^2 v|}{n - a_i} \, dH^{n-1} \, dt \leq C_1 \sqrt{n^2 + n} \int_{\Theta((P \cap B_R(0)) \times (-R,R))} |\nabla^2 v| \, dx.$$  

It remains to examine the set $\Theta((P \cap B_R(0)) \times (-R,R))$. Recall that we have the assumption $\sup_{R^n} |v| \leq M$ in Lemma 11. Thus (6) implies that

$$|\Xi_{n+1}(p, \phi(p, t), t)| \leq M.$$  

Since

$$\Xi_{n+1}(p, \phi(p, t), t) = \frac{t + \sqrt{n} \phi(p, t)}{\sqrt{n + 1}},$$

this means that

$$|\phi(p, t)| \leq M \sqrt{\frac{n + 1}{n} + \frac{R}{\sqrt{n}}}$$

when $t \in (-R, R)$. Hence there exists a constant $C_2 = C_2(n)$ such that

$$|\Theta(p, t)| \leq C_2(M + R)$$

for all $p \in P \cap B_R(0)$ and all $t \in (-R, R)$. Thus (10) implies the second inequality of Lemma 11.

### 7 Proof of Theorem 4

In this section we combine the previous results to prove the second main theorem. We first consider a function $u \in BV^2_{loc}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ such that graph($u$) is close to the graph of $\lambda_i \wedge \lambda_j$ or $\lambda_i \vee \lambda_j$ in a cube in $\mathbb{R}^{n+1}$ with edges parallel to $\nu_1, \ldots, \nu_{n+1}$. We will give a condition which implies that such a function actually coincides with $\lambda_i \wedge \lambda_j$ or $\lambda_i \vee \lambda_j$ up to a constant in part of the domain.

For $i, j \in \mathbb{Z}_{n+1}$ with $i \neq j$ and for $r, R > 0$, we define

$$Q_{ij}(r, R) = \left\{ \sum_{k \in \mathbb{Z}_{n+1}} c_k \nu_k : c_i, c_j \in (-r, r) \text{ and } c_k \in (-R, R) \text{ for } k \notin \{i, j\} \right\}.$$  

Again we consider the map $U : \mathbb{R}^n \to \mathbb{R}^{n+1}$ with $U(x) = (\frac{x}{u(x)})$ for $x \in \mathbb{R}^n$. The following is the key statement for the proof of Theorem 4.

**Proposition 13.** Let $n \in \mathbb{N}$. For any $\delta > 0$ there exist $\epsilon > 0$ with the following properties. Let $i, j \in \mathbb{Z}_{n+1}$ with $i \neq j$. Suppose that $|u(0)| \leq \epsilon$ and either

$$|u - \lambda_i \wedge \lambda_j| \leq \epsilon \text{ in } U^{-1}(Q_{ij}(1, 1))$$  

(10)
\[ |u - \lambda_i \vee \lambda_j| \leq \epsilon \quad \text{in } U^{-1}(Q_{ij}(1,1)). \]  \hspace{1cm} (11)

Then
\[ \mathcal{H}^{n-1}(E^* \cap Q_{ij}(\frac{1}{2},1)) \geq 2^{n-1}(1 - \delta). \]  \hspace{1cm} (12)

If, in addition,
\[ \mathcal{H}^{n-1}(E^* \cap Q_{ij}(1,1)) \leq 2^{n-1}(1 + \epsilon), \]  \hspace{1cm} (13)

then there exist \( \alpha, \beta \in \mathbb{R} \) such that
\[ u = (\lambda_i + \alpha) \wedge (\lambda_j + \beta) \quad \text{in } U^{-1}(Q_{ij}(\frac{1}{2}, \frac{1}{2})). \]  \hspace{1cm} (14)

Let

\[ \text{Lemma 15.} \]

Then
\[ \mathcal{H}^{n-1}(E^* \cap Q_{ij}(1,1)) \leq 2^{n-1}(1 + \epsilon). \]

Proof. Let \( p \in P \) and \( t \in \mathbb{R} \). Set
\[ \mathcal{L}^{-1} \left( p + f_i(p,t)\sigma \right). \]

If \( x \in E^* \), then Proposition 13 implies that graph(u) coincides with a hyperplane in a neighbourhood of x. If that hyperplane is perpendicular to \( \nu_i \), then \( p + f_i(p,t)\sigma \in \Gamma_i(t) \) and \( t \) belongs to the null set identified in Proposition 8. Otherwise, the function \( f_i(\cdot, t) \) is affine near \( p \), and hence \( \Phi_i(x) \) cannot belong to \( D_i'(t) \times \{ t \} \). This implies the first claim.

The second claim is now a consequence of the coarea formula [1, Theorem 2.93].

\[ \text{Let } k \in \{ 1, \ldots, n \}. \] Suppose that \( z \in \mathbb{R}^{n-1} \) with \( z_s < \pi \). For \( z \in \mathbb{R}^{n-1} \), define \( \ell_k(s) = (z_1, \ldots, z_k-1, s, z_{k+1}, \ldots, z_n) \) for \( s \in [z, \pi] \) and \( L_z = \{ \ell_k(s): \underline{z} \leq s \leq \pi \} \). Fix \( i \in \mathbb{N} \). Then for \( \mathcal{H}^{n-1} \)-almost every \( z \in \mathbb{R}^{n-1} \), either
\[ q_{\underline{z}}(y) = \Phi_i(y) = q_{\underline{z}}(y') = \Phi_i(y') \]

for all \( y, y' \in L_z \), or there exist \( y \in L_z \times \mathbb{R} \) such that
\[ \Phi_i(\ell_z(\pi)) \leq y_{n+1} \leq g(\ell_z(\underline{z})) \]

and \( y \in \Phi_i(E^*) \).
Proof. Consider the projection $\Pi: \mathbb{R}^{n+1} \to \mathbb{R}^n$ given by $\Pi(y) = y$ for $y \in \mathbb{R}^n$. Set $\Psi_i = \Pi \circ \Phi_i$. Then for $j \in \mathbb{Z}_{n+1}$ with $j \neq i$ and for $x \in F_j^*$, it is clear that $J_{\mathcal{F}_i} \Psi_i(x) = 0$. Hence the area formula gives $\mathcal{H}^n(\Psi_i(F_j^*)) = 0$. This means that for $\mathcal{H}^{n-1}$-almost every $z \in \mathbb{R}^{n-1}$,

$$\mathcal{H}^1(L_z \cap \Psi_i(F_j^*)) = 0 \quad (16)$$

for all $j \neq i$. Furthermore, since $\mathcal{E}^*$ is an $\mathcal{H}^{n-1}$-rectifiable set and $\mathcal{H}^{n-1}(\mathcal{E}^*) = 0$, we also know that for $\mathcal{H}^{n-1}$-almost every $z \in \mathbb{R}^{n-1}$,

$$\mathcal{H}^1(L_z \cap \Psi_i(\mathcal{E}^*)) = 0 \quad (17)$$

and

$$L_z \cap \Phi^*_i = \emptyset. \quad (18)$$

Consider a point $z \in \mathbb{R}^{n-1}$ such that (16), (17), and (18) hold true. Recall that by Lemma 7, a point $y \in \mathbb{R}^{n+1}$ belongs to $\Phi_i(\text{graph}(u))$ if, and only if, $

$$g_i(y) \leq y_{n+1} \leq \overline{g}_i(y).$$

Also recall that

$$\text{graph}(u) = \mathcal{E}^* \cup \mathcal{N}^* \cup \bigcup_{j \in \mathbb{Z}_{n+1}} F_j^*.$$ 

From (16)–(18) we therefore infer that for $\mathcal{H}^1$-almost all $y \in L_z$,

$$(y, t) \in \Phi_i(F_i^*) \quad \text{for all } t \in [g_i(y), \overline{g}_i(y)]. \quad (19)$$

Consider $y \in \Phi_i(F_i^*)$ with $y \in L_z$. Then, setting $x = \Phi_i^{-1}(y)$, we have the locally uniform convergence $u_{x, \rho} \to \lambda_i$ as $\rho \to 0$. Hence for any compact set $K \subseteq \mathbb{R}^{n+1}$ and any $\epsilon > 0$ there exists $\rho_0 > 0$ such that

$$\frac{1}{\rho}(\text{graph}(u) - x) \cap K \subseteq \{ \tilde{x} \in \mathbb{R}^{n+1} : \text{dist}(\tilde{x}, \text{graph}(\lambda_i)) < \epsilon/2 \}$$

for all $\rho \in (0, \rho_0]$. Recall that $e_1, \ldots, e_n$ are the standard basis vectors in $\mathbb{R}^n$. It follows that there exists $r_0 > 0$ such that for all $r \in (0, r_0]$,

$$|g_i(y \pm re_i) - g_i(y)| \leq \epsilon \quad \text{and} \quad |\overline{g}_i(y \pm re_i) - \overline{g}_i(y)| \leq \epsilon$$

and $|g_i(y) - \overline{g}_i(y)| \leq \epsilon$. Thus

$$\frac{\partial}{\partial y_k} g_i(y) = 0 \quad \text{and} \quad \frac{\partial}{\partial y_k} \overline{g}_i(y) = 0$$

and $g_i(y) = \overline{g}_i(y)$. Since this is true for $\mathcal{H}^1$-almost all $y \in L_z$, Lemma 7(vi) implies that

$$g_i(L_z) \geq \overline{g}_i(L_z) \geq g_i(\ell_2(\mathcal{N}))$$

(20)

for all $y \in L_z$.

If (19) holds for all $y \in L_z$, then we immediately conclude that $g_i$ and $\overline{g}_i$ are constant and coincide on $L_z$, i.e., we have the first alternative from the statement of the lemma. If there exists $y \in L_z$ such that (19) does not hold true, then by the above observations, we know that

$$(y, t) \notin \Phi_i(F_i^*)$$

21
holds in fact for all \( t \in [g_0(y), \overline{g}_i(y)] \). Moreover, because \((19)\) still holds true almost everywhere on \( L_z \), there exists a sequence \((\tilde{y}_m)_m \in \mathbb{N} \) in \( L_z \) such that 
\[ y = \lim_{m \to \infty} \tilde{y}_m \]
and such that \((19)\) holds for every \( \tilde{y}_m \). We may then choose \( \tilde{t}_m \) such that 
\[ \tilde{t}_m \in [g_0(\tilde{y}_m), \overline{g}_i(\tilde{y}_m)] \]. Extracting a subsequence if necessary, we may assume that 
\[ y_{n+1} = \lim_{m \to \infty} \tilde{t}_m \]exists. Set \( y = (y_{n+1}) \). Then \( \Phi_i^{-1}(y) \) belongs to the boundary of \( \mathcal{F}_n^* \) relative to \( \text{graph}(u) \).

Proposition \([16]\) implies that \( \mathcal{F}_n^* \) is an open set relative to \( \text{graph}(u) \), and its relative boundary is contained in \( \mathcal{E}^* \cup \mathcal{N}^* \). Because of \((18)\), it follows that \( \Phi_i^{-1}(y) \in \mathcal{E}^* \). Moreover, \((20)\) implies that
\[ \overline{g}_i(\ell_z(x)) \leq y_{n+1} \leq g_0(\ell_z(x)). \]
Thus \( y \) has the properties from the second alternative in the statement.

**Lemma 16.** Let \( i \in \mathbb{Z}_{n+1} \). Suppose that \( G \subseteq \mathbb{R}^n \) is a connected set such that \( G \cap \Gamma_i(t) = \emptyset \) for all \( t \in (-1,1) \). Then either \( g_0(y) \geq 1 \) for all \( y \in G \) or \( \overline{g}_i(y) \leq -1 \) for all \( y \in G \).

**Proof.** Assume that there exists \( y_0 \in G \) such that \( g_0(y_0) < 1 \). Since \( G \cap \Gamma_i(t) = \emptyset \) for all \( t \in (-1,1) \), this implies that
\[ -1 \geq \overline{g}_i(y_0) \geq g_0(y_0) \]
by Lemma \([iv]\).

Given \( t \in (-1,1) \), define
\[ H_t = \{ y \in G : \overline{g}_i(y) \geq t \}. \]
Because \( \overline{g}_i \) is upper semicontinuous by Lemma \([iv]\) this is a closed set relative to \( G \). Moreover, if \( y \in H_t \), it follows that
\[ \overline{g}_i(y) \geq g_0(y) \geq 1, \]
because \( G \cap \Gamma_i(t') = \emptyset \) for all \( t' \in (-1,1) \). By the lower semicontinuity of \( g_0 \), this means that there exists \( \rho > 0 \) such that \( \overline{g}_i \geq g_0 \geq \rho \) in \( B_{\rho}(y) \). Hence \( H_t \) is also open relative to \( G \). Since \( G \) is connected and \( y_0 \notin H_t \), it follows that \( H_t = \emptyset \). This is true for all \( t \in (-1,1) \), so \( \overline{g}_i(y) \leq -1 \) for all \( y \in G \).

We now have everything in place for the proof of Proposition \([13]\).

**Proof of Proposition \([13]\).** We use induction over \( n \). The statement is clear for \( n = 1 \). We now assume that \( n \geq 2 \) and the statement holds true for \( n - 1 \).

For simplicity, we assume that \( i = 1 \) and \( j = 2 \). We also assume that \((10)\) holds true; the proof is similar under the assumption \((11)\).

Let
\[ \Lambda = (\{0\} \times (-\infty,0] \times \mathbb{R}^{n-2}) \cup (\{0\} \times (0,\infty) \times \mathbb{R}^{n-2}) \]
Then
\[ \Phi_0(\text{graph}(\lambda_1 \cap \lambda_2)) = \Lambda \times \mathbb{R}. \]

Let
\[ \varepsilon' = \varepsilon \sqrt{\frac{n}{n+1}}. \]
Under the assumptions of the proposition, the set \( \Phi_0(\text{graph}(u)) \cap (-1,1)^n \) is between \((\Lambda - \epsilon') \times \mathbb{R}\) and \((\Lambda + \epsilon') \times \mathbb{R}\), i.e.,

\[
\Phi_0(\text{graph}(u)) \cap (-1,1)^{n+1} \subseteq \bigcup_{-\epsilon' \leq s \leq \epsilon'} (\Lambda + s) \times \mathbb{R}.
\]

Set \( s_0 = \sqrt{\frac{4}{n+1}} u(0) \). Then \(|s_0| \leq \epsilon'\) by the assumption that \(|u(0)| \leq \epsilon\).

Moreover, we compute

\[
\Phi_0 \left( \frac{0}{u(0)} \right) = \frac{u(0)}{\sqrt{n+1}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = s_0 \left( \frac{\sigma}{\sqrt{n}} \right).
\]

Assuming that \( \epsilon < \sqrt{n+1} \), we infer that \( \tau_{\sigma}(s_0\sigma) > -1 \) and \( g_{\sigma}(s_0\sigma) < 1 \). Using Lemma \([7](v)\) and Lemma \([16]\), we conclude that

\[
g_{\sigma}(y) \geq 1 \quad \text{for} \ y \in (-1,1)^n \cap \bigcup_{s < -\epsilon'} (\Lambda + s) \sigma
\]

and

\[
g_{\sigma}(y) \leq -1 \quad \text{for} \ y \in (-1,1)^n \cap \bigcup_{s > \epsilon'} (\Lambda + s) \sigma.
\]

Now consider the function \( f_0 : P \times \mathbb{R} \to \mathbb{R} \) from Proposition \([5]\). For almost every \( t \in (-1,1) \), the graph of \( f_0(\cdot,t) \), which is given by \( \Gamma_0(t) \), is between \( \Lambda - \epsilon' \) and \( \Lambda + \epsilon' \) in the hypercube \((-1,1)^n\).

Define \( \mu_1, \mu_2 : P \to \mathbb{R} \) by \( \mu_1(p) = b_1 \cdot p \) and \( \mu_2(p) = b_2 \cdot p \) for \( p \in P \) (where \( b_1 \) and \( b_2 \) are the vectors defined on page \([11]\)). Let \( F_t : P \to \mathbb{R}^n \) be the map with \( F_t(p) = p + f_0(p,t)\sigma \) for \( p \in P \). Then it follows that

\[
|f_0(\cdot,t) - \mu_1 \wedge \mu_2| \leq \epsilon' \quad \text{in} \ F_t^{-1}((-1,1)^n).
\]

Moreover, the condition \(|f_0(0,t)| \leq \epsilon'\) is clearly satisfied. Hence we may apply the induction hypothesis to the function \( f_0(\cdot,t) \). We thereby obtain the inequality

\[
\mathcal{H}^{n-2}(\mathcal{D}_0^1(t) \cap \{(\beta_1, \beta_2) \in (-1,1)^{n-2}\}) \geq 2^{n-2}(1 - \delta) \tag{21}
\]

for almost all \( t \in (-1,1) \), provided that \( \epsilon \) is sufficiently small. Using Lemma \([14]\) we therefore obtain inequality \([12]\). This proves the first statement of Proposition \([13]\).

In order to prove the second statement, assume now that \([13]\) holds true. Then

\[
\int_{-1}^1 \mathcal{H}^{n-2}(\mathcal{D}_0^1(t) \cap (-1,1)^n) \, dt \leq 2^{n-1}(1 + \epsilon).
\]

Recall that we also have inequality \([21]\), and we may now assume that \( \delta \) is arbitrarily small. Hence there exist \( t_- \in (-1, -\frac{1}{2}) \) and \( t_+ \in (\frac{1}{2}, 1) \) such that

\[
\mathcal{H}^{n-2}(\mathcal{D}_0^1(t_{\pm}) \cap (-1,1)^n) \leq 2^{n-2}(1 + 3\delta + 4\epsilon).
\]

By the induction hypothesis, if \( \delta \) and \( \epsilon \) are sufficiently small, then

\[
f_0(\cdot,t_{\pm}) = (\mu_1 + \alpha_{\pm}) \wedge (\mu_2 + \beta_{\pm}) \quad \text{in} \ F_t^{-1}((-1,1)^n)
\]

in
for certain numbers $\alpha_-, \alpha_+, \beta_-, \beta_+ \in \mathbb{R}$. Therefore, there exist $y_-, y_+ \in \mathbb{R}^2 \times \{0\}^{n-2}$ such that

$$\Gamma_0(t_{\pm}) \cap (-\frac{1}{2}, \frac{1}{2})^n = (y_{\pm} + A) \cap (-\frac{1}{2}, \frac{1}{2})^n.$$ Clearly, by the above observations on $\Phi_0(\text{graph}(u))$, this implies that $y_{\pm} \in B_r(0)$. We assume that $\epsilon' \leq \frac{1}{4}$.

If $y_- = y_+$, then by Lemma 7

$$\Gamma_0(t_{\pm}) \cap (-\frac{1}{2}, \frac{1}{2})^n = (y_+ + A) \cap (-\frac{1}{2}, \frac{1}{2})^n$$

for every $t \in (t_-, t_+)$ as well. In this case, we conclude that (13) holds true. Thus it now suffices to show that $y_+ = y_-$. We argue by contradiction here. Suppose that $y_+ \neq y_-$. We assume that in fact the first components $y_{1-}$ and $y_{1+}$ are different. The arguments are similar if $y_{2-} \neq y_{2+}$.

If $y_{1-} \neq y_{1+}$, then for any $z \in (-\frac{1}{2}, -\frac{1}{4}) \times (-\frac{1}{2}, \frac{1}{2})^{n-2}$, it follows that

$$g_0 \left( \frac{y_{1-}}{z} \right) \leq t_- \leq g_0 \left( \frac{y_{1-}}{z} \right)$$

and

$$g_0 \left( \frac{y_{1+}}{z} \right) \leq t_+ \leq g_0 \left( \frac{y_{1+}}{z} \right).$$

Since $t_- < t_+$, it is therefore not true that $g_2 \leq g_0$ and $g_0$ are constant with $g_0 = g_0$ on $[y_{1+}, y_{1-}] \times \{z\}$. Lemma 13 now implies that for $H^{n-1}$-almost every $z \in (-\frac{1}{2}, -\frac{1}{4}) \times (-\frac{1}{2}, \frac{1}{2})^{n-2}$, the set $[y_{1+}, y_{1-}] \times \{z\} \times [t_-, t_+]$ intersects $\Phi_0(\mathcal{E}^*)$. It follows that

$$H^{n-1}(\Phi_0(\mathcal{E}^*) \cap ((-1, 1) \times (-\frac{1}{2}, -\frac{1}{4}) \times (-1, 1)^{n-1})) \geq \frac{1}{4}.$$ 

Furthermore, because of (12), we obtain the estimate

$$H^{n-1}(\mathcal{E}^* \cap Q_{12}(1, 1)) \geq 2^{n-1}(1 - \delta) + \frac{1}{4}.$$ If $\delta + \epsilon < 2^{-n-1}$, then this contradicts the hypothesis.

Finally we can prove the second main result with the help of Proposition 5 and Proposition 12.

**Proof of Theorem 3** Suppose that $A \subseteq \mathbb{R}^n$ is affinely independent. Then $A$ contains at most $n + 1$ elements. If there are fewer, then we can add additional elements to $A$ such that it remains affinely independent. Thus we may assume without loss of generality that the size of $A$ is exactly $n + 1$.

Now suppose that $A = \{\tilde{a}_0, \ldots, \tilde{a}_n\}$. Consider $M \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}^n$ such that $Ma_i + c = a_i$ for $i = 0, \ldots, n$. Then the function $v: \mathbb{R}^n \to \mathbb{R}$ with $v(x) = u(M^Tx + c)$ has the property that $\nabla v(x) \in \{a_0, \ldots, a_n\}$ for almost all $x \in \mathbb{R}^n$. Hence we may assume that $A$ consists of the vectors $a_0, \ldots, a_n$.

Now for the sets $\mathcal{F}$, $\mathcal{E}$, and $\mathcal{N}$ as defined in Section 4, Proposition 5 implies that $\mathcal{F} \subseteq \mathcal{R}(u)$ with the same arguments as in the proof of Theorem 3.

For $x \in \mathcal{E}$, the functions $u_{x, \rho}$ converge locally uniformly to $\lambda_i \land \lambda_j$ or to $\lambda_i \lor \lambda_j$ as $\rho \searrow 0$ for some $i, j \in \mathbb{Z}_{n+1}$ with $i \neq j$. Moreover, the approximate
tangent space of $\mathcal{E}^*$ exists at the point $U(x)$. Clearly this approximate tangent space is $\text{graph}(\lambda_i) \cap \text{graph}(\lambda_j)$. Hence for $\rho$ sufficiently small, the function $u_{x,\rho}$ satisfies the hypotheses of Proposition 13, including (13). It follows that $u_{x,\rho}$ satisfies (14) or (15). In particular, it is regular near 0, and hence $x \in \mathcal{R}(u)$.

Thus $\mathcal{S}(u) \subseteq \mathcal{N}$, which is an $\mathcal{H}^{n-1}$-null set.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000.

[2] J. M. Ball and R. D. James, *Fine phase mixtures as minimizers of energy*, Arch. Rational Mech. Anal. 100 (1987), 13–52.

[3] M. Cicalese, M. Forster, and G. Orlando, *Variational analysis of a two-dimensional frustrated spin system: emergence and rigidity of chirality transitions*, SIAM J. Math. Anal. 51 (2019), 4848–4893.

[4] S. Conti, I. Fonseca, and G. Leoni, *A $\Gamma$-convergence result for the two-gradient theory of phase transitions*, Comm. Pure Appl. Math. 55 (2002), 857–936.

[5] E. Davoli and M. Friedrich, *Two-well linearization for solid-solid phase transitions*, arXiv:2005.03892 [math.AP], 2020.

[6] E. Davoli and M. Friedrich, *Two-well rigidity and multidimensional sharp-interface limits for solid-solid phase transitions*, Calc. Var. Partial Differential Equations 59 (2020), Article No. 44.

[7] F. Liu and H. Metiu, *Dynamics of phase separation of crystal surfaces*, Phys. Rev. B 48 (1993), 5808.

[8] L. Modica and S. Mortola, *Il limite nella $\Gamma$-convergenza di una famiglia di funzionali ellittici*, Boll. Un. Mat. Ital. A (5) 14 (1977), 526–529.

[9] L. Modica and S. Mortola, *Un esempio di $\Gamma^{-}$-convergenza*, Boll. Un. Mat. Ital. B (5) 14 (1977), 285–299.

[10] R. Moser, *Structure and rigidity of functions in $\text{BV}^2_{\text{loc}}(\mathbb{R}^2)$ with gradients taking only three values*, Proc. Lond. Math. Soc. 116 (2018), 813–846.

[11] S. Müller, *Variational models for microstructure and phase transitions*, Calculus of variations and geometric evolution problems (Cetraro, 1996), Lecture Notes in Math., vol. 1713, Springer, Berlin, 1999, pp. 85–210.

[12] H. R. Parks and D. C. Wills, *An elementary calculation of the dihedral angle of the regular $n$-simplex*, Amer. Math. Monthly 109 (2002), 756–758.

[13] J. Stewart and N. Goldenfeld, *Spinodal decomposition of a crystal surface*, Phys. Rev. A 46 (1992), 6505–6512.

[14] S. J. Watson and S. A. Norris, *Scaling theory and morphometrics for a coarsening multiscale surface, via a principle of maximal dissipation*, Phys. Rev. Lett. 96 (2006), 176103.