Calc. Var. (2021) 60:86
https://doi.org/10.1007/s00526-021-01981-z

Calculus of Variations

On the constancy theorem for anisotropic energies through differential inclusions

Jonas Hirsch[1] · Riccardo Tione[2]

Received: 7 December 2020 / Accepted: 5 April 2021
© The Author(s) 2021

Abstract
In this paper we study stationary graphs for functionals of geometric nature defined on currents or varifolds. The point of view we adopt is the one of differential inclusions, introduced in this context in the recent papers (De Lellis et al. in Geometric measure theory and differential inclusions, 2019. arXiv:1910.00335; Tione in Minimal graphs and differential inclusions. Commun Part Differ Equ 7:1–33, 2021). In particular, given a polyconvex integrand f, we define a set of matrices C_f that allows us to rewrite the stationarity condition for a graph with multiplicity as a differential inclusion. Then we prove that if f is assumed to be non-negative, then in C_f there is no T'_N configuration, thus recovering the main result of De Lellis et al. (Geometric measure theory and differential inclusions, 2019. arXiv:1910.00335) as a corollary. Finally, we show that if the hypothesis of non-negativity is dropped, one can not only find T'_N configurations in C_f, but it is also possible to construct via convex integration a very degenerate stationary point with multiplicity.

Mathematics Subject Classification 35B65 · 49Q15 · 49Q20

1 Introduction

In this paper we continue the study started in [5,26] of functionals arising from geometric variational problems from the point of view of differential inclusions. The energies we consider are of the form

$$\Sigma_\Psi(T) \doteq \int_E \Psi(\vec{T}(x))\theta(x)d\mathcal{H}^m(x), \tag{1.1}$$

Communicated by J. M. Ball.

✉ Riccardo Tione
riccardo.tione@epfl.ch

Jonas Hirsch
hirsch.jonas@math.uni-leipzig.de

1 Mathematisches Institut, Universität Leipzig, Augustusplatz 10, 04109 Leipzig, Germany

2 EPFL, B, Station 8, 1015 Lausanne, Switzerland

Published online: 21 April 2021

defined on m-dimensional rectifiable currents (resp. varifolds) $T = [\![E, \vec{T}, \theta]\!]$ of $\Omega \times \mathbb{R}^n$, where $\Omega \subset \mathbb{R}^m$ is a convex and bounded open set, and the integrand Ψ is defined on the oriented (resp. non-oriented) Grassmanian space. In order to keep the technicalities at a minimum level, we defer all the definitions of these geometric objects to Section A. The main interest is the regularity of stationary points for energies as in (1.1) satisfying suitable *ellipticity* conditions. From the celebrated regularity theorem of Allard of [2], it is known that an *ε-regularity theorem* holds for stationary points of the area functional, namely the case in which $\Psi \equiv 1$. Since then, the question of extending this result to more general energies has been an important open problem in Geometric Measure Theory, see [1] for a result in this direction and [6,8,9] for more recent contributions. On the other hand, the situation is more understood for minimizers of energies of the form (1.1), where similar partial regularity theorems are known, see for instance [11, Ch. 5], [22].

In [5], the second author togheter with C. De Lellis, G. De Philippis and B. Kirchheim already approached this regularity problem through the viewpoint of differential inclusions. The theory of differential inclusions has a rich history, we refer the reader to [17] for an overview and to [18,19] for more recent results. Since this work is also based on that viewpoint, let us briefly explain what this means. The strategy of [5] consisted first in rewriting (1.1) on a special class of geometric objects, namely multiplicity one graphs of Lipschitz maps, and study the differential inclusion associated to the system of PDEs arising from the stationarity condition. Namely, it can be shown that, see [5, Sec. 6] or Sect. A.5, to a C^k integrand Ψ as the one appearing in (1.1), one can naturally associate a C^k function $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ with the property that

$$\mathbb{E}_f(u) \doteq \int_\Omega f(Du(x))dx = \Sigma_\Psi(T_u). \tag{1.2}$$

where $T_u = [\![\Gamma_u, \vec{\xi}_u, 1]\!]$ is the current associated to the graph of u i.e. if $v(x) \doteq (x, u(x))$ is the graph map we have $T_u = v_\#[\![\Omega]\!]$. In particular, it is possible to prove, see [5, Prop. 6.8] that T_u is stationary for the energy (1.1) if and only if u solves the following equations:

$$\int_\Omega \langle Df(Du), Dv\rangle dx = 0, \quad \forall v \in C^1_c(\Omega, \mathbb{R}^n) \tag{1.3}$$

and

$$\int_\Omega \langle Df(Du), Du D\phi\rangle dx - \int_\Omega f(Du)\,\mathrm{div}\,\phi\; dx = 0, \quad \forall \phi \in C^1_c(\Omega, \mathbb{R}^m). \tag{1.4}$$

The Euler–Lagrange equation (1.3) corresponds to variations of the form

$$\frac{d}{d\varepsilon}|_{\varepsilon=0}\mathbb{E}_f(u + \varepsilon v) = 0,$$

usually called *outer* variations, and (1.4) corresponds to variations of the form

$$\frac{d}{d\varepsilon}|_{\varepsilon=0}\mathbb{E}_f(u \circ (x + \varepsilon\Phi)) = 0,$$

called *inner* (or *domain*) variations. The second step is to study (1.3) and (1.4) from the point of view of differential inclusions. This amounts to rewrite (1.3)–(1.4) equivalently as

$$\begin{pmatrix} Du \\ A \\ B \end{pmatrix} \in K_f \doteq \left\{ C \in \mathbb{R}^{(2n+m)\times m} : C = \begin{pmatrix} X \\ Df(X) \\ X^T Df(X) - f(X)\,\mathrm{id} \end{pmatrix} \right\}, \tag{1.5}$$

for $A \in L^\infty(\Omega, \mathbb{R}^{n\times m})$, $B \in L^\infty(\Omega, \mathbb{R}^m)$ with $\mathrm{div}(A) = 0$, $\mathrm{div}(B) = 0$.

This paper focuses on the same problem as [5], i.e. regularity of stationary points for geometric integrands, but with the addition of considering graphs with arbitrary positive multiplicity. This of course enlarges the class of competitors and might allow for more flexibility in the regularity of solutions. In particular, we consider *polyconvex* functions f, i.e.

$$f(X) = g(X, \Phi(X)),$$

where $g \in C^1(\mathbb{R}^k)$ is a convex function and $\Phi : \mathbb{R}^{n\times m} \to \mathbb{R}^k$ is the vector containing all the minors (subdeterminants) of order larger than or equal to 2 of $X \in \mathbb{R}^{n\times m}$. In analogy with (1.3)–(1.4), we will be interested in the following system of PDEs

$$\begin{cases} \displaystyle\int_\Omega \langle Df(Du), Dv\rangle \beta dx = 0 & \forall v \in C^1_c(\Omega, \mathbb{R}^n) \\ \displaystyle\int_\Omega \langle Df(Du), Du D\phi\rangle \beta\, dx - \int_\Omega f(Du)\,\mathrm{div}\,\phi\beta\, dx = 0 & \forall \phi \in C^1_c(\Omega, \mathbb{R}^m). \end{cases} \tag{1.6}$$

for a Lipschitz map $u \in \mathrm{Lip}(\Omega, \mathbb{R}^n)$, and a Borel function $\beta \in L^\infty(\Omega, \mathbb{R}^+)$. The study of objects with multiplicity is rather natural in the context of stationary rectifiable varifolds or currents. When dealing with these objects, one is interested in showing a so-called *constancy theorem*, see [23, Theorem 8.4.1]. A constancy theorem in the sense of [23, Theorem 8.4.1] asserts that if a stationary (for the area) varifold of dimension m has support contained in a C^2 manifold of the same dimension, then the varifold must be given by a fixed multiple of the manifold, so that in particular the multiplicity must be constant. In [10], it was shown that instead of C^2, even Lipschitz regularity of the manifold is sufficient to guarantee the validity of the Constancy Theorem. This is connected to the following algebraic fact. If a C^2 map u solves (1.3), then it necessarily solves also (1.4), hence the system (1.3)-(1.4) reduces to equation (1.3). Nonetheless, if $u \in C^2$ and solves (1.6) for a bounded multiplicity β, then it is not anymore true that u automatically solves the first. One therefore would like to show a priori that the multiplicity is constant and subsequently one is again in the situation given by (1.3)-(1.4). As for regularity theorems, no general constancy result is known at the moment for general functionals, except for the codimension one case, see [7].

As said, the tools we use are the same as the ones of [5], namely we rewrite (1.6) as

$$\begin{pmatrix} Du \\ A \\ B \end{pmatrix} \in C_f \doteq \left\{ C \in \mathbb{R}^{(2n+m)\times m} : C = \begin{pmatrix} X \\ \beta Df(X) \\ \beta X^T Df(X) - \beta f(X)\,\mathrm{id} \end{pmatrix}, \text{ for some } \beta > 0 \right\}, \tag{1.7}$$

again for $A \in L^\infty(\Omega, \mathbb{R}^{n\times m})$, $B \in L^\infty(\Omega, \mathbb{R}^m)$ with $\mathrm{div}(A) = 0$, $\mathrm{div}(B) = 0$. Our result is twofold. First, we will show that, if f is assumed to be non-negative, then the same result as [5, Theorem 1] holds, namely in C_f there are no T'_N configurations. Secondly, we show the optimality of this result by proving that if we drop the hypothesis on the positivity of f, one can not only embed a special family of matrices in C_f, but one can actually construct a

stationary current for the energy give in (1.1) whose support lies on the graph of a Lipschitz and nowhere C^1 map. In order to formulate properly these results, we need some terminology concerning differential inclusions.

Differential inclusions are relations of the form

$$M(x) \in K \subset \mathbb{R}^{n\times m} \text{ a.e. in } \Omega \tag{1.8}$$

for $M \in L^\infty(\Omega, \mathbb{R}^{n\times m})$ satisfying $\mathscr{A}(M) = 0$ in the weak sense for some constant coefficients, linear differential operator $\mathscr{A}(\cdot)$. To every operator $\mathscr{A}(\cdot)$, one can associate a *wave cone*, denoted with $\Lambda_{\mathscr{A}}$, that is made of those directions A in which it is possible to have plane wave solution, i.e. $A \in \Lambda_{\mathscr{A}}$ if and only if there exists $\xi \in \mathbb{R}^m$ such that

$$\mathscr{A}(h((x,\xi))A) = 0, \quad \forall h \in C^1(\mathbb{R}).$$

In this work, we will not need to consider various differential operators, as we will only work with the mixed div-curl operator introduced in (1.5). In that case, we denote the cone with Λ_{dc} and we will introduce it in detail in Sect. 2.1. Due to the connection of the wave-cone to the existence of oscillatory solutions of (1.8), a very first step to exclude wild solutions of (1.8) is to check that

$$A - B \notin \Lambda_{\mathscr{A}}, \quad \forall A, B \in K. \tag{1.9}$$

This is usually quite simple to verify, and indeed we will show in Proposition 3.1 that, if f is positive, then (1.9) holds with $\Lambda_{\mathscr{A}} = \Lambda_{dc}$ and K replaced by C_f. Property (1.9) is in general not sufficient to guarantee good regularity properties of solution of (1.8). Indeed, in [20], S. Müller and V. Šverák constructed a striking counterexample to elliptic regularity for solutions of

$$Dv(x) \in K'_f \doteq \left\{ C \in \mathbb{R}^{4\times 2} : C = \begin{pmatrix} X \\ Df(X)J \end{pmatrix} \right\} \subset \mathbb{R}^{4\times 2}, \tag{1.10}$$

where the function $f \in C^\infty(\mathbb{R}^{2\times 2})$ is quasiconvex (for the definition of quasiconvex function, we refer the reader to [20]), and J is a matrix satisfying $J = -J^T$ and $J^2 = -\mathrm{id}$. In particular, they were able to show that there exists a Lipschitz and nowhere C^1 function $v : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^4$ satisfying the differential inclusion (1.10). Their strategy was subsequently improved by L. Székelyhidi in [24] showing that f can be chosen polyconvex. In both cases, K'_f does not contain rank one connections, i.e.

$$\mathrm{rank}(A - B) = 2, \quad \forall A, B \in K'_f,$$

and this can proved to be equivalent to (1.9) in the case $\mathscr{A} = \mathrm{curl}$. Their strategy was based on showing that in K'_f other suitable families of matrices could be embedded, the so-called T_N configurations. In our situation, since we are dealing with mixed div-curl operators, we need to consider a slightly different version of T_N configurations, that we have named T'_N configurations in [5]. We postpone the definition of T_N and T'_N configurations to Sect. 2, but we are finally able to formally state our main positive result:

Theorem *If $f \in C^1(\mathbb{R}^{n\times m})$ is a strictly polyconvex function, then C_f does not contain any set $\{A_1, \dots, A_N\} \subset \mathbb{R}^{(2n+m)\times m}$ which induces a T'_N configuration, provided that $f(X_1) \geq 0, \dots, f(X_N) \geq 0$, if*

$$A_i = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \quad X_i, Y_i \in \mathbb{R}^{n\times m}, Z_i \in \mathbb{R}^{m\times m}, \forall i \in \{1, \dots, N\}.$$

This result, as [5, Theorem 1], shows that it is not possible to apply the convex integration methods of [20,24] to prove the existence of an irregular solution of the system (1.6). This theorem is stronger than [5, Theorem 1], in the sense that we are able to show [5, Theorem 1] as a corollary:

Corollary *If $f \in C^1(\mathbb{R}^{n\times m})$ is a strictly polyconvex function (not necessarily non-negative), then K_f does not contain any set $\{A_1, \dots, A_N\}$ which induces a T'_N configuration.*

Finally, in Sect. 4, we show the optimality of the hypothesis of non-negativity of the previous theorem by proving the following:

Theorem *There exists a smooth and elliptic integrand $\Psi : \Lambda_2(\mathbb{R}^4) \to \mathbb{R}$ such that the associated energy Σ admits a stationary point T whose (integer) multiplicities are not constant. Moreover the rectifiable set supporting T is given by a graph of a Lipschitz map $u : \Omega \to \mathbb{R}^2$ that fails to be C^1 in any open subset $\mathcal{V} \subset \Omega$.*

The last Theorem is obtained by embedding in the differential inclusion (1.7) what has been named in [13] *large* T_N configuration. Following the strategy of [24], we do not a priori choose a polyconvex $f \in C^\infty(\mathbb{R}^{2\times 2})$, but rather we construct it in such a way that C_f already contains this special family of matrices. Once the polyconvex function f has been built, we prove an extension result for f to the Grassmanians, thus obtaining the integrand Ψ of the statement of the Theorem. The extension results are quite simple and might be of independent interest. The construction of our counterexample can not be carried out in the varifold setting. The reason is quite elementary, as the integrand Ψ we would need to construct in the varifold case should be even, convex and positively 1-homogeneous, hence positive. We refer the reader to Remark 5.6 for more details. Moreover, let us point out that positivity of the integrand is a necessary assumption when studying existence of minima, but to the best of our knowledge there is no available example for it to be a necessary assumption also when studying regularity properties of stationary points.

The paper is organized as follows. In Sect. 2, we recall the statements of our main results in the case of non-negative integrands f and we collect some crucial preliminary results of [5]. The proof of the main results in the positive case, i.e. Proposition 3.1, Theorem 3.3 and Corollary 3.4, will be given in Sect. 3. In Sect. 4, we provide a counterexample to regularity when dropping the hypothesis of positivity of the integrand. Some lemmas of Sect. 4 concerning the extension of polyconvex functions to the Grassmaniann manifold can be easily extended to general dimension and codimensions. Therefore, we give the proof of these general versions in Sect. 5. Finally, the appendix contains a concise introduction to the tools of geometric measure theory used along the paper.

2 Positive case: absence of T_N configurations

In this section we collect some preliminary results proved in [5], that will be essential for the proofs of the next section.

2.1 Div-curl differential inclusions, wave cones and inclusion sets

In this subsection, we explain how to rephrase the system (1.3)–(1.4) as a differential inclusion. As recalled in the introduction, the Euler–Lagrange equations defining stationary points

for energies $\mathbb{E}_f$ are the couple of equations (1.3), (1.4), that can be written in the classical form:

$$\begin{cases} \operatorname{div}(Df(Du)) = 0 \\ \operatorname{div}(Du^T Df(Du) - f(Du)\,\mathrm{id}) = 0 \end{cases}$$

Thus we are lead to study the following *div-curl differential inclusion* for a triple of maps $X, Y \in L^\infty(\Omega, \mathbb{R}^{n\times m})$ and $Z \in L^\infty(\Omega, \mathbb{R}^{m\times m})$:

$$\operatorname{curl} X = 0, \qquad \operatorname{div} Y = 0, \qquad \operatorname{div} Z = 0\,, \tag{2.1}$$

$$W \doteq \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \in K_f = \left\{ A \in \mathbb{R}^{(2n+m)\times m} : A = \begin{pmatrix} X \\ Df(X) \\ X^T Df(X) - f(X)\,\mathrm{id} \end{pmatrix} \right\}, \tag{2.2}$$

where $f \in C^1(\mathbb{R}^{n\times m})$ is a fixed function.

Moreover, we also consider the following more general system of PDEs, for $u \in \operatorname{Lip}(\Omega, \mathbb{R}^n)$ and a Borel map $\beta \in L^\infty(\Omega, (0, +\infty))$:

$$\begin{cases} \displaystyle\int_\Omega \langle Df(Du), Dv\rangle \beta dx = 0 & \forall v \in C^1_c(\Omega, \mathbb{R}^n) \\ \displaystyle\int_\Omega \langle Df(Du), Du D\phi\rangle \beta dx - \int_\Omega f(Du) \operatorname{div} \phi \beta\, dx = 0 & \forall \phi \in C^1_c(\Omega, \mathbb{R}^m). \end{cases} \tag{2.3}$$

This system is equivalent to the stationarity in the sense of varifolds of the varifold $V = [\![\Gamma_u, \beta]\!]$, where Γ_u is the graph of u. This is discussed in Sect. A.5. The div-curl differential inclusion associated to this system is, again for a triple of maps $X, Y \in L^\infty(\Omega, \mathbb{R}^{n\times m})$ and $Z \in L^\infty(\Omega, \mathbb{R}^{m\times m})$:

$$\operatorname{curl} X = 0, \qquad \operatorname{div} Y = 0, \qquad \operatorname{div} Z = 0\,, \tag{2.4}$$

$$W \doteq \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \in C_f \tag{2.5}$$

where

$$C_f = \left\{ C \in \mathbb{R}^{(2n+m)\times m} : C = \begin{pmatrix} X \\ \beta Df(X) \\ \beta X^T Df(X) - \beta f(X)\,\mathrm{id} \end{pmatrix}, \text{ for some } \beta > 0 \right\}, \tag{2.6}$$

This discussion proves the following

Lemma 2.1 *Let $f \in C^1(\mathbb{R}^{n\times m})$. A map $u \in \operatorname{Lip}(\Omega, \mathbb{R}^n)$ is a stationary point of the energy (1.2) if and only if there are matrix fields $Y \in L^\infty(\Omega, \mathbb{R}^{n\times m})$ and $Z \in L^\infty(\Omega, \mathbb{R}^{m\times m})$ such that $W = (Du, Y, Z)$ solves the div-curl differential inclusion (2.1)–(2.2).*

Moreover, the couple $(u, \beta) \in \operatorname{Lip}(\Omega, \mathbb{R}^n) \times L^\infty(\Omega, (0, +\infty))$ solves (2.3) if and only there are matrix fields $Y \in L^\infty(\Omega, \mathbb{R}^{n\times m})$ and $Z \in L^\infty(\Omega, \mathbb{R}^{m\times m})$ such that $W = (Du, Y, Z)$ solves the div-curl differential inclusion (2.4)–(2.5).

Finally, we introduce here the wave-cone associated to the mixed *div-curl* operator that is relevant for us.

Definition 2.2 The cone $\Lambda_{dc} \subset \mathbb{R}^{(2n+m)\times m}$ consists of the matrices in block form

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

with the property that there is a direction $\xi \in \mathbb{S}^{m-1}$ and a vector $u \in \mathbb{R}^n$ such that $X = u \otimes \xi$, $Y\xi = 0$ and $Z\xi = 0$.

2.2 T_N configurations and T'_N configurations

We start defining T_N configurations for *classical* curl-type differential inclusions.

Definition 2.3 An ordered set of $N \geq 2$ distinct matrices $\{X_i\}_{i=1}^N \subset \mathbb{R}^{n\times m}$ is said to *induce a T_N configuration* if there exist matrices $P, C_i \in \mathbb{R}^{n\times m}$ and real numbers $k_i > 1$ such that:

(a) each C_i belongs to the wave cone of curl $X = 0$, namely $\operatorname{rank}(C_i) \leq 1$ for each i;
(b) $\sum_i C_i = 0$;
(c) $X_1, \dots, X_N$, P and $C_1, \dots, C_N$ satisfy the following N linear conditions

$$\begin{aligned} X_1 &= P + k_1 C_1, \\ X_2 &= P + C_1 + k_2 C_2, \\ &\dots \\ &\dots \\ X_N &= P + C_1 + \dots + k_N C_N\,. \end{aligned} \tag{2.7}$$

In the rest of the chapter we will use the word T_N configuration for the data

$$P, C_1, \dots, C_N, k_1, \dots k_N.$$

We will moreover say that the configuration is *nondegenerate* if $\operatorname{rank}(C_i) = 1$ for every i.

As in [5], we give a slightly more general definition of T_N configuration than the one usually given in the literature (cf. [20,24,25]), in that we drop the requirement that there are no rank-one connections between distinct X_i and X_j. We refer the reader to [5] for discussions concerning T_N configurations.

Adapted to the div-curl operator we introduce T'_N configurations, originally introduced in [5].

Definition 2.4 A family $\{A_1, \dots, A_N\} \subset \mathbb{R}^{(2n+m)\times m}$ of $N \geq 2$ *distinct*

$$A_i \doteq \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}$$

induces a T'_N *configuration* if there are matrices $P, Q, C_i, D_i \in \mathbb{R}^{n\times m}$, $R, E_i \in \mathbb{R}^{m\times m}$ and coefficients $k_i > 1$ such that

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} P \\ Q \\ R \end{pmatrix} + \begin{pmatrix} C_1 \\ D_1 \\ E_1 \end{pmatrix} + \dots + \begin{pmatrix} C_{i-1} \\ D_{i-1} \\ E_{i-1} \end{pmatrix} + k_i \begin{pmatrix} C_i \\ D_i \\ E_i \end{pmatrix} \tag{2.8}$$

and the following properties hold:

(a) each element (C_i, D_i, E_i) belongs to the wave cone Λ_{dc} of (2.1);
(b) $\sum_\ell C_\ell = 0$, $\sum_\ell D_\ell = 0$ and $\sum_\ell E_\ell = 0$.

We say that the T'_N configuration is *nondegenerate* if $\operatorname{rank}(C_i) = 1$ for every i.

We collect here some simple consequences of the definition above.

Proposition 2.5 *Assume $A_1, \dots, A_N$ induce a T'_N configuration with P, Q, R, C_i, D_i, E_i and k_i as in Definition 2.4. Then:*

(i) $\{X_1, \dots, X_N\}$ induce a T_N configuration of the form (2.7), if they are distinct; moreover the T'_N configuration is nondegenerate if and only if the T_N configuration induced by $\{X_1, \dots, X_N\}$ is nondegenerate;
(ii) For each i there is an $n_i \in \mathbb{S}^{m-1}$ and a $u_i \in \mathbb{R}^n$ such that $C_i = u_i \otimes n_i$, $D_i n_i = 0$ and $E_i n_i = 0$;
(iii) $\operatorname{tr} C_i^T D_i = \langle C_i, D_i \rangle = 0$ for every i.

2.3 Strategy

Before starting with the proof of the main result of this chapter, it is convenient to explain the strategy we intend to follow. In order to do so, let us consider here the case $n = m = 2$, $N = 5$. Suppose by contradiction that there exists a strictly polyconvex function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$, $f(X) = g(X, \det(X))$ and a T'_5 configuration A_1, A_2, A_3, A_4, A_5,

$$A_i = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \quad \forall i \in \{1, \dots, 5\},$$

where X_i, Y_i, Z_i fulfill the relations of (2.8), i.e.

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} P \\ Q \\ R \end{pmatrix} + \begin{pmatrix} C_1 \\ D_1 \\ E_1 \end{pmatrix} + \cdots + \begin{pmatrix} C_{i-1} \\ D_{i-1} \\ E_{i-1} \end{pmatrix} + k_i \begin{pmatrix} C_i \\ D_i \\ E_i \end{pmatrix}.$$

We will see below that we can without loss of generality assume that $P = 0$. The first part of the strategy follows the same lines of the one of [5]. Indeed, we think the relations $A_i \in C_f$, $\forall i$, where C_f has been defined in (2.6), as two separate pieces of information:

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \in K'_f = \left\{ A \in \mathbb{R}^{4\times 2} : A = \begin{pmatrix} X \\ \beta Df(X) \end{pmatrix}, \beta > 0, X \in \mathbb{R}^{2\times 2} \right\} \tag{2.9}$$

and

$$Z_i = X_i^T Y_i - \beta_i f(X_i)\,\mathrm{id}\,. \tag{2.10}$$

Let us denote with $c_i \doteq f(X_i)$. As in [5], we use (2.9) to obtain inequalities involving X_i, Y_i and quantities involving f. These are deduced from the polyconvexity of f, analogously to [24, Lemma 3]. In particular, (2.9) is rewritten as

$$c_i - c_j + \frac{1}{\beta_i}\langle Y_i, X_j - X_i \rangle - d_i \det(X_i - X_j) < 0, \tag{2.11}$$

for $d_i \doteq \partial_{y_5} g(y_1, y_2, y_3, y_4, y_5)|_{(X_i, \det(X_i))}$. This is proved in Proposition 2.9. The final goal is to prove that these inequalities can not be fulfilled at the same time. As in [5], we can simplify (2.11) using the structure result on T_N configurations in $\mathbb{R}^{2\times 2}$ of [25, Proposition 1]. This

asserts, in the specific case of the ongoing example, the existence of 5 vectors $(t_1^i, \dots, t_5^i)$, $i \in \{1, \dots, 5\}$ with positive components, such that

$$\sum_{j=1}^{5} t_j^i \det(X_j - X_i) = 0, \quad \forall i \in \{1, \dots, 5\}. \tag{2.12}$$

If we use this result in (2.11), we can eliminate from the expression the variable d_i, thus obtaining

$$\begin{aligned} \nu_i &\doteq \sum_{j=1}^{5} t_j^i \left(c_i - c_j + \frac{1}{\beta_i} \langle Y_i, X_j - X_i \rangle - d_i \det\left(X_i - X_j\right) \right) \\ &= \sum_{j=1}^{5} t_j^i \left(c_i - c_j + \frac{1}{\beta_i} \langle Y_i, X_j - X_i \rangle \right) < 0, \quad \forall i \in \{1, \dots, 5\}, \end{aligned}$$

compare Corollary 2.10. In [5], [25, Proposition 1] was extended to T_N configurations in $\mathbb{R}^{n \times m}$, so that relations (2.12) remain true in every dimension and target dimension. This extension is recalled in Proposition 2.8. Despite being very useful, the last simplification can not conclude the proof. Indeed, up to now we have exploited (2.9) and the fact that $\{X_1, \dots, X_5\}$ induce a T_5 configuration, but, if $\beta_i = 1, \forall i$, this is the exact same situation of [24]. Since from that paper we know the existence of T_5 configurations in K'_f, clearly we can not reach a contradiction at this point of the strategy. This is where the inner variations come into play. We rewrite (2.10) using the definition of T'_5 configuration and, after some manipulations, we find that the numbers

$$\mu_i \doteq \sum_{j=1}^{5} t_j^i (\langle X_i - X_j, Y_i \rangle - \beta_i c_i + \beta_j c_j)$$

must all be 0. For the index I such that $\beta_I = \min_i \beta_i$, and essentially using the positivity of c_j, we find that

$$\begin{aligned} 0 = -\mu_I &= \sum_{j=1}^{5} t_j^I (\langle X_j - X_I, Y_I \rangle + \beta_I c_I - \beta_j c_j) \\ &\leq \sum_{j=1}^{5} t_j^I (\langle X_j - X_I, Y_I \rangle + \beta_I c_I - \beta_I c_j) = \beta_I \nu_I, \end{aligned}$$

which is in contradiction with the negativity of ν_I.

2.4 Preliminary results: T_N configurations

To follow the strategy explained in Sect. 2.3, we need to recall the extension of [25, Proposition 1] proved in [5]. Here we will only recall the essential results without proof, we refer the interested reader to [5] for the details. First, it is possible to associate to a set T_N-configuration

of the form (2.7), i.e.

$$\begin{aligned} X_1 &= P + k_1 C_1, \\ X_2 &= P + C_1 + k_2 C_2, \\ &\dots \\ &\dots \\ X_N &= P + C_1 + \dots + k_N C_N\,, \end{aligned}$$

a *defining vector* $(\lambda, \mu) \in \mathbb{R}^{N+1}$, see [5, Definition 3.7], defined as follows:

$$\mu \doteq \frac{k_1 \dots k_N}{(k_1 - 1) \dots (k_N - 1)} \text{ and } \lambda_i \doteq \frac{k_1 \dots k_{i-1}}{(\mu - 1)(k_1 - 1) \dots (k_{i-1} - 1)}. \tag{2.13}$$

These relations can be inverted, in fact one can express

$$k_i = \frac{\mu\lambda_1 + \dots + \mu\lambda_i + \lambda_{i+1} \dots + \lambda_N}{(\mu - 1)\lambda_i}\,. \tag{2.14}$$

Since $k_i > 1, \forall i \in \{1, \dots, N\}$, (2.13) implies that $\lambda_i > 0, \forall i, \mu > 1$ and also

$$\sum_i \lambda_i = 1.$$

As in [25, Proposition 1], we define N vectors of $\mathbb{R}^N$ with positive components

$$t^i \doteq \frac{1}{\xi_i}(\mu\lambda_1, \dots, \mu\lambda_{i-1}, \lambda_i, \dots, \lambda_N), \text{ for } i \in \{1, \dots, N\}, \tag{2.15}$$

where $\xi_i > 1$ are normalization constants chosen in such a way that $\|t^i\|_1 = 1$. For a vector $v = (v_1, \dots, v_N) \in \mathbb{R}^N$,

$$\|v\|_1 = \sum_{j=1}^N |v_j|.$$

The importance of these vectors t^i comes from [25, Proposition 1], where it is proved that, for a T_N configuration of the form (2.7) in $\mathbb{R}^{2\times 2}$,

$$\sum_{j=1}^N t^i_j X_j = P + C_1 + \dots + C_{i-1} \tag{2.16}$$

Moreover, the following relation holds for every i:

$$\det\left(\sum_{j=1}^N t^i_j X_j\right) = \sum_{j=1}^N t^i_j \det(X_j)\,. \tag{2.17}$$

We need to state the generalization of the previous relations for T_N configurations of any size. In [5, Lemma 3.10] it was proved this general Linear Algebra result:

Lemma 2.6 *Assume the real numbers $\mu > 1, \lambda_1, \dots, \lambda_N > 0$ and $k_1, \dots, k_N > 1$ are linked by the formulas (2.13). Assume $v, v_1, \dots, v_N, w_1, \dots, w_N$ are elements of a vector space satisfying the relations*

$$w_i = v + v_1 + \dots + v_{i-1} + k_i v_i \tag{2.18}$$

$$0 = v_1 + \dots + v_N\,. \tag{2.19}$$

If we define the vectors t^i as in (2.15), then

$$\sum_j t^i_j w_j = v + v_1 + \ldots + v_{i-1}\,. \tag{2.20}$$

This lemma allows to generalize (2.16) and (2.17), compare [5, Proposition 3.8]. To state this result, we need some notation concerning multi-indexes. We will use I for multi-indexes referring to ordered sets of rows of matrices and J for multi-indexes referring to ordered sets of columns. In our specific case, where we deal with matrices in $\mathbb{R}^{n\times m}$ we will thus have

$$\begin{aligned} &I = (i_1, \ldots, i_r), \qquad 1 \le i_1 < \cdots < i_r \le n\,,\\ \text{and}\qquad &J = (j_1, \ldots, j_s), \qquad 1 \le j_1 < \cdots < j_s \le m \end{aligned}$$

and we will use the notation $|I| \doteq r$ and $|J| \doteq s$. In the sequel we will always have $r = s$.

Definition 2.7 We denote by $\mathcal{A}_r$ the set

$$\mathcal{A}_r = \{(I, J) : |I| = |J| = r\}, \qquad 1 \le r \le \min(n, m).$$

For a matrix $M = (m_{ij}) \in \mathbb{R}^{n\times m}$ and for $Z \in \mathcal{A}_r$ of the form $Z = (I, J)$, we denote by M^Z the squared $r \times r$ matrix obtained by M considering just the elements m_{ij} with $i \in I$, $j \in J$ (using the order induced by I and J).

We are finally in position to state [5, Proposition 3.8].

Proposition 2.8 *Let $\{X_1, \ldots, X_N\} \subset \mathbb{R}^{n\times m}$ induce a T_N configuration as in (2.7) with defining vector (λ, μ). Define the vectors $t^1, \ldots, t^N$ as in (2.15) and for every $Z \in \mathcal{A}_r$ of order $1 \le r \le \min\{n, m\}$ define the minor $\mathcal{S} : \mathbb{R}^{n\times m} \ni X \mapsto \mathcal{S}(X) \doteq \det(X^Z) \in \mathbb{R}$. Then*

$$\sum_{j=1}^N t^i_j \mathcal{S}(X_j) = \mathcal{S}\left(\sum_{j=1}^N t^i_j X_j\right) = \mathcal{S}(P + C_1 + \cdots + C_{i-1})\,. \tag{2.21}$$

and $A^\mu_Z \lambda = 0$.

It is clear that the previous result extends (2.16) and (2.17) to all the minors.

2.5 Preliminary results: inclusion set associated to polyconvex functions

As in [5, Section 4], we write a necessary condition for a set of distinct matrices $A_i \in \mathbb{R}^{2n\times m}$

$$A_i \doteq \begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \tag{2.22}$$

to belong to a set of the form

$$K'_f \doteq \left\{ \begin{pmatrix} X \\ Df(X) \end{pmatrix} : X \in \mathbb{R}^{n\times m} \right\} \tag{2.23}$$

for some strictly polyconvex function $f : \mathbb{R}^{n\times m} \to \mathbb{R}$. First, introduce the following notation, that is the same as in [5]. Let $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ be a strictly polyconvex function of the form $f(X) = g(\Phi(X))$, where $g \in C^1(\mathbb{R}^k)$ is strictly convex and Φ is the vector of all the subdeterminants of X, i.e.

$$\Phi(X) = (X, v_1(X), \ldots, v_{\min(n,m)}(X)),$$

and

$$v_s(X) = (\det(X_{Z_1}), \dots, \det(X_{Z_{\#\mathcal{A}_s}}))$$

for some fixed (but arbitrary) ordering of all the elements $Z \in \mathcal{A}_s$. Variables of $\mathbb{R}^k$, and hence partial derivatives in $\mathbb{R}^k$, are labeled using the ordering induced by Φ. The first nm partial derivatives, corresponding in $\Phi(X)$ to X, are collected in a $n \times m$ matrix denoted with $D_X g$. The j-th partial derivative, $mn + 1 \leq j \leq k$, is instead denoted by $\partial_Z g$, where Z is the element of $\mathcal{A}_s$ corresponding to the j-th position of Φ. Let us make an example in low dimension: if $n = 3$, $m = 2$, then $k = 9$, and we choose the ordering of Φ to be

$$\Phi(X) = (X, \det(X_{(12,12)}), \det(X_{(13,12)}), \det(X_{(23,12)})).$$

In this case, $y \in \mathbb{R}^k$ has coordinates

$$y = (y_{11}, y_{12}, y_{21}, y_{22}, y_{31}, y_{32}, y_{(12,12)}, y_{(13,12)}, y_{(23,12)}).$$

The partial derivatives with respect to the first 6 variables are collected in the 3×2 matrix:

$$D_X g = \begin{pmatrix} \partial_{11} g & \partial_{12} g \\ \partial_{21} g & \partial_{22} g \\ \partial_{31} g & \partial_{32} g \end{pmatrix}$$

The partial derivatives with respect to the remaining variables are denoted as $\partial_{(12,12)} g$, $\partial_{(13,12)} g$ and $\partial_{(23,12)} g$, i.e. following the ordering induced by Φ. Finally, for a matrix $A \in \mathbb{R}^{r \times r}$, we denote with $\operatorname{cof}(A)$ the matrix defined as

$$\operatorname{cof}(A)_{ij} = (-1)^{i+j} \det(M_{ji}(A)),$$

where $M_{ji}(A)$ denotes the $(n-1) \times (n-1)$ submatrix of A obtained by eliminating from A the j-th row and the i-th column. In particular, the following relation holds

$$\operatorname{cof}(A)A = A \operatorname{cof}(A) = \det(A) \operatorname{id}_r .$$

We are ready to state the following:

Proposition 2.9 *Let $f : \mathbb{R}^{n \times m} \to \mathbb{R}$ be a strictly polyconvex function of the form $f(X) = g(\Phi(X))$, where $g \in C^1$ is strictly convex and Φ is the vector of all the subdeterminants of X, i.e.*

$$\Phi(X) = (X, v_1(X), \dots, v_{\min(n,m)}(X)),$$

and

$$v_s(X) = (\det(X_{Z_1}), \dots, \det(X_{Z_{\#\mathcal{A}_s}}))$$

for some fixed (but arbitrary) ordering of all the elements $Z \in \mathcal{A}_s$. If $A_i \in K'_f$ and $A_i \neq A_j$ for $i \neq j$, then X_i, $Y_i = Df(X_i)$ and $c_i = f(X_i)$ fulfill the following inequalities for every $i \neq j$:

$$c_i - c_j + \langle Y_i, X_j - X_i \rangle - \sum_{r=2}^{\min(m,n)} \sum_{Z \in \mathcal{A}_r} d_Z^i \left(\langle \operatorname{cof}(X_i^Z)^T, X_j^Z - X_i^Z \rangle - \det(X_j^Z) + \det(X_i^Z) \right) < 0, \tag{2.24}$$

where $d_Z^i = \partial_Z g(\Phi(X_i))$.

This result was proved in [5, Proposition 4.1]. We now introduce the set

$$C'_f \doteq \left\{ C' \in \mathbb{R}^{2n\times m} : C' = \begin{pmatrix} X \\ \beta Df(X) \end{pmatrix}, \text{ for some } \beta > 0 \right\}.$$

Notice that C'_f is the projection of C_f on the first $2n \times m$ coordinates. We immediately obtain from the previous proposition and the definition of C'_f that

$$A_i \in C'_f, \quad \forall i \in \{1, \dots, N\}$$

if and only if there exist numbers $\beta_i > 0, \forall i$, such that

$$c_i - c_j + \frac{1}{\beta_i}\langle Y_i, X_j - X_i\rangle - \sum_{r=2}^{\min(m,n)} \sum_{Z\in\mathcal{A}_r} d^i_Z \left(\langle \operatorname{cof}(X^Z_i)^T, X^Z_j - X^Z_i\rangle - \det(X^Z_j) + \det(X^Z_i)\right) < 0. \tag{2.25}$$

The expressions in (2.25) can be simplified when the matrices $X_1, \dots, X_N$ induce a T_N configuration:

Corollary 2.10 *Let f be a strictly polyconvex function and let $A_1, \dots, A_N$ be distinct elements of K'_f with the additional property that $\{X_1, \dots, X_N\}$ induces a T_N configuration of the form (2.7) with defining vector (μ, λ). Then,*

$$c_i - \sum_j t^i_j c_j - \frac{k_i}{\beta_i}\langle Y_i, C_i\rangle < 0, \quad \forall i \in \{1, \dots, N\}, \tag{2.26}$$

where the t^i's are given by (2.15).

This corresponds to [5, Corollary 4.3], and concludes the list of preliminary results needed for the results of this paper.

3 Positive case: proof of the main results

Before checking whether the inclusion set C_f contains T_N or T'_N configurations, we need to exclude more basic building block for wild solutions, such as rank-one connections or, as in this case, Λ_{dc}-connections in C_f. It is rather easy to see, compare for instance [24], that if f is strictly polyconvex, then for $A, B \in K_f$ it is not possible to have

$$A - B \in \Lambda_{dc}.$$

Indeed the same result holds even considering K'_f. To prove this, it is sufficient to observe that if $X, Y \in \mathbb{R}^{n\times m}$ are rank-one connected, i.e. for some $u \in \mathbb{S}^{m-1}$

$$(X - Y)v = 0, \ \forall v \perp u, \tag{3.1}$$

and

$$(Df(X) - Df(Y))u = 0, \tag{3.2}$$

then

$$\begin{aligned}\langle Df(X) - Df(Y), X - Y\rangle &= \sum_{i=1}^{m} ((Df(X) - Df(Y))u_i, (X - Y)u_i) \\ &\overset{(3.1)}{=} ((Df(X) - Df(Y))u, (X - Y)u) \overset{(3.2)}{=} 0,\end{aligned}$$

where $\{u_1, \ldots, u_m\}$ is an orthonormal basis of $\mathbb{R}^m$ with $u_1 = u$. On the other hand, since f is strictly polyconvex, it is easy to see that

$$\langle Df(X) - Df(Y), X - Y\rangle > 0$$

if $\operatorname{rank}(X - Y) = 1$. The first result of this section shows that this result holds also for C_f, provided f is positive.

Proposition 3.1 *Let f be strictly polyconvex. If*

$$A = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \ B = \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} \in C_f,$$

and $f(X) \geq 0$, $f(X') \geq 0$, then

$$A - B \notin \Lambda_{dc}.$$

Proof Suppose by contradiction that there exist

$$A = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \in C_f, \ B = \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} X + C \\ Y + D \\ Z + E \end{pmatrix} \in C_f,$$

with $c \doteq f(X) \geq 0$, $c' \doteq f(X') \geq 0$, and there is a vector $\xi \in \mathbb{R}^m$ with $\|\xi\| = 1$ such that for every $v \perp \xi$,

$$Cv = 0, \ D\xi = 0, \ E\xi = 0.$$

Now we can use the so-called Matrix Determinant Lemma 3.2 to see that the expressions found in (2.25) evaluated at

$$A_1 = \begin{pmatrix} X \\ Y \end{pmatrix}, \ A_2 = \begin{pmatrix} X + C \\ Y + D \end{pmatrix},$$

yield the following inequalities:

$$c - c' - \frac{1}{\beta}\langle X - X', Y\rangle < 0, \tag{3.3}$$

$$c' - c - \frac{1}{\beta'}\langle X' - X, Y'\rangle < 0. \tag{3.4}$$

Moreover by assumption $(Z' - Z)\xi = 0$, i.e.

$$(Z' - Z)\xi = 0 = (X')^T Y'\xi - X^T Y\xi - (c'\beta' - c\beta)\xi.$$

Thus, using $(Y' - Y)\xi = 0$,

$$0 = (X' - X)^T Y'\xi - (c'\beta' - c\beta)\xi = \langle C, Y\rangle\xi - (c'\beta' - c\beta)\xi,$$

that yields, since $\|\xi\| = 1$,

$$\langle C, Y\rangle = c'\beta' - c\beta. \tag{3.5}$$

In the previous lines we have used the fact that

$$(X' - X)^T Y'\xi = C^T(Y + D)\xi = C^T Y\xi,$$

and, since C is of rank one with $Cv = 0, \forall v \perp \xi$,

$$C^T Y\xi = \langle C, Y\rangle \xi.$$

Exploiting (3.5), we rewrite (3.3) as

$$c - c' - \frac{1}{\beta}\langle X - X', Y\rangle = c - c' + \frac{1}{\beta}\langle C, Y\rangle = c - c' + \frac{1}{\beta}(c'\beta' - c\beta) < 0, \tag{3.6}$$

and (3.4) as

$$c' - c - \frac{1}{\beta'}\langle C, Y\rangle = c' - c - \frac{1}{\beta'}(c'\beta' - c\beta) < 0 \tag{3.7}$$

From (3.6), we infer

$$\beta c - \beta c' + (c'\beta' - c\beta) < 0 \Leftrightarrow c'(\beta' - \beta) < 0$$

and from (3.7)

$$\beta' c' - \beta' c - (c'\beta' - c\beta) < 0 \Leftrightarrow c(\beta - \beta') < 0.$$

Since $c \geq 0$ and $c' \geq 0$, we get a contradiction. □

Let us recall the Matrix Determinant Lemma used in the proof of the last proposition:

Lemma 3.2 *Let A, B be matrices in $\mathbb{R}^{m\times m}$, and let* $\operatorname{rank}(B) \leq 1$. *Then,*

$$\det(A + B) = \det(A) + \langle \operatorname{cof}(A)^T, B\rangle.$$

Now that we have excluded Λ_{dc}-connections, we can ask ourselves the same question concerning T'_N configurations. In particular we want to prove the main Theorem of this part of the paper:

Theorem 3.3 *If $f \in C^1(\mathbb{R}^{n\times m})$ is a strictly polyconvex function, then C_f does not contain any set $\{A_1, \ldots, A_N\} \subset \mathbb{R}^{(2n+m)\times m}$ which induces a T'_N configuration, provided that $f(X_1) \geq 0, \ldots, f(X_N) \geq 0$, if*

$$A_i = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \quad X_i, Y_i \in \mathbb{R}^{n\times m}, Z_i \in \mathbb{R}^{m\times m}, \forall i \in \{1, \ldots, N\}.$$

At the end of the section we will show the following

Corollary 3.4 *If $f \in C^1(\mathbb{R}^{n\times m})$ is strictly polyconvex, then K_f does not contain any set $\{A_1, \ldots, A_N\}$ which induces a T'_N configuration.*

Let us fix the notation. We will always consider T'_N configurations of the following form:

$$A_i \doteq \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix}, \quad X_i, Y_i \in \mathbb{R}^{n\times m}, Z_i \in \mathbb{R}^{m\times m}, \tag{3.8}$$

with:

$$X_i = P + \sum_{j=1}^{i-1} C_j + k_i C_i, \; Y_i = Q + \sum_{j=1}^{i-1} D_j + k_i D_i, \; Z_i = R + \sum_{j=1}^{i-1} E_j + k_i E_i, \tag{3.9}$$

and we denote with $n_i \in \mathbb{S}^{m-1}$ the vectors such that

$$D_i n_i = 0, E_i n_i = 0, C_i v = 0, \quad \forall v \perp n_i, \; \forall 1 \leq i \leq N.$$

3.1 Idea of the proof

Before proving the theorem, let us give an idea of the key steps of the proof. First of all, in Lemma 3.5, we will see that without loss of generality we can choose $P = 0$. As already explained in Sect. 2.3, we want to prove that the system of inequalities

$$\nu_i \doteq \beta_i c_i - \sum_j \beta_i t_j^i c_j - k_i \langle Y_i, C_i \rangle < 0, \forall i \,, \tag{3.10}$$

cannot be fulfilled at the same time. This gives a contradiction with Corollary 2.10. In particular, we show that for the index σ such that $\beta_\sigma = \min_j \beta_j$,

$$\nu_\sigma \geq 0.$$

To do so, we prove that the quantities

$$\mu_i \doteq -\beta_i c_i + \sum_j \beta_j t_j^i c_j + k_i \langle Y_i, C_i \rangle \tag{3.11}$$

equal to 0 for every i. Then, choosing σ as above and exploiting the positivity of $c_j, \forall j$, we estimate

$$0 = -\mu_\sigma = \beta_\sigma c_\sigma - \sum_j \beta_j t_j^\sigma c_j - k_\sigma \langle Y_\sigma, C_\sigma \rangle \leq \beta_\sigma c_\sigma - \sum_j \beta_\sigma t_j^\sigma c_j - k_\sigma \langle Y_\sigma, C_\sigma \rangle = \nu_\sigma. \tag{3.12}$$

This will then yield the required contradiction. In order to show $\mu_i = 0, \forall 1 \leq i \leq N$, we consider N matrices M_i defined as

$$M_i \doteq \mu \sum_{j \leq i-1} \alpha_j C_j^T D_j + \sum_{j \geq i} \alpha_j C_j^T D_j,$$

where $\mu > 1$ is part of the defining vector of the T_N configuration $\{X_1, \ldots, X_N\}$, compare 2.13, and α_j are real numbers. We prove that for numbers $\xi_j > 0$, a subset $\mathcal{I}_i \subset \{\xi_1 \mu_1, \ldots, \xi_N \mu_N\}$ is made of generalized eigenvalues of M_i, see (3.24). This is achieved thanks to Lemma 3.6. Since M_i is trace-free, as can be seen by the structure of C_j and D_j, we will find N relations of the form

$$\sum_{\xi_j \mu_j \in \mathcal{I}_i} \xi_j \mu_j = 0.$$

This can be read as the equations for the kernel for a specific matrix $N \times N$ matrix, W. Proving that W has trivial kernel will yield $\xi_j \mu_j = 0, \forall j$, and thus $\mu_j = 0$ since $\xi_j > 0$. The proof of the invertibility of W is the content of the last Lemma 3.10.

3.2 Proof of Theorem 3.3

Lemma 3.5 *If f is a strictly polyconvex function such that $A_i \in C_f$, $\forall 1 \leq i \leq N$ and $f(X_i) \geq 0, \forall 1 \leq i \leq N$, then there exists another strictly polyconvex function F such that the T'_N configuration B_i defined as*

$$B_i = \begin{pmatrix} X_i - P \\ Y_i \\ Z_i - P^T Y_i \end{pmatrix}$$

satisfies $B_i \in C_F$, for every $1 \leq i \leq N$ and moreover $F(X_i - P) \geq 0, \forall i$.

Proof Simply define the new polyconvex function $F(X)$ by $F(X) \doteq f(X+P)$. Clearly the newly defined family $\{B_1, \dots B_N\}$ still induces a T'_N configuration, and it is straightforward that $B_i \in C_F$. Moreover, this does not affect positivity, in the sense that $F(X_i - P) = f(X_i - P + P) = f(X_i) \geq 0$. □

Lemma 3.6 *Suppose $A_i \in C_f$, $\forall i$, and $P = 0$. Then, for every $i \in \{1, \dots, N\}$:*

$$\sum_{j=1}^{N} k_j(k_j-1)t_j^i C_j^T D_j n_i = \left(k_i \langle C_i, Y_i \rangle - \beta_i c_i + \sum_{j=1}^{N} \beta_j t_j^i c_j \right) n_i \overset{(3.11)}{=} \mu_i n_i, \quad \forall i = 1, \dots, N,$$

where t^i is the vector defined in (2.15).

Proof We need to compute the following sums:

$$\sum_j t_j^i Z_j = \sum_j t_j^i X_j^T Y_j - \sum_j t_j^i c_j \beta_j \,\mathrm{id}\,. \tag{3.13}$$

Let us start computing the sum for $i = 1$, $\sum_j \lambda_j X_j^T Y_j$. First, notice that

$$\sum_j \lambda_j X_j^T Y_j = \sum_j \lambda_j X_j^T (Y_j - Q) + \sum_j \lambda_j X_j^T Q = \sum_j \lambda_j X_j^T (Y_j - Q),$$

since, by Lemma 2.6 or (2.21),

$$\sum_j \lambda_j X_j^T Q = P^T Q = 0.$$

We rewrite it in the following way:

$$\begin{aligned}
\sum_j \lambda_j X_j^T Y_j &= \sum_j \lambda_j X_j^T (Y_j - Q) \\
&= \sum_{j=1}^{N} \lambda_j \left(\sum_{1 \leq a,b \leq j-1} C_a^T D_b + k_j \sum_{1 \leq a \leq j-1} C_a^T D_j + k_j \sum_{1 \leq b \leq j-1} C_j^T D_b + k_j^2 C_j^T D_j \right) \\
&= \sum_{i,j} g_{ij} C_i^T D_j,
\end{aligned} \tag{3.14}$$

where we collected in the coefficients g_{ij} the following quantities:

$$g_{ij} = \begin{cases} \lambda_i k_i + \sum_{r=i+1}^{N} \lambda_r, & \text{if } i \neq j \\ \lambda_i k_i^2 + \sum_{r=i+1}^{N} \lambda_r, & \text{if } i = j. \end{cases}$$

Using (2.14), we have, if $i \neq j$:

$$g_{ij} = g_{ji} = \lambda_i k_i + \sum_{r=i+1}^{N} \lambda_r = \frac{\mu}{\mu - 1}.$$

On the other hand, again using (2.14),

$$g_{ii} = k_i^2 \lambda_i + \sum_{r=i+1}^{N} \lambda_r = k_i(k_i - 1)\lambda_i + \frac{\mu}{\mu - 1}.$$

Using the equalities $\sum_\ell C_\ell = 0 = \sum_\ell D_\ell$, then also $\sum_{i,j} C_i^T D_j = 0$, and so $\sum_{i\neq j} C_i^T D_j = -\sum_i C_i^T D_i$. Hence, (3.14) becomes

$$\sum_{i,j} g_{ij} C_i^T D_j = \frac{\mu}{\mu-1}\sum_{i\neq j} C_i^T D_j + \sum_i \left(k_i(k_i-1)\lambda_i + \frac{\mu}{\mu-1}\right) C_i^T D_i = \sum_i k_i(k_i-1)\lambda_i C_i^T D_i.$$

We just proved that

$$\sum_j \lambda_j X_j^T Y_j = \sum_j k_j(k_j-1)\lambda_j C_j^T D_j. \tag{3.15}$$

Recall the definition of t^i, namely

$$t^i = \frac{1}{\xi_i}(\mu\lambda_1, \dots, \mu\lambda_{i-1}, \lambda_i, \dots, \lambda_N).$$

By the previous computation ($i = 1$), it is convenient to rewrite (3.13) using (3.15) as

$$R + \sum_{j=1}^{i-1} E_j = \frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j + (\mu-1)\sum_{j=1}^{i-1}\lambda_j X_j^T Y_j\right) - \sum_j t_j^i c_j \beta_j \,\mathrm{id}. \tag{3.16}$$

In the previous equation, we have used the equality

$$\sum_{j=1}^{N} t_j^i Z_j = R + \sum_{j=1}^{i-1} E_j, \quad \forall i \in \{1, \dots, N\}, \tag{3.17}$$

that easily follows from Lemma 2.6. Once again, let us express the sum up to $i-1$ in the following way:

$$\sum_{j=1}^{i-1}\lambda_j X_j^T Y_j = \sum_{j=1}^{i-1}\lambda_j X_j^T Q + \sum_{j=1}^{i-1}\lambda_j X_j^T (Y_j - Q) = \sum_{j=1}^{i-1}\lambda_j X_j^T Q + \sum_{k,j}^{i-1} s_{kj} C_k^T D_j.$$

A combinatorial argument analogous to the one in the previous case gives

$$\begin{aligned} s_{\ell\ell} &= k_\ell^2\lambda_\ell + \dots + \lambda_{i-1} \\ &= (k_\ell^2 - k_\ell)\lambda_\ell + k_\ell\lambda_\ell + \dots + \lambda_{i-1}, \\ s_{\alpha\beta} &= k_\alpha\lambda_\alpha + \dots + \lambda_{i-1}, \qquad \alpha \neq \beta \end{aligned}$$

Now

$$k_r\lambda_r + \dots + \lambda_{i-1} = \frac{\mu(\sum_{j=1}^{i-1}\lambda_j) + \sum_{j=i}^{N}\lambda_j}{\mu-1}$$

and so

$$k_r\lambda_r + \dots + \lambda_{i-1} = \frac{(\mu-1)(\sum_{j=1}^{i-1}\lambda_j) + 1}{\mu-1} = \frac{\xi_i}{\mu-1} =: b_{i-1}$$

Hence

$$\begin{aligned}\sum_{j=1}^{i-1}\lambda_j X_j^T Y_j &= \sum_{j=1}^{i-1}\lambda_j X_j^T Q \\ &+ \sum_{k,j}^{i-1} s_{kj} C_k^T D_j = \sum_{j=1}^{i-1}\lambda_j X_j^T Q + b_{i-1}\sum_{k,j}^{i-1} C_k^T D_j \\ &+ \sum_{j=1}^{i-1} k_j(k_j-1)\lambda_j C_j^T D_j .\end{aligned}$$

We rewrite (3.16) as

$$\begin{aligned}R + \sum_{j=1}^{i-1} E_j = \frac{1}{\xi_i}\Bigg(&\sum_j k_j(k_j-1)\lambda_j C_j^T D_j + \xi_i \sum_{k,j}^{i-1} C_k^T D_j \\ &+(\mu-1)\sum_{j=1}^{i-1}(k_j(k_j-1)\lambda_j C_j^T D_j + \lambda_j X_j^T Q)\Bigg) \\ &- \sum_j \beta_j t_j^i c_j \,\mathrm{id}\end{aligned} \tag{3.18}$$

Now we substitute (3.18) in the definition (3.9) of Z_i in order to compute E_i:

$$\begin{aligned}&k_i E_i + \frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j + \xi_i \sum_{k,j}^{i-1} C_k^T D_j + (\mu-1)\sum_{j=1}^{i-1}(k_j(k_j-1)\lambda_j C_j^T D_j + \lambda_j X_j^T Q)\right) \\ &\quad - \sum_j \beta_j t_j^i c_j \,\mathrm{id} = X_i^T Y_i - \beta_i c_i \,\mathrm{id}\,.\end{aligned}$$

Multiply by n_i the previous expression and recall that $E_i n_i = 0$ to find:

$$\begin{aligned}&\frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j n_i + \xi_i \sum_{k,j}^{i-1} C_k^T D_j n_i + (\mu-1)\sum_{j=1}^{i-1}(k_j(k_j-1)\lambda_j C_j^T D_j n_i + \lambda_j X_j^T Q n_i)\right) \\ &\quad - \sum_j \beta_j t_j^i c_j n_i = X_i^T Y_i n_i - \beta_i c_i n_i .\end{aligned} \tag{3.19}$$

Now notice that, since $D_i n_i = 0$,

$$\begin{aligned}X_i^T Y_i n_i &= X_i^T Q n_i + \sum_{j,k}^{i-1} C_k^T D_j n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i + k_i \sum_{j=1}^{i-1} C_j^T D_i n_i + k_i^2 C_i^T D_i n_i \\ &= X_i^T Q n_i + \sum_{j,k}^{i-1} C_k^T D_j n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i .\end{aligned}$$

Thus (3.19) becomes

$$\frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j n_i + (\mu-1)\sum_{j=1}^{i-1}(k_j(k_j-1)\lambda_j C_j^T D_j n_i + \lambda_j X_j^T Q n_i)\right)$$
$$-\sum_j \beta_j t_j^i c_j n_i = X_i^T Q n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i - \beta_i c_i n_i. \tag{3.20}$$

Now we need to compute

$$\sum_{j=1}^{i-1}\lambda_j X_j = \sum_{j=1}^{i-1} y_j C_j,$$

and

$$y_j = k_j\lambda_j + \cdots + \lambda_{i-1} = \frac{\xi_i}{\mu-1},\ \forall j \in \{1,\ldots,i-1\}.$$

Using this computation, (3.20) reads as:

$$\frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j n_i + (\mu-1)\sum_{j=1}^{i-1} k_j(k_j-1)\lambda_j C_j^T D_j n_i\right)$$
$$+\sum_{j=1}^{i-1} C_j^T Q n_i - \sum_j \beta_j t_j^i c_j n_i = \tag{3.21}$$
$$X_i^T Q n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i - \beta_i c_i n_i.$$

Exploiting the definition of t_j^i, we see that we can rewrite

$$\frac{1}{\xi_i}\left(\sum_j k_j(k_j-1)\lambda_j C_j^T D_j n_i + (\mu-1)\sum_{j=1}^{i-1} k_j(k_j-1)\lambda_j C_j^T D_j n_i\right) = \sum_j k_j(k_j-1)t_j^i C_j^T D_j n_i,$$

and

$$X_i^T Q n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i - \sum_{j=1}^{i-1} C_j^T Q n_i$$
$$= \sum_{j=1}^{i-1} C_j^T Q n_i + k_i C_i^T Q n_i + k_i \sum_{j=1}^{i-1} C_i^T D_j n_i - \sum_{j=1}^{i-1} C_j^T Q n_i = k_i C_i^T Y_i n_i.$$

Thus (3.21) becomes

$$\sum_j k_j(k_j-1)t_j^i C_j^T D_j n_i - \sum_j \beta_j t_j^i c_j n_i = k_i \sum_{j=1}^{i-1} C_i^T Y_i n_i - \beta_i c_i n_i.$$

Since $C_i v = 0, \forall v \perp n_i$, we have $C_i^T Y_i n_i = \langle C_i, Y_i \rangle n_i$, and we finally obtain the desired equalities:

$$\sum_{j=1}^{N} k_j(k_j-1)t_j^i C_j^T D_j n_i = \left(k_i \langle C_i, Y_i \rangle - \beta_i c_i + \sum_{j=1}^{N} \beta_j t_j^i c_j \right) n_i, \quad \forall i = 1, \dots, N.$$

□

We are finally in position to prove the main Theorem.

Proof of Theorem 3.3 Assume by contradiction the existence of a T'_N configuration induced by matrices $\{A_1, \dots, A_N\}$ of the form (3.8) which belong to the inclusion set C_f of some stictly polyconvex function $f \in C^1(\mathbb{R}^{n\times m})$ and $f(X_i) \geq 0$ for every i. We can assume, without loss of generality by Lemma 3.5, that

$$P = 0\,.$$

Using Lemma 3.6, we find

$$\sum_{j=1}^{N} k_j(k_j-1)t_j^i C_j^T D_j n_i = \left(k_i \langle C_i, Y_i \rangle - \beta_i c_i + \sum_{j=1}^{N} \beta_j t_j^i c_j \right) n_i = \mu_i n_i, \quad \forall i. \tag{3.22}$$

Define $\alpha_j \doteq k_j(k_j-1)\lambda_j > 0$, and

$$M_i \doteq \mu \sum_{j\leq i-1} \alpha_j C_j^T D_j + \sum_{j\geq i} \alpha_j C_j^T D_j,$$

for $i \in \{1, \dots, N\}$. Also set

$$M_i \doteq \mu M_{i-N}, \quad \forall i \in \{N+1, \dots, 2N\}.$$

Then, (3.22) can be rewritten as

$$M_i n_i = \xi_i \mu_i n_i, \quad \forall i \in \{1, \dots, N\}. \tag{3.23}$$

We define $n_s \doteq n_{s-N}$, for $s \in \{N+1, \dots, 2N\}$. As explained in Sect. 3.1, the idea is to show that a subset of the vectors n_j are *generalized eigenvectors* and a subset of $\xi_j \mu_j$ *are generalized eigenvalues* of M_i. In particular, for every $i \in \{1, \dots, N\}$, we want to show the following equalities:

$$\begin{cases} M_i n_{i+a} = \xi_{i+a}\mu_{i+a} n_{i+a} + v_{i,a}, & \text{if } a : i \leq i+a \leq N \\ M_i n_{i+a} = \mu \xi_{i+a}\mu_{i+a} n_{i+a} + v_{i,a}, & \text{if } a : N+1 \leq i+a \leq N+i-1, \end{cases} \tag{3.24}$$

where $v_{i,a} \in \operatorname{span}\{n_i, \dots, n_{i+a-1}\}$. From now on, we fix $i \in \{1, \dots, N\}$. To prove (3.24), first we rewrite

$$M_i n_{i+a} = (M_i - M_{i+a}) n_{i+a} + M_{i+a} n_{i+a}, \tag{3.25}$$

and then we use (3.23) to obtain

$$(M_i - M_{i+a}) n_{i+a} + M_{i+a} n_{i+a} = \begin{cases} \xi_{i+a}\mu_{i+a} n_{i+a} + (M_i - M_{i+a}) n_{i+a}, & \text{if } i+a \leq N, \\ \mu \xi_{i+a}\mu_{i+a} n_{i+a} + (M_i - M_{i+a}) n_{i+a}, & \text{if } i+a > N. \end{cases}.$$

To conclude the proof of (3.24), we only need to show that

$$(M_i - M_{i+a}) n_{i+a} \in \operatorname{span}\{n_i, \dots, n_{i+a-1}\}, \quad \forall a \in \{0, \dots, N-1\}. \tag{3.26}$$

To do so, we compute $M_i - M_{i+a}$. Let us start from the case $1 \le i + a \le N$:

$$\begin{aligned} M_i - M_{i+a} &= \mu \sum_{j<i} \alpha_j C_j^T D_j + \sum_{j\ge i} \alpha_j C_j^T D_j - \mu \sum_{j<i+a} \alpha_j C_j^T D_j - \sum_{j\ge i+a} \alpha_j C_j^T D_j \\ &= \sum_{i\le j<i+a} \alpha_j C_j^T D_j - \mu \sum_{i\le j<i+a} \alpha_j C_j^T D_j. \end{aligned}$$

On the other hand, if $N + 1 \le i + a \le i + N - 1$, then

$$\begin{aligned} M_i - M_{i+a} &= M_i - \mu M_{i+a-N} \\ &= \mu \sum_{j<i} \alpha_j C_j^T D_j + \sum_{j\ge i} \alpha_j C_j^T D_j - \mu^2 \sum_{j<i+a-N} \alpha_j C_j^T D_j - \mu \sum_{j\ge i+a-N} \alpha_j C_j^T D_j \\ &= \mu \sum_{j\ge i+a-N} \alpha_j C_j^T D_j - \mu \sum_{j<i} \alpha_j C_j^T D_j + \sum_{j\ge i} \alpha_j C_j^T D_j - \mu^2 \sum_{j<i+a-N} \alpha_j C_j^T D_j \\ &= \mu \sum_{j\ge i} \alpha_j C_j^T D_j + \sum_{j\ge i} \alpha_j C_j^T D_j - \mu^2 \sum_{j<i+a-N} \alpha_j C_j^T D_j. \end{aligned}$$

Now the crucial observation is that, due to the fact that $C_j v = 0$ for every $v \perp n_j$, the image of $C_j^T D_j$ is contained in the line $\mathrm{span}(n_j)$, for every $j \in \{1, \dots, N\}$. Therefore, the previous computations prove (3.26) and hence (3.24). Now we introduce

$$V_i \doteq \{n_i, n_{i+1}, n_{i+2}, \dots, n_N, n_{N+1}, \dots, n_{N+i-1}\}.$$

We can extract a basis for $\mathrm{span}(V_i)$ in the following way. First, choose indexes

$$\overline{S}_i \doteq \{k : k = i \text{ or } i < k \le N + i - 1, n_k \notin \mathrm{span}(n_i, \dots, n_{k-1})\}. \tag{3.27}$$

Then, consider the basis $\mathcal{B}_i \doteq \{n_k : k \in \overline{S}_i\}$ for $\mathrm{span}(V_i)$. Since

$$\mathrm{span}(\mathcal{B}_i) = \mathrm{span}(\{n_1, \dots, n_N\}), \quad \forall i,$$

then $\#\overline{S}_i = C \le \min\{m, N\}, \forall i$. Indexes in $\overline{S}_i$ lie in the set $\{1, \dots, 2N\}$. For technical reasons, we also need to consider the *modulo N* counterpart of $\overline{S}_i$, that is

$$S_i \doteq \{k \in \{1, \dots, N\} : k \in \overline{S}_i \text{ or } k + N \in \overline{S}_i\}. \tag{3.28}$$

In S_i, consider furthermore $S_i' \doteq S_i \cap \{i, \dots, N\}$, $S_i'' \doteq S_i \cap \{1, \dots, i-1\}$. If necessary, complete $\mathcal{B}_i$ to a basis of $\mathbb{R}^m$ made with elements γ_j orthogonal to the ones of $\mathcal{B}_i$. Note that, since $\mathrm{Im}(C_i^T D_i) \subset \mathrm{span}(n_i)$, then $\mathrm{Im}(M_i) \subset \{n_1, \dots, n_N\}$. Then, the associated matrix to M_i with respect to $\mathcal{B}_i$ is

$$\left(\begin{array}{ccccc|c} a_{1i} & * & * & \dots & * & \\ 0 & a_{2i} & * & \dots & * & \mathbf{T} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & a_{Ci} & \\ \hline & \mathbf{0}_{m-C,C} & & & & \mathbf{0}_{m-C,m-C} \end{array}\right). \tag{3.29}$$

We denoted with $\mathbf{0}_{c,d}$ the zero matrix with c rows and d columns, with $\mathbf{T}$ the $C \times (m - C)$ matrix of the coefficients of $M_i \gamma_j$ with respect to $\{n_s : s \in \overline{S}_i\}$, and with $*$ numbers we are not interested in computing explicitely. Finally, we have chosen an enumeration $s_1 < s_2 < \dots < s_\ell < \dots < s_C$ of the elements of $\overline{S}_i$, and we have defined

$$a_{\ell i} = \begin{cases} \xi_{s_\ell} \mu_{s_\ell}, & \text{if } s_\ell \in S_i', \\ \mu \xi_{s_\ell - N} \mu_{s_\ell - N}, & \text{if } s_\ell - N \in S_i''. \end{cases}$$

The triangular form of the matrix representing M_i is exactly due to (3.24). Now, $\mathrm{tr}(M_i) = 0, \forall i$, since $C_i^T D_i$ is trace-free for every i. This implies that the matrix in (3.29) must be trace-free, hence:

$$0 = \mathrm{tr}(M_i) = \sum_{\ell=1}^{C} a_{\ell i} = \sum_{a \in S_i'} \xi_a \mu_a + \mu \sum_{b \in S_i''} \xi_b \mu_b. \tag{3.30}$$

We have thus reduced the problem to the following simple Linear Algebra statement: we wish to show that, if W is the $N \times N$ matrix defined as

$$W_{ij} = \begin{cases} 1, & \text{if } j \in S_i', \\ \mu, & \text{if } j \in S_i'', \\ 0, & \text{if } j \notin S_i, \end{cases}$$

then, $Wx = 0 \Rightarrow x = 0$. By (3.30), the vector $x \in \mathbb{R}^N$ defined as $x_j \doteq \xi_j \mu_j, \forall 1 \leq j \leq N$, is such that $Wx = 0$, thus if the statement is true we get $\xi_j \mu_j = 0, \forall 1 \leq j \leq N$, and since $\xi_j > 1$, also $\mu_j = 0, \forall 1 \leq j \leq N$. By (3.12), this is sufficient to reach a contradiction. Therefore, we only need to show that $Wx = 0 \Rightarrow x = 0$. This proof will be given in Lemma 3.10. □

Before giving the proof of the final Lemma, let us make some examples of possible matrices W arising from the previous construction. For the sake of illustration, let us take N to be as small as possible, i.e. $N = 4$.

Example 3.7 Consider the case in which $C = 2$. This corresponds, for instance, to the case $m = 2$. Then, by Proposition 3.1 and (3.27), the only possible form of W is

$$W = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ \mu & 0 & 0 & 1 \end{pmatrix}, \quad Wx = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_4 \\ \mu x_1 + x_4 \end{pmatrix} = 0.$$

Let W_i be the i-th row of W. We notice that for $i = 1, 2, 3$, W_{i+1} differs from W_i by exactly two elements, while W_4 does not differ with W_1 by only two elements. It does, though, with μW_1. Hence we rewrite equivalently the system $Wx = 0$ as $(W_i - W_{i+1}, x) = 0$, $(W_4 - \mu W_1, x) = 0$:

$$0 = \begin{pmatrix} x_1 - x_3 \\ x_2 - x_4 \\ x_3 - \mu x_1 \\ x_4 - \mu x_2 \end{pmatrix}, \text{ i.e. } x_i = a_i x_{h(i)}, a_i = \begin{cases} 1, & \text{if } h(i) > i, \\ \mu, & \text{if } h(i) \leq i, \end{cases}$$

For a function $h : \{1, \ldots, 4\} \to \{1, \ldots, 4\}$. Since $\mu > 1$, this immediately implies $x_i = 0, \forall i$.

Example 3.8 Consider the case in which $C = 4$, corresponding to n_1, n_2, n_3, n_4 linearly independent. Then,

$$W = \begin{pmatrix} 1 & 1 & 1 & 1 \\ \mu & 1 & 1 & 1 \\ \mu & \mu & 1 & 1 \\ \mu & \mu & \mu & 1 \end{pmatrix}, \quad Wx = \begin{pmatrix} x_1 + x_2 + x_3 + x_4 \\ \mu x_1 + x_2 + x_3 + x_4 \\ \mu x_1 + \mu x_2 + x_3 + x_4 \\ \mu x_1 + \mu x_2 + \mu x_3 + x_4 \end{pmatrix} = 0.$$

As in the previous example, for $i = 1, 2, 3$, W_{i+1} differs from W_i by exactly one element, while W_4 does the same with μW_1. Thus as before we rewrite equivalently the system $Wx = 0$ as $(W_i - W_{i+1}, x) = 0$, $(W_4 - \mu W_1, x) = 0$:

$$0 = \begin{pmatrix} (\mu-1)x_1 \\ (\mu-1)x_2 \\ (\mu-1)x_3 \\ (\mu-1)x_4 \end{pmatrix}, \text{ i.e. } x_i = a_i x_{h(i)}, a_i = \begin{cases} 1, & \text{if } h(i) > i, \\ \mu, & \text{if } h(i) \leq i, \end{cases}$$

In this case, $h(i) = i, \forall i \in \{1, \dots, 4\}$. Clearly also in this case $\mu > 1$, implies $x_i = 0, \forall i$.

Finally, let us show a less symmetric example:

Example 3.9 Consider the case in which $C = 3$. Then, a possible matrix is:

$$W = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ \mu & 0 & 1 & 1 \\ \mu & \mu & 0 & 1 \end{pmatrix}, \quad Wx = \begin{pmatrix} x_1 + x_2 + x_4 \\ x_2 + x_3 + x_4 \\ \mu x_1 + x_3 + x_4 \\ \mu x_1 + \mu x_2 + x_4 \end{pmatrix} = 0.$$

First, let us comment on the fact that this is a possible matrix appearing in the proof of the previous Theorem. Indeed, let us consider the first two lines:

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}.$$

The fact that $W_{13} = 0$ means that $n_3 \in \text{span}(n_1, n_2)$, since $3 \notin S_1$. On the other hand, Proposition 3.1 ensures that n_3 is not a multiple of n_2, hence $n_3 \in S_2$, and $W_{23} = 1 \neq 0$. For this reason, the matrix

$$W = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ \mu & 0 & 1 & 1 \\ \mu & \mu & 0 & 1 \end{pmatrix}$$

would for instance have been non-admissible. Now, in order to prove $Wx = 0 \Rightarrow x = 0$, we work as in the previous examples, by noticing that for $i = 1, 2, 3$, W_{i+1} differs from W_i by at most two elements, while W_4 must be compared with μW_1. Thus we write $W_i - W_{i+1}$, $W_4 - \mu W_1$:

$$0 = \begin{pmatrix} x_1 - x_3 \\ x_2 - \mu x_1 \\ x_3 - \mu x_2 \\ (\mu-1)x_4 \end{pmatrix}, \text{ i.e. } x_i = a_i x_{h(i)}, a_i = \begin{cases} 1, & \text{if } h(i) > i, \\ \mu, & \text{if } h(i) \leq i. \end{cases}$$

It is an elementary computation to show that $x_i = 0, \forall i$.

Even though the examples we have given are too simple to appreciate the usefulness of the function h such that $x_i = a_i x_{h(i)}$, this will be crucial in the proof of the Lemma.

Lemma 3.10 *Let W be the matrix defined in the proof of Theorem 3.3. Then,* $\text{Ker}(W) = \{0\}$.

Proof Throughout the proof, we always consider a given vector $x \in \mathbb{R}^N$ such that $Wx = 0$. The proof, partially suggested by the previous examples, consists in the following steps. First, we show that the rows of W, W_i and W_{i+1} (if $i = N$, we compare W_N with μW_1) differ for

at most two elements, and one of them is always x_i. This immediately yields the existence of a function $h : \{1, \dots, N\} \to \{1, \dots, N\}$ such that $x_i = a_i x_{h(i)}$. We will then use this and the crucial fact that $\mu > 1$ to conclude that $x_i = 0, \forall i$. Let us make the following claims, and see from them how to conclude the proof of the present Lemma. We will use freely the notation introduced at the end of the proof of Theorem 3.3.

Claim 1: Let $i \in \{1, \dots, N\}$. Then $\overline{S}_i$ differs from $\overline{S}_{i+1}$ (if $i = N$, $\overline{S}_{i+1} = \overline{S}_1$) of at most two elements, in the sense that

$$\overline{S}_i \Delta \overline{S}_{i+1} \doteq \overline{S}_i \setminus \overline{S}_{i+1} \cup \overline{S}_{i+1} \setminus \overline{S}_i$$

contains at most 2 elements. Moreover, if $\overline{S}_i \Delta \overline{S}_{i+1} \neq \emptyset$, then $\overline{S}_i \Delta \overline{S}_{i+1} = \{i, I(i)\}$, with $i \in \overline{S}_i \setminus \overline{S}_{i+1}$, and $I(i) \in \overline{S}_{i+1} \setminus \overline{S}_i$.

Claim 2: Let $i \in \{1, \dots, N-1\}$. The couple of rows W_i,W_{i+1} and μW_1, W_N differ at most by two elements, in the sense that if $W_i = (W_{i1}, \dots, W_{iN})$ and $W_{i+1} = (W_{(i+1)1}, \dots, W_{(i+1)N})$, then there are at most two indexes j_1, j_2 such that $W_{ij_1} - W_{(i+1)j_1} \neq 0$ and $W_{ij_2} - W_{(i+1)j_2} \neq 0$ (and analogously for μW_1 and W_N).

Finally, with this claim at hand, we are going to prove

Claim 3: There exists a function $h : \{1, \dots, N\} \to \{1, \dots, N\}$ and numbers $a_i, i \in \{1, \dots, N\}$, such that

$$x_i = a_i x_{h(i)} \tag{3.31}$$

with the property

$$a_i = \begin{cases} 1, & \text{if } h(i) > i, \\ \mu, & \text{if } h(i) \leq i. \end{cases}$$

Let us show how the proof of the Lemma follows from Claim 3, and postpone the proofs of the claims. Fix $i \in \{1, \dots, N\}$ and use (3.31) recursively to find

$$x_i = a_i a_{h(i)} \dots a_{h^{(n-1)}(i)} x_{h^{(n)}(i)},$$

where $h^{(n)}$ denotes the function obtained by applying h to itself n times. We also use the notation $h^{(0)}$ to denote the identity function: $h^{(0)}(i) = i, \forall i \in \{1, \dots, N\}$. By the properties of a_j, we have, $\forall r \in \{1, \dots, n\}$,

$$a_{h^{(r)}(i)} = \begin{cases} 1, & \text{if } h^{(r)}(i) > h^{(r-1)}(i), \\ \mu, & \text{if } h^{(r)}(i) \leq h^{(r-1)}(i). \end{cases}$$

Fix $k \in \mathbb{N}$, and let $r \in \{k+1, \dots, k+N+1\}$. Then, $h^{(r)}(i) > h^{(r-1)}(i)$ can occur at most N times in this range, since otherwise we would find

$$1 \leq h^{(k)}(i) < h^{(k+1)}(i) < h^{(k+2)}(i) < \dots < h^{(k+N+1)}(i) \leq N,$$

and this is impossible since we would have $N+1$ distinct elements in the set $\{1, \dots, N\}$. Now clearly this observation implies that for every fixed $l \in \mathbb{N}$, there exists $s \in \mathbb{N}$ such that

$$x_i = \mu^t x_{h^{(s)}(i)}, \text{ for some } t \geq l.$$

This can only happen if $x_i = 0$. Since i is arbitrary, the conclusion follows. □

Let us now turn to the proof of the claims.

Proof of claim 1*:* To prove the claim, we need to use the definition of $\overline{S}_i$. Let us recall the definition of $\overline{S}_i$, given in (3.27). To build $\overline{S}_i$ what we do is consider the ordered set $\{n_i, n_{i+1}, \dots, n_{i+1-N}\}$ and select from it a basis of $\text{span}\{n_1, \dots, n_N\}$ starting from n_i and then at step $1 \leq k \leq N-1$ deciding whether to insert the vector n_{i+k} in our collection based on the fact that it is linear dependent or not from the previous ones. Recall also that S_i is the *modulo* N version of $\overline{S}_i$, see (3.28), and that we define $n_j \doteq n_{j-N}$, for $j \in \{N+1, \dots, 2N\}$. Hence now fix $i \in \{1, \dots, N\}$ and consider S_i. If $S_i = \{1, \dots, N\}$, then $\#S_i = N$, thus $S_j = \{1, \dots, N\}, \forall 1 \leq j \leq N$ and the claim holds. Otherwise, let $i+1 < I = I(i) \leq i+N-1$ be the first element in $(\overline{S_i})^c$. There are two cases:

(1) $n_I \in \text{span}(n_i, \dots, n_{I-1}) \setminus \text{span}(n_{i+1}, \dots, n_{I-1})$;
(2) $n_I \in \text{span}(n_{i+1}, \dots, n_{I-1})$.

At the same time, consider what happens in $\overline{S}_{i+1}$: the span in the $(i+1)$-th case starts with one vector less than the one of the i-th case, simply because the collection of indexes in $\overline{S}_{i+1}$ starts from n_{i+1}. Hence, since I is the first missing index in $\overline{S}_i$, I is also the first possible missing index for $\overline{S}_{i+1}$. Therefore, consider the first case

$$n_I \in \text{span}(n_i, \dots, n_{I-1}) \setminus \text{span}(n_{i+1}, \dots, n_{I-1}).$$

This implies that $I \in \overline{S}_{i+1}$. Moreover, we are now adding n_I to the set of vectors $n_{i+1}, \dots, n_{I-1}$, and $n_I \in \text{span}(n_i, \dots, n_{I-1}) \setminus \text{span}(n_{i+1}, \dots, n_{I-1})$, hence n_I adds to the previous vectors the component relative to n_i, in the sense that

$$\text{span}(n_{i+1}, \dots, n_I) = \text{span}(n_i, \dots, n_{I-1}).$$

This moreover implies that $j \in \overline{S}_i \Leftrightarrow j \in \overline{S}_{i+1}, \forall I \leq j < N+i-1$. Since $n_i \in \text{span}(n_{i+1}, \dots, n_I), i \notin \overline{S}_{i+1}$. Thus $\overline{S}_i$ and $\overline{S}_{i+1}$ differ by at most two elements, and we have $i \in \overline{S}_i \setminus \overline{S}_{i+1}$ and $I = I(i) \in \overline{S}_{i+1} \setminus \overline{S}_i$. This concludes the case

$$n_I \in \text{span}(n_i, \dots, n_{I-1}) \setminus \text{span}(n_{i+1}, \dots, n_{I-1}).$$

If instead $n_I \in \text{span}(n_{i+1}, \dots, n_{I-1})$, then we see that $I \notin \overline{S}_{i+1}$, and we can iterate this reasoning from there, in the sense that we look for the next index I' such that $I' \notin \overline{S}_i$ and divide again into the two cases above. Clearly, for the indexes $i+1 \leq j < I'_1$, we have $j \in \overline{S}_{i+1}$ and $j \in \overline{S}_i$. Either this iteration enters in case 1 of the previous subdivision for some element $I \notin \overline{S}_i$, or we conclude $\overline{S}_i = \overline{S}_{i+1}$. This concludes the proof of the claim. □

Proof of claim 2*:* Note that nonzero elements of W_i are found in positions corresponding to elements of S_i. Hence now fix $i \in \{1, \dots, N-1\}$ and consider W_i and W_{i+1}. If $S_i = S_{i+1}$, then $W_{ij} = 0 \Leftrightarrow W_{(i+1)j} = 0$. Moreover, we introduce the *modulo* N counterpart of the number $I(i)$ found in Claim 2, i.e. $I'(i) = I(i)$ if $I(i) \in \{1, \dots, N\}$, and $I'(i) = I(i) - N$ if $I(i) \in \{N+1, \dots, 2N\}$. Thus using the definition of W, we can deduce, if $S_i = S_{i+1}$,

$$\begin{cases} W_{(i+1)j} = W_{ij} = 0, & \text{if } j \notin S_i \\ W_{(i+1)j} = W_{ij} = \mu, & \text{if } j \in S_i,\ j < i \\ W_{(i+1)j} = W_{ij} = 1, & \text{if } j \in S_i,\ j > i \\ W_{(i+1)i} = \mu,\ W_{ii} = 1, \end{cases} \tag{3.32}$$

and the claim holds in this case. Finally, if $S_i \Delta S_{i+1} = \{i, I'(i)\}$, then:

$$\begin{cases} W_{(i+1)j} = W_{ij} = 0, & \text{if } j \notin S_i, j \neq I'(i) \\ W_{(i+1)j} = 1, W_{ij} = 0, & \text{if } j = I'(i) > i+1 \\ W_{(i+1)j} = \mu, W_{ij} = 0, & \text{if } j = I'(i) < i+1 \\ W_{(i+1)j} = W_{ij} = \mu, & \text{if } j \in S_i, j < i \\ W_{(i+1)j} = W_{ij} = 1, & \text{if } j \in S_i, j > i \\ W_{(i+1)i} = 0, W_{ii} = 1. \end{cases} \tag{3.33}$$

This concludes the proof of the claim if $i \in \{1, \dots, N-1\}$. If $i = N$, then we need to compare W_N with μW_1, and we obtain two cases, in analogy with the previous situation:

$$\text{if } S_N \Delta S_1 = \emptyset, \text{ then : } \begin{cases} \mu W_{1j} = W_{Nj} = 0, & \text{if } j \notin S_N \\ \mu W_{1j} = W_{Nj} = \mu, & \text{if } j \in S_N, j < N \\ \mu W_{1N} = \mu, W_{NN} = 1, & \text{otherwise,} \end{cases} \tag{3.34}$$

and

$$\text{if } S_N \Delta S_1 = \{N, I'(N)\}, \text{ then : } \begin{cases} \mu W_{1j} = W_{Nj} = 0, & \text{if } j \notin S_N, j \neq I'(N) \\ \mu W_{1j} = \mu, W_{Nj} = 0, & \text{if } j \notin S_N, j = I'(N) \\ \mu W_{1j} = W_{Nj} = \mu, & \text{if } j \in S_N, j < N \\ \mu W_{1N} = 0, W_{NN} = 1, & \text{otherwise.} \end{cases} \tag{3.35}$$

□

Proof of Claim 3***:*** Fix $i \in \{1, \dots, N\}$. We consider the equations given by

$$(W_{i+1} - W_i, x) = 0, \text{ if } i \in \{1, \dots, N-1\}, \text{ and } (W_N - \mu W_1, x) = 0.$$

If we consider $i \in \{1, \dots, N-1\}$, we see from (3.32) and (3.33) that

$$\begin{aligned} 0 = (W_i - W_{i+1}, x) &= \sum_{j=1}^{N} (W_{ij} - W_{(i+1)j}) x_j \\ &= \begin{cases} (1-\mu) x_i, & \text{if } S_i \Delta S_{i-1} = \emptyset \\ x_i - x_{I'(i)}, & \text{if } S_i \Delta S_{i-1} = \{i, I'(i)\}, I'(i) > i+1 \\ x_i - \mu x_{I'(i)}, & \text{if } S_i \Delta S_{i-1} = \{i, I'(i)\}, I'(i) < i-1 \end{cases} \end{aligned}$$

and from (3.34) and (3.35) we infer

$$0 = (W_N - \mu W_1, x) = \begin{cases} (1-\mu) x_N, & \text{if } S_N \Delta S_1 = \emptyset \\ x_N - \mu x_{I'(N)}, & \text{if } S_N \Delta S_1 = \{N, I'(N)\}. \end{cases}$$

From these equations we see that (3.31) holds with the choice $h(i) \doteq I'(i)$, when i is such that $S_i \Delta S_{i+1} \neq \emptyset$, and $h(i) \doteq i$ otherwise. □

3.3 Proof of Corollary 3.4

We end this section by showing that Theorem 3.3 implies Theorem 3.4. Assume by contradiction that there exists a family of matrices

$$\{A_1, \dots, A_N\} \subset K_f$$

inducing a T'_N configuration of the form (3.8). We show that then there exists another T'_N configuration $\{B_1, \dots, B_N\}$ such that $B_i \in K_F \subset C_F, \forall 1 \leq i \leq N$ for some strictly polyconvex F with

$$F(X'_i) \geq 0, \ \forall 1 \leq i \leq N,$$

if

$$B_i = \begin{pmatrix} X'_i \\ Y'_i \\ Z'_i \end{pmatrix}, \ \forall 1 \leq i \leq N.$$

This is a contradiction with Theorem 3.3. To accomplish this, it is is sufficient to define $F(X) \doteq f(X) - \min_i f(X_i)$. This function is clearly strictly polyconvex, since f is. Moreover, we define

$$X'_i \doteq X_i, \ Y'_i \doteq Y_i \text{ and } Z'_i \doteq Z_i + \min_j f(X_j)\,\mathrm{id}\,.$$

In this way, B_i is still a T'_N configuration. Moreover, $B_i \in K_F, \ \forall 1 \leq i \leq N$. To see this, it is sufficient to notice that, since $A_i \in K_f$,

$$Y'_i = Y_i = Df(X_i) = DF(X'_i), \ \forall 1 \leq i \leq N,$$

and

$$Z'_i = Z_i + \min_i f(X_i)\,\mathrm{id} = X_i^T Y_i - f(X_i)\,\mathrm{id} + \min_i f(X_i)\,\mathrm{id} = (X'_i)^T Y'_i - F(X_i)\,\mathrm{id}\,.$$

This finishes the proof.

4 Sign-changing case: the counterexample

In this section, we construct a counterexample to regularity in the case in which the hypothesis of non-negativity on f is dropped. Let us explain the strategy, that follows the one of [24]. First of all, we consider the following equivalent formulation of the differential inclusion of div-curl type considered in the previous sections. Indeed, due to the fact that for $a \in \mathrm{Lip}(\mathbb{R}^2, \mathbb{R}^2)$,

$$\mathrm{div}(a) = \mathrm{curl}(aJ),$$

if

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

one easily sees that (2.3) holds if and only if

$$\begin{cases} \mathrm{curl}(Df(Du)J) = 0, \\ \mathrm{curl}(Du^T Df(Du)J - f(Du)J) = 0, \end{cases}$$

in the weak sense. Since Ω is convex, the latter allows us to say that (2.3) holds for $u \in \mathrm{Lip}(\Omega, \mathbb{R}^2)$ if and only if there exist $w_1, w_2 : \Omega \to \mathbb{R}^2$ such that

$$w \doteq \begin{pmatrix} u \\ w_1 \\ w_2 \end{pmatrix}$$

solves a.e. in Ω:

$$Dw \in \tilde{C}_f \doteq \left\{ C \in \mathbb{R}^{(2n+m)\times m} : C = \begin{pmatrix} X \\ \beta Df(X)J \\ \beta X^T Df(X)J - \beta f(X)J \end{pmatrix}, \text{ for some } \beta > 0 \right\}. \tag{4.1}$$

From now on, we will always use this reformulation of the problem. Let us also introduce

$$\tilde{C}'_f \doteq \left\{ C \in \mathbb{R}^{(2n+m)\times m} : C = \begin{pmatrix} X \\ \beta Df(X)J \end{pmatrix}, \text{ for some } \beta > 0 \right\}.$$

In order to construct the counterexample, we want to find a set of *non-rigid* matrices $\{A_1, A_2, A_3, A_4, A_5\}$, $A_i \in \mathbb{R}^{6\times 2}$, $\forall i$, satisfying

$$A_i = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} \in \tilde{C}_f. \tag{4.2}$$

Roughly, non-rigidity means that there exists a non-affine solution of the problem

$$Dw \in \{A_1, \dots, A_5\},$$

see Lemma 4.3. The integrand f is of the form

$$f(X) = \varepsilon \mathcal{A}(X) + g(X, \det(X)), \tag{4.3}$$

for some convex and smooth $g : \mathbb{R}^5 \to \mathbb{R}$ and

$$\mathcal{A}(X) = a(X, \det(X)), \quad \text{where } a(X, d) \doteq \sqrt{1 + \|X\|^2 + d^2}, \tag{4.4}$$

is the area function. As in [24], f is not fixed from the beginning, but rather becomes another unknown of the problem. In particular, in order to find f, it is sufficient for the following condition to be fulfilled:

Condition 1 There exist 2×2 matrices $\{X_1, \dots, X_5\}$, $\{Y_1, \dots, Y_5\}$, real numbers $c_1, \dots, c_5, d_1, \dots, d_5$ and positive integers $\beta_1, \dots, \beta_5$ such that for $Q_{ij} \doteq c_i - c_j + d_i \det(X_i - X_j) + \dfrac{1}{\beta_i}\langle X_i - X_j, Y_i J\rangle$, one has

$$Q_{ij} < 0, \forall i \neq j. \tag{4.5}$$

If this condition is satisfied, then one has

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \in \tilde{C}'_f, \quad \text{i.e. } Y_i = \beta_i Df(X_i)J. \tag{4.6}$$

The construction of f is the content of Lemma 4.2. Moreover, we will be able to build f in such a way that for some large $R > 0$,

$$g(z) = M\sqrt{1 + \|z\|^2} - L = Ma(z) - L, \quad \forall z \in \mathbb{R}^5, \|z\| \geq R, \tag{4.7}$$

and constants $M, L > 0$. The *non-rigidity* of $A_1, \dots, A_5$ stems from the fact that we choose $\{X_1, \dots, X_5\}$ forming a *large* T_5-configuration, in the terminology of [13]. Therefore we introduce:

Condition 2 $\{X_1, X_2, X_3, X_4, X_5\}$ form a large T_5 configuration, i.e. there exists at least three permutations $\sigma_1, \sigma_2, \sigma_3 : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$ such that the ordered set $[X_{\sigma_i(1)}, X_{\sigma_i(2)}, \dots, X_{\sigma_i(5)}]$ is a T_5 configuration and moreover $\{C_{\sigma_1(i)}, C_{\sigma_2(i)}, C_{\sigma_3(i)}\}$ are linearly independent for every $i \in \{1, \dots, 5\}$.

Once this condition is guaranteed, by [13, Theorem 1.2], we find a non-affine Lipschitz map $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$Du \in \{X_1, X_2, X_3, X_4, X_5\}$$

almost everywhere in Ω. Furthermore, we can choose u with the property that for any subset $\mathcal{V} \subset \Omega$, Du attains each of these matrices on a set of positive measure. This is proved in Lemma 4.3.

In order to find Lipschitz maps $w_1, w_2 : \Omega \to \mathbb{R}^2$ such that

$$w = \begin{pmatrix} u \\ w_1 \\ w_2 \end{pmatrix} : \Omega \to \mathbb{R}^6$$

satisfies

$$Dw \in \tilde{C}_f \quad \text{a.e. in } \Omega,$$

we simply consider $w_1 = Au + B$, $w_2 = Cu + D$, for suitable 2×2 matrices A, B, C, D. We therefore get our last

Condition 3 Y_i and Z_i can be chosen of the form

$$Y_i = AX_i + B, \quad Z_i = CX_i + D,$$

and $Z_i = X_i^T Y_i - \beta_i c_i J$, where $c_i = f(X_i)$.

In Sect. (4.1), we will give an explicit example of values such that the Conditions 1-2-3 are fulfilled.

Once this is achieved, we need to extend the energy $\mathbb{E}_f$ to an energy defined on integral currents of dimension 2 in $\mathbb{R}^4$. Some of the results we present in this section in our specific case can be easily generalized to more general polyconvex integrands. Therefore, we defer their proofs to Sect. 5.

In order to extend our polyconvex function f to a *geometric* functional, we first recall (4.3), i.e.

$$f(X) = \varepsilon \mathcal{A}(X) + g(X, \det(X)),$$

for $g : \mathbb{R}^5 \to \mathbb{R}$ convex and smooth, and introduce the convex function $h : \mathbb{R}^5 \to \mathbb{R}$:

$$h(z) \doteq \varepsilon \sqrt{1 + \|z\|^2} + g(z).$$

We consider the *perspective function* of h:

$$G(z, t) \doteq yh\left(\frac{z}{y}\right), \quad \forall z \in \mathbb{R}^5, y > 0. \tag{4.8}$$

It is a standard result in convex analysis that G is convex on $\mathbb{R}^5 \times \mathbb{R}_+$ as soon as h is convex on $\mathbb{R}^5$, compare [4, Lemma 2]. Property (4.7) reads as

$$h(z) = (M + \varepsilon)\sqrt{1 + \|z\|^2} - L, \quad \forall z \in B_R^c(0), \tag{4.9}$$

therefore we also find that the *recession* function of G is

$$h^*(z) = \lim_{y \to 0^+} G(z, t) = M\|z\|, \quad \forall z \in \mathbb{R}^5.$$

Hence, G can be extended to the hyperplane $y = 0$ as

$$G(z, 0) \doteq h^*(z).$$

In Lemma 4.4, we will prove that $G(z, t)$ admits a finite, positively 1-homogeneous convex extension $\mathcal{G}$ to the whole space $\mathbb{R}^6$. We are finally able to define an integrand on the space of 2-vectors of $\mathbb{R}^4$, $\Lambda_2(\mathbb{R}^4)$. For a more thorough introduction to k-vectors, see Sect. A.1. Recall that

$$\Lambda_2(\mathbb{R}^4) = \text{span}\{v_1 \wedge v_2 : v_1, v_2 \in \mathbb{R}^4\}.$$

A basis for $\Lambda_2(\mathbb{R}^4)$ is given by the six elements $e_i \wedge e_j$, $1 \le i < j \le 4$, where e_1, e_2, e_3, e_4 is the canonical basis of $\mathbb{R}^4$. Recall moreover that this vector space can be endowed with a scalar product that acts on simple vectors as

$$\langle v_1 \wedge v_2, w_1 \wedge w_2 \rangle \doteq \det \begin{pmatrix} (v_1, w_1) & (v_1, w_2) \\ (v_2, w_1) & (v_2, w_2) \end{pmatrix},$$

where (u, v) denotes as usual the standard scalar product of $\mathbb{R}^4$. The integrand

$$\Psi : \Lambda_2(\mathbb{R}^4) \to \mathbb{R},$$

is thus defined as, for $\tau \in \Lambda_2(\mathbb{R}^4)$,

$$\Psi(\tau) \doteq \mathcal{G}(\langle \tau, e_3 \wedge e_2 \rangle, \langle \tau, e_4 \wedge e_2 \rangle, \langle \tau, e_1 \wedge e_3 \rangle, \langle \tau, e_1 \wedge e_4 \rangle, \langle \tau, e_3 \wedge e_4 \rangle, \langle \tau, e_1 \wedge e_2 \rangle). \tag{4.10}$$

Consequently, we define an energy on $\mathcal{I}_2(\mathbb{R}^4)$ as

$$\Sigma(T) \doteq \int_E \Psi(\vec{T}(z))\theta(z) d\mathcal{H}^2(z),$$

if $T = [\![E, \vec{T}, \theta]\!]$. For the notation concerning rectifiable currents and graphs, we refer the reader to Sect. A.4. The energy defined in this way satisfies Almgren's ellipticity condition (A.11), as we will prove in Lemma 4.5. Finally, in Lemma 4.6, we will prove that the current

$$T_{u,\theta} = [\![\Gamma_u, \vec{\xi}_u, \theta]\!] \tag{4.11}$$

is stationary for the energy Σ. The definition of stationarity for geometric functionals is recalled in Sect. A.5. In (4.11), Γ_u is the graph of u, $\vec{\xi}_u$ is its orientation, see (A.5), and $\theta(y)$ is a multiplicity, defined as $\theta(x, u(x)) = \beta_i$ if $x \in \Omega$ is such that

$$Dw(x) = \begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} X_i \\ \beta_i Df(X_i) \\ \beta_i X_i^T Df(X_i) J - \beta_i f(X_i) J \end{pmatrix}.$$

This discussion constitutes the proof of the following:

Theorem 4.1 *There exists a smooth and elliptic integrand $\Psi : \Lambda_2(\mathbb{R}^4) \to \mathbb{R}$ such that the associated energy Σ admits a stationary point T whose (integer) multiplicities are not constant. Moreover the rectifiable set supporting T is given by a graph of a Lipschitz map $u : \Omega \to \mathbb{R}^2$ that fails to be C^1 in any open subset $\mathcal{V} \subset \Omega$.*

Lemma 4.2 *There exists a smooth function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ of the form*

$$f(X) \doteq \varepsilon \mathcal{A}(X) + g(X, \det(X))$$

with $g : \mathbb{R}^5 \to \mathbb{R}$ convex and smooth, such that

(1) (4.6) is fulfilled;
(2) $g(X) = M\mathcal{A}(X) - L$ for constants $M, L > 0$, if $\|X\| \geq R$.

Proof We will follow roughly the strategy of [24, Lemma 3]. At first we construct the function g in several steps. Let $\{(X_i, Y_i, Z_i, \beta_i)\}_{i=1}^5$ the set of admissible matrices. For $\varepsilon > 0$ consider for each i the perturbed values

$$\begin{aligned} Y_i^\varepsilon &\doteq Y_i J - \varepsilon \beta_i D\mathcal{A}(X_i) J \\ c_i^\varepsilon &\doteq c_i - \varepsilon \mathcal{A}(X_i) \\ d_i^\varepsilon &\doteq d_i - \partial_y a(X_i, \det(X_i)) \end{aligned} \tag{4.12}$$

where $a(X, d) = \sqrt{1 + |X|^2 + d^2}$ and $\mathcal{A}(X) = a(X, \det(X))$, as defined in (4.4). Furthermore we introduce the perturbed matrix

$$Q_{ij}^\varepsilon \doteq c_i^\varepsilon - c_j^\varepsilon + d_i^\varepsilon \det(X_i - X_j) + \frac{1}{\beta_i}\langle X_i - X_j, Y_i^\varepsilon J\rangle.$$

Thanks to the strict inequality in (4.5) we can fix $\varepsilon, \sigma > 0$ such that $Q_{ij}^\varepsilon \leq -\sigma < 0$ for all i, j. Let us define the linear functions

$$l_i(X, d) \doteq c_i^\varepsilon - \frac{1}{\beta_i}\langle Y_i^\varepsilon J, X - X_i\rangle + d_i \left(\langle \mathrm{cof}(X_i)^T, X_i - X\rangle + d - \det(X_i)\right)$$

and the convex function

$$g_1(X, d) \doteq \max_{1\leq i\leq 5} l_i(X, d).$$

Note that $l_j(X_j, \det(X_j)) = c_j^\varepsilon$ and

$$l_i(X_j, \det(X_j)) = c_j^\varepsilon + Q_{ij}^\varepsilon < c_j^\varepsilon.$$

Hence there is $\delta > 0$ such that $l_i(X, d) < l_j(X, d)$ for all $(X, d) \in B_\delta(X_j, \det(X_j))$ for all $i \neq j$ which implies that $g_1 = l_j$ on $B_\delta(X_j, \det(X_j))$. Choosing a radial symmetric, non-negative smoothing kernel on $\mathbb{R}^5$, ρ_ε, $0 < \varepsilon << \delta$ we have that $g_2 \doteq \rho_\varepsilon \star g_1$ satisfies

(1) g_2 is smooth and convex
(2) $g_2 = l_j$ in a neighbourhood of $(X_j, \det(X_j))$ for all $j \in \{1, \ldots, 5\}$.
(3) $|g_2(X, d)| \leq C\|(1, X, d)\|$ for all (X, d) for some $C > 0$.

We choose any $R > 2\max_{1\leq i\leq 5}\{\|X_i\| + |\det(X_i)|\}$, and any $M > C$. Now we may choose $L > 0$ such that

$$F(X, d) \doteq Ma(X, d) - L < g_2(X, d) \text{ on } B_R. \tag{4.13}$$

Since $M > C$ we have that

$$F(X, d) = M\|(1, X, d)\| - L > g_2(X, d) \tag{4.14}$$

for all $(X, d) \notin B_{R_2}$, for some $R_2 > R$. Now let us fix a smooth approximation of the max function, say

$$m(a, b) \doteq (\phi_\varepsilon \star \max)(a, b),$$

where ϕ_ε is a radial symmetric, non-negative smoothing kernel in $\mathbb{R}^2$. Note that $m(a, b) = \max(a, b)$ outside a neighborhood of $\{a = b\}$. In particular if we choose ε sufficiently small we can ensure that

$$g(X, d) \doteq m(F(X, d), g_2(X, d))$$

agrees with g_2 on $B_{\frac{2R}{3}}$ by (4.13), and that it agrees with $F(X, d)$ outside B_{2R_2} by (4.14). It remains to check that $g(X, d)$ is still convex. First note that $\partial_a m \geq 0$ and $\partial_b m \geq 0$ since $\partial_a \max = \mathbf{1}_{\{a>b\}} \geq 0$, $\partial_b \max = \mathbf{1}_{\{b>a\}} \geq 0$. Now it is a direct computation on the Hessian to see that if $f_1, f_2 \in C^2(\mathbb{R}^N)$ are two convex functions and $\tilde{m} \in C^2(\mathbb{R}^2)$ is convex with $\partial_a \tilde{m}(a, b), \partial_b \tilde{m}(a, b) \geq 0$, then the composition $k(x) \doteq \tilde{m}(f_1(x), f_2(x))$ is convex. Thus we conclude that g is convex. Let us summarize the properties of g and the related polyconvex integrand $f_1(X) \doteq g(X, \det(X))$

(1) g is a smooth, convex function;
(2) $g = M\,a - L$ outside a ball B_{R_3}
(3) $g = g_2$ on a ball B_{R_0}, that implies that $f_1(X_i) = c_i^\varepsilon$ and $\beta_i Df_1(X_i)J = Y_i^\varepsilon$ for all i.

In particular from the last conditions and (4.12) we conclude that $h(X, d) \doteq \varepsilon a(X, d) + g(X, d)$ is convex, $f(X) \doteq \varepsilon \mathcal{A}(X) + f_1(X)$ is smooth, polyconvex and satisfies the desired properties, in particular $f(X_i) = c_i$, $\beta_i Df(X_i)J = Y_i$ for all i and $f = (\varepsilon + M)\mathcal{A} - L$ outside a ball centered at 0. □

Lemma 4.3 *Given a large T_5 configuration $\{X_1, \ldots, X_5\} \subset \mathrm{Sym}(2)$, where $\mathrm{Sym}(2)$ is the space of symmetric matrices of $\mathbb{R}^{2\times 2}$, there exists a map $u \in \mathrm{Lip}(\Omega, \mathbb{R}^2)$ such that*

$$Du \in \{X_1, \ldots, X_5\} \tag{4.15}$$

and such that for every open $\mathcal{V} \subset \Omega$,

$$|\{x \in \Omega : Du(x) = X_i\} \cap \mathcal{V}| > 0, \quad \forall i \in \{1, \ldots, 5\}. \tag{4.16}$$

Proof This statement is well-known, so we will only sketch its proof and give references where to find the relevant results. As shown in [13, Theorem 2.8], if $K \doteq \{X_1, \ldots, X_5\}$ forms a large T_5 configuration, then there exists an in-approximation of K inside $\mathrm{Sym}(2)$. This means, compare [13, Definition 1.3], that there exists a sequence of sets $\{U_k\}_{k\in\mathbb{N}}$, open relatively to $\mathrm{Sym}(2)$, such that

- $\sup_{X\in U_k} \mathrm{d}(X, K) \to 0$ as $k \to \infty$;
- $U_k \subset U_{k+1}^{rc}, \forall k \in \mathbb{N}$.

For a compact $C \subset \mathbb{R}^{2\times 2}$, the rank-one convex hull is defined as

$$C^{rc} \doteq \{P \in \mathbb{R}^{2\times 2} : f(P) \leq 0, \forall f \text{ rank-one convex such that } \sup_{X\in C} f(X) \leq 0\},$$

where $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is said to be rank-one convex if

$$f(tA + (1-t)B) \leq tf(A) + (1-t)f(B), \quad \forall A, B \in \mathbb{R}^{2\times 2}, \det(A - B) = 0, t \in [0, 1].$$

For an open set $U \subset \mathbb{R}^{2\times 2}$,

$$U^{rc} \doteq \bigcup_{C\subset U, C \text{ compact}} C^{rc}.$$

In this way, if U is open, then U^{rc} is open as well. The existence of a in-approximation for K implies the existence of a non-affine map u such that $Du \in \{X_1, \dots, X_5\}$, hence (4.15). This is proved in [13, Theorem 1.1]. To show (4.16), there are two ways. Either, one can use the same proof of [20, Theorem 4.1] or [24, Proposition 2] to show that the essential oscillation of Du is positive on any open subset of Ω. Since there is rigidity for the four gradient problem, see [3], this implies (4.16). Another way to show (4.16) is to use the Baire Theorem approach of convex integration as introduced by Kirchheim in [16]. In particular, in [16, Corollary 4.15], it is proved the following. Define

$$\mathcal{U} \doteq \bigcup_{k\in\mathbb{N}} U_k^{rc},$$

we fix $A \in \mathcal{U}$, and we also set

$$\mathcal{P} \doteq \{v \in \mathrm{Lip}(\Omega, \mathbb{R}^2) : Dv \in \mathcal{U},\ v \text{ piecewise affine},\ v|_{\partial\Omega} = A\},$$

then the typical (in the sense of Baire) map

$$u \in \mathcal{P}^{\|\cdot\|_\infty}$$

has the property that $Du \in K$. Then, we can use [26, Lemma 7.4] to show that actually the typical map is non-affine on any open set, hence again by the rigidity for the four gradient problem, we conclude (4.16). □

Lemma 4.4 *Let $G : \mathbb{R}^5 \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ be the convex function defined in (4.8). Then, there exists a positively 1-homogeneous, convex function $\mathcal{G} \in C^\infty(\mathbb{R}^6 \setminus \{0\}) \cap \mathrm{Lip}(\mathbb{R}^6)$ such that*

$$\mathcal{G}(z, t) = G(z, t),$$

if $z \in \mathbb{R}^5$, $t \in \mathbb{R}_+$.

Proof To prove the statement, it is sufficent to notice that the convexity of h and (4.9) tells us that h has property (P), see the beginning of Sect. 5, and therefore we can simply apply Proposition 5.2. The smoothness is a consequence of the smoothness of h, property (4.9) and Corollary 5.3. □

Lemma 4.5 *The energy Σ_Ψ satisfies the uniform Almgren ellipticity condition (A.11).*

Proof By construction, it is immediate to see that also

$$\mathcal{G}_\varepsilon(z, t) \doteq \mathcal{G}(z, t) - \frac{\varepsilon}{2}\sqrt{t^2 + \|z\|^2}$$

is still convex and positively 1-homogenous. Define Ψ_ε as in (4.10) by substituting $\mathcal{G}_\varepsilon$ to $\mathcal{G}$. By the general Proposition 5.5, we see that $\Sigma_{\Psi_\varepsilon}$ satisfies Almgren condition, hence Σ_Ψ satisfies (A.11) with constant $\frac{\varepsilon}{2}$. □

Lemma 4.6 *The current $T_{u,\theta} = [\![\Gamma_u, \vec{\xi}_u, \theta]\!]$ defined in (4.11) is stationary in $\Omega \times \mathbb{R}^2$ for the energy Σ_Ψ.*

Proof A direct computation shows that f and Ψ fulfill

$$f(X) = \Psi(W(X))\mathcal{A}(X), \forall X \in \mathbb{R}^{2\times 2},$$

where $W(X) = M^1(X) \wedge M^2(X)$ and M^i are the columns of the matrix

$$M(X) \doteq \begin{pmatrix} \mathrm{id}_m \\ X \end{pmatrix}.$$

Once this is checked, the proof is entirely analogous to the one of [5, Proposition 6.8], and will be sketched in the appendix, see Proposition 5.8. □

4.1 Explicit values

The following values were found using Maple 2020. Define the following quantities:

$$\begin{aligned}
&(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5) \doteq (2, 5, 10, 1, 2);\\
&(d_1, d_2, d_3, d_4, d_5) \doteq \left(-\frac{1204}{828115}, 0, \frac{-1309}{454800}, \frac{-10097}{2546880}, 0\right);\\
&(c_1, c_2, c_3, c_4, c_5) \doteq \left(0, 0, -\frac{2929}{1137000}, \frac{5233}{113700}, -\frac{33}{15160}\right).
\end{aligned}$$

The large T_5 configuration is given by:

$$A_1 \doteq \begin{pmatrix} \frac{8}{5} & -2 \\ -2 & \frac{8}{5} \\ -\frac{8}{1137} & \frac{7361}{454800} \\ \frac{267}{151600} & \frac{8}{1137} \\ \frac{-3361}{227400} & \frac{3361}{284250} \\ \frac{4801}{284250} & -\frac{4801}{227400} \end{pmatrix}; A_2 \doteq \begin{pmatrix} \frac{8}{5} & 2 \\ 2 & \frac{8}{5} \\ \frac{8}{1137} & \frac{7361}{454800} \\ \frac{267}{151600} & -\frac{8}{1137} \\ \frac{3361}{227400} & \frac{3361}{284250} \\ \frac{4801}{284250} & \frac{4801}{227400} \end{pmatrix}; A_3 \doteq \begin{pmatrix} \frac{2}{5} & 0 \\ 0 & -\frac{18}{5} \\ 0 & -\frac{959}{454800} \\ \frac{907}{151600} & 0 \\ 0 & -\frac{10083}{379000} \\ \frac{4801}{1137000} & 0 \end{pmatrix};$$

$$A_4 \doteq \begin{pmatrix} -\frac{18}{5} & 0 \\ 0 & \frac{2}{5} \\ 0 & \frac{5441}{454800} \\ \frac{9121}{454800} & 0 \\ 0 & \frac{3361}{1137000} \\ -\frac{14403}{379000} & 0 \end{pmatrix}; A_5 \doteq \begin{pmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{4} \\ 0 & \frac{6001}{454800} \\ \frac{2161}{454800} & 0 \\ 0 & \frac{3361}{606400} \\ \frac{4801}{606400} & 0 \end{pmatrix}.$$

Define $X_i, Y_i, Z_i \in \mathbb{R}^{2\times 2}$ through the relations

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = A_i.$$

The matrices A, B, C, D appearing in Condition 3 are given by:

$$A \doteq \begin{pmatrix} 0 & \frac{4}{1137} \\ -\frac{4}{1137} & 0 \end{pmatrix}; B \doteq \begin{pmatrix} 0 & \frac{4801}{454800} \\ \frac{3361}{454800} & 0 \end{pmatrix}; C \doteq \begin{pmatrix} 0 & \frac{3361}{454800} \\ \frac{4801}{454800} & 0 \end{pmatrix}, D \doteq 0.$$

These values fulfill Conditions 1, 2, 3. In particular, the three permutations in the definition of large T_5 configuration of Condition 2 are: [1, 2, 3, 5, 4], [1, 2, 4, 5, 3], [1, 2, 5, 3, 4].

5 Extension of polyconvex functions

Let $\Phi : \mathbb{R}^{n\times m} \to \mathbb{R}^k$ be the usual map that, to a matrix $X \in \mathbb{R}^{n\times m}$, associates the vector of the subdeterminants of Φ. Consider a polyconvex function

$$f(X) = h(\Phi(X)),$$

$h : \mathbb{R}^k \to \mathbb{R}$ being[1] C^1. The purpose of this section is to generalize the arguments of the previous section to arbitrary n, m, and hence to prove some of the lemmas of that section. Consider the following set of assumptions

(i) h is convex;
(ii) h has linear growth, i.e. $|h(z)| \leq A\|z\| + B, \forall z \in \mathbb{R}^k$, for $A, B \geq 0$;
(iii) $\lambda \doteq \inf\{h(z) - (Dh(z), z) : z \in \mathbb{R}^k\} > -\infty$;
(iv) $(Dh(z_2), z_2 - z_1) \leq h(z_1) + h(z_2), \quad \forall z_1, z_2 \in \mathbb{R}^k$.

If h fulfills (i)-(ii)-(iii), we will say it has property (P). If, in addition, h satisfies (iv), we will say that h fulfills property (PE).

Remark 5.1 Notice that (iii) is a consequence of (iv), indeed if (iv) holds we can write, for $z_1 = 0$ and for any $z_2 = z \in \mathbb{R}^k$:

$$(Dh(z), z) \leq h(0) + h(z),$$

hence

$$-h(0) \leq h(z) - (Dh(z), z), \quad \forall z \in \mathbb{R}^k,$$

that implies (iii).

We denote with h^* the *recession function* of h:

$$h^*(x) \doteq \lim_{y\to 0^+} yh\left(\frac{x}{y}\right), \quad \forall x \in \mathbb{R}^k.$$

It is not difficult to prove that the limit above always exists and is finite for a function h satisfying (P). To show it, one can use the fact that the function

$$y \mapsto yh\left(\frac{x}{y}\right)$$

defined for $y > 0$ is convex for every fixed $x \in \mathbb{R}^k$, see [4, Lemma 2].

[1] This hypothesis on the regularity of h is not necessary, and one could simply consider $h \in \mathrm{Lip}(\mathbb{R}^k)$. Indeed, all the results of this section would work with simple modifications in the Lipschitz case. Nonetheless, we prefer to assume C^1 regularity in order to avoid further technicalities.

As above, we define the *perspective function*

$$G(x, y) \doteq yh\left(\frac{x}{y}\right), \quad \text{if } y > 0.$$

We consider the smallest convex extension of G to the whole $\mathbb{R}^{k+1}$:

$$\mathcal{G}(z, t) \doteq \sup\{G(x, y) + (DG(x, y), (z, t) - (x, y)) : (x, y) \in \mathbb{R}^k \times (0, +\infty)\}.$$

By 1-homogeneity of G, we can write

$$\mathcal{G}(z, t) = \sup\{(DG(x, y), (z, t)) : (x, y) \in \mathbb{R}^k \times (0, +\infty)\} \tag{5.1}$$

First, we prove

Proposition 5.2 *Let $\mathcal{G}$ be defined as in (5.1). Then, if h satisfies (P), $\mathcal{G}$*

(1) is convex and extends G on $\mathbb{R}^k \times (0, +\infty)$;
(2) is positively 1-homogeneous;
(3) is finite everywhere.

Conversely, if there exists a function $\mathcal{G}$ that fulfills (1)–(2)–(3), then h fulfills (P).

Furthermore, we can prove the following characterization of $\mathcal{G}$:

Corollary 5.3 *Let h fulfill property (P), and let $\mathcal{G}$ be defined as in (5.1). Assume further that there exists $\lambda' \in \mathbb{R}$ and $R > 0$ such that*

$$h(z) = h^*(z) + \lambda', \quad \text{for } \|z\| \geq R. \tag{5.2}$$

Then, $\lambda' = \lambda$ and for $t < 0$, we have

$$\mathcal{G}(z, t) = h^*(z) + \lambda' t,$$

where λ is the quantity appearing in (iii).

Before starting with the proof of the proposition, we need to recall some results concerning the notion of subdifferential at $x \in \mathbb{R}^N$ of a convex function $f : \mathbb{R}^N \to \mathbb{R}$.

5.1 Subdifferentials

The subdifferential of f at x, denoted with $\partial f(x)$, is the collection of those vectors $v \in \mathbb{R}^N$ such that

$$(v, y - x) \leq f(y) - f(x), \quad \forall y \in \mathbb{R}^N.$$

We will use the following facts concerning the subdifferential. For a convex function with finite values, $\partial f(x) \neq \emptyset$ at all $x \in \mathbb{R}^N$, see [21, Theorem 23.4]. Conversely, if $f : \mathbb{R}^N \to \mathbb{R}$ is such that $\partial f(x) \neq \emptyset$ at every $x \in \mathbb{R}^N$, then f is convex, since in that case

$$f(x) = \sup_{y \in \mathbb{R}^N} \sup_{v \in \partial f(y)} \{(v, x - y) + f(y)\}.$$

As can be seen from the definition of subdifferential,

$$|f(x) - f(y)| \leq \max\left\{\sup_{v \in \partial f(x)} \|v\|, \sup_{w \in \partial f(y)} \|w\|\right\} \|x - y\|.$$

This, together with the fact that if K is compact, then $\partial f(K) \doteq \bigcup_{x\in K} \partial f(x)$ is compact, see [12, Lemma A.22], yields the fact that every convex function is locally Lipschitz. Moreover, if f is positively 1-homogeneous, a simple application of the definition of subdifferential shows that

$$v \in \partial f(x) \Leftrightarrow v \in \partial f(\lambda x), \quad \forall \lambda > 0, x \in \mathbb{R}^N. \tag{5.3}$$

In particular, combining (5.3) with the local Lipschitz property of convex functions, we infer that if $f : \mathbb{R}^N \to \mathbb{R}$ is convex and positively-1 homogeneous, f must be globally Lipschitz. Furthermore, using the definition of subdifferential and (5.3) for f convex and positively-1 homogeneous, it is easy to see that the following *generalized Euler's formula* holds

$$(v, x) = f(x), \quad \forall v \in \partial f(x), \forall x \in \mathbb{R}^N. \tag{5.4}$$

Finally, we recall that at x, the convex function $f : \mathbb{R}^N \to \mathbb{R}$ is differentiable if and only if

$$\partial f(x) = \{Df(x)\},$$

see [12, Lemma A.20-A.21] and references therein. We can now start the proof of the proposition.

5.2 Proof of Proposition 5.2

First we assume that h has property (P). $\mathcal{G}$ is convex since it is supremum of linear functions. Moreover, the convexity of h yields the convexity of G on $\mathbb{R}^k \times (0, +\infty)$. Having established that G is convex, the fact that $\mathcal{G}$ as in (5.1) extends G is a classical fact. This proves (1). Since G was positively 1-homogeneous, we have that $\mathcal{G}$ is as well homogeneous. Therefore (2) is checked, and we only need to prove (3). By (5.1) we see that in order to conclude we only need to show that, for fixed $(z, t) \in \mathbb{R}^{k+1}$,

$$(DG(x, y), (z, t)) \leq L < +\infty, \quad \forall (x, y) \in \mathbb{R}^k \times (0, +\infty),$$

where L possibly depends on (z, t). Let us compute DG. Firstly we have that

$$\partial_{x_i} G(x, y) = \partial_{x_i} h\left(\frac{x}{y}\right).$$

Now, exploiting the convexity of h, we can choose any $v \in \mathbb{R}^k$ with $\|v\| = 1$ and write

$$(Dh(x), v) \leq \frac{h(x + sv) - h(x)}{s}, \quad \forall s \in \mathbb{R}^+.$$

Using the linear growth of h, i.e. (ii), we bound:

$$(Dh(x), v) \leq \frac{A\|x + sv\| + B + A\|x\| + B}{s}.$$

Letting $s \to +\infty$, the previous expression yields

$$(Dh(x), v) \leq A, \quad \forall x, v \in \mathbb{R}^n, \|v\| = 1. \tag{5.5}$$

Thus, if we can show that $\partial_y G(x, y)$ is uniformly bounded, then we conclude the proof. We compute explicitly, for every $(x, y) \in \mathbb{R}^k \times (0, +\infty)$

$$\partial_y G(x, y) = h\left(\frac{x}{y}\right) - \left(Dh\left(\frac{x}{y}\right), \frac{x}{y}\right).$$

We are therefore left to study the boundedness (from below) of the function $z \mapsto h(z) - (Dh(z), z)$, but this is a consequence of (iii) of property (P).

Finally, let us show the necessity of (P). If $\mathcal{G}$ is convex and extends G, then in particular

$$\mathcal{G}(z, 1) = G(z, 1) = h(z), \quad \forall z \in \mathbb{R}^k,$$

hence h is convex. By the discussion of Sect. 5.1, we know that $\mathcal{G}$ is globally Lipschitz with constant $L > 0$. Since

$$\mathcal{G}(z, 1) = h(z), \quad \forall z \in \mathbb{R}^k,$$

we infer that h has linear growth, i.e. it enjoys property (ii). Finally, we need to show (iii). Since $\mathcal{G}$ extends G in the upper half-space, we obtain

$$|\partial_y G(x, y)| \leq L, \quad \forall (x, y) \in \mathbb{R}^k \times (0, +\infty).$$

By the definition of G, we deduce

$$|\partial_y G(z, 1)| = |h(z) - (Dh(z), z)| \leq L, \quad \forall z \in \mathbb{R}^k,$$

hence (iii).

5.3 Proof of Corollary 5.3

First we show that $\lambda = \lambda'$. To see this, consider for any $z \neq 0$ the auxiliary function $g(t) \doteq h(tz) - (Dh(tz), tz)$, for $t > 0$. Then, g is non-increasing. Indeed,

$$g\left(\frac{1}{t}\right) = \partial_y \mathcal{G}(z, t),$$

and we can use that $\mathcal{G}$ is convex to deduce that $t \mapsto \partial_y \mathcal{G}(z, t)$ is non-decreasing, hence that $t \mapsto g(t)$ is non-increasing. Now, for any t sufficiently large, by assumption (5.2), we have that

$$h(tz) - (Dh(tz), tz) = \lambda'.$$

This shows that

$$\lambda' = \lim_{t \to +\infty} [h(tz) - (Dh(tz), tz)] = \inf_{t > 0} [h(tz) - (Dh(tz), tz)] \geq \lambda.$$

In particular, notice that $h(0) = \lim_{t \to 0^+} g(t) \geq \lambda'$. To show the equality between λ and λ', consider now a sequence $z_n \in \mathbb{R}^k$ such that $a_n \doteq h(z_n) - (Dh(z_n), z_n) \to \lambda$ as $n \to \infty$. If $z_n = 0$ for infinitely many n, we can write $\lambda = \lim_{n \to \infty} a_n = h(0) \geq \lambda'$ and the proof is concluded. Otherwise, by the computation above, we have, for every $t \geq 1$

$$a_n \geq h(tz_n) - (Dh(tz_n), tz_n).$$

By choosing t in dependence of z_n, we can ensure through assumption (5.2) that

$$h(tz_n) - (Dh(tz_n), tz_n) = \lambda'.$$

Therefore,

$$\lambda = \lim_n a_n \geq \lambda'$$

and the proof of the first part of the Corollary is finished.

Now we wish to show the characterization of $\mathcal{G}$. Fix $(z,t) \in \mathbb{R}^k \times (-\infty, 0)$. Let $(x,y) \in \mathbb{R}^k \times (0,+\infty)$. Then, using the definition $G(x,y) = yh\left(\frac{x}{y}\right)$

$$(DG(x,y),(z,t)) = \left(Dh\left(\frac{x}{y}\right), z\right) + \left(h\left(\frac{x}{y}\right) - \left(Dh\left(\frac{x}{y}\right), \frac{x}{y}\right)\right)t. \tag{5.6}$$

By (iii), we get

$$h\left(\frac{x}{y}\right) - \left(Dh\left(\frac{x}{y}\right), \frac{x}{y}\right) \geq \lambda,$$

hence, since $t < 0$, then

$$(D\mathcal{G}(x,y),(z,t)) \leq \left(Dh\left(\frac{x}{y}\right), z\right) + \lambda t.$$

We now show that

$$\left(Dh\left(\frac{x}{y}\right), z\right) \leq h^*(z). \tag{5.7}$$

Let $a, b \in \mathbb{R}^k$, $r > 0$. Then, using the convexity of h,

$$0 \leq \left(Dh(a) - Dh\left(\frac{b}{r}\right), a - \frac{b}{r}\right),$$

or

$$0 \leq \left(Dh(a) - Dh\left(\frac{b}{r}\right), ra - b\right). \tag{5.8}$$

To conclude (5.7), we might use assumption (5.2), but let us use a slightly more general argument in order to use the same inequality below. By (5.5), we have that

$$\left\{Dh\left(\frac{b}{r}\right)\right\}_{r>0}$$

is an equibounded family of vectors, hence up to subsequences it admits a limit $\lim_{j\to\infty} Dh\left(\frac{b}{r_j}\right) = w \in \mathbb{R}^k$, where $\lim_{j\to\infty} r_j = 0$. Hence,

$$0 \leq \lim_{j\to\infty}\left(Dh(a) - Dh\left(\frac{b}{r_j}\right), r_j a - b\right) = -(Dh(a) - w, b). \tag{5.9}$$

Now, $w \in \partial h^*(b)$, in fact using the convexity of h we can write

$$\left(Dh\left(\frac{b}{r_j}\right), \frac{a}{r_j} - \frac{b}{r_j}\right) \leq h\left(\frac{a}{r_j}\right) - h\left(\frac{b}{r_j}\right).$$

Multiplying by r_j and letting $j \to \infty$, we find that $w \in \partial h^*(b)$. By (5.4), (5.7) now follows from (5.9). Therefore, we can conclude that, for $t < 0$,

$$\mathcal{G}(z,t) \leq h^*(z) + \lambda t.$$

To conclude the assertion, we consider for any $t > 0$:

$$(D\mathcal{G}(z,y),(z,t)) = \left(Dh\left(\frac{z}{y}\right), z\right) + \left(h\left(\frac{z}{y}\right) - \left(Dh\left(\frac{z}{y}\right), \frac{z}{y}\right)\right)t.$$

If we choose y sufficiently small (in dependence of z), once again using (5.2), we see that

$$(D\mathcal{G}(z,y),(z,t)) = h^*(z) + \lambda' t = h^*(z) + \lambda t,$$

the latter being true by the first part of the proof. This concludes the proof of the corollary.

5.4 Symmetric extension

Now we show the link between (PE) and a symmetric extension. Notice that imposing that h admits a 1-homogeneous and even extension such that $\mathcal{G}(z, 1) = h(z)$ forces this extension to have the form

$$\mathcal{G}(z, t) = |t| h\left(\frac{z}{t}\right) \tag{5.10}$$

for $t \neq 0$. If we require that $\mathcal{G}$ is convex too, then it is continuous, hence it becomes uniquely determined on $\{(x, y) : y = 0\}$ as $\mathcal{G}(z, 0) = h^*(z)$. Therefore, instead of considering a general convex extension as in (5.1), we are going to work with the function $\mathcal{G}$ obtained in (5.10).

Proposition 5.4 *h satisfies (PE) if and only if $\mathcal{G} : \mathbb{R}^{k+1} \to \mathbb{R}$ defined as*

$$\mathcal{G}(z, t) = \begin{cases} |y| h\left(\frac{z}{y}\right), & \text{if } y \neq 0 \\ h^*(z), & \text{if } y = 0 \end{cases} \tag{5.11}$$

is even and convex.

Proof Assume that h satisfies (PE). First we prove that $\mathcal{G}$ is even. This amounts to show that

$$h^*(z) = \lim_{t \to 0^+} \frac{h(tz)}{t} = \lim_{t \to 0^+} \frac{h(-tz)}{t} = h^*(-z). \tag{5.12}$$

To see this, we simply evaluate (iv) at $z_1 = \frac{-z}{t}$ and $z_2 = \frac{z}{t}$ for any $z \in \mathbb{R}^k$, $t > 0$ to find

$$2\left(Dh\left(\frac{z}{t}\right), \frac{z}{t}\right) \leq h\left(\frac{z}{t}\right) + h\left(-\frac{z}{t}\right). \tag{5.13}$$

We now use the same argument to prove (5.7) to see that for a sequence of positive numbers $\{t_j\}_{j \in \mathbb{N}}$ with $\lim_j t_j = 0$, $\lim_{j \to \infty} Dh\left(\frac{z}{t_j}\right) = w \in \partial h^*(z)$. Therefore, multiplying by t in (5.13) and passing to the limit along this subsequence, we get

$$2(w, z) \leq h^*(z) + h^*(-z), \quad \forall z \in \mathbb{R}^k.$$

By (5.4), $(w, z) = h^*(z)$, and in this way we see that, using the last equation,

$$2h^*(z) \leq h^*(z) + h^*(-z) \Rightarrow h^*(z) \leq h^*(-z), \quad \forall z \in \mathbb{R}^k,$$

that implies (5.12).

Now we show that $\mathcal{G}$ is convex. We rely on the results of Sect. 5.1, and we aim to show that at every point $p = (z, t) \in \mathbb{R}^{k+1}$,

$$\partial \mathcal{G}(p) \neq \emptyset.$$

Let first $t > 0$. Since at p the function is differentiable, the only possible candidate for an element of the subdifferential is $v \doteq D\mathcal{G}(p)$. Notice moreover that by the 1-homogeneity of $\mathcal{G}$, $(D\mathcal{G}(p), p) = \mathcal{G}(p)$. Thus we have, for any $q = (x, y) \in \mathbb{R}^{k+1}$:

$$(D\mathcal{G}(p), q - p) \leq \mathcal{G}(q) - \mathcal{G}(p) \Leftrightarrow (D\mathcal{G}(p), q) \leq \mathcal{G}(q). \tag{5.14}$$

If we establish $(D\mathcal{G}(p), q) \leq \mathcal{G}(q)$ for any $y \neq 0$, then we can use the pointwise convergence

$$\lim_{y\to 0^+} \mathcal{G}(x, y) = \lim_{y\to 0^+} yh\left(\frac{x}{y}\right) = h^*(x) = \mathcal{G}(x, 0)$$

to infer that the inequality holds also for $y = 0$. We therefore compute, for any $y \neq 0$:

$$(D\mathcal{G}(p), q) = \left(Dh\left(\frac{z}{t}\right), x\right) + h\left(\frac{z}{t}\right) y - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t}\right) y = y\left[h\left(\frac{z}{t}\right) - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t} - \frac{x}{y}\right)\right]$$

Using (5.14), $v = D\mathcal{G}(p)$ is a supporting hyperplane if and only if

$$(D\mathcal{G}(p), q) = y\left[h\left(\frac{z}{t}\right) - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t} - \frac{x}{y}\right)\right] \leq \mathcal{G}(q). \tag{5.15}$$

Following the same argument of the beginning of Proposition 5.2, since h is convex, $\mathcal{G}$ is convex on $\mathbb{R}^k \times (0, +\infty)$. Thus (5.15) is surely fulfilled if $y > 0$. If $y < 0$, (5.15) becomes

$$y\left[h\left(\frac{z}{t}\right) - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t} - \frac{x}{y}\right)\right] \leq \mathcal{G}(q) = -yh\left(\frac{x}{y}\right),$$

that can be rewritten as

$$h\left(\frac{z}{t}\right) - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t} - \frac{x}{y}\right) \geq -h\left(\frac{x}{y}\right), \forall (z, t) \in \mathbb{R}^k \times (0, +\infty), (x, y) \in \mathbb{R}^k \times (-\infty, 0).$$

The last condition is equivalent to (iv). Now we need to prove that also for points $p = (z, t)$ with $t < 0$ an element in the subdifferential exists. This is anyway a consequence of the evenness of $\mathcal{G}$ and the proof above, indeed the evenness of $\mathcal{G}$ yields

$$D\mathcal{G}(p) = -D\mathcal{G}(-p), \quad \forall p \in \mathbb{R}^k \times (\mathbb{R} \setminus \{0\}).$$

Therefore, for any $q = (x, y) \in \mathbb{R}^{k+1}$,

$$\begin{aligned}(D\mathcal{G}(p), q - p) &= -(D\mathcal{G}(-p), q - p) = (D\mathcal{G}(-p), p - q) = (D\mathcal{G}(-p), -q - (-p)) \\ &\leq \mathcal{G}(-q) - \mathcal{G}(-p) = \mathcal{G}(q) - \mathcal{G}(p),\end{aligned}$$

where we exploited the fact that $D\mathcal{G}(-p) \in \partial\mathcal{G}(-p)$, as proved above. Finally, we need to produce an element in the subdifferential at points $p = (z, 0)$. To do so, we again use the fact that for any $p' = (z, t)$ with $t > 0$, $q = (x, y) \in \mathbb{R}^{k+1}$,

$$(D\mathcal{G}(p'), q - p') \leq \mathcal{G}(q) - \mathcal{G}(p').$$

We only need to observe that $\{D\mathcal{G}(z, t)\}_{t>0}$ is an equibounded family of vectors. This allows us to choose a sequence $t_j > 0$ convergent to 0 such that $\{D\mathcal{G}(z, t_j)\}_{j\in\mathbb{N}}$ converges to a vector $w \in \mathbb{R}^{k+1}$. Since $\lim_{t\to 0^+} \mathcal{G}(z, t) = \mathcal{G}(z, 0)$, we have

$$(w, q - p) = \lim_{j\to\infty} (D\mathcal{G}(p_j), q - p_j) \leq \lim_{j\to\infty} (\mathcal{G}(q) - \mathcal{G}(p_j)) = \mathcal{G}(q) - \mathcal{G}(p).$$

where $p_j = (z, t_j), \forall j \in \mathbb{N}$. To show the equi-boundedness of $\{D\mathcal{G}(z, t)\}_{t>0}$, we observe that

$$D_z\mathcal{G}(z, t) = Dh\left(\frac{z}{t}\right),$$

that is equibounded in z and t by (5.5). Exactly as in the proof of Proposition 5.2, we use (iii) to say that

$$\lambda \leq \partial_t \mathcal{G}(z,t), \quad \forall (z,t) \in \mathbb{R}^k \times (0,+\infty).$$

Hence we only need to provide a bound from above. To show it, we use the convexity of h to estimate

$$\begin{aligned}\partial_t \mathcal{G}(z,t) &= h\left(\frac{z}{t}\right) - \left(Dh\left(\frac{z}{t}\right), \frac{z}{t}\right) = h\left(\frac{z}{t}\right) + \left(Dh\left(\frac{z}{t}\right), 0 - \frac{z}{t}\right)\\ &\leq h\left(\frac{z}{t}\right) + h(0) - h\left(\frac{z}{t}\right) = h(0),\end{aligned}$$

that provides the desired bound. This finishes the proof of the convexity of $\mathcal{G}$.

To conclude, we need to show the converse statement, i.e. that if $\mathcal{G}$ is even and convex, then h fulfills (PE). The fact that h fulfills (i)-(ii)-(iii) can be proved in a completely analogous way as in Proposition 5.2. By Remark 5.1, one could also infer (iii) as a corollary of (iv). Finally, to see (iv), one can simply follow the chain of logical equivalences of the previous part of the proof. This proves that (PE) is also necessary to the existence of the even extension. □

5.5 Extension to geometric functionals

Consider an orthonormal basis of $\Lambda_m(\mathbb{R}^{m+n})$, denoted with $E_1, \dots, E_{\binom{m+n}{m}}$, where

$$E_1 \doteq e_1 \wedge \cdots \wedge e_m,$$

as done in (A.1). We define, for every $\tau \in \Lambda_m(\mathbb{R}^{m+n})$

$$\Psi(\tau) \doteq \mathcal{G}(\langle \tau, E_2\rangle, \dots, \langle \tau, E_{\binom{m+n}{m}}\rangle, \langle \tau, E_1\rangle), \tag{5.16}$$

and consequently the energy

$$\Sigma_\Psi(T) \doteq \int_E \Psi(\vec{T}(x))\theta(x)d\mathcal{H}^m(x),$$

for $T = [\![E, \vec{T}, \theta]\!] \in \mathcal{R}_m(\mathbb{R}^{n+m})$. For convenience, let us denote

$$\phi(\tau) \doteq \left(\langle \tau, E_2\rangle, \dots, \langle \tau, E_{\binom{m+n}{m}}\rangle, \langle \tau, E_1\rangle\right).$$

We have

Proposition 5.5 *Let $\mathcal{G}$ be positively-1 homogeneous and convex, and define Ψ as in (5.16). Then, Σ_Ψ fulfills Almgren's condition (A.10).*

Proof Let $R, S \in \mathcal{R}_m(\mathbb{R}^{n+m})$, $\partial R = \partial S$, spt S is contained in the vectorsubspace of $\mathbb{R}^n$ associated with a simple m vector $\vec{S}_0$ of $\mathbb{R}^n$, and $\vec{S}(z) = \vec{S}_0$ for $\|S\|$-almost all z. Since $\partial R = \partial S$ we have that

$$\int \vec{R}\, d\|R\| = \int \vec{S}\, d\|S\| = \mathbb{M}(S)\, \vec{S}_0,$$

compare [11, 5.1.2]. Note that this implies by the linearity of ϕ that

$$\int \phi \circ \vec{R}\, d\|R\| = \int \phi \circ \vec{S}\, d\|S\| = \mathbb{M}(S)\, \phi \circ \vec{S}_0.$$

Now we may use Jensen inequality and the 1-homogeneity of $\mathcal{G}$ to deduce that

$$\int \Psi \circ \vec{R} d\|R\| = \int \mathcal{G} \circ \phi \circ \vec{R} d\|R\| \geq \mathcal{G}\left(\int \phi \circ \vec{R}\, d\|R\|\right) = \mathcal{G}\left(\mathbb{M}(S)\, \phi \circ \vec{S}_0\right)$$
$$= \mathbb{M}(S)\, \mathcal{G} \circ \phi \circ \vec{S}_0 = \int \Phi \circ \vec{S} d\|S\|\,,$$

where we used again in the last line that $\mathcal{G}$ is 1-homogeneous. □

Remark 5.6 If $\mathcal{G}$ is even, then Σ_Ψ is a well-defined energy on varifolds. Notice that in this case, $\mathcal{G}$ is convex, even and 1-homogeneous. A simple computation in convex analysis shows that this imposes for $\mathcal{G}$ to be positive. This observation is what makes it impossible to extend an integrand f as the one constructed in Sect. 4 to an integrand defined on varifolds using the methods introduced here.

Acknowledgements The authors would like to thank Camillo De Lellis for his interest in the problem and some preliminary discussions. This work was developed while R. T. was finishing his PhD at the University of Zürich, and is now supported by the SNF Grant 200021_182565. J. H. was partially supported by the German Science Foundation DFG in the context of the Priority Program SPP 2026 *Geometry at Infinity*.

Funding Open Access funding provided by EPFL Lausanne.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A. Currents, varifolds and geometric functionals

In this section we give the main definitions concerning currents and varifolds we have used throughout the paper. One can give more general definitions, namely flat, normal currents and general varifolds, see for instance [11,23], but we limit ourselves to rectifiable currents and varifolds in order to keep the exposition as concise as possible.

A.1. Multilinear algebra

Let $n \in \mathbb{N}, m \geq 0$. We denote with $\Lambda_m(\mathbb{R}^{n+m})$ the space of m-vectors of $\mathbb{R}^{n+m}$, i.e. the vector space given by finite linear combinations of elements of the form

$$v_1 \wedge \cdots \wedge v_m, \quad v_i \in \mathbb{R}^{n+m}, \forall 1 \leq i \leq m.$$

We also let $\Lambda_m^s(\mathbb{R}^{n+m}) \subset \Lambda_m(\mathbb{R}^{n+m})$ be the space of non-zero simple m-vectors, i.e. all elements $\tau \in \Lambda_m(\mathbb{R}^{n+m})$ such that

$$\tau = v_1 \wedge \cdots \wedge v_m$$

for $v_1, \ldots, v_m \in \mathbb{R}^{n+m}$. We define a *canonical* basis of $\Lambda_m(\mathbb{R}^{n+m})$ as follows. Let $e_1, \ldots, e_{n+m}$ be the vectors of the canonical basis of $\mathbb{R}^{n+m}$. Consider any multivector of

length m, $I = (i_1, \dots, i_m)$, with $1 \le i_1 < \cdots < i_m \le n+m$. There are $\binom{n+m}{m}$ of these multivectors, and each one defines a simple m-vector

$$E_j = e_{i_1} \wedge \cdots \wedge e_{i_m}, \quad j \in \left\{1, \dots, \binom{n+m}{m}\right\}.$$

It is easy to check that $\left\{E_1, \dots, E_{\binom{n+m}{m}}\right\}$ is a basis for $\Lambda_m(\mathbb{R}^{n+m})$. We also set

$$E_1 \doteq e_1 \wedge \cdots \wedge e_m, \tag{A.1}$$

while the ordering of the other indexes is arbitrary (but fixed).

The vector space $\Lambda_m(\mathbb{R}^{n+m})$ can be endowed with a scalar product that is defined on simple vectors as

$$(v_1 \wedge \cdots \wedge v_m, w_1 \wedge \cdots \wedge w_m) \doteq \det(X),$$

where $X \in \mathbb{R}^{m\times m}$ is defined as $X_{ij} = (v_i, w_j)$. We define a norm on $\Lambda_m(\mathbb{R}^{n+m})$ by setting $\|\tau\| = \sqrt{(\tau, \tau)}$. Analogously, one introduces the space of m-covectors of $\mathbb{R}^{n+m}$, $\Lambda_m^*(\mathbb{R}^{n+m})$, as the linear space generated by the wedge product of m covectors of $\mathbb{R}^{n+m}$. An element $\eta \in \Lambda_m^*(\mathbb{R}^{n+m})$ acts by duality on elements of $\Lambda_m(\mathbb{R}^{n+m})$ in the following way. Let $\eta = \eta^1 \wedge \cdots \wedge \eta^m$ and $\tau = v_1 \wedge \cdots \wedge v_m$ (the general case follows by linearity). Then,

$$\eta(\tau) = \det(Y),$$

where $Y \in \mathbb{R}^{m\times m}$ is the matrix defined as $Y_{ij} \doteq \eta_i(v_j)$, $\forall 1 \le i, j \le m$.

With these definition at hand, we can consider the space $\mathcal{D}^m(\mathbb{R}^{n+m})$ as the space of smooth m-forms of $\mathbb{R}^{n+m}$ with compact support, namely

$$\mathcal{D}^m(\mathbb{R}^{n+m}) \doteq C_c^\infty(\mathbb{R}^{n+m}, \Lambda_m^*(\mathbb{R}^{n+m})).$$

We endow $\Lambda_m^*(\mathbb{R}^{n+m})$ with a norm given by

$$\|\eta\| \doteq \sup_{\tau \in \Lambda_m^s(\mathbb{R}^{n+m})\setminus\{0\}} \frac{|\langle \eta, \tau\rangle|}{\|\tau\|},$$

hence we can consider on $\mathcal{D}^m(\mathbb{R}^{n+m})$ the norm

$$\|\omega\|_\infty \doteq \sup_{x\in\mathbb{R}^{n+m}} \|\omega(x)\|.$$

A.2. Planes and rectifiable sets

We denote with $\mathbb{G}(m, n+m)$ the space of unoriented m-planes of $\mathbb{R}^{n+m}$. In [5], we used the identification of $\mathbb{G}(m, n+m)$ with the space of orthogonal projections on m-planes

$$\left\{P \in \mathbb{R}^{(m+n)\times(m+n)} : P = P^T, P^2 = P, \operatorname{rank}(P) = \operatorname{tr}(P) = m\right\}.$$

It is not difficult to show that $\Lambda_m^s(\mathbb{R}^{n+m})$ can be identified with the space of *oriented* m-dimensional planes of $\mathbb{R}^{m+n}$, see [14, Section 2.1]. It is thus natural to introduce the two-to-one map

$$f : \Lambda_m^s(\mathbb{R}^{n+m}) \to \mathbb{G}(m, n+m) \tag{A.2}$$

that takes $v_1 \wedge \cdots \wedge v_m \in \Lambda_m^s(\mathbb{R}^{n+m})$ to the projection on the m-plane spanned by $v_1, \dots, v_m$. Notice that f is not injective since $f(\tau) = f(-\tau)$, $\forall \tau \in \Lambda_m^s(\mathbb{R}^{n+m})$.

We recall that a set $E \subset \mathbb{R}^{m+n}$ is called rectifiable of dimension m if

$$E = E_0 \cup \bigcup_{j \geq 1} F_j(E_j),$$

where $\mathcal{H}^m(E_0) = 0$, $F_j \in C^1(\mathbb{R}^m, \mathbb{R}^{n+m})$, and $E_j \subset \mathbb{R}^{m+n}$ is Borel. To such a set E it is possible to associate naturally a notion of approximate tangent plane, i.e. a map

$$x \mapsto T_x E \in \mathbb{G}(m, n+m).$$

For the definition of $T_x E$, we refer the reader to [23, Section 3.1]. An *orientation* $x \mapsto \vec{T}(x)$ of $T_x E$ is a Borel map $\vec{T} \in L^\infty(E, \Lambda^s_m(\mathbb{R}^{n+m}))$ with $\|\vec{T}(x)\| = 1$ for $\mathcal{H}^m \llcorner E$ a.e. $x \in \mathbb{R}^{n+m}$, and $\vec{T}(x) \in f^{-1}(T_x E)$, for $\mathcal{H}^m \llcorner E$ a.e. $x \in \mathbb{R}^{n+m}$, where f is the map defined in (A.2).

A.3. Varifolds

A m-dimensional rectifiable varifold V is a measure on $\mathbb{R}^{n+m} \times \mathbb{G}(n+m, m)$ given by

$$V(g) \doteq \int_E g(x, T_x E)\theta(x) d\mathcal{H}^m(x), \quad \forall g \in \mathbb{R}^{n+m} \times \mathbb{G}(n+m, m) \tag{A.3}$$

where E is m dimensional rectifiable set and $\theta \in L^1(E; \mathcal{H}^m \llcorner E)$. The varifold is called *integer rectifiable* if in addition θ has values in $\mathbb{N} \setminus \{0\}$. The notation for the varifold V defined as in (A.3) is

$$V = [\![E, \theta]\!].$$

A.4. Currents

A rectifiable current of dimension m, denoted by T, is a linear functional over $\mathcal{D}^m(\mathbb{R}^{n+m})$ represented as:

$$T(\omega) \doteq \int_E \langle \omega(x), \vec{T}(x) \rangle \theta(x) d\mathcal{H}^m(x),$$

where E is an m-rectifiable subset of $\mathbb{R}^{n+m}$, $\vec{T}(x)$ is an orientation of $T_x E$, and $\theta \in L^1(E; \mathcal{H}^m \llcorner E)$. Such a current T is denoted as

$$T = [\![E, \vec{T}, \theta]\!].$$

The mass of the current T is defined as

$$\mathbb{M}(T) \doteq \int_E |\theta(x)| d\mathcal{H}^m(x),$$

and we introduce the notion of boundary ∂T of a rectifiable current T as the $m-1$ dimensional current

$$\partial T(\omega) \doteq T(d\omega).$$

We restrict our attention to the space of *integer rectifiable currents* of dimension m, $\mathcal{R}_m(\mathbb{R}^{n+m})$, defined as the space of m-dimensional rectifiable currents T with finite mass and for which θ has values in $\mathbb{N} \setminus \{0\}$.

Given $T = [\![E, \vec{T}, \theta]\!]$ and an injective vector-field $X \in C^1(\mathbb{R}^{n+m}, \mathbb{R}^{n+m})$, we define the pushforward of T as the current $X_\#(T) \in \mathcal{R}_m(\mathbb{R}^{n+m})$ defined by

$$X_\#(T) \doteq [\![X(E), \vec{\xi}, \theta \circ X^{-1}]\!],$$

where

$$\vec{\xi}(X(x)) \doteq \frac{DX(x)v_1(x) \wedge \cdots \wedge DX(x)v_m(x)}{\|DX(x)v_1(x) \wedge \cdots \wedge DX(x)v_m(x)\|}, \quad \forall x \in E, \tag{A.4}$$

if $\vec{T}(x) = v_1(x) \wedge \cdots \wedge v_m(x)$. Analogously, for a varifold $V = [\![E, \theta]\!]$,

$$X_\# V \doteq [\![X(E), \theta \circ X^{-1}]\!].$$

Notice that to every current $T \in \mathcal{R}_m(\mathbb{R}^{n+m})$ one can associate an integer rectifiable varifold V_T in the obvious way

$$V_T(g) \doteq \int_E g(x, f(\vec{\tau}(x)))\theta(x) d\mathcal{H}^m(x),$$

where f is the map defined in (A.2).

Let us explain how to give to a graph of a Lipschitz map a structure of current, hence also of varifold. We essentially follow the theory developed in [14,15]. We also refer the reader to [5], where this discussion was made for giving the graph a structure of varifold. Let $\Omega \subset \mathbb{R}^m$ be open and bounded and let $u \in \mathrm{Lip}(\Omega, \mathbb{R}^n)$. Then, the graph of u defined as

$$\Gamma_u = \{(x, u(x)) : x \in \Omega\}$$

is m-rectifiable. Furthermore, as proved in [14, Sec. 1.5, Th. 5] its approximate tangent plane is, at a.e. $x_0 \in \Omega$, given by the orthogonal projection on

$$\begin{aligned}\pi(x_0) &\doteq \mathrm{span}\{\partial_1(x, u(x))|_{x=x_0}, \ldots, \partial_m(x, u(x))|_{x=x_0}\} \\ &= \mathrm{span}\{(f_1, \partial_1 u(x_0))^T, \ldots, (f_m, \partial_m u(x_0))^T\},\end{aligned}$$

where $f_1, \ldots, f_m$ are the elements of the canonical basis of $\mathbb{R}^m$. Define $v_i(x) \doteq (f_i, \partial u_i(x))^T$. The orientation we define on $\pi(x_0)$ is the natural one:

$$\vec{\xi}_u(x) \doteq \frac{v_1(x) \wedge \cdots \wedge v_m(x)}{\|v_1(x) \wedge \cdots \wedge v_m(x)\|}. \tag{A.5}$$

Given a Borel function $\theta \in L^1(\Gamma_u; \mathcal{H}^m \llcorner \Gamma_u)$, we define the current $T_{u,\theta} = [\![\Gamma_u, \vec{\xi}_u, \theta]\!]$ and $\beta(x) \doteq \theta(x, u(x))$. Through the area formula, see for instance [5, Proposition 6.4], we have

$$\mathbb{M}(T_{u,\theta}) = \int_\Omega \mathcal{A}(Du(x))\beta(x) dx,$$

where

$$\mathcal{A}(X) \doteq \sqrt{\det(\mathrm{id}_{\mathbb{R}^{m\times m}} + X^T X)} \tag{A.6}$$

is the *area function*. Notice that in the case $n = m = 2$, $\mathcal{A}$ has the form (4.4). In particular, by the definition of norm of a m-vector, we notice that

$$\mathcal{A}(Du) = \|v_1 \wedge \cdots \wedge v_m\|, \tag{A.7}$$

where we have used the notation of (A.5).

A.5. Geometric functionals

Given a smooth and 1-homogeneous function $\Psi : \Lambda_s(\mathbb{R}^{n+m}) \to \mathbb{R}$, we can define the functional on rectifiable currents $T = [\![E, \vec{T}, \theta]\!]$

$$\Sigma_\Psi(T) \doteq \int_E \Psi(\vec{T}(x))\theta(x)d\mathcal{H}^m(x) = \int \Psi(\vec{T})\, d\|T\|, \tag{A.8}$$

as done in (1.1). On varifolds, given a smooth integrand $F : \mathbb{G}(m, n+m) \to \mathbb{R}$, the counterpart of the previous energy has the form

$$\Sigma'_F(V) \doteq \int_E F(T_x E)\theta(x)d\mathcal{H}^m(x), \tag{A.9}$$

if $V = [\![E, \theta]\!]$. In particular, any even integrand $\Psi : \Lambda_s(\mathbb{R}^{n+m}) \to \mathbb{R}$ as above allows us to define a functional on varifold too. The minimal hypotheses that one requires on an integrand Ψ to get lower semicontinuity of the energy Σ_Ψ, see [11, Section 5.1], is that Ψ has positive values and it satisfies *Almgren's ellipticity condition*, i.e.

$$\Sigma_\Psi(T) - \Sigma_\Psi(Q) \geq 0, \tag{A.10}$$

whenever $\partial T = \partial Q$, Q has support contained in an m subspace of $\mathbb{R}^{n+m}$ whose orienting m-vector is $\vec{\tau} = v_1 \wedge \cdots \wedge v_m$ and the orientation of Q is given by τ. We say it satisfies a uniform Almgren ellipticity condition if there exists $\varepsilon > 0$ such that

$$\Sigma_\Psi(T) - \Sigma_\Psi(Q) \geq \varepsilon(\mathbb{M}(T) - \mathbb{M}(Q)). \tag{A.11}$$

We give now the definition of stationarity in the sense of currents (or varifolds). Fix an energy Σ_Ψ and let $\mathcal{U} \subset \mathbb{R}^{m+n}$ be open. Given any function $g \in C_c^\infty(\mathcal{U}, \mathbb{R}^{n+m})$, we define the flow $X_\varepsilon(x) \doteq \gamma_x(\varepsilon)$, where γ_x is the solution of the ODE

$$\begin{cases} \gamma'(t) = g(\gamma(t)) \\ \gamma(0) = x. \end{cases} \tag{A.12}$$

We define the variation of T with respect to the vector field $g \in C_c^1(\mathcal{U}; \mathbb{R}^{m+n})$ as

$$[\delta_\Psi T](g) \doteq \lim_{\varepsilon \to 0} \frac{\Sigma_\Psi((X_\varepsilon)_\# T) - \Sigma_\Psi(T)}{\varepsilon}. \tag{A.13}$$

Finally, the current T is said to be *stationary in* $\mathcal{U}$ if $[\delta_\Psi T](g) = 0, \forall g \in C_c^1(\mathcal{U}; \mathbb{R}^{m+n})$. With obvious modifications, this definition holds for varifolds as well.

A simple computation shows the following characterization of the first variation of a geometric functional

Lemma 5.7 *Let $T = [\![T, \vec{\tau}, \theta]\!]$ with $\vec{\tau} = \tau_1 \wedge \cdots \wedge \tau_m$ and $\|\vec{\tau}(x)\| = 1$. For any, $g \in C_c^1(\mathcal{U}, \mathbb{R}^{m+n})$,*

$$[\delta_\Psi T](g) = \sum_{i=1}^m \int_E \langle d\Psi(\vec{\tau}(x)), \tau_1(x) \wedge \cdots \wedge Dg(x)\tau_i(x) \wedge \cdots \wedge \tau_m(x)\rangle \theta(x)d\mathcal{H}^m(x).$$

A.6. Functionals on graphs

It was shown in [5, Section 6] that from a functional defined on varifolds, one can define a functional on graphs, simply using the area formula. To do so, we introduced the map $h : \mathbb{R}^{n\times m} \to \mathbb{R}^{(n+m)\times(n+m)}$ defined as

$$h(X) \doteq M(X)S(X)M(X)^T$$

where

$$M(X) \doteq \begin{pmatrix} \mathrm{id}_m \\ X \end{pmatrix} \quad \text{and} \quad S(X) \doteq (M(X)^T M(X))^{-1}, \tag{A.14}$$

or, more explicitely,

$$h(X) = \left[\begin{array}{c|c} S(X) & S(X)X^T \\ \hline XS(X) & XS(X)X^T \end{array}\right]. \tag{A.15}$$

This represents the orthogonal projection on the plane

$$\tau(X) = \mathrm{span}\{M(X)^T e_1, \dots, M(X)^T e_{n+m}\},$$

and is the parametrization of one chart of $\mathbb{G}(m, n+m)$. If F is an integrand as in (A.9), we can define

$$f(X) \doteq F(h(X))\mathcal{A}(X), \tag{A.16}$$

where $\mathcal{A}$ is the area element defined in (A.6). The following holds

$$\Sigma'_F(V_{u,\theta}) = \int_\Omega f(Du(x))\beta(x)dx, \tag{A.17}$$

for every $V_{u,\theta} = [\![\Gamma_u, \theta]\!]$, where $\beta(x) = \theta(x, u(x))$. In [5, Proposition 6.6], we proved the previous equality in the case $\theta \equiv 1$, but the case with multiplicity holds with the same proof. One can do the same for functionals defined on currents in the following way. Let $\Psi \in C^\infty(\Lambda^s_m(\mathbb{R}^{n+m}))$ and associate, to $X \in \mathbb{R}^{n\times m}$, the simple vector

$$W(X) \doteq M^1(X) \wedge \cdots \wedge M^m(X),$$

where $M^i(X)$ denotes the i-th column of the matrix $M(X)$ defined in (A.14). If we define

$$f(X) \doteq \Psi(W(X))\mathcal{A}(X), \tag{A.18}$$

the area formula once again yields the equality

$$\Sigma_\Psi(T_{u,\theta}) = \int_\Omega f(Du(x))\beta(x)dx, \tag{A.19}$$

for every $T_{u,\theta} = [\![\Gamma_u, \vec{\xi}_u, \theta]\!]$, where $\beta(x) = \theta(x, u(x))$.

Finally, let us discuss the link between stationarity for geometric objects and stationarity in the graph sense. We refer the interested reader to [5, Proposition 6.8] for a more precise statement in the case of multiplicity 1 graphs.

Proposition 5.8 *Let $F : \mathbb{G}(m, n+m) \to \mathbb{R}$ or $\Psi : \Lambda^s_m(\mathbb{R}^{n+m}) \to \mathbb{R}$ be given and define f through formula (A.16) or (A.18), respectively. Let Ω be a Lipschitz, bounded, open subset of $\mathbb{R}^m$ and $\beta \in L^1(\Omega, \mathbb{R}^+)$. A map $u \in \mathrm{Lip}(\Omega, \mathbb{R}^n)$ satisfies*

$$\begin{cases} \displaystyle\int_\Omega \langle Df(Du), Dv\rangle\beta(x)dx = 0 & \forall v \in C^1_c(\Omega, \mathbb{R}^n) \\ \displaystyle\int_\Omega \langle Df(Du), Du D\phi(x)\rangle\beta(x)dx - \int_\Omega f(Du)\,\mathrm{div}(\phi)\beta(x)dx = 0 & \forall \phi \in C^1_c(\Omega, \mathbb{R}^m), \end{cases}$$

if and only if the rectifiable varifold $V_{u,\theta} = [\![\Gamma_u, \theta]\!]$ *or the rectifiable current* $T_{u,\theta} = [\![\Gamma_u, \vec{\xi}_u, \theta]\!]$, *where* $\theta(x, y) = \beta(x), \forall (x, y) \in \mathbb{R}^{m+n}$, *are stationary with respect to* Σ'_F *or* Σ_Ψ, *respectively.*

Proof Since the proof is essentially the same of [5, Proposition 6.8], we only sketch it. In [5, Proposition 6.8], only the varifold case was considered, hence let us consider the case of functionals defined on currents here.

Step 1: Reduction to special vector fields.
Define, for any $g \in C^1_c(\Omega \times \mathbb{R}^n, \mathbb{R}^{m+n})$, $g = (g_1, \dots, g_{n+m},)$, two fields $g^1 \doteq (g_1, \dots, g_m, 0, \dots, 0)$ and $g^2 \doteq (0, 0, g_{m+1}, \dots, g_{n+m})$, so that $g = g^1 + g^2$. From now on, consider g fixed. From Lemma 5.7, we see that the first variation $[\delta_\Psi T]$ (see the notation introduced in (A.13)) enjoys the following properties:

$$[\delta_\Psi T_u](g) = [\delta_\Psi T_u](g^1) + [\delta_\Psi T_u](g^2). \tag{A.20}$$

and

$$[\delta_\Psi T_u](g) = [\delta_\Psi T_u](h), \qquad \forall g, h \in C^1_c(\Omega \times \mathbb{R}^n, \mathbb{R}^{m+n}), g|_{\mathrm{spt}(\Gamma_u)} = h|_{\mathrm{spt}(\Gamma_u)}. \tag{A.21}$$

(A.20) is trivial, while to show (A.21), simply notice that if

$$g|_{\mathrm{spt}(\Gamma_u)} = h|_{\mathrm{spt}(\Gamma_u)},$$

then

$$Dg(x, u(x))w = Dh(x, u(x))w, \text{ for a.e. } x \in \Omega \text{ and } w \in T_{(x,u(x))}\Gamma_u. \tag{A.22}$$

By exploiting the explicit form of the first variation written in Lemma 5.7, (A.21) follows at once. From (A.21) we conclude that it suffices to consider the first variation of the current T_u for vector fields g of the form

$$g(x, y) = \chi(y)G(x, u(x)), \tag{A.23}$$

for $G \in C^1_c(\Omega \times \mathbb{R}^n, \mathbb{R}^{n+m})$, and $\chi \in C^\infty_c(\mathbb{R}^n)$, $\chi(y) \equiv 1$ on $B_{2M}(0) \subset \mathbb{R}^n$ and $\chi(y) \equiv 0$ on $B_{3M+1}(0)$, where $M \doteq \max_{x \in \pi(\mathrm{spt}(G))} \|u(x)\|$. $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ here denotes the projection $\pi(x, y) \doteq x, \forall (x, y) \in \mathbb{R}^{m+n}$, $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$.

Step 2: Inner variations.
We let X_ε be the flow generated by g^1, for g as in (A.23). It is easy to see that

$$X_\varepsilon(x, u(x)) = (Z_\varepsilon(x), u(x)), \forall x \in \Omega,$$

where Z_ε is the flow generated by the field $x \mapsto G^1(x, u(x))$. Using this information, one readily checks that

$$(X_\varepsilon)_\# T_u = [\![X_\varepsilon(\Gamma_u), \vec{\tau}_\varepsilon, \theta \circ (X_\varepsilon)^{-1}]\!] = [\![\Gamma_{u \circ Z_{-\varepsilon}(\cdot)}, \vec{\xi}_{u \circ Z_{-\varepsilon}}, \theta \circ (\cdot, Z_{-\varepsilon}(\cdot))]\!] = T_{u \circ Z_{-\varepsilon}, \theta \circ (\cdot, Z_{-\varepsilon}(\cdot))}.$$

Through formula (A.19), we see that

$$\begin{aligned}\Sigma_\Psi(T_{u \circ Z_{-\varepsilon}, \theta \circ (\cdot, Z_{-\varepsilon}(\cdot))}) &= \int_\Omega f(D(u \circ Z_{-\varepsilon}))\beta \circ Z_{-\varepsilon}(x)\, dx \\ &= \int_\Omega f(Du(x)DZ_{-\varepsilon}(Z_\varepsilon(x)))\, \beta(x) \det(DZ_\varepsilon)\, dx\,.\end{aligned}$$

By taking the derivative at $\varepsilon = 0$ of the previous expression, we get

$$[\delta_\Psi T](g^1) = \int_\Omega \langle Df(Du), Du D(G^1(x, u(x)))\rangle \beta(x)dx - \int_\Omega f(Du)\,\mathrm{div}(G^1(x, u(x)))\beta(x)dx \tag{A.24}$$

Step 3: Outer variations.
Similarly to the case above, consider the flow Y_ε generated by g^2, for g as in (A.23). Then, one checks that

$$Y_\varepsilon(x, u(x)) = (x, u(x) + \varepsilon G^2(x, u(x)),$$

and hence

$$(Y_\varepsilon)_\# T_u = [\![Y_\varepsilon(\Gamma_u), \vec{\tau}_\varepsilon, \theta \circ Y_{-\varepsilon}]\!] = [\![\Gamma_{v_\varepsilon}, \vec{\xi}_{v_\varepsilon}, \theta \circ Y_{-\varepsilon}]\!] = T_{v_\varepsilon, \theta \circ Y_{-\varepsilon}}.$$

By (A.19), we write

$$\Sigma((Y_\varepsilon)_\# T_u) = \int_\Omega f(Du + \varepsilon D(G^2(x, u(x))))\beta\, dx,$$

whose derivative at $\varepsilon = 0$ yields

$$[\delta_\Psi T](g^2) = \int_\Omega \langle Df(Du), D(G^2(x, u(x)))\rangle \beta(x)dx. \tag{A.25}$$

Now the Proposition follows at once from (A.20)-(A.21)-(A.24)-(A.25). □

References

1. Allard, W.K.: An integrality theorem and a regularity theorem for surfaces whose first variation with respect to a parametric elliptic integrand is controlled, Geometric measure theory and the calculus of variations (1986)
2. Allard, W.K.: On the first variation of a varifold. Ann. Math. Second Series (1972)
3. Chlebik, M., Kirchheim, B.: Rigidity for the four gradient problem. Journal für die reine und angewandte Mathematik (Crelles Journal) **551** (200201), 1-9
4. Dacorogna, B., Maréchal, P.: The role of perspective functions in convexity, polyconvexity, rank-one convexity and separate convexity. J. Convex Anal. **15**(2), 271–284 (2008)
5. De Lellis, C., De Philippis, G., Kirchheim, B., Tione, R.: Geometric measure theory and differential inclusions, Accepted Paper: Annales de la Faculté des Sciences de Toulouse, arXiv:1910.00335 (2019)
6. De Philippis, G., De Rosa, A., Ghiraldin, F.: Rectifiability of Varifolds with Locally Bounded First Variation with Respect to Anisotropic Surface Energies. Commun. Pure Appl. Math. **71**(6), 1123–1148 (2017)
7. De Philippis, G., De Rosa, A., Hirsch, J.: The area blow up set for bounded mean curvature submanifolds with respect to elliptic surface energy functionals. Discrete Contin. Dyn. Syst. **39**(12), 7031–7056 (2019)
8. De Rosa, A., Kolasiñski, S.: Equivalence of the ellipticity conditions for geometric variational problems. Commun. Pure Appl. Math. **73**(11), 2473–2515 (2020)
9. De Rosa, A., Tione, R.: Regularity for graphs with bounded anisotropic mean curvature, arXiv:2011.09922 (2020)
10. Duggan, J.P.: Regularity theorems for varifolds with mean curvature, Ph.D. Thesis, (1986)
11. Federer, H.: Geometric Measure Theory. Springer, Berlin (1969)
12. Figalli, A.: The Monge-Ampère Equation and Its Applications, Zurich Lectures in Advanced Mathematics, European Mathematical Society, (2017)
13. Förster, C., Székelyhidi, L.: T5-Configurations and non-rigid sets of matrices. Calc. Var. Partial Differ. Equ. **57**(1), 19 (2017)
14. Giaquinta, M., Modica, G., Soucek, J.: Cartesian Currents in the Calculus of Variations, vol. I. Springer, Berlin (1998)
15. Giaquinta, M., Modica, G., Soucek, J.: Cartesian Currents in the Calculus of Variations, vol. II. Springer, Berlin (1998)

16. Kirchheim, B.: Rigidity and Geometry of Microstructures. Habilitation thesis, University of Leipzig (2003)
17. Kirchheim, B., Müller, S., Šverák, V.: Studying nonlinear pde by geometry in matrix space. In: Hildebrandt, S., Karcher, H. (eds.) Geometric Analysis and Nonlinear Partial Differential Equations. Springer, Berlin, Heidelberg (2003). https://doi.org/10.1007/978-3-642-55627-2_19
18. Lorent, A., Peng, G.: Null Lagrangian measures in subspaces, compensated compactness and conservation laws. Arch. Rational Mech. Anal. **234**(2), 857–910 (2019)
19. Lorent, A., Peng, G.: On the Rank-1 convex hull of a set arising from a hyperbolic system of Lagrangian elasticity. Calc. Var. Part. Differ. Equ. **59** (August 2020), no. 5
20. Müller, S., Šverák, V.: Convex integration for Lipschitz mappings and counterexamples to regularity. Ann. Math. Second Ser. **157**(3), 715–742 (2003)
21. Rockafellar, R.T.: Convex Analysis, Princeton Landmarks in Mathematics and Physics. Princeton University Press (1970)
22. Schoen, R., Simon, L.: A new proof of the regularity theorem for rectifiable currents which minimize parametric elliptic functionals. Indiana Univ. Math. J. **31**(3), 415–434 (1982)
23. Simon, L.: Lectures on Geometric Measure Theory. Australian National University (2008)
24. Székelyhidi Jr., L.: The regularity of critical points of polyconvex functionals. Arch. Ration. Mech. Anal. **172**(1), 133–152 (2004)
25. Székelyhidi Jr., L.: Rank-one convex hulls in $\mathbb{R}^2 \times 2$. Calc. Variations Part. Differ. Equ. **28**(4), 545–546 (2007)
26. Tione, R.: Minimal graphs and differential inclusions. Commun Part Differ. Equ. (2021). https://doi.org/10.1080/03605302.2020.1871367

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.