MINIMAL CURRENTS, GEODESICS AND RELAXATION OF VARIATIONAL INTEGRALS ON MAPPINGS OF BOUNDED VARIATION

By

Patricio Aviles

and

Yoshikazu Giga

IMA Preprint Series # 883

October 1991

Minimal currents, geodesics and relaxation of variational integrals on mappings of bounded variation

Patricio Aviles
Department of Mathematics
University of Illinois
Urbana, Illinois 61801

Yoshikazu Giga
Department of Mathematics
Hokkaido University
Sapporo 060, JAPAN

1. Introduction

We consider a functional $\mathcal{F}$ of C^1 mapping u from an open set Ω in $\mathbb{R}^n$ into $\mathbb{R}^m$:

$$\mathcal{F}(u) = \int_\Omega f(x, u(x), \nabla u(x))dx, \tag{1.1}$$

with the density function

$$f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \to \mathbb{R}, \quad f = f(x, y, \xi) \tag{1.2}$$

which is nonnegative, continuous, convex in ξ and linear growth in ξ. We are interested in extending $\mathcal{F}$ to the space of mappings of bounded variation $BV(\Omega, \mathbb{R}^m)$ in a natural way. It often happens that in variational problems no minimizers exist in the space of C^1 mappings, so it is necessary to extend $\mathcal{F}$ to a wider class so that minimizers exist. There is a natural but abstract way to extend $\mathcal{F}$ to $BV(\Omega, \mathbb{R}^m)$ which is called the L^1_{loc}– lower semicontinuous *relaxation* of $\mathcal{F}$. It is defined by

$$\overline{\mathcal{F}}(u) = \inf\left\{ \liminf_{\ell\to\infty} \mathcal{F}(u_\ell); \quad u_\ell \in C^1(\Omega, \mathbb{R}^m), \quad u_\ell \to u \text{ in } L^1_{\text{loc}}(\Omega, \mathbb{R}^m)\right\}$$

for $u \in BV(\Omega, \mathbb{R}^m)$. In other words $\overline{\mathcal{F}}$ is the greatest lower semicontinuous function on $BV(\Omega, \mathbb{R}^m)$ less than $\mathcal{F}$ on $C^1(\Omega, \mathbb{R}^m)$. The idea of relaxation goes back to H. Lebesgue in the definition of area of non parametric surfaces. It was strengthened by Serrin [Se 1,2] and extended to a very general setting by De Giorgi. Since the value of $\overline{\mathcal{F}}$ is implicitly defined, the problem is to find an explicit integral representation of $\overline{\mathcal{F}}$. This problem is posed by De Giorgi [DG] when $m > 1$ and f depends on y. The goal of this paper is to solve this problem for a class of f, which we call isotropic densities (see (1.5)).

Finding a representation of relaxation is also important in many other context for example in the theory of harmonic maps [BBC] [GMS3] and in the theory of elasticity [GMS2].

We briefly explain our representation formula. In this section we restrict ourselves in the case when f is positively homogeneous of degree one in ξ, for simplicity (see §4 for general results). Our main representation formula reads:

$$\begin{aligned} \overline{\mathcal{F}}(u) = &\int_{\Omega\setminus\Sigma} f\left(x, u(x), \frac{d\nabla u}{d|\nabla u|}(x)\right)|\nabla u| \\ &+ \int_{\Sigma} D_x(u^-(x), u^+(x), \nu(x))d\mathcal{H}^{n-1}(x) \end{aligned} \tag{1.3}$$

provided that f is coercive and satisfies a growth condition:

$$\begin{aligned} &(a) \qquad C_0|\xi| \le f(x,y,\xi) \qquad \text{with } C_0 > 0 \\ &(b) \qquad f(x,y,\xi) \le C|\xi| \quad \text{with } C > 0 \end{aligned} \tag{1.4}$$

and that f satisfies an isotropic condition

$$f(x,y,\xi) \ge f(x,y,\xi q \otimes q), \quad |q| = 1, \quad q \in \mathbb{R}^n, \tag{1.5}$$

where ξ is identified with an $m \times n$ matrix. Here $D_x(a,b,q)$ is a distance like function defined by

$$D_x(a,b,q) = \inf\{\int_0^1 f(x,\gamma(t),\dot\gamma(t)\otimes q)dt;$$

$$\gamma : [0,1] \to \mathbb{R}^m \text{ is Lipschitz and } \gamma(0) = a,\ \gamma(1) = b\}$$

for $a, b \in \mathbb{R}^m, q \in \mathbb{R}^n$. The set Σ is the set of (approximate) jump discontinuities of u and ν represents a unit normal to Σ. The functions $u^\pm$ are the trace of u on Σ defined by

$u^{\pm}(x) = \underline{\lim}_{\varepsilon\downarrow 0} u(x \pm \varepsilon\nu(x))$ and $\mathcal{H}^{n-1}$ denotes the $n-1$ dimensional Hausdorff measure; $d\mu/d|\mu|$ denotes the Radon–Nikodym derivative of μ with respect to its total variation measure $|\mu|$. For the definition of these terminology see [F1, Giu, Si].

An explicit representation of $\overline{\mathcal{F}}$ is previously obtained when f does not depend on ξ by Goffman and Serrin [GS]; when f is independent of y by Reshetnyak [Re] and Giaquinta, Modica and Souček [GMS1]; when $m = 1$ by Dal Maso [DM]; when $n = 1$ by Rockaffelar [Ro]. Our formula (1.3) is a first natural extension of these results for $m > 1$ and $n > 1$ when f depends on y. One may interpret that (1.3) is an extension of Dal Maso's representation to the vector valued case.

Our results are based on [AG1] where the authors estimate $\overline{\mathcal{F}}$ from below by using a *minimal graph* (current) associated with the graph of u (even without assuming (1.4), (1.5)). In the present paper we further investigate the minimal graph and obtain the inequality '$\geq$' in (1.3), assuming (1.4) and (1.5). This process conceptually resembles characterization of Young measure in the theory of elasticity [D], [JK]. The opposite inequality is shown by Ambrosio, Mortola and Tortorelli [AMT]. Moreover, they proved that the inequality '$\leq$' in (1.3) is not true without (1.5).

Our crucial step generalizes the elementary fact that the straight line from a point a to b in $\mathbb{R}^m$ minimizes mass among all one *currents* T whose boundary consists of oriental two points a and b. We shall prove that the same fact is true for general 'metric' F if we replace straight lines by F-geodesic and mass by 'F-weighted' mass under certain nondegeneracy assumptions of F (Theorem 3.1). Our result generalizes a result of Federer [F2, 5. 12] (Remark 3.2). Moreover, our proof is different from his. We approximate minimizing current by real polygonal chain (§2) and reduce the problem to network flow problem. By solving the network flow problem we conclude that the minimum is attained at an integral current. We note that area minimizing integral current may not be a minimizer among all currents when dimension and codimension of currents are more than one. This is first shown by L. C. Young [Y]. For further development see [F2], [W] and [Mor].

The assumption (1.4a) is too restrictive to apply to the theory of phase transitions. A typical example from the phase transition theory is

$$f(x, y, \xi) = |y - \alpha||y - \beta||\xi|, \quad \alpha, \beta \in \mathbb{R}^m, \tag{1.6}$$

which corresponds to the energy density of a bi-stable phase. In the context of the van der Waals–Cahn–Hilliard theory of phase transitions $\mathcal{F}$ with f in (1.6) is obtained as a singular limit of some energy under an integral constraint; see [M, St1] for $m = 1$ and [Ba, FT, St2] for $m > 1$. For further reference on phase transitions see references cited in [AG1]. Their results in particular yields a representation of $\overline{\mathcal{F}}(u)$ for u with two values α and β. Although our assumptions exclude (1.6) it is very likely that our theory will apply to (1.6) by extending our approximate Lemma 2.2 for weighted (degenerate) mass.

The results in this paper have been announced in [AG2].

After this work was completed, we were informed of recent work of I. Fonseca and P. Rybka [FR]. They obtain a representation formula without assuming the isotropy condition (1.5) but for a bounded domain Ω with Lipschitz boundary. Their definition of $\overline{\mathcal{F}}$ apparently differs from ours but for a bounded domain Ω it agrees with ours. Their method is quite different from ours.

2. Strong polyhedral approximation

This section constructs a real polygonal approximation of a 1-current. There is a standard result in this direction [F1, 4.1.23] where a flat chain is approximated by real polyhedral chains together with its mass. In our case we should arrange the approximation so that it has the same boundary as the original one. Moreover,our original current is not necessarily compactly supported. Although our results may be known (among a few top specialists), it seems there is no literature including our approximation results, so we state our results with a proof for the reader's convenience. For the approximation by compactly supported currents (Lemma 2.2) we shall give an analytic proof without using geometric measure theory although it is possible to prove this fact just by using geometric measure theory. We shall give the second proof in the Appendix. After this approximation we further approximate by polygonal chains (Lemma 2.3). A general version (without dimension restriction) is given in Appendix. This proof is due to R.Hardt. An elementary proof of a weaker version of Lemma 2.3 is given in [AG2] when the space dimension $m = 2$.

We first establish conventions of notation ([F1]). Let $\mathcal{D}^k$ denote the space of smooth k-forms with compact support in $\mathbb{R}^m$. The space $\mathcal{D}_k$ of k-currents is the topological dual of $\mathcal{D}^k$ equipped with the usual locally convex topology. For an oriented Lipschitz curve Γ parametrized by

$$\gamma : \ [0,1] \to \Gamma \subset \mathbb{R}^m,$$

there is a corresponding 1-current $[\![\Gamma]\!] \in \mathcal{D}_1$ defined by

$$[\![\Gamma]\!](\omega) := \int_\Gamma \omega \ \Big(= \sum_{j=1}^m \int_0^1 \omega_j(\gamma(t))\dot{\gamma}_j(t)dt \Big), \tag{2.1}$$

$$\omega = \sum_{j=1}^m \omega_j(y)dy_j \in \mathcal{D}^1,$$

where $\gamma = (\gamma_1, \cdots, \gamma_m)$ and $\dot{\gamma}_j = d\gamma_j/dt$. Let d denote the exterior derivative operator. For $T \in \mathcal{D}_k$ the boundary $\partial T \in \mathcal{D}_{k-1}$ is defined by

$$\partial T(\omega) = T(d\omega), \quad \omega \in \mathcal{D}^{k-1}.$$

For example, $\partial[\![\Gamma]\!] \in \mathcal{D}_0$ is expressed as

$$\partial[\![\Gamma]\!](\varphi) = [\![\Gamma]\!](d\varphi) = \varphi(b) - \varphi(a), \quad \varphi \in \mathcal{D}^o,$$

where $\gamma(0) = a, \ \gamma(1) = b$. In other expression

$$\partial[\![\Gamma]\!] = \delta_b - \delta_a, \tag{2.2}$$

where δ_a denotes the Dirac measure supported at a point $a \in \mathbb{R}^m$. As expected the boundary of $[\![\Gamma]\!]$ consists of two oriented points.

Let $\mathbf{M}(T)$ be a mass of current $T \in \mathcal{D}_k$ as in [F1]. If $T \in \mathcal{D}_1$, we have

$$\mathbf{M}(T) = \sup\Big\{T(\omega); \ |\omega| \leqslant 1, \ \omega \in \mathcal{D}^1\Big\},$$

where $|\cdot|$ is the norm of the space $\wedge^1 \cong \mathbb{R}^m$ of 1-covectors associated with the standard inner product. The value $\mathbf{M}(T)$ is called the *total mass* of T. For example, it is well known that

$$\mathbf{M}([\![\Gamma]\!]) = \text{the length of } \Gamma = \int_0^1 |\dot{\gamma}(t)|dt.$$

If Γ is an oriented segment from a to b, $[\![\Gamma]\!]$ is denoted $[\![a,b]\!]$. A *real polygonal chain* (or polyhedral 1-chain) is a current of a finite linear combination of $[\![a,b]\!]$. In other words the space of real polygonal chains is

$$\mathbf{P}_1 = \Big\{ \sum_{\ell=1}^{r} c_\ell [\![a_\ell, b_\ell]\!];\ r \text{ is a positive integer}$$
$$\text{and } c_\ell \in \mathbb{R},\ a_\ell, b_\ell \in \mathbb{R}^m \text{ for } \ell = 1, 2, \cdots, r \Big\}.$$

of course $P \in \mathbf{P}_1$ is compactly supported and $\mathbf{M}(P) < \infty$. We say $T \in \mathcal{D}_1$ is a *locally finite mass current* if

$$\sup \Big\{ T(\omega);\ |\omega| \leqslant 1,\ \omega \in \mathcal{D}_1, \operatorname{spt}\omega \subset W \Big\}$$

is finite for every open W whose closure $\overline{W}$ is compact in $\mathbb{R}^m$. Let $\mathcal{M}_1$ be the space of locally finite mass 1-current. Of course $\mathbf{P}_1 \subset \mathcal{M}_1$. It is convenient to recall a flat norm of $T \in \mathcal{D}_k$:

$$\mathbf{F}(T) = \inf \Big\{ \mathbf{M}(R) + \mathbf{M}(S);\ T = \partial R + S \Big\}.$$

We now state our key approximation result.

THEOREM 2.1. *Suppose that $T_0 \in \mathcal{M}_1$ with $\partial T_0 = 0$ has a finite total mass $\mathbf{M}(T_0)$. For every $\varepsilon > 0$ there are a polygonal chain P_ε, a 2-current S_ε with compact support and $R_\varepsilon \in \mathcal{M}_1$ with $\partial R_\varepsilon = 0$ such that T_0 is of the form*

$$T_0 = P_\varepsilon + R_\varepsilon + \partial S_\varepsilon$$

with

$$\mathbf{M}(P_\varepsilon) \leqslant \mathbf{M}(T_0) + \varepsilon \tag{2.3}$$
$$\mathbf{M}(R_\varepsilon) + \mathbf{M}(S_\varepsilon) \leqslant \varepsilon.$$

In particular

$$\mathbf{F}(T_0 - P_\varepsilon) \leqslant \varepsilon. \tag{2.4}$$

This is a combination of the following two approximation lemmas.

Lemma 2.2 (Approximation by compactly supported currents). *Suppose that $T_0 \in \mathcal{M}_1$ with $\partial T_0 = 0$ has a finite total mass $\mathbf{M}(T_0)$. For every $\varepsilon > 0$ there is $T_1 \in \mathcal{M}_1$ with $\partial T_1 = 0$ such that T_1 is compactly supported in $\mathbb{R}^m$ and that*

$$\mathbf{M}(T_0 - T_1) \leqslant \varepsilon.$$

Lemma 2.3 (Approximation by polygonal chains). *Suppose that $T_1 \in \mathcal{M}_1$ is compactly supported with $\partial T_1 = 0$. For every $\varepsilon > 0$ there are a polygonal chain $P \in \mathbf{P}_1$ and a 2-current S with compact support such that*

$$T_1 = P + \partial S$$

with $\mathbf{M}(P) \leqslant \mathbf{M}(T_1) + \varepsilon$ and $\mathbf{M}(S) \leqslant \varepsilon$.

Theorem 2.1 now follows from Lemmas 2.2 and 2.3 with ε replaced by $\varepsilon/2$ by setting $P_\varepsilon = P$, $S_\varepsilon = S$, $R_\varepsilon = T_0 - T_1$.

We shall give the proofs of Lemmas 2.2 and 2.3 based on geometric measure theory (and its generalizations for higher dimensional currents) in the Appendix. Since we have a short analytic proof of Lemma 2.2, we shall present it at the end of this section. An application of Theorem 2.1 is the following approximation result which we use in §3.

Theorem 2.4. *Suppose that $T \in \mathcal{M}_1$ satisfies $\partial T = \delta_b - \delta_a$ and $\mathbf{M}(T)$ is finite. There is a sequence $\{Q_\varepsilon\}$ of real polyhedral chains with $\partial Q_\varepsilon = \delta_b - \delta_a$ such that*

$$\lim_{\varepsilon \to 0} \mathbf{M}(Q_\varepsilon) = \mathbf{M}(T) \tag{2.5}$$

and

$$\lim_{\varepsilon \to 0} \mathbf{F}(T - Q_\varepsilon) = 0. \tag{2.6}$$

In particular $Q_\varepsilon \rightharpoonup T$, i.e.,

$$\lim_{\varepsilon \to 0} Q_\varepsilon(\omega) = T(\omega) \quad \text{for all} \quad \omega \in \mathcal{D}^1. \tag{2.7}$$

A sketch of a proof of Theorem 2.4 without (2.6) for two dimensional space $\mathbb{R}^2$ is found in a short paper [AG2].

Theorem 2.4 easily follows from Theorem 2.1. To see this we need to recall some properties of $T \in \mathcal{M}_1$. By the Riesz representation theorem $T \in \mathcal{M}_1$ is (uniquely) expressed as

$$T(\omega) = \int_{\mathbb{R}^m} \omega_j(y)\xi_{Tj}(y)d\mu_T(y), \quad \omega = \sum_{j=1}^{m} \omega_j(y)dy_j \in \mathcal{D}^1,$$

where $\xi_T = (\xi_{T1}, \cdots, \xi_{Tm})$ is μ_T-measurable and $|\xi_T| = 1,\ \mu - a.e.$ and μ_T is a nonnegative Radon measure on $\mathbb{R}^m$ (see, e.g. [Si, Theorem 4.1]). The measure μ_T is called the total variation measure associated with T and

$$\mu_T(W) = \sup\left\{T(\omega);\ |\omega| \leqslant 1,\ \operatorname{spt}\omega \subset W\right\}, \quad W \subset\subset U.$$

We thus observe that $T \in \mathcal{M}_1$ is identified with a $\mathbb{R}^m$-valued Radon measure $(T^1, \cdots, T^m)$ on $\mathbb{R}^m$, or a vector field on $\mathbb{R}^m$ with measures as coefficients. The function ξ_T in (2.8) agrees with the Radon-Nikodym derivative $dT/d\mu_T$ and the measure μ_T agrees with usual total variation measure defined for vector-valued Radon measure. For example, if $T = [\![\Gamma]\!]$ is defined by (2.1), it is easy to see that ξ_T is the unit tangent vector to Γ (consistent with the orientation of Γ) and $\mu_T = \mathcal{H}^1 \lfloor \Gamma$. Here $\mathcal{H}^k$ denotes the k-dimensional Hausdorff measure and for a measure μ and a set $A \subset \mathbb{R}^m$ a new measure $\mu \lfloor A$ is defined by

$$(\mu \lfloor A)(B) := \mu(A \cap B), \quad B \subset \mathbb{R}^m.$$

PROOF OF THEOREM 2.4: We observe that there is a piecewise linear curve Γ with $\partial[\![\Gamma]\!] = \delta_b - \delta_a$ such that

$$\mathbf{M}(T) = \mathbf{M}([\![\Gamma]\!]) + \mathbf{M}(T - [\![\Gamma]\!]). \tag{2.9}$$

Indeed, let H be a hyperplane orthogonal to the segment $[a,b]$ such that the middle point $(a+b)/2$ belongs to H. For each $h \in H$ we consider a piecewise linear curve Γ_h consisting of two oriented segment $[a,h]$ and $[h,b]$. Such Γ_h's are mutually disjoint outside a, b. We see $\mu_T \lfloor \Gamma_h = 0$ except for countably many h, otherwise $\mathbf{M}(T) = \infty$. We take h such that $\mu_T \lfloor \Gamma_h = 0$ and observe

$$\begin{aligned}\mathbf{M}(T) = \mathbf{M}(T\lfloor \Gamma_h^c) &= \mathbf{M}([\![\Gamma_h]\!]) + \mathbf{M}(T\lfloor \Gamma_h^c - [\![\Gamma_h]\!]) \\ &= \mathbf{M}([\![\Gamma_h]\!]) + \mathbf{M}(T - [\![\Gamma_h]\!]), \quad \Gamma_h^c = \mathbb{R}^m \setminus \Gamma_h\end{aligned}$$

which yields (2.9). Here $T\lfloor A$ denotes the measure defined by

$$T\lfloor A = (T^1 \lfloor A, \cdots, T^m \lfloor A).$$

We now take a piecewise linear curve Γ such that (2.2) and (2.9) hold. From (2.2) it follows that

$$\partial T_0 = 0 \quad \text{with} \quad T_0 = T - [\![\Gamma]\!].$$

We apply Theorem 2.1 to T_0 and obtain the approximation P_ε. If we modify P_ε slightly, we may assume that P_ε with $\partial P_\varepsilon = 0$ satisfies

$$\mathbf{M}(P_\varepsilon + [\![\Gamma]\!]) = \mathbf{M}(P_\varepsilon) + \mathbf{M}([\![\Gamma]\!]) \tag{2.10}$$

as well as (2.3) and (2.4). Since $\mathbf{M}$ is lower semicontinuous in $\mathbf{F}$-topology, (2.3) yields

$$\lim_{\varepsilon \downarrow 0} \mathbf{M}(P_\varepsilon) = \mathbf{M}(T_0).$$

By (2.9) and (2.10) we now observe that $Q_\varepsilon = P_\varepsilon + [\![\Gamma]\!]$ satisfies (2.5). The convergence (2.6) follows from (2.4). Since $\mathbf{F}$-convergence is stronger than weak convergence of currents, we have (2.7) from (2.6). ∎

REMARK 2.5: In Theorem 2.4 the assumption

$$\partial T = \delta_b - \delta_a$$

can be replaced by

$$\partial T = \sum_{l=1}^{r} c_\ell \delta_{a_\ell}, \quad c_\ell \in \mathbb{R}, \quad a_\ell \in \mathbb{R}^m \quad \text{for} \quad 1 \leqslant \ell \leqslant r.$$

Notice that $\sum_\ell c_\ell$ is automatically zero since $\partial T(1) = 0$. In fact it is not difficult to construct $P \in \mathbf{P}_1$ such that

$$\partial P = \partial T \quad \text{and}$$
$$\mathbf{M}(T) = \mathbf{M}(T - P) + \mathbf{M}(P).$$

The approximation T_ε is now constructed in the same way as in the proof of Theorem 2.4 where $[\![\Gamma]\!]$ shall be replaced by P.

In the rest of this section we give an analytic proof of Lemma 2.2. We recall Bogovski's improvement of Poincaré's lemma for the equation $\operatorname{div} u = f$.

LEMMA 2.6 (Bogovski [Bo]). *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^m$. There is a $\mathbb{R}^m$-valued function $k = k_\Omega(x,y)$ on $\Omega \times \Omega$ such that*

(i) k is smooth except $x = y$ and $|k(x,y)| \leqslant C|x-y|^{1-m}$ with C depending only on Ω and m.

(ii) Let K be an operator defined by

$$\Big(Kf\Big)(x) = \int_\Omega k(x,y)f(y)dy.$$

Then K is a linear operator from $C_0^\infty(\Omega)$ to $\Big(C_0^\infty(\Omega)\Big)^m$. More precisely, if the support of f (shortly, $\operatorname{spt} f$) is contained in a compact set $A \in \Omega$, then, there is a compact set A' depending only on A and Ω such that $\operatorname{spt} Kf \subset A'$.

(iii) $\int_\Omega f dx = 0$, then $\operatorname{div} Kf = f$ in Ω.

(iv) If f is a finite Radon measure on Ω, then $Kf \in L^1$ and

$$\|Kf\|_1 \leqslant C'\|f\|_1$$

with $C' = C'(m,\Omega)$, where $\|\cdot\|_1$ denotes the total variation in Ω.

(v) Even if f is merely a finite Radon measure in Ω, $\operatorname{spt} f \subset A \subset \Omega$ implies $\operatorname{spt} Kf \subset A' = A'(\Omega, A)$ as in (ii). Also (iii) holds for a finite Radon measure f in Ω.

PROOF: In [Bo] Bogovski constructed $k = k_\Omega$ satisfying (i)–(iii). If we look over his construction we see that (v) holds. The property (i) yields (iv) since Ω is bounded. For the detail see [BS] ∎

REMARK 2.7: (Scaling) If k_Ω satisfies (i)–(v), then the rescaled kernel

$$k_R(x,y) = R^{1-m}k\Big(\frac{x}{R}, \frac{y}{R}\Big), \quad R > 0$$

satisfies (i)–(v) with domain

$$\Omega_R = \{Rx;\ x \in \Omega\}.$$

Here C in (i) is taken independent of R and C' in (iv) is of the form

$$C' = C_0 R, \quad C_0 = C_0(\Omega).$$

Of course, A' also depends on R. These properties are easily checked by rescaling x and y.

PROOF OF LEMMA 2.2: Let $\varphi \in C_0^\infty(\mathbb{R}^m)$ a cut-off function of $B_{3/2}$ supported in B_2 such that $0 \leqslant \varphi \leqslant 1$, $\varphi = 1$ on $B_{3/2}$, $|\nabla \varphi| \leqslant 2$ in B_2, where B_R is the open ball of radius $R > 0$ centered at the origin. We set $\varphi_R(x) = \varphi(x/R)$.

We regard $T_0 \in \mathcal{M}_1$ as a $\mathbb{R}^m$-valued Radon measure g. Let $\Omega = B_2 \setminus B$, and k_Ω be kernel as in Remark 2.7. We set

$$g_R = \varphi_R g - w_R \quad \text{with} \quad w_R = K_R(\nabla \varphi_R \cdot g),$$

where

$$K_R f = \int_{\Omega_R} k_R(x,y) f(y) dy \quad \text{with} \quad \Omega_R = B_{2R} \setminus \overline{B_R}.$$

Since $\nabla \varphi_R$ is compactly supported in Ω_R, so is w_R by Lemma 2.6. By extending w_R to be zero outside Ω_R, one may regard w_R as an L^1 function defined in $\mathbb{R}^m$ by Lemma 2.6 (iv). By Lemma 2.5 (iii) we see $\operatorname{div} g_R = 0$ since

$$\operatorname{div}(\varphi_R g) = \nabla \varphi_R \cdot g.$$

We now estimate w_R. Since $k_R(x,y) \leqslant C|x-y|^{1-m}$ independent of R and $|\nabla \varphi_R| \leqslant 2/R$, we see

$$\begin{aligned} \|w_R\|_1 &\leqslant \int_{\Omega_R} \int_{\Omega_R} |k_R(x,y)||\nabla \varphi_R(y) \cdot) g(dy)| \\ &\leqslant \frac{2C}{R} \int_{\Omega_R} |g|(dy) \Big(\sup_{y \in \Omega} \int_{\Omega_R} \frac{dx}{|x-y|^{m-1}} \Big). \end{aligned}$$

Since

$$\int_{\Omega_R} \frac{dx}{|x-y|^{m-1}} \leqslant \int_{B_{2R}} \frac{dx}{|x|^{m-1}} = CR \quad \text{with} \quad c = c(m),$$

we end up with

$$\|w_R\|_1 \leqslant 2cC \int_{\Omega_{2R}} |g|(dy).$$

This estimate yields

$$\begin{aligned} \|g - g_R\|_1 &\leqslant \int_{\mathbb{R}^m} (1 - \varphi_R)|g|(dy) + 2cC \int_{\Omega_R} |g|(dy) \\ &\leqslant (1 + 2cC) \int_{\mathbb{R}^m \setminus B_R} |g|(dy) \to 0 \quad \text{as} \quad R \to \infty. \end{aligned}$$

For $\varepsilon > 0$ take $T_1 \in \mathcal{M}_1$ as a current corresponding to g_R such that

$$\|g - g_R\|_1 \leqslant \varepsilon.$$

Since $\partial T_1 = -\operatorname{div} g_R = 0$ and T_1 is compactly supported, T_1 satisfies all properties in Lemma 2.2. ∎

3. Minimal currents and geodesics

We consider a 'metric' density $F(y, \eta)$ defined on $\mathbb{R}^m \times \mathbb{R}^m$. We assume that F satisfies

(F1) $F : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is continuous and nonnegative.

(F2) $F(y, \eta)$ is convex in η.

(F3) $F(y, \eta)$ is positively homogeneous of degree one, i.e.,

$$F(y, \lambda\eta) = \lambda F(y, \eta) \quad \text{for all} \quad \lambda > 0.$$

(F4) $k|\eta| \leqslant F(y, \eta) \leqslant K|\eta|$ holds for all $(y, \eta) \in \mathbb{R}^m \times \mathbb{R}^m$ with $K > k > 0$ independent of y, η.

Note that we do not assume *evenness of* F in η. For an oriented Lipschitz curve Γ parametrized by

$$\gamma : [0, 1] \to \Gamma \subset \mathbb{R}^m$$

we set

$$\ell_F(\gamma) := \int_0^1 F(\gamma(t), \dot\gamma(t))dt, \quad \dot\gamma(t) = \frac{d\gamma}{dt}. \tag{3.1}$$

If $F(y, \eta) = |\eta|$, $\ell_F(\gamma)$ is nothing but the Euclidean length of γ. Since we do not assume evenness of F in η, the value $\ell_F(\gamma)$ may depend on the orientation of the curve Γ.

For a 1-current $T \in \mathcal{M}_1$ we set

$$I_F(T) := \int_{\mathbb{R}^m} F(y, \xi_T(y))d\mu_T(y), \tag{3.2}$$

where ξ_T, μ_T are defined by (2.8). If the measure $T = (T^1, \cdots, T^m)$ has a density $\widehat{T}(y)$ with respect to Lebesgue measure, then, by (F3), we see

$$I_F(T) = \int_{\mathbb{R}^m} F(y, \widehat{T}(y))d\mathcal{L}^m(y),$$

where $\mathcal{L}^m$ denotes the Lebesgue measure. We thus observe that our definition of $I_F(T)$ is consistent with the case when T is an $\mathbb{R}^m$-valued function.

Our goal is to compare two minima

$$d_F(a,b) := \inf\Big\{\ell_F(\gamma);\ \gamma:\ [0,1] \to \mathbb{R}^m \text{ is Lipschitz and } \gamma(0)=a,\ \gamma(1)=b\Big\} \tag{3.3}$$

$$\tilde{d}_F(a,b) := \inf\Big\{I_F(T);\ T \in \mathcal{M}_1,\ \partial T = \delta_b - \delta_a\Big\} \tag{3.4}$$

for given two points $a,\ b \in \mathbb{R}^m$. The function $d_F(a,b)$ is a distance function from a to b with respect to the metric density F if F satisfies the evenness $F(y,\eta) = F(y,-\eta)$ in addition to (F1)–(F4). For this kind of F it is also possible to prove that $\tilde{d}_F$ is a distance function.

For a given oriented Lipschitz curve Γ from a to b we see

$$[\![\Gamma]\!] \in \mathcal{M}_1 \quad \text{with} \quad \partial[\![\Gamma]\!] = \delta_b - \delta_a$$

as in (2.1)–(2.2). Comparing (3.1) with (3.2) we see

$$I_F\Big([\![\Gamma]\!]\Big) = \ell_F(\gamma)$$

since $\mu_{[\![\Gamma]\!]} = \mathcal{H}^1 \lfloor \Gamma$ and $\xi_{[\![\Gamma]\!]}$ is the unit tangent vector to Γ (consistent with the orientation of Γ) as observed in §2. We now conclude from (3.3) and (3.4) that

$$d_F(a,b) \geqslant \tilde{d}_F(a,b).$$

We shall prove the converse. We say $S \in \mathcal{M}_1$ is a *minimal current* from a to b if

$$\tilde{d}_F(a,b) = I_F(S).$$

In other words S is a minimizer of $I_F(T)$ with $\partial T = \delta_b - \delta_a$. The minimizer of d_F may be called a geodesic from a to b.

THEOREM 3.1. *Suppose that f satisfies (F1)–(F4). There exists a current from a to b representing a simple Lipschitz curve from a to b which is a minimal current from a to b if $a \neq b$.In particular*

$$d_F(a,b) = \tilde{d}_F(a,b) \quad \textit{for all} \quad a,\ b \in \mathbb{R}^m.$$

REMARK 3.2: It is interesting to compare our Theorem 3.1 with Federer's result in [F2, 5.12]. His result may be interpreted as a localized version of our Theorem 2.1 assuming in addition the evenness of $F(y,\eta)$ in η, so our result is not included in his result. Since the statement in [F2, 5.12] is written by using terminology from geometric measure theory, we reproduce his result (with our boundary constraint) for the reader's convenience.

'Suppose that F satisfies (F1)–(F3) and the localized version of (F4). Suppose that F is even in η, i.e.,

$$F(y,\eta) = F(y,-\eta).$$

Let A be a compact set in $\mathbb{R}^m$ and $a,\ b \in A$. Then

$$\begin{aligned}&\inf\left\{I_F(T);\ \partial T = \delta_b - \delta_a,\ \ \operatorname{spt} T \subset A\right\}\\ =&\inf\left\{\ell_F(\gamma);\ \gamma:\ [0,1] \to A \text{ is Lipschitz and } \gamma(0) = a,\ \gamma(1) = b\right\}.\text{'}\end{aligned}$$

In [F2, 5.12] he used integral currents instead of Lipschitz curve. But by the structure theory of 1-integral currents [F1, 4.2.25] one observes that this statement is equivalent to [F2, 5.12] with $U = \mathbb{R}^m$, $\partial T = \delta_b - \delta_a$, $B = \emptyset$ and $\Psi = F$.

Our proof of Theorem 3.1 is different from that of [F2, 5.12] and we think our proof is very elementary compared with [F2, 5.12] where it is studied along the line of a general framework. Let us sketch the idea of our proof. We take a minimizing sequence of $\widetilde{d}_F$ and approximate it by polygonal chains as in Theorem 2.4. When T in (3.4) is restricted to *real* polygonal chains with a fixed support, the minimizing problem is reduced to a linear programming problem called a network flow problem. It is well known such a minimum is attained at multiplicity one polygonal chain. This leads to $\widetilde{d}_F \geqslant d_F$ and we conclude $d_F = \widetilde{d}_F$.

The remaining part of this section is devoted to the proof of Theorem 3.1. We recall a standard result from network flow problem; see e.g. [I]. Let $\mathcal{N}$ be a finite subset in $\mathbb{R}^m$ and $\mathcal{A}$ be a finite subset of

$$\mathcal{N} \times \mathcal{N} \setminus \{(x,x);\ x \in \mathbb{R}^m\}.$$

For $i \in \mathcal{N}$ we write

$$A(i) = \{j \in \mathcal{N};\ (i,j) \in \mathcal{A}\}$$

$$B(i) = \{j \in \mathcal{N};\ (j,i) \in \mathcal{A}\}$$

The set $\mathcal{N}$ is called vorteces of the *network* $\mathcal{A}$. Let d_{ij} for $(i,j) \in \mathcal{A}$ be a given positive real number. For

$$x = \{x_{ij} \in \mathbb{R};\ (i,j) \in \mathcal{A}\}$$

we consider a function

$$I(x) = \sum_{(i,j)\in\mathcal{A}} d_{ij} x_{ij}.$$

The vector x is called a *flow* on the network $\mathcal{A}$.

LEMMA 3.3. *Let a, b be two given different points in $\mathcal{N}$. Suppose that there is $x = \{x_{ij}\}$ satisfying*

$$\sum_{j\in B(i)} x_{ji} - \sum_{j\in A(i)} x_{ij} = \begin{cases} -1, & i = a \\ 0, & i \neq a \text{ and } i \neq b \\ 1, & i = b \end{cases} \tag{3.5}$$

$$x_{ij} \geqslant 0 \quad \textit{for} \quad (i,j) \in \mathcal{A}. \tag{3.6}$$

Then the minimum of $I(x)$ under (3.5) (3.6) is attained at x of the form

$$x_{ij} = 1 \quad \text{or} \quad 0 \quad \textit{for} \quad (i,j) \in \mathcal{A}.$$

The condition (3.5) implies the flow x begins from a and ends at b. The minimizing problem for $I(x)$ under the constraint (3.5), (3.6) is called a *minimal flow* problem. Lemma 3.3 is known as a variant of integrity of solutions in the theory of linear programming, see e.g. [I]. We reformulate this lemma in terminology of real polygonal chains.

LEMMA 3.4. *Suppose that F satisfies (F1), (F3), (F4). For a, $b \in \mathbb{R}^m$ with $a \neq b$ let $P_0 \in \mathbf{P}_1$ be given such that $\partial P_0 = \delta_b - \delta_a$. Let K be the support of P_0. There is a multiplicity one current $S \in \mathbf{P}_1$ minimizing $I_F(P)$ under the constraints*

$$P \in \mathbf{P}_1, \quad \operatorname{spt} P \subset K \quad \textit{with} \quad \partial P = \delta_b - \delta_a. \tag{3.7}$$

PROOF: Let $P_0 \in \mathbf{P}_1$ denote

$$P_0 = \sum_{l=1}^{r} c_\ell [\![a_\ell, b_\ell]\!] \quad \text{with} \quad c_\ell > 0, \quad a_\ell \neq b_\ell \ (1 \leqslant \ell \leqslant r).$$

We set

$$\mathcal{A} = \{(a_\ell, b_\ell);\ 1 \leqslant \ell \leqslant r\} \cup \{(b_\ell, a_\ell);\ 1 \leqslant \ell \leqslant r\}$$

and $\mathcal{N} = \{a_\ell;\ 1 \leqslant \ell \leqslant r\} \cup \{b_\ell;\ 1 \leqslant \ell \leqslant r\}$. We identify $P \in \mathbf{P}_1$ with a flow $x = \{x_{ij}\}$ on $\mathcal{A}$ satisfying (3.5), (3,6) by

$$P = \sum_{(i,j)\in\mathcal{A}} x_{ij}[\![i,j]\!]. \tag{3.8}$$

By (3.5) we see $\partial P = \delta_b - \delta_a$. The correspondence $x \mapsto P$ gives a mapping

$$\sigma :\ A \to B,$$

where

$$A = \{x = \{x_{ij}\};\ x_{ij} \text{ satisfies } (3.5), (3.6)\}$$
$$B = \{P \in \mathbf{P}_1;\ P \text{ satisfies } (3.7)\}.$$

One easily observe that σ is a surjection (but not a injection because $[\![i,j]\!] = 2[\![i,j]\!] + [\![j,i]\!]$). We consider the subset of A:

$$A_0 = \{x = \{x_{ij}\} \in A;\ \text{for } (i,j) \in A,\ \text{either } x_{ij} = 0 \text{ or } x_{ji} = 0\}.$$

Then $\sigma :\ A_0 \to B$ is now bijection and for P expressed as (3.8) with $x \in A_0$, i.e. $P = \sigma(x)$ we have

$$\begin{aligned} I_F(P) &= \sum_{(i,j)\in\mathcal{A}} \int_{L_{ij}} F(y, x_{ij}e_{ij})d\mathcal{H}^1 \\ &= \sum_{(i,j)\in\mathcal{A}} d_{ij}x_{ij} = I(x), \quad d_{ij} = \int_{L_{ij}} F(y, e_{ij})d\mathcal{H}^1, \end{aligned} \tag{3.9}$$

where L_{ij} denotes the line segment $[i,j]$ oriented by a unit tangent vector e_{ij}.

Since the existence of P_0 guarantees that A is not empty and since (F4) in particular implies $d_{ij} > 0$, we now apply Lemma 3.3 and conclude that $\inf_A I$ is attained at $\overline{x} \in A$ such that

$$\overline{x}_{ij} = 1 \quad \text{or} \quad 0 \quad \text{for} \quad (i,j) \in A. \tag{3.10}$$

By $d_{ij} > 0$ we observe that $\overline{x} \in A_0$ since otherwise it contradicts the minimality of $\overline{x}$. We set

$$S = \sigma(x) \in B$$

and observe that

$$I_F(S) = I(\overline{x}) \leqslant \inf_A I \leqslant \inf_{A_0} I = \inf_B I_F$$

since $\sigma : A_0 \to B$ is bijection and $I_F(P) = I(x),\ P = \sigma(x)$ on A_0 by (3.9). The current S is multiplicity one by (3.10) so S fulfills all desired properties in Lemma 3.4. ∎

PROOF OF THEOREM 3.1: Since $d_F(a,a) = 0 = \widetilde{d}_F(a,a)$, we may assume $a \neq b$. Let $\{T_j\}_{j=1}^{\infty}$ be a minimizing sequence of $\widetilde{d}_F$ in (3.4). Since $I_F([\![a,b]\!])$ is finite by (F4), we see $\widetilde{d}_F$ is finite. The condition (F4) guarantees that $\mathbf{M}(T_j)$ is finite. Let $\{Q_{j\varepsilon}\}_{\varepsilon>0}$ be an approximate sequence of polygonal chains defined in Theorem 2.4 with T replaced by T_j. Let K denote the support of $Q_{j\varepsilon}$. From Lemma 3.4 it follows that there is a multiplicity one current $S_{j\varepsilon} \in \mathbf{P}_1$ satisfying (3.7) with $P = S_{j\varepsilon}$ which attains the minimum of $I_F(P)$ under (3.7), i.e.,

$$I_F(S_{j\varepsilon}) = \inf\{I_F(P);\ P \in \mathbf{P}_1,\ \operatorname{spt} P \subset K,\ \partial P = \delta_b - \delta_a\}. \tag{3.10}$$

By (F1), (F2), (F3) one applies Reshetnyak's continuity theorem [Re] to get

$$I_F(T_{j\varepsilon}) \to I_F(T_j) \quad \text{as} \quad \varepsilon \downarrow 0 \tag{3.11}$$

since $\mathbf{M}(T_{j\varepsilon}) \to \mathbf{M}(T_j)$ by (2.5), (2.7). By (3.10) and (3.11) we observe that there is $\varepsilon_j \to 0$ such that $\{S_{j\varepsilon_j}\}_{j=1}^{\infty}$ is a minimizing sequence of $\widetilde{d}_F$ in (3.4). Since $S_{j\varepsilon}$ may be regarded as a Lipschitz curve from a to b we have

$$\widetilde{d}_F(a,b) \geqslant d_F(a,b).$$

The converse inequality is trivial so we have proved $\widetilde{d}_F = d_F$ in (3.3)–(3.4).

It remains to prove that the value $d_F(a,b)$ is attained at some simple Lipschitz curve from a to b. This is nothing but the following standard lemma. ∎

LEMMA 3.5. *Suppose that F satisfies (F1)–(F4). For $a \neq b$ there is a simple Lipschitz curve γ from a to b such that*

$$d_F(a,b) = \ell_F(\gamma). \tag{3.12}$$

PROOF: The proof is standard but we give it here for the reader's convenience. For a Lipschitz curve $\gamma: [0,1] \to \mathbb{R}^m$, we introduce the arc length parameter

$$s = \int_0^t F(\gamma(t), \dot\gamma(t))dt.$$

If t is the arc length parameter, clearly

$$1 = F(\gamma(s), \dot\gamma(s)).$$

From (F4) it follows that $d_F(a,b) = d > 0$. Replace F by F/d we may assume $d = 1$. Suppose that γ_j is a minimizing sequence of $d_F(a,b)$ defined in(3.3), say,

$$\int_0^1 F(\gamma_j, \dot\gamma_j)dt = 1 + \varepsilon_j \quad \text{with} \quad \varepsilon_j \downarrow 0.$$

Let s denote the arc length parameter of γ_j. We introduce a parameter t, $0 \leqslant t \leqslant 1$ such that

$$t = s/(1+\varepsilon_j), \qquad 0 \leqslant s \leqslant 1 + \varepsilon_j.$$

We observe that for $p > 1$

$$\begin{aligned}\int_0^1 F(\gamma_j(t), \dot\gamma_j(t))^p dt &= (1+\varepsilon_j)^p \int_0^{1+\varepsilon_j} F(\gamma_j, d\gamma_j/ds)^p ds \\ &= (1+\varepsilon_j)^{p+1},\end{aligned}$$

where we have used (F3) and (3.13). By coerciveness, that is (F4), we see

$$\left(\int_0^1 |\dot\gamma_j(t)|^p dt\right)^{1/p} \leqslant \frac{1}{k}(1+\varepsilon_j)^{1+1/p}$$

which yields

$$\|\dot\gamma_j\|_{L^\infty(0,1)} \leqslant \frac{1+\varepsilon_j}{k}$$

by sending $p \to \infty$. Since $\gamma_j(0) = a$, $\gamma_j(1) = b$ is fixed, this implies

$$\sup_j \|\dot\gamma_j\|_{L^\infty(0,1)} < \infty.$$

By Ascoli-Arzela's theorem we find a subsequence (still denoted by γ_j) and a Lipschitz function γ such that

$$\gamma_j \to \gamma \quad \text{uniformly in} \quad [0,1]. \tag{3.14}$$

The convexity (F2) guarantees the lower semicontinuity of the integral $\ell_F(\gamma)$ under (3.14). We thus have

$$\ell_F(\gamma) \leqslant \liminf_{j\to\infty} \ell_F(\gamma_j) = \lim_{j\to\infty}(1+\varepsilon_j) = 1.$$

Since $\gamma(0) = a$ and $\gamma(1) = b$, this yields (3.12). ∎

4. Representation of the relaxation

Our main goal is to obtain a general representation of $\overline{\mathcal{F}}$ including (1.3).

We consider a functional $\mathcal{F}$ of C^1 mapping u: $\Omega \to \mathbb{R}^m$

$$\mathcal{F}(u) = \int_\Omega f(x, u, \nabla u(x))dx,$$

where Ω is an open set in $\mathbb{R}^n$. Here $\nabla u(x)$ denotes the Jacobi matrix at x and is identified with an element of $\mathbb{R}^{nm}$. The energy density f is assumed to satisfy the conditions listed below.

(f 1) $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm}$ is continuous and nonnegative.

(f 2) (Convexity) $f(x,y,\xi)$ is convex in $\xi \in \mathbb{R}^{nm}$ for all $(x,y) \in \Omega \times \mathbb{R}^m$.

(f 3) (Linear growth) $0 \le f(x,y,\xi) \le K(1+|\xi|)$ with K independent of $(x,y,\xi) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm}$.

(f 4) (Coerciveness) $f(x,y,\xi) \ge k|\xi|$ with $k > 0$ independent of $(x,y,\xi) \in \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm}$.

(f 5) For every $(x_0, y_0) \in \Omega \times \mathbb{R}^m$ and $\varepsilon > 0$, there is a positive constant δ such that $|x - x_0| < \delta, |y - y_0| < \delta$ implies

$$|f(x,y,\xi) - f(x_0,y_0,\xi)| \le \varepsilon(1+|\xi|) \text{ for all } \xi \in \mathbb{R}^{nm}.$$

(f 6) (Isotropy condition)

$$f(x,y,\xi) \ge f(x,y,\xi(\nu\otimes\nu)) \text{ for all } \nu \in \mathbb{R}^n, \text{ with } |\nu| = 1,$$

where ξ is identified with an $m \times n$ matrix.

Our goal is to derive an explicit representation of $L^1_{\rm loc}$– lower semicontinuous relaxation $\overline{\mathcal{F}}$ of $\mathcal{F}$ on $BV(\Omega, \mathbb{R}^m)$:

$$\overline{\mathcal{F}}(u) = \inf\left\{ \liminf_{\ell\to\infty} \mathcal{F}(u_\ell);\ u_\ell \in C^1(\Omega, \mathbb{R}^m),\ u_\ell \to u \ \text{ in }\ L^1_{\rm loc}(\Omega, \mathbb{R}^m)\right\}$$

for $u \in BV(\Omega, \mathbb{R}^m)$. For $u \in BV(\Omega, \mathbb{R}^m)$ it is well-known [F1, Giu, Si] that ∇u is a matrix Radon measure decomposed as

$$\nabla u = \nabla u \lfloor \Omega_0 + \nabla u \lfloor (\Omega - \Omega_0 - \Sigma) + (u^+ - u^-) \otimes \nu \mathcal{H}^{n-1} \lfloor \Sigma.$$

Here Σ denotes the set of approximate jump discontinuities of u and ν represents a unit normal to Σ. The functions $u^\pm$ are the trace of u on Σ defined by

$$\lim_{\varepsilon\downarrow 0} u(x \pm \varepsilon\nu(x)).$$

By $\mu\lfloor A$ we mean a measure on Ω defined by

$$(\mu\lfloor A)(B) = \mu(A \cap B) \text{ for } B \subset \Omega,$$

where μ is a measure. For $a, b \in \mathbb{R}^m$ and $q \in \mathbb{R}^n$ we introduce a distance like function:

$$\begin{aligned} (4.1) \qquad D_x(a, b, q) =& \inf\{\int_0^1 f_\infty(x, \gamma(t), \dot\gamma(t) \otimes q)dt; \\ & \gamma : [0, 1] \to \mathbb{R}^m \text{ is Lipschitz and } \gamma(0) = a, \gamma(1) = b\}, \end{aligned}$$

where f_∞ is the recession function defined by

$$(4.2) \qquad f_\infty(x, y, \xi) = \lim_{t\downarrow 0} f(x, y, \xi/t)t.$$

By $|\mu|$ we mean the total variation measure of μ and $d\mu/d|\mu|$ denotes the Radon – Nikodym derivative. A combination of Theorem 3.1 and results in [AG1] yields an explicit estimate of $\overline{\mathcal{F}}$ from below under assumptions (f 1) – (f 6). Combining the other side estimate of $\overline{\mathcal{F}}$ by [AMT] we find a representation of $\overline{\mathcal{F}}$.

THEOREM 4.1. *Assume that f satisfies (f 1) – (f 6). For $u \in BV(\Omega, \mathbb{R}^m)$*

$$\begin{aligned} (4.3) \qquad \overline{\mathcal{F}}(u) =& \int_{\Omega_0} f(x), u(x), \nabla u(x) d\mathcal{L}^n(x) \\ &+ \int_{\Omega\setminus\Omega_0\setminus\Sigma} f_\infty(x, u(x), \frac{d\nabla u}{d|\nabla u|}(x))|\nabla u| \\ &+ \int_\Sigma D_x(u^-(x), u^+(x), \nu(x)) d\mathcal{H}^{n-1}(x). \end{aligned}$$

REMARK 4.2: The representation (4.3) is a natural extension of Dal Maso's results [DM] to the case $m > 1$. As pointed out in [AMT], (4.3) may not hold for general f without assuming the isotropy condition (f 6) although (f 6) is unnecessary for $m = 1$ (cf. [DM]).

REMARK 4.3: If f is positively homogeneous of degree one in ξ, (f 1), (f 2), (1.4), (1.5) implies (f 1) – (f 6). Since (4.3) yields (1.3) if f is positively homogeneous of degree one in ξ, Theorem 4.1 deduces (1.3).

PROOF: (i) (Estimate of $\bar{\mathcal{F}}$ from below). We first recall an estimate of $\overline{\mathcal{F}}$ from below given in [AG1]. Let $\widetilde{D}_x$ denote

$$\widetilde{D}_x(a,b,q) = \inf\{\int_{\mathbb{R}^m} f_\infty(x,y,(S_i^j));\ \partial S_i = \nu_i(\delta_b - \delta_a),\ S_i \in \mathcal{M}_1,\ 1 \le i \le n\}$$

where $a, b \in \mathbb{R}^m$, $q \in \mathbb{R}^n$ and f_∞ is defined by (4.2). The main results in [AG1, Theorems 5.1 and 8.1] yields

$$\begin{aligned}\overline{\mathcal{F}}(u) \ge \int_{\Omega_0} f(x,u,\nabla u)d\mathcal{L}^n(x) + \int_{\Omega\setminus\Omega_0\setminus\Sigma} f_\infty(x,u,\frac{d\nabla u}{d|\nabla u|})|\nabla u| \\ + \int_\Sigma \widetilde{D}_x(u^-(x),u^+(x),\nu)d\mathcal{H}^n(x).\end{aligned} \tag{4.4}$$

So far we have only invoked assumptions (f 1) – (f 3), (f 5). If f satisfies (f 6), so does f_∞. Applying (f 6) we observe

$$f_\infty(x,y,(S_i^j)) \ge f_\infty(x,y,(\sum_{k=1}^n S_k^j \nu_k)\nu_i)$$

which yields

$$\widetilde{D}_x(a,b,\nu) \ge \widetilde{d}_F(a,b) \tag{4.5}$$

with

$$F(y,\eta) = f_\infty(x,y,\eta \otimes \nu(x)) \tag{4.6}$$

where $\widetilde{d}_F$ is defined by (3.4), since $\partial S_i = \nu_i(\delta_b - \delta_a)$ implies

$$\partial T = \delta_b - \delta_a \quad \text{with} \quad T = \sum_{k=1}^n S_k \nu_k.$$

By (f 1) – (f 5) our F in (4.6) satisfies all assumptions in Theorem 3.1 which yields

$$\tilde{d}_F(a,b) \geq D_x(a,b,\nu), \tag{4.7}$$

where D_x is defined in (4.1). Combining (4.4), (4.5) and (4.7) we have shown that $\overline{\mathcal{F}}$ is estimated from below by the right hand side (RHS) of (4.3).

(ii) (Estimate of $\overline{\mathcal{F}}$ from above). If $f = f(x,y,\xi)$ is positively homogeneous of degree one in ξ, Theorem 4.9 in [AMT] implies that $\overline{\mathcal{F}}$ is bounded from above by an integral which is less than or equal to the RHS of (4.3). Note that their integral may be strictly less than the RHS of (4.3) if we do not assume the isotropy condition (f 6).

Although we do not assume homogeneuity of f in ξ the method in [AMT] works to get the estimate of $\overline{\mathcal{F}}$ from above by the RHS of (4.3). Let us briefly sketch how to apply their proof without assuming the homogeneuity of f.

Let $\hat{f}$ be the homogenization of f defined by

$$\hat{f}(x,y,\xi_0,\xi) = \begin{cases} f_\infty(x,y,\xi) & \text{if} \quad \xi_0 = 0 \\ f(x,y,\xi/\xi_0)\xi_0 & \quad\quad \xi_0 > 0 \end{cases}$$

By (f 1) – (f 3), (f 5) we see $\hat{f}$ is continuous in $\Omega \times \mathbb{R}^m \times [0,\infty) \times \mathbb{R}^{nm}$ convex in ξ, positively homogeneous of degree one in (ξ_0, ξ) and satisfies a growth estimate

$$0 \leq \hat{f}(x,y,\xi_0,\xi) \leq C(1 + |\xi_0| + |\xi|)$$

(ef. [DM]). For $u \in C^1(\Omega, \mathbb{R}^m)$ we set

$$\tilde{u} : I \times \Omega \to \mathbb{R}^m, \quad (s,x) \mapsto (s, u(x)),$$

where I is a unit interval. Clearly

$$\int_\Omega f(x,u,\nabla u) d\mathcal{L}^n = \int_{I\times\Omega} \hat{f}(x,u,\tilde{\nabla}\tilde{u}) d\mathcal{L}^{n+1}, \quad \tilde{\nabla} = (\partial_s, \nabla) \tag{4.8}$$

As in [AMT] for $u \in BV(\Omega, \mathbb{R}^m)$ we set

$$\overline{\mathcal{F}}(u,A) = \inf\{\lim \int_A f(x,u_h,\nabla u_h); \quad u_h \in C^1(A,\mathbb{R}^m); u_h \to u \text{ in } L^1_{\text{loc}}(A,\mathbb{R}^m)\}$$

where A is an open set in Ω. As in [AMT, Theorem 4.3] one can prove that $A \mapsto \overline{\mathcal{F}}(u, A)$ is regarded as a Borel regular measure in Ω.

Noting the relation (4.8) one next proves

$$\overline{\mathcal{F}}(u, A\backslash\Sigma) \leq \int_{I\times(A\backslash\Sigma)} \hat{f}(x, u, \tilde{\nabla}\tilde{u})d\mathcal{L}^{n+1} \tag{4.9}$$

$$\overline{\mathcal{F}}(u, A\cap\Sigma) \leq \int_{\Sigma} D_x(u^-, u^+, \nu)d\mathcal{H}^{n-1} \tag{4.10}$$

at least for $u \in BV(\Omega, \mathbb{R}^m) \cap L^\infty(\Omega, \mathbb{R}^m)$. The estimate (4.9) corresponds the estimate outside the jump discontinuity Σ of u and one may approximate u by a standard mollifier technique. To get (4.10) one approximate Σ by polyhedra and reduce the problem that u has a two valued function whose discontinuity is a C^1 hypersurface. These proofs parallel Proposition 4.6 – Proposition 4.8 in [AMT]. The extra assumption $u \in L^\infty(\Omega, \mathbb{R}^m)$ can be removed as in Theorem 3.9 in [AMT]. We note that in [AMT] they derive a sharper estimate than (4.10).

Our desired estimate is a combination of (4.9) and (4.10) since

$$\begin{aligned}&\int_{\Omega\backslash\Sigma} \hat{f}(x, u(x), 1, \nabla u)\\ &= \int_{\Omega_0} f(x, u(x), \nabla u(x))d\mathcal{L}^n(x) + \int_{\Omega\backslash\Omega_0\backslash\Sigma} f_\infty(x, u(x), \frac{d\nabla u}{d|\nabla u|}(x))|\nabla u|;\end{aligned}$$

see e.g. [GMS1].

Acknowledgements. Most of the work of the first author was done while he visited the Mathematics Department of Hokkaido University during 1989, supported by a grant from the Japan Association of Mathematical Sciences. Additional support from NSF grant no. 89–01695. The work of the second author was completed while he visited the Institute for Mathematics and its Applications, University of Minnesota during 1990. Its hospitality is gratefully acknowledged, as is support from the Japanese Ministry of Education, Science, and Culture through grant no. 01740076 and 01540092 for scientific research.

Appendix

We use some standard notation in [F1]. Let $\mathbf{P}_k = \mathbf{P}_k(\mathbb{R}^m)$ denote the space of real polygonal k-chain in $\mathbb{R}^m$. Let $\mathbf{N}_k = \mathbf{N}_k(\mathbb{R}^m)$ denote the space of k-normal current (with compact support in $\mathbb{R}^m$), i.e.,

$$\mathbf{N}_k = \{T \in \mathcal{D}_k;\ \mathbf{M}(T) + \mathbf{M}(\partial T) < \infty, \quad \operatorname{spt} T \text{ is compact}\},$$

when $\mathbf{M}$ denotes the mass in [F1]. We first prove Lemma 2.2 (and its generalization to higher dimensional currents).

LEMMA A. *Suppose that $T_0 \in \mathcal{D}_k$ with $\partial T_0 = 0$ has a finite total mass $\mathbf{M}(T_0)$. For every $\varepsilon > 0$ there is $T_1 \in \mathbf{N}_k$ with $\partial T_1 = 0$ such that T_1 is compactly supported in $\mathbb{R}^m$ and that $\mathbf{M}(T_0 - T_1) \leqslant \varepsilon$.*

PROOF: According to the deformation theorem ([F1, 4.2.9] or [Si, §29]) there is $P \in \mathbf{P}_k$ and $S \in \mathcal{D}_{k+1}$ with $\mathbf{M}(S) < \infty$ such that

$$T_0 = P + \partial S \tag{1}$$

since $\partial T_0 = 0$ and $\mathbf{M}(T_0) < \infty$. Let φ_R be a cut-off function of $B_{3R/2}$ supported in B_{2R} as is defined in the proof of Lemma 2.2 in §2. Take R large so that $\operatorname{spt} P \subset B_R$ and set

$$T_R = P + \partial(S \lfloor \varphi_R),$$

where

$$(S \lfloor \varphi_R)(\omega) = S(\varphi_R \wedge \omega) = S(\varphi_R \omega), \quad \omega \in \mathcal{D}^{k+1}.$$

Applying Leibniz' rule yields

$$T_0 - T_R = \partial(S \lfloor (1 - \varphi_R)) = \partial S \lfloor (1 - \varphi_R) + (-1)^{k+1} S \lfloor d\varphi_R.$$

Since $\operatorname{spt} P \subset B_R$, it follows from (1) that

$$\partial S \lfloor (1 - \varphi_R) = T_0 \lfloor (1 - \varphi_R)$$

which now yields

$$T_0 - T_R = T_0 \lfloor (1 - \varphi_R) + (-1)^{k+1} S \lfloor d\varphi_R.$$

Since both $\mathbf{M}(T_0)$ and $\mathbf{M}(S)$ are finite, we observe that

$$\begin{aligned}\mathbf{M}(T_0 \lfloor (1-\varphi_R)) &\leqslant \mu_{T_0}(\mathbb{R}^m \setminus B_R) \to 0 \\ \mathbf{M}(S \lfloor d\varphi_R) &\leqslant \sup |\nabla \varphi_R| \mu_S(\mathbb{R}^m \setminus B_R) \to 0\end{aligned}$$

as $R \to \infty$, where μ_S denotes the total variation measure $\|S\|$ in [F1]. This shows

$$\mathbf{M}(T_0 - T_R) \to a \quad \text{as} \quad R \to 0.$$

From (1) it follows that $\partial P = 0$ which implies $\partial T_R = 0$. We now obtain a desired approximation T_R of T_0. ∎

We next prove Lemma 2.3 (and its generalization to higher dimensional currents). We thank Robert Hardt who let us know the following lemma with a proof.

LEMMA B. *Suppose that $T \in \mathbf{N}_k$ with $\partial T = 0$. For every $\varepsilon > 0$ there are $P \in \mathbf{P}_k$ and $S \in \mathbf{N}_{k+1}$ such that*

$$T = P + \partial S$$

with $\mathbf{M}(P) \leqslant \mathbf{M}(T) + \varepsilon$ and $\mathbf{M}(S) \leqslant \varepsilon$.

PROOF: We may assume $\varepsilon < 1$. By the strong approximation theorem [F1, 4.2.24, 4.1.24] T is expressed as

$$T = P_1 + R_1 + \partial S_1$$

with some $P_1 \in \mathbf{P}_k$, $R_1 \in \mathcal{D}_k$, $S_1 \in \mathcal{D}_{k+1}$ such that

$$\begin{aligned}\mathbf{M}(P_1) + \mathbf{M}(\partial P_1) &\leqslant \mathbf{M}(T) + \frac{\varepsilon}{2} \\ \mathbf{M}(R_1) + \mathbf{M}(S_1) &\leqslant \frac{\varepsilon}{4\gamma}, \quad \gamma = 2m^{2k+2}.\end{aligned}$$

Applying the deformation theorem [F1, 4.2.9] to R with ε replaced by $\varepsilon/(4\gamma(\mathbf{M}(T)+1))$ we find that R_1 is of the form

$$R_1 = P_2 + Q + \partial S_2$$

with $P_2 \in \mathbf{P}_k$, $Q \in \mathcal{D}_k$, $S_2 \in \mathcal{D}_{k+1}$ such that

$$\mathbf{M}(P_2) \leqslant \gamma\mathbf{M}(R_1) + \frac{\gamma\varepsilon}{4\gamma(\mathbf{M}(T)+1)} \leqslant \frac{\varepsilon}{4}$$
$$\mathbf{M}(Q) \leqslant \frac{\varepsilon\gamma\mathbf{M}(\partial R_1)}{4\gamma(\mathbf{M}(T)+1)} = \frac{\varepsilon\mathbf{M}(\partial P_1)}{4(\mathbf{M}(T)+1)} \leqslant \frac{\varepsilon}{4} \cdot \frac{(\mathbf{M}(T)+\varepsilon/2)}{\mathbf{M}(T)+1} < \frac{\varepsilon}{4}$$
$$\mathbf{M}(S_2) \leqslant \varepsilon\mathbf{M}(R_1) \leqslant \frac{\varepsilon^2}{4\gamma} < \frac{\varepsilon}{2}.$$

Here we have applied $\partial R_1 = \partial P_1 \in \mathbf{P}_k$, and by this property the construction of Q in the deformation theory guarantees that $Q \in \mathbf{P}_k$. We set

$$P = P_1 + P_2 + Q, \quad S = S_1 + S_2$$

and observe that

$$T = P + \partial S$$

with $P \in \mathbf{P}_k$,

$$\mathbf{M}(P) \leqslant \mathbf{M}(T) + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4}$$
$$\mathbf{M}(S) \leqslant \frac{\varepsilon}{4\gamma} + \frac{\varepsilon}{2} < \varepsilon. \blacksquare$$

References

[AMT]. L.Ambrosio, S.Mortola and V.M.Tortorelli, *Functionals with linear growth defined on vector valued BV functions*, preprint.

[AG1]. P.Aviles and Y.Giga, *Variational integrals on mappings of bounded variation and their lower semicontinuity*, (to appear), Arch. Rational Mech. Anal..

[AG2]. ——, *Minimal currents and relaxation of variational integrals on mappings of bounded variation*, Proc. Japan Acad. Ser.A. **66** (1990), 68–71.

[Ba]. S.Baldo, *Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids*, Ann. Inst. Henri Poincaré, Analyse non linéaire **7** (1990), 67–90.

[BBC]. F. Bethuel, H. Brezis and J.M. Coron, *Relaxed energy for harmonic maps*, In 'Variational Methods', Birkhauser (1990), pp.37–52 (eds. H. Berestycki, J.M. Coron, J. Ekeland).

[Bo]. M.E.Bogovski, *Solution of the first boundary value problem for the equation of continuity of an incompressible medium*, Soviet Math. Dokl. **20** (1979), 1094–1098.

[BS]. W.Bochers and H.Sohr, *On the equations* $\operatorname{rot} v = g$ *and* $\operatorname{div} u = f$ *with zero boundary conditions*, Hokkaido Math. J. **19** (1990), 67–87.

[D]. B.Dacorogna, *Quasiconvexity and relaxation of nonconvex problems in the calculus of variations*, J. Func. Anal. **46** (1982), 102–118.

[DM]. G.Dal Maso, *Integral representation of* $BV(\Omega)$ *of* Γ*-limits of variational integrals*, Manuscripta Math. **30** (1980), 387–417.

[DG]. E.De Giorgi, *G-operators and* Γ*-convergence*, Proceeding of ICM Poland, August 16–24, 1983, *Warszawa* **2** (1983), 1175-1191.

[F1]. H. Federer, "Geometric Measure Theory," Springer, Berlin–Heidelberg–New York, 1969.

[F2]. ——, *Real flat chains, cochains and variational problems*, Indiana Univ. Math. J. **24** (1974), 351–407.

[FR]. I.Fonseca and P.Rybka, *Relaxation of multiple integrals in the space* $BV(\Omega; \mathbb{R}^p)$, preprint.

[FT]. I.Fonseca and L.Tartar, *The gradient theory of phase transitions for systems with two potential wells*, Proc. Roy. Soc. Edinburgh **111A** (1989), 89–102.

[GMS1]. M.Giaquinta, G.Modica and J.Souček, *Functionals with linear growth in the calculus of variation I*, Comment. Math. Univ. Carolinae **20** (1979), 143–156.

[GMS2]. ——, *The Dirichlet energy of mappings with values into the sphere*, Manuscripta Math. **65** (1989), 489–507.

[Giu]. E.Giusti, "Minimal Surfaces and Functions of Bounded Variation," Birkhauser, Boston–Basel–Stuttgart, 1984.

[GS]. C.Goffman and J.Serrin, *Sublinear functions of measures and variational integrals*, Duke Math. J. **31** (1964), 159–178.

[I]. M.Iri, "Network flow, transportation and scheduling," Academic Press, New York–

London, 1969.

[JK]. R.James and D.Kinderlehrer, *Theory of diffusionless phase transitions*, In 'PDEs and continuum models of phase transitions', Lecture Note in Physics **344** (1989), Springer, pp.51–84 (eds. M.Rascle, D.Serre and M.Slemrod).

[M1]. L.Modica, *The gradient theory of phase transitions and the minimal surface criterion*, Arch. Rational Mech. Anal. **98** (1987), 123–142.

[Mor]. Frank Morgan, *Area minimizing currents bounded by higher multiples of curves*, Rend. Circ. Mat. Palermo **33** (1984).

[Re]. Yu.G.Reshetnyak, *Weak convergence of completely additive vector functions on a set*, Siberian Math. J. **9** (1968), 1039–1045 (translation of *Sibirsk. Math. Ž.* 9 (1968), 1386–1394).

[Ro]. R.T.Rockaffelar, *Dual problems of Lagrange for arcs of bounded variation*, Calculus of Variations and Control Theory, Academic Press (1976), 139–192. (ed. D.L. Russel).

[Se1]. J.Serrin, *A new definition of the integral for non-parametric problems in the calculus of variations*, Acta. Math. **102** (1959), 23–32.

[Se2]. ——, *On the definition and properties of certain variational integrals*, Trans. Amer. Math. Soc. **101** (1961), 139–167.

[S]. L.Simon, *Lectures on Geometric Measure Theory*, Proceeding of the Centre for Mathematical Analysis, Australian National University **3**, p. 1983.

[St1]. P.Sternberg, *The effect of a singular perturbation on nonconvex variational problems*, Arch. Rational Mech. Anal. **101** (1988), 209–260.

[St2]. ——, *Vector-valued minimizers of nonconvex variational problems*, Proc. of Workshop on Nonlinear P.D.E.'s, Provo, (1987) eds. P.Bates and P.Fife (to appear).

[W]. B.White, *The least area bounded by multiples of a curve*, Proc. of AMS **90** (1984), 230–232.

[Y]. L.C.Young, *Some extremal questions for simplicial complexes, V. The relative area of a Klain bottle*, Rend. Circ. Mat. Palermo (2) **12** (1963), 257–274.